

Introdução aos Espaços de Escala (EDPs em Processamento de Imagens)

Publicações Matemáticas

**Introdução aos Espaços de Escala
(EDPs em Processamento de Imagens)**

Ralph Teixeira
PUC-RJ

impa



23^o Colóquio Brasileiro de Matemática

Copyright © 2001 by Ralph Teixeira
Direitos reservados, 2001 pela Associação Instituto
Nacional de Matemática Pura e Aplicada - IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger

23^a Colóquio Brasileiro de Matemática

- Aspectos de Modelagem Matemática em Dinâmica dos Fluidos – André Nachbin
- Cálculo Variacional e Controle Ótimo – Antonio Leitão
- Computational Methods in the Local Theory of Curves – Abramo Hefez e Marcelo Escudeiro Hernandes
- Contornos, Conjuntos Convexos e Autômatos Celulares – André Toom
- Inteiros Quadráticos e o Grupo de Classes – Antonio J. Engler e Paulo Brumatti
- Introduction to Toric Varieties – Jean-Paul Brasselet
- **Introdução aos Espaços de Escala (EDPs em Processamento de Imagens) – Ralph Teixeira**
- Índices de Campos Holomorfos e Aplicações – Marcio G. Soares e Rogério S. Mol
- Notes on Morse Theory – Daniel V. Tausk, Francesco Mercuri e Paolo Piccione
- One Dimensional Dynamics: the Mathematical Tools – Edson de Faria e Wellington de Melo
- Partial Regularity of Solutions of the 3-D Incompressible Navier-Stokes Equations – Hermano Frid e Mikhail Perepelitsa
- Perfect Simulation of Spatial Processes – Nancy L. Garcia
- Riemannian and Submanifold Geometry – Carlos Olmos
- Tópicos em Combinatória Contemporânea – Carlos Gustavo Moreira e Yoshiharu Kohayakawa
- Uma Introdução Sucinta a Algoritmos de Aproximação – Cristina G. Fernandes et al

Distribuição:

IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
e-mail: dic@impa.br
<http://www.impa.br>

ISBN: 85-244-0176-1

Conteúdo

Prefácio	vii
Introdução	ix
0.1 Escala e Ciência	ix
0.2 Escala em Computação Gráfica	xi
0.2.1 Por que e como?	xi
0.2.2 Onde?	xiii
0.3 Estrutura do Livro	xiv
0.3.1 Parte I	xiv
0.3.2 Parte II	xv
0.3.3 Parte III	xvi
0.3.4 Parte IV	xvi
0.3.5 Parte V	xvii
I Ferramentas	1
1 Transformadas de Sinais Não-Discretos	3
1.1 Convoluções	3
1.2 Transformada de Fourier	5
1.3 Transformada Bilateral de Laplace	11
1.4 Transformadas Bidimensionais	13
1.5 Exercícios	16
2 Transformadas de Sinais Discretos	19
2.1 Convoluções	19
2.2 Transformada de Fourier de Tempo Discreto (DTFT)	21
2.3 Transformada Z	24
2.4 Transformadas Bidimensionais	28
2.5 Exercícios	30
3 Transformadas de Sinais Finitos	33
3.1 Convoluções Circulares	33
3.2 Transformada de Fourier Discreta (DFT)	34

3.3	Transformadas Bidimensionais	38
3.4	Exercícios	40
4	A Função Gaussiana e a Equação do Calor	43
4.1	A Função Gaussiana	43
4.2	A Equação do Calor	46
4.3	Caso Bidimensional	49
4.4	Exercícios	51
II	Espaço de Escala Gaussiano	53
5	Espaço de Escala Gaussiano	55
5.1	Caso Unidimensional	56
5.2	Caso Bidimensional	59
5.3	Exercícios	62
6	Discretização Espacial	65
6.1	Caso Unidimensional	66
6.2	Caso Bidimensional	69
6.3	Exercícios	72
7	Discretização Temporal	75
7.1	Caso Unidimensional	76
7.1.1	Equação Explícita	76
7.1.2	Equação Implícita	78
7.2	Caso Bidimensional	82
7.2.1	Equação Explícita	83
7.2.2	Equação Implícita	85
7.3	Discretização Somente no Tempo*	88
7.3.1	Equação Explícita	88
7.3.2	Equação Implícita	89
7.4	Exercícios	91
8	Invariantes geométricos	93
8.1	Motivação: detecção de arestas	93
8.2	Invariantes diferenciais	96
8.3	Detecção de detalhes	98
8.3.1	Arestas	98
8.3.2	Junções	100
8.3.3	Arestas por Laplaciano	100
8.4	Exercícios	100

9	Derivadas Discretas	103
9.1	Introdução	103
9.2	Discretização	103
9.2.1	Máscaras para derivadas 1D	105
9.2.2	Máscaras para derivadas 2D	106
9.3	Séries de Taylor e Derivadas Discretas*	107
9.3.1	Caso 1D	107
9.3.2	Caso 2D	112
9.4	Exercícios	121
10	Abordagens Axiomáticas	123
10.1	Via Equações Diferenciais Parciais	123
10.1.1	Caso contínuo	123
10.1.2	Caso Discreto	126
10.2	Via Convoluções	130
10.2.1	Caso contínuo	130
10.2.2	Caso Discreto	132
10.3	Exercícios	133
11	Conclusão da Parte II	135
III	EDPs Não-Lineares	137
12	Suavização Fora das Arestas	139
12.1	O modelo de Perona-Malik	139
12.2	Detecção de Arestas via Funcionais de Energia	146
13	Suavização Morfológica	149
13.1	Movimento por Curvatura	150
13.1.1	Definição e Propriedades	150
13.2	Formas Multi-Escala	152
13.3	Computando Movimentos por Curvatura	154
13.4	A Equação do Calor Intrínseca	155
13.4.1	Filtro Mediano	158
13.5	Espaço de Escala Morfológico Afim	160
13.6	Exercícios	162
IV	Esboço Primitivo	165
14	Esboço Primitivo	167
14.1	Bolhas	167
14.1.1	Árvores de bolhas	172
14.2	Hierarquia multi-escala	172
14.3	Importância de bolhas multi-escala	175
14.3.1	Tempo de vida efetivo de uma bolha multi-escala	176

14.3.2	Volume efetivo de uma bolha	177
14.3.3	Conclusão	178
14.4	Exercícios	178
15	Singularidades*	179
15.1	O Teorema da Função Implícita	179
15.2	Trajétórias de Pontos Críticos	180
15.3	Classificação de singularidades	183
15.3.1	Caso 1D	184
15.3.2	Caso ND	185
15.4	Genericidade de Singularidades	185
15.5	Famílias de Funções	187
15.5.1	Teorema de Classificação de Thom	188
15.6	Aplicação aos Espaços de Escala	189
15.7	Exercícios	194
16	Algoritmos	197
16.1	Detecção de bolhas	197
16.1.1	Bolha a partir de máximo local	197
16.1.2	Bolha a partir de um ponto qualquer	198
16.1.3	Todas as bolhas de uma imagem	198
16.1.4	Árvores de bolhas	199
16.2	Detecção de bolhas multi-escala	199
16.2.1	Como determinar pares candidatos?	200
16.2.2	Problemas e refinamentos	202
16.3	Estrutura de Dados	203
V	Referências	207
17	Implementação de convoluções	209
17.0.1	Extensões de f	209
17.0.2	Algoritmos para convoluções	210
18	Curvatura de Curvas Planas	217
18.1	Expressão em coordenadas cartesianas	217
18.2	Parametrização por comprimento de arco	218
18.3	Curvatura de curvas de nível	219
18.4	Curvatura das curvas integrais do gradiente	220
19	Funções Modificadas de Bessel	223

Prefácio

Este livro é baseado inicialmente nas notas do curso de Espaços de Escala lecionado no IMPA em Abril-Junho de 1999 e novamente em Março-Junho de 2000. Muito da estrutura deste livro foi herdada de [28], já que este foi o livro-texto inicialmente utilizado no curso. O livro amadureceu quando foi escrito como livro-texto para a Escola de Computação 2000 (o leitor atento notará que vários capítulos estão intactos desde então). As principais mudanças desde aquele livro são o acréscimo dos capítulos sobre Equações Diferenciais Não-Lineares e algum material mais avançado (e bastante teórico) sobre o uso da Teoria das Singularidades em Espaços de Escala. O objetivo é ambicioso: que este livro seja útil tanto para o matemático interessado na belíssima teoria ao redor dos Espaços de Escala como também para os que queiram diretamente as suas aplicações em Visão Computacional e Processamento de Imagens.

Portanto, é natural que o livro apresente mais material do que o estritamente necessário num curso introdutório aos Espaços de Escala. Por exemplo, os capítulos da primeira parte, sobre Transformadas de Sinais e sobre a Gaussiana, apresentam material que talvez seja bem conhecido do público alvo. Também, quase todas as “contas” necessárias para demonstrar os vários resultados estão aqui presentes; **o leitor que não esteja interessado nesses cálculos pode ignorar todos os lemas apresentados no livro assim como as demonstrações das proposições.** No entanto, tais cálculos tendem a reforçar o método que está por trás da teoria de Espaços de Escala; sem este método, há o perigo de a teoria parecer ser um bando de resultados dispersos.

Seções que não são necessárias para o prosseguimento do curso são marcadas com * e podem ser sacrificadas dependendo do tempo disponível. No entanto, algum do material nestas seções é extremamente enriquecedor; parte dele é, acreditamos, inédito. O instrutor encarregado do curso deve analisá-las com muito carinho.

Por outro lado, este livro não menciona alguma das conexões da teoria de espaços de escala com outras subáreas de processamento de imagens. Enquanto esta versão é mais completa do que a original, ainda há ausências lamentáveis: faltam a relação entre espaços de escala e *wavelets* (espaços de escala tradicionais correspondem a uma wavelet base de primeiro momento nulo) e, como sempre, mais aplicações práticas em visão.

O autor gostaria de agradecer aos colegas Jonas de Miranda Gomes e Luiz Velho pela motivação de escrever este livro e por várias sugestões essenciais a

respeito de sua organização; ao aluno Moacyr Alvim Silva pela ajuda e paciência com revisões do texto e discussões esclarecedoras; ao aluno Anderson Mayrink da Cunha pela implementação de grande parte dos algoritmos aqui descritos e pela confecção de algumas das figuras; e a todos os alunos do projeto Visgraf que implementaram os algoritmos aqui expostos durante os cursos de 1999 e 2000. No entanto, quaisquer idéias estapafúrdias presentes no texto são responsabilidade exclusiva do autor¹, fora os exatos 97 erros tipográficos e de cálculo que foram deixados propositalmente para o entretenimento do leitor.

O autor contou também com a ajuda financeira indispensável do CNPq (processo 300956/98-3), com a fantástica infra-estrutura do IMPA e com o apoio do Departamento de Matemática da PUC-RJ. A estas organizações, sinceros agradecimentos.

¹O autor também deve um agradecimento à NovalginaTM que proporcionou intervalos são entre acessos febris durante os dias finais de preparo desse manuscrito. Apesar disso, o leitor deve ter cuidado ao atacar as partes mais herméticas do texto – o autor está completamente certo de que tudo que está mal escrito é claramente culpa de seu estado alterado. :)

Introdução

Neste capítulo, descreveremos brevemente como o conceito de “escala” aparece em várias ciências; em particular, notamos que o conceito de multi-escala (ou processamento em várias escalas) já tem aparecido de diversas formas em Processamento de Imagens e Visão Computacional. Brevemente, mostraremos ao leitor que a teoria de Espaços de Escala deve ser considerada já nos primeiros módulos de Visão Computacional, de acordo com o modelo de Marr em [30]. Por fim, na última seção, listamos os capítulos deste livro notando a sua fraca interdependência.

0.1 Escala e Ciência

A percepção de um objeto depende profundamente da escala em que ele é observado. Observando a sua mão com o braço estendido, você provavelmente perceberá o seu formato geral, e algumas das texturas maiores de sua pele – as junções dos dedos, o “M” da palma da sua mão e aquela marca de nascença que você tanto odeia. Aproxime sua mão de seus olhos e você começará a perceber estruturas que não pareciam estar lá – impressões digitais e outras pequenas texturas. Em geral, à medida que um observador se afasta do objeto observado, detalhes do objeto tendem a ficar “borrados” unindo-se para formar novas estruturas².

Um ótimo exemplo de como a percepção do mundo muda com a escala é o vídeo de 9 minutos “Powers of Ten” ([8]), que nos leva numa viagem do limite longínquo do universo até o interior de um átomo através de mudanças de escala (visite o website [34] para uma amostra). Outro exemplo é o mosaico da figura 1 que apresenta uma estrutura global bem distinta da estrutura observada quando examinada em escala menor (baseado em [11]; veja também o website do Projeto Visgraf [48]).

Muitos associam imediatamente a palavra “escala” à **cartografia**. A obtenção de mapas em escalas maiores segue o princípio da *generalização cartográfica* (como em [31]), e mapas são úteis exatamente porque *omittem* informações

²Se o observador for míope como o autor, o processo será grandemente acelerado; por outro lado, observadores hipermétropes notarão o efeito inverso para um pequeno intervalo de distâncias...

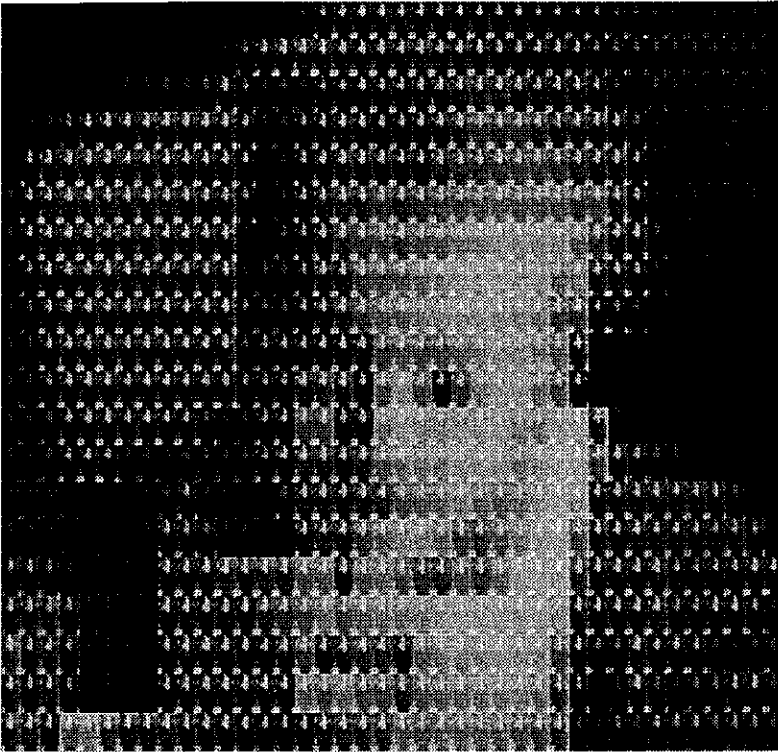


Figura 1: Um bebê composto por várias fotos de uma única pessoa.

indesejáveis em escalas grandes! As seguintes propriedades básicas de mapas cartográficos (planos) nos perseguirão pelo resto deste livro:

- Isotropia (não há direções preferenciais) – um detalhe é representado da mesma maneira independentemente de sua orientação;
- Invariância por translações (não há locações preferenciais) – um detalhe é representado da mesma maneira independentemente de sua localização;
- Causalidade (não há detalhes espúrios) – detalhes presentes em escalas maiores têm causa em escalas menores; à medida que a escala cresce, “novos” detalhes são introduzidos apenas como consequência de detalhes já presentes (que, em geral, desaparecem para criar os novos).

O estudo dos fractais é um outro ramo da ciência que tem recebido bastante atenção onde a noção de escala parece desempenhar um papel fundamental. De fato, uma das propriedades mais interessantes dos fractais é que eles possuem “auto-semelhança”, isto é, parecem formar “figuras semelhantemente

complicadas” independentemente da escala em que são analisados. É desta multitude de detalhes que surge a complexidade de tais estruturas – e, de certa forma, a noção de que fractais podem ser modelos melhores para *praticamente todos* os objetos do mundo natural (incluindo objetos de estudo da própria cartografia, como linhas de costa e montanhas, por exemplo).

Por outro lado, os elementos fundamentais da **geometria euclideana**, como “ponto” e “reta” não têm escala, isto é, são representados de maneira única através de escalas. Isto quer dizer que tais elementos não são apropriados para uma representação multi-escala para sinais e imagens – precisaremos de algum outro modelo matemático.

0.2 Escala em Computação Gráfica

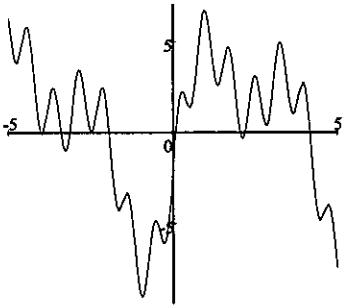
0.2.1 Por que e como?

A idéia de processar um sinal ou uma imagem em diversas escalas não é nova em Processamento de Sinais. O primeiro trabalho em detecção de arestas que sugere o uso de operadores de diversos tamanhos parece ser [39]. Outro exemplo simples de representação multi-escala de um sinal é tomar sucessivas subamostragens como níveis de uma **pirâmide**. A esperança destes esquemas de representação é realizar parte do processamento em níveis mais altos, onde a imagem é pequena e simples, e só descer aos níveis mais complexos se estritamente necessário. No mesmo espírito, as **quad-trees** e **oct-trees** procuram representar objetos gráficos dividindo-o recursivamente em pedaços menores até que cada um deles seja “simples” o suficiente, como em [25]. Infelizmente, tais representações não costumam apresentar invariância por translações – translate o objeto a ser representado e a sua representação muda violentamente. Uma outra alternativa bastante estudada nos últimos anos é a representação por **wavelets** (veja [5], por exemplo), que praticamente gera a mesma teoria dos espaços de escala sob uma abordagem diferente – esperamos abordar mais explicitamente a relação entre wavelets e espaços de escala num trabalho futuro.

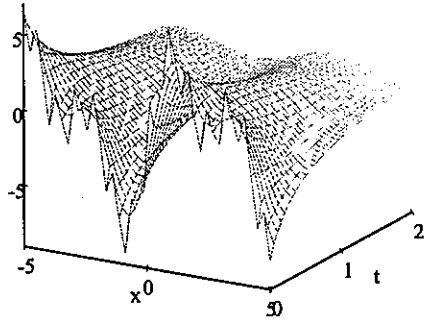
Em suma, o processamento de sinais em diferentes escalas simplifica a computação; é mais fácil detectar detalhes em escalas maiores, onde o processamento não é distraído por ruído ou textura, e depois, se necessário, segui-los para escalas menores para melhorar a sua localização. Assim, queremos uma teoria que represente um objeto gráfico em diversas escalas; que o objeto seja simplificado à medida que a escala aumenta; e que a representação seja razoavelmente robusta frente a algumas transformações básicas (translações e rotações).

A invariância por translações praticamente impede o uso de subamostragens; o **espaço de escala** de um sinal será representado por algo como a figura a seguir, enquanto o espaço de escala de uma imagem será algo semelhante à figura 2; sem querer estragar a surpresa, as figuras são simplesmente suavizações com filtros Gaussianas de variância crescente. Mas a idéia não é simplesmente suavizar o sinal em diversas escalas – já que criamos uma estrutura adicionando uma dimensão extra ao nosso sinal, procuraremos analisá-la como um todo,

através das escalas.



Um sinal unidimensional $f(x)$...



...e seu espaço de escala $F_t(x)$.

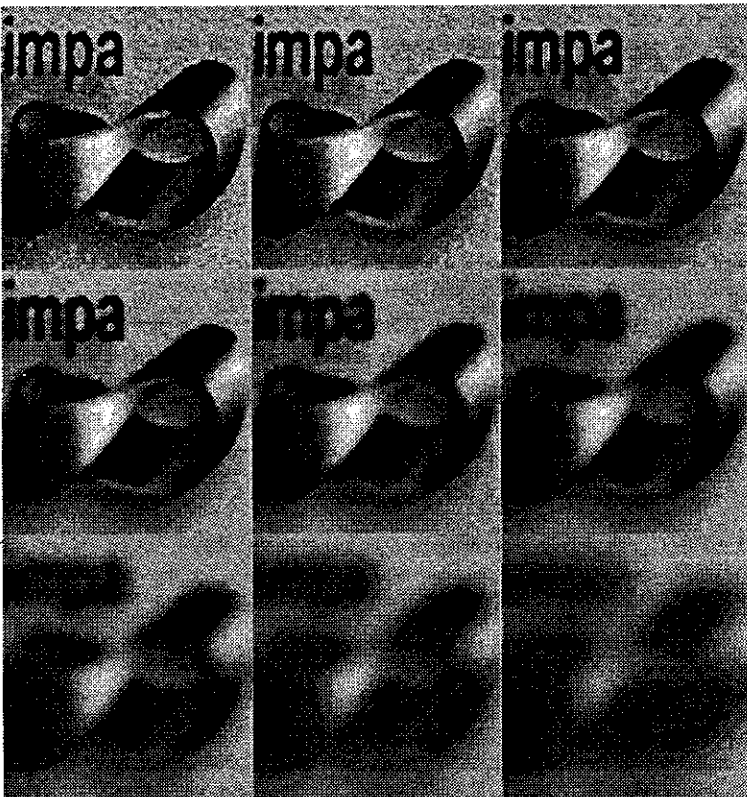


Figura 2: Um espaço de escala em 9 escalas distintas.

0.2.2 Onde?

Antes de respondermos onde entra a teoria de espaços de escala em Visão Computacional, devemos explicar o que entendemos por Visão Computacional. O processo de Visão é o problema inverso do processo de síntese de imagens. A síntese é um problema direto e razoavelmente bem entendido, enquanto a visão é um problema mal-posto: não só pode haver mais de uma descrição válida para uma mesma imagem, mas também pequenas variações na imagem podem levar a descrições radicalmente diferentes. O diagrama da figura 3 ilustra esses conceitos (tanto para Visão Computacional como para Visão Natural).

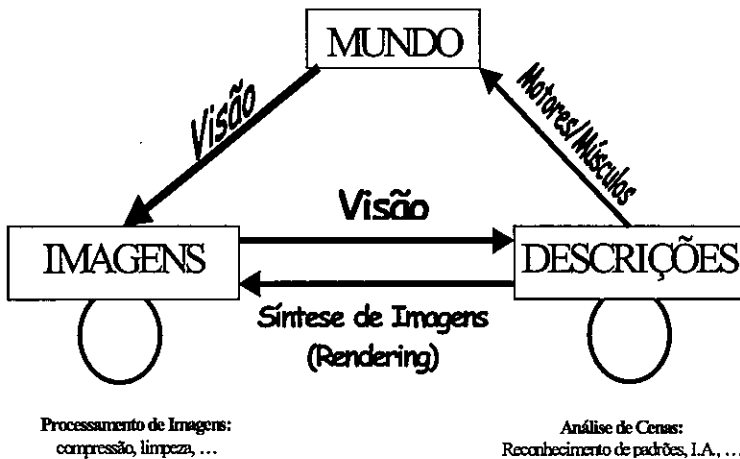


Figura 3: Panorama da Computação Gráfica

Um dos modelos mais bem aceitos para o processo de Visão Computacional é o modelo de David Marr (de [30]), esboçado no diagrama da figura 4. Este modelo assume que processos de “baixo nível” repassam informação analisada a processos de “alto nível”, formando uma espécie de “pipeline”. De fato, há razões fortes para acreditar que o “pipeline” não deve ser unidirecional: acredita-se que, pelo menos no caso da visão natural, é comum que os processos mais inteligentes retornem dados aos processos mais simples para verificar hipóteses. De qualquer forma, a teoria de Espaços de Escala deve entrar bem cedo no processamento, antes do processo de detecção de arestas. Em particular, existem indicações de que a visão natural realiza uma análise multi-escala a nível equivalente ou mais baixo do que os primeiros processamentos de detecção de arestas.

Note também que um sistema genérico de visão não tem conhecimento a priori da relação entre a escala mínima presente nos dados de entrada (chamada *escala interna*) e a escala dos objetos importantes da imagem – mais uma motivação para criar uma teoria que encompasse várias escalas. Pelo mesmo

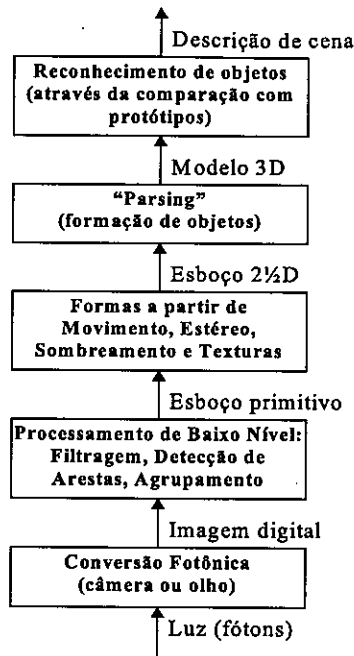


Figura 4: Modelo de Marr para Visão Computacional

motivo, é razoável supor que nenhuma escala seja tratada diferentemente das demais, ou seja, queremos criar uma certa "homogeneidade" também através das escalas.

0.3 Estrutura do Livro

0.3.1 Parte I

Os conceitos principais da maioria dos capítulos deste livro podem ser entendidos independentemente dos demais. No entanto, para que o leitor entenda profundamente todo o método por trás dos espaços de escala, recomendamos uma boa experiência com as ferramentas utilizadas. A parte I se dedica a apresentar tais ferramentas e estabelecer gradativamente a notação a ser usada no resto do livro.

Assim, os capítulos 1, 2 e 3 apresentam *convoluções, transformadas de Fourier, a transformada de Laplace e a transformada Z* sob o ponto de vista da Álgebra Linear; mesmo o leitor que tenha um bom conhecimento destes tópicos pode descobrir que nossa abordagem é instrutiva e revela algumas conexões "ocultas" entre estes objetos. Optamos por escrever os casos contínuo, discreto e finito separadamente – mesmo considerando as várias analogias entre eles, há

uma surpresa ou outra. Há vários exercícios interessantes, alguns dedicados a criar familiaridade com o assunto, outros que podem surpreender até mesmo leitores mais calejados.

O capítulo 4 apresenta a *função Gaussiana*, suas propriedades e o papel que ela desempenha na solução da *equação do calor* (ou equação de difusão linear).

0.3.2 Parte II

A parte II trata do espaço de escala Gaussiano em si. Iniciamo-la com o capítulo 5, que define o *espaço de escala Gaussiano* e procura formalizar as suas principais propriedades, como linearidade, invariância por translações, semi-grupo e causalidade (também isotropia no caso multidimensional).

Já com um olho na implementação, o capítulo 6 trata da *discretização nas coordenadas espaciais* do espaço de escala. A discretização é feita de forma a manter-se as boas propriedades encontradas no caso contínuo. O *núcleo de Poisson simétrico* é apresentado como a versão discreta da Gaussiana, e mostramos como usá-lo também no caso multidimensional para gerar espaços de escala discretos no espaço.

No capítulo 7 discutimos o que fazer se precisarmos de uma *discretização em escala* – veremos que a equação do calor pode ser discretizada de maneira explícita ou implícita, e resolvemo-la em todos os casos, apresentando os *núcleos “binomiais” simétricos* (que são de fato trinomiais) e os *núcleos geométricos simétricos*. Em particular, é possível discretizar o espaço de escala *somente* na escala, apesar de não sabermos aplicações diretas deste caso – esta situação gera o *núcleo exponencial simétrico*.

O capítulo 8 discute como obter *invariantes geométricos* – isto é, expressões dependendo das derivadas que procuram encontrar diretamente os detalhes desejados de uma imagem. Apresentamos três aplicações imediatas: detecção de arestas, junções e o uso do Laplaciano. Como exercício, notamos um outro fato que parece um pouco escondido na literatura de Visão Computacional: o Laplaciano é um péssimo detector de arestas quando estas têm alta curvatura!

O capítulo 9 mostra como gerar o espaço de escala das *derivadas* de um sinal, tão necessárias para o cômputo dos invariantes geométricos. Como bônus, acrescentamos uma discussão rigorosa de como obter derivadas discretas que pode ser estendida para o caso multidimensional. Esta discussão gera máscaras para derivadas parciais que são corretamente assimétricas, e que surpreendentemente não são exatamente as mais comumente encontradas na literatura.

No capítulo 10, partimos para uma discussão mais *axiomática* que essencialmente diz que o espaço de escala Gaussiano é a única maneira de definir um espaço de escala que satisfaz as condições de linearidade, invariância por translação, causalidade e simetria (rotacional). De um ponto de vista estritamente teórico, esta discussão viria antes da definição do espaço de escala Gaussiano; preferimos adia-la até aqui para que o leitor possa chegar mais rapidamente aos resultados e para que ele adquira mais familiaridade com os espaços de escala antes de lidar com tais assuntos de interesse mais abstrato.

Finalmente, o capítulo 11 apresenta em uma página uma visão geral do que seria a detecção de detalhes multi-escala utilizando nosso método – talvez o leitor deva vislumbrá-lo agora para começar a se familiarizar com o *panorama da teoria*.

0.3.3 Parte III

Na parte III, apresentamos algumas idéias iniciais sobre Espaços de Escala não-lineares. Esta área tem sido objeto de intensa pesquisa recentemente e limitamo-nos apenas a motivar a introdução de alguns modelos não-lineares sem entrar a fundo na matemática envolvida (ao leitor interessado em mais detalhes, recomendamos entusiasticamente o livro [9] e as notas de aula preliminares [17]). Assim, no capítulo 12 abordamos brevemente a idéia de modificar a condutância na equação do calor para obter equações diferenciais que “borrem” pouco a imagem perto de suas *arestas*, para que estas sejam mais facilmente detectadas em escalas maiores. Esta é a idéia principal de equações como a de *Perona-Malik* ali apresentada.

No capítulo 13, apresentamos algumas equações diferenciais que são *morfológicas*, isto é, equações diferenciais *invariantes por mudança de contraste*. A principal destas equações é a de *movimento por curvatura*, que tem propriedades matemáticas belíssimas e já foi sugerida para gerar *espaços de escala de formas* (“shapes”). Mencionaremos a conexão desta equação com o *filtro mediano* e apenas mencionaremos uma outra EDP que é invariante por transformações afins (chamado de *AMSS*).

0.3.4 Parte IV

A parte IV trata de uma outra possível aplicação dos espaços de escala: o cálculo de um *esboço primitivo* baseado no conceito de *bolhas* (que esperançosamente correspondem a estruturas de uma imagem). O capítulo 14 oferece uma possível definição para tais bolhas e mostra como estruturá-las e ligá-las através das escalas, criando *bolhas multi-escala*. Discutimos brevemente como calcular a importância de bolhas multi-escala distintas. A maioria dos resultados (que esperamos não ser contra o senso comum) é apresentada inicialmente sem provas.

É no capítulo 15 que oferecemos um gosto da *Teoria das Singularidades* que está por trás dos resultados sobre bolhas. Na verdade, vários resultados podem ser obtidos usando apenas o *Teorema da Função Implícita* que também apresentamos neste capítulo. Acreditamos que esta é uma ótima oportunidade para apreender um pouco sobre essa teoria tendo como meta a classificação dos eventos de bolhas de um espaço de escala Gaussiano.

O capítulo 16 mostra mais explicitamente os *algoritmos e estruturas de dados* a serem usados no cálculo das tais bolhas multi-escala. Estes foram os algoritmos implementados pelos alunos do projeto Visgraf em 1999 e 2000. Como o cômputo das bolhas multi-escala pode gerar massivas estruturas de dados, estamos sempre procurando maneiras mais eficientes (e robustas) de realizá-lo.

Assim, sugerimos ao leitor interessado que verifique freqüentemente o website do Projeto Visgraf (www.visgrafimpa.br) onde as bibliotecas serão apresentadas – esperamos oferecê-las como software de domínio público ao final de nosso projeto.

0.3.5 Parte V

O primeiro apêndice descreve alguns detalhes sobre como obter *extensões de um sinal finito e como implementar a convolução discreta* em cada caso. Um caso especial é o dos *núcleos recursivos*, que podem ser implementados de maneira extremamente eficiente. Tais núcleos podem ser usados para aproximar outros núcleos; por exemplo, apresentamos brevemente a aproximação do núcleo Gaussiano de *Deriche* (de [7]).

Os outros dois apêndices apresentam material de referência sobre a *curvatura de curvas planas* (um conhecido invariante geométrico) e sobre as *funções modificadas de Bessel* (que são cruciais na definição do núcleo de Poisson simétrico). A *bibliografia* que fecha o livro está longe de refletir todo o trabalho que tem sido feito na área de espaços de escala – ela tende apenas a refletir algumas escolhas pessoais do autor.

Parte I

Ferramentas

Capítulo 1

Transformadas de Sinais Não-Discretos

1.1 Convoluções

Seja um operador (ou filtro) L que recebe um sinal unidimensional $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ e retorna um novo sinal unidimensional $(Lf)(x) : \mathbb{R} \rightarrow \mathbb{R}$.

Definição 1.1 L é linear se

$$L\{f + \lambda.g\}(x) = Lf(x) + \lambda.Lg(x)$$

para quaisquer sinais f e g e qualquer constante $\lambda \in \mathbb{R}$.

Definição 1.2 L é invariante por translações se

$$L\{f(x - \alpha)\} = (Lf)(x - \alpha)$$

para qualquer sinal f e qualquer constante $\alpha \in \mathbb{R}$.

Se definirmos o operador translação dado por $T_\alpha g(x) = g(x - \alpha)$ então esta definição equivale a

$$L \circ T_\alpha = T_\alpha \circ L$$

Filtros lineares invariantes por translações são muito comuns em processamento de sinais, e portanto merecem uma caracterização mais explícita. Para tanto, utilizaremos uma “função” bem conhecida em processamento de sinais.

Definição 1.3 A função delta de Dirac (também denominada impulso unitário) é a função com a propriedade de amostragem

$$\int_{-\infty}^{+\infty} f(u) \delta(u) du = f(0) \text{ para qualquer } f : \mathbb{R} \rightarrow \mathbb{R}$$

que também satisfaz

$$\delta(x) = 0 \text{ para qualquer } x \neq 0$$

A definição acima não é rigorosa; de fato, estritamente falando, é impossível definir uma função $\delta : \mathbb{R} \rightarrow \mathbb{R}$ que satisfaça a definição acima. No entanto, é possível definir “funções generalizadas” tal como δ acima em Análise Funcional¹ e, por incrível que pareça, tais “funções” podem ser operadas da maneira usual. Note que em particular tomando $f \equiv 1$

$$\int_{-\infty}^{+\infty} \delta(u) du = 1$$

Portanto, por assim dizer, a função delta é nula em qualquer ponto exceto na origem, onde seu valor é infinito o suficiente para que a área sob seu gráfico seja exatamente 1.

Note que uma simples mudança de variáveis nos dá

$$\int_{-\infty}^{+\infty} f(u) \delta(x-u) du = \int_{-\infty}^{+\infty} f(x-v) \delta(v) dv = f(x-0) = f(x)$$

e as funções $\delta_u(x) = \delta(x-u)$ formam uma base natural do espaço de funções; para tanto, leia a expressão acima: $f(x)$ é um vetor que pode ser decomposto como uma combinação linear (não uma soma, mas uma integral) dos vetores $\delta_u(x)$, cada um com coeficiente dado por $f(u)$.

Proposição 1.4 *Toda transformação linear L invariante por translações é dada por uma convolução*

$$(Lf)(x) = (f * h)(x) \doteq \int_{-\infty}^{+\infty} f(u) h(x-u) du$$

onde $h = L\delta$ é a resposta de L ao impulso unitário.

Demonstração. De fato, seja $h(x) = (L\delta)(x)$. Como L é linear

$$\begin{aligned} (Lf)(x) &= L \left\{ \int_{-\infty}^{+\infty} f(u) \delta(x-u) du \right\} = \int_{-\infty}^{+\infty} L\{f(u) \delta(x-u)\} du = \\ &= \int_{-\infty}^{+\infty} f(u) L\{\delta(x-u)\} du \end{aligned}$$

e portanto, usando a invariância por translações

$$(Lf)(x) = \int_{-\infty}^{+\infty} f(u) (L\delta)(x-u) du = \int_{-\infty}^{+\infty} f(u) h(x-u) dx = (f * h)(x)$$

¹ A idéia da Análise Funcional é definir $\delta(x)$ como o limite de uma seqüência de funções $f_n(x)$, todas elas positivas, satisfazendo $\int f_n(x) dx = 1$ e com suportes cada vez menores com interseção $\{0\}$. Tal limite não existe no sentido clássico, mas pode ser bem definido num espaço de funcionais que inclui de maneira natural o espaço das funções $f : \mathbb{R} \rightarrow \mathbb{R}$. Para o leitor interessado, recomendamos [19].

A função $h = L\delta$ é chamada *função de espalhamento pontual* da transformação linear L . Frequentemente, usaremos a notação $L_h f = f * h$.

Dados dois sinais quaisquer, nem sempre a convolução entre eles está bem definida. No entanto, se nos limitarmos a sinais de quadrado integrável, isto é, tais que

$$\int_{-\infty}^{+\infty} |f(u)|^2 du < +\infty$$

então a integral da convolução converge. Mais ainda, não é difícil provar as seguintes propriedades da convolução

Proposição 1.5 *Dados sinais f , g e h de quadrado integrável*

$$\begin{aligned} f * g &= g * f \\ (f * g) * h &= f * (g * h) \\ f * \delta &= f \\ \frac{d(f * g)}{dx} &= \frac{df}{dx} * g = f * \frac{dg}{dx} \end{aligned}$$

sendo esta última expressão válida sempre que as derivadas envolvidas forem válidas.

Comentário 1.6 *Note que muitas funções que não são diferenciáveis passam a ter derivadas quando introduzimos funções generalizadas (como a função delta de Dirac). Por exemplo, a função degrau unitário*

$$u(x) = \begin{cases} 0, & \text{se } x < 0 \\ 1, & \text{se } x \geq 0 \end{cases}$$

passa a ter como derivada a função delta

$$u'(x) = \delta(x)$$

Estas “novas derivadas” têm as mesmas propriedades das derivadas “usuais”.

1.2 Transformada de Fourier

Uma boa maneira de entender uma transformação linear é estudar a sua diagonalização. Se $\{\delta_\alpha\}$ é uma base do espaço de funções, que outra base diagonalizará a convolução? Em outras palavras, que vetores são autovetores da convolução?

Se estendermos o nosso espaço de funções para incluir funções complexas, a resposta é surpreendente.

Proposição 1.7 *A função $e_w(x) = e^{2\pi i w x}$ é um autovetor do operador L_h (qualquer que seja h).*

Demonstração. De fato,

$$\begin{aligned} L_h e_w(x) &= h * e_w = \int_{-\infty}^{+\infty} h(u) e^{2\pi i w(x-u)} du = \\ &= \left(\int_{-\infty}^{+\infty} e^{-2\pi i w u} h(u) du \right) \cdot e^{2\pi i w x} = \hat{h}(w) e_w(x) \end{aligned}$$

mostra que e_w é um autovetor com autovalor dado pela expressão

$$\hat{h}(w) = \int_{-\infty}^{+\infty} h(u) e^{-2\pi i w u} du$$

Note que a escolha de uma convolução (de um h) não muda os autovetores que são sempre os mesmos e_w , mas corresponde apenas a uma escolha dos autovalores $\hat{h}(w)$. A função $\hat{h}(w)$ é chamada *função de transferência* do filtro L_h .

Assim, a operação $L_h f = h * f$ pode ser realizada da seguinte forma (novamente, compare esta descrição às transformações lineares em \mathbb{R}^n):

- Decomponha f na base e_w ; neste contexto, ao invés da combinação linear usual com um somatório, temos

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(w) e_w(x) dw$$

onde os coeficientes $\hat{f}(w)$ de cada vetor e_w têm de ser determinados de alguma forma;

- Multiplique cada $\hat{f}(w)$ (que é o coeficiente do vetor e_w) pelo autovalor correspondente $\hat{h}(w)$;
- Reconponha a combinação linear para obter $L_h f$

$$L_h f = \int_{-\infty}^{+\infty} \hat{h}(w) \hat{f}(w) e_w(x) dw$$

E o nosso grande problema é: como obter os tais $\hat{f}(w)$? Felizmente, podemos definir neste espaço de funções o seguinte produto interno entre duas funções f e g

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(u) \overline{g(u)} du$$

Note que a definição acima satisfaz todas as propriedades usuais de um produto interno (lembrando que $\langle f, g \rangle = \overline{\langle g, f \rangle}$).

Afirmção 1.8

$$\langle e_{w_1}, e_{w_2} \rangle = \delta(w_1 - w_2)$$

Esta afirmação não pode ser demonstrada apenas com o material que expusemos até aqui. De fato, o leitor que tentar prová-la diretamente de nossas definições encontrará integrais divergentes que não podem ser resolvidas com o cálculo tradicional. No entanto, este resultado pode ser entendido da seguinte forma: duas funções do tipo $e_{w_1}(x) = e^{2\pi i w_1 x}$ e $e_{w_2}(x) = e^{2\pi i w_2 x}$ não se correlacionam, a menos, é claro, que $w_1 = w_2$ (caso em que elas se correlacionam e muito).

Mais importante, esta afirmação diz que a base formada pelos vetores e_w é ortogonal com este produto interno! Assim, os $\hat{f}(w)$ que queremos serão simplesmente a projeção do vetor f na direção de e_w , isto é

$$\hat{f}(w) = \langle f, e_w \rangle = \int_{-\infty}^{+\infty} f(u) e^{-2\pi i w u} du$$

Definição 1.9 A Transformada de Fourier de uma função $f(t)$ é dada por

$$\hat{f}(w) = \int_{-\infty}^{+\infty} f(u) e^{-2\pi i w u} du$$

sempre que esta integral for válida.

Note que a notação apresentada para a função de transferência \hat{h} é consistente com esta definição.

Proposição 1.10 (Transformada de Fourier Inversa) É possível decompor uma função $f(x)$ como uma "combinação linear" dos vetores e_w da seguinte forma

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(w) e^{2\pi i w x} dw$$

Demonstração. Usando a definição de \hat{f} e a afirmação 1.8 com e_x e e_u

$$\begin{aligned} \int_{-\infty}^{+\infty} \hat{f}(w) e^{2\pi i w x} dw &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(u) e^{-2\pi i w u} e^{2\pi i w x} du dw = \\ &= \int_{-\infty}^{+\infty} f(u) \left(\int_{-\infty}^{+\infty} e^{2\pi i (x-u)w} dw \right) du = \int_{-\infty}^{+\infty} f(u) \delta(x-u) du = f(x) \end{aligned}$$

Definição 1.11 A Transformada Inversa de Fourier de um sinal $g(w)$ é denotada por

$$\check{g}(x) = \int_{-\infty}^{+\infty} g(w) e^{2\pi i w x} dw$$

Corolário 1.12 *A Transformada Inversa de Fourier nada mais é do que a Transformada de Fourier de um sinal “refletido”, isto é, dada $g(w)$*

$$\check{g} = \widehat{g(-w)}$$

ou em outras palavras, dada $f(x)$,

$$\widehat{\check{f}}(x) = f(-x)$$

Uma maneira comum de encarar a Transformada de Fourier em processamento de sinais: $\hat{f}(w)$ mede o quanto de uma frequência w está “presente” no sinal $f(x)$. De fato, o sinal $e_w(x) = e^{2\pi i w x}$, cuja parte real é $\cos(2\pi w x)$, é o sinal “puro” de frequência w e $\hat{f}(w)$ é o peso de e_w na reconstrução de $f(x)$. Por esse mesmo motivo, diz-se que $f(x)$ é a representação de f no domínio do tempo ou domínio do espaço, enquanto $\hat{f}(w)$ é a representação do mesmo sinal no domínio da frequência.

Note também como a intensidade da frequência w em $L_h f$ é $\hat{f}(w) \hat{h}(w)$, isto é, há um ganho de $|\hat{h}(w)|$ (possivelmente com defasagem dada pela fase de $\hat{h}(w)$) na frequência w quando f passa pelo filtro L_h . Por esse motivo, quando o ganho é alto perto de $w = 0$ mas baixo para $|w|$ grande, dizemos que L_h é um filtro de passa-baixa (apenas as frequências baixas sobrevivem a esse filtro). Por outro lado, se $|\hat{h}(w)|$ é alto para $|w|$ grande e baixo para w perto de 0, L_h é um filtro de passa-alta.

Algumas das propriedades básicas da Transformada de Fourier seguem abaixo:

Afirmção 1.13 *As seguintes propriedades são válidas (f e h são sinais quaisquer; $\mathbb{I}(x) \equiv 1$)*

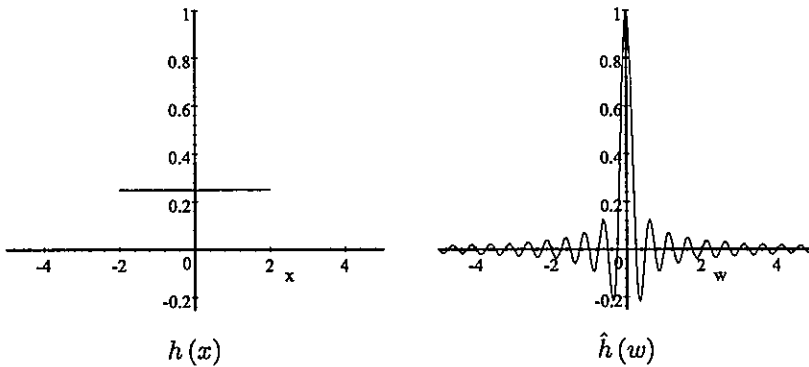
$$\begin{aligned} \hat{\delta} &= \mathbb{I}(w) \text{ (e portanto } \widehat{\mathbb{I}} = \delta(w) \text{)} \\ \widehat{f * h} &= \hat{f} \cdot \hat{h} \text{ (e portanto } \widehat{f \hat{h}} = \hat{f} * \hat{h} \text{)} \\ g(x) = \delta_a * f = f(x - a) &\Rightarrow \hat{g}(w) = \hat{f}(w) e^{-2\pi i w a} \text{ (e assim, } |\hat{g}(w)| = |\hat{f}(w)| \text{)} \\ g(x) = f(ax) &\Rightarrow \hat{g}(w) = \frac{1}{a} \hat{f}\left(\frac{w}{a}\right) \\ g(x) = f'(x) = \delta' * f &\Rightarrow \hat{g}(w) = (\delta') \cdot \hat{f} = 2\pi i w \hat{f}(w) \end{aligned}$$

A seguir, apresentamos alguns exemplos de Transformadas de Fourier.

Exemplo 1.14

$$\begin{aligned} h(x) &= \begin{cases} \frac{1}{2a}, & -a \leq x \leq a \\ 0, & \text{caso contrário} \end{cases} \Rightarrow \\ &\Rightarrow \hat{h}(w) = \frac{1}{2a} \int_{-a}^a e^{-2\pi i w u} du = \frac{1}{2a} \frac{e^{-2\pi i w u}}{-2\pi i w} \Big|_{u=-a}^a = \frac{\sin(2\pi w a)}{2\pi w a} \end{aligned}$$

cujos gráficos seguem abaixo (para $a = 2$):



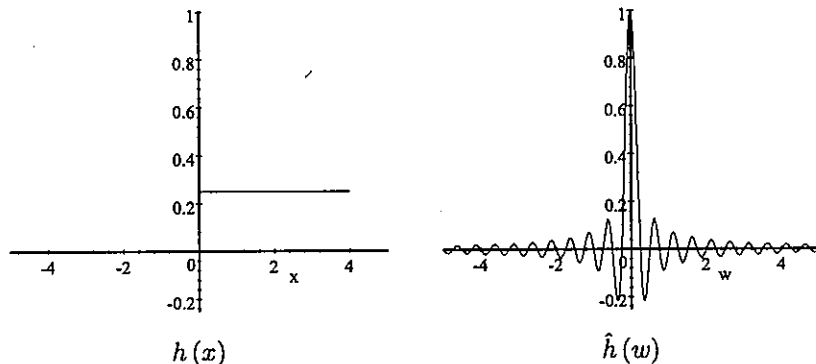
Note como a Transformada de Fourier é maior em baixas freqüências (como $w = 0$, onde $\hat{f}(0) = 1$), indicando que h é um filtro de passa-baixa (não muito ideal). Quanto maior o valor de a , mais saliente é este pico. Também vale a pena notar que quando $a \rightarrow 0$, $h \rightarrow \delta$ e $\hat{h} \rightarrow \mathbb{I}$, como esperado. Aliás, a função δ é em geral definida através de um limite deste tipo.

Exemplo 1.15

$$h(x) = \begin{cases} \frac{1}{2a}, & 0 \leq x \leq 2a \\ 0, & \text{caso contrário} \end{cases} \Rightarrow$$

$$\Rightarrow \hat{h}(w) = \frac{1}{2a} \int_0^{2a} e^{-2\pi i w u} du = \frac{1}{2a} \frac{e^{-2\pi i w u}}{-2\pi i w} \Big|_{u=0}^{2a} = \frac{\sin 2\pi w a}{2\pi w a} e^{-2\pi i w a}$$

Como $\hat{h}(w)$ é complexa, mostramos abaixo os gráficos de $h(x)$ e $|\hat{h}(w)| = \frac{\sin 2\pi w a}{2\pi w a}$ para $a = 2$:



Este exemplo resalta uma propriedade já indicada: transladar h no tempo

(a unidades para a direita) simplesmente muda a fase de \hat{h} (multiplica \hat{h} por $e^{-2\pi i w a}$) mas não muda o seu ganho².

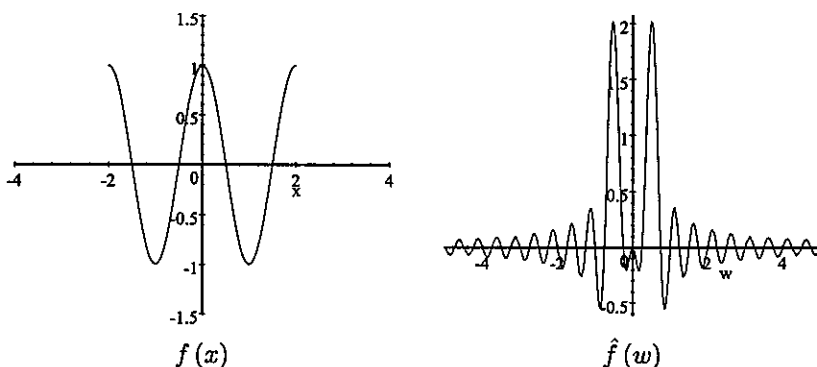
Exemplo 1.16

$$f(x) = \begin{cases} \cos(2\pi w_0 x), & -a \leq x \leq a \\ 0, & \text{caso contrário} \end{cases} \Rightarrow$$

$$\Rightarrow \hat{f}(w) = \int_{-a}^a \cos(2\pi w_0 u) e^{-2\pi i w u} du =$$

$$= \frac{1}{2\pi} \left(\frac{\sin(2a\pi(w+w_0))}{w+w_0} + \frac{\sin(2a\pi(w-w_0))}{w-w_0} \right)$$

cujos gráficos seguem abaixo (para $a = 2$ e $w_0 = 0.5$):



Note como a Transformada de Fourier acusa as fortes presenças das frequências $w = \pm 0.5$. Estes picos têm altura aproximadamente igual a a e ficam mais salientes à medida que a cresce. Note que como f não foi “normalizada”, o valor de $\hat{f}(w)$ perto de $w = \pm w_0$ irá explodir quando a vai para infinito. Esta observação está associada à propriedade de manutenção de energia a seguir.

Afirmção 1.17

$$\int_{-\infty}^{\infty} |f(u)|^2 du = \int_{-\infty}^{\infty} |\hat{f}(w)|^2 dw$$

Em outras palavras, a energia do sinal $f(u)$ no domínio do tempo é a energia de $\hat{f}(w)$ no domínio da frequência. Isto não é surpreendente já que os coeficientes $\hat{f}(w)$ e os coeficientes $f(u)$ nada mais são que decomposições de f em bases ortogonais distintas.

²Literalmente, $|\hat{h}(w)| = \left| \frac{\sin 2\pi w a}{2\pi w a} \right|$, mas usamos esta função sem o módulo a título de comparação com o exemplo anterior. As partes negativas simplesmente indicam uma fase de π .

1.3 Transformada Bilateral de Laplace

Definição 1.18 Dada uma função $f : \mathbb{R} \rightarrow \mathbb{R}$, definimos a sua Transformada Bilateral de Laplace como

$$\mathcal{L}_f(s) = \int_{-\infty}^{+\infty} e^{-st} f(t) dt$$

A Transformada Bilateral de Laplace nada mais é do que a Transformada de Fourier onde escrevemos s ao invés de $2\pi iw$. A diferença é que em geral s é real, enquanto $2\pi iw$ é complexo; se permitirmos que s e w sejam complexos, é imediato verificar que

$$\begin{aligned}\hat{f}(w) &= \mathcal{L}_f(2\pi iw) \\ \mathcal{L}_f(s) &= \hat{f}\left(\frac{s}{2\pi i}\right)\end{aligned}$$

A integral da Transformada de Laplace pode muito bem não convergir para uma gama de valores de s . Em geral, a sua expressão só faz sentido se $f(t)$ apresenta decaimento rápido quando $t \rightarrow \pm\infty$. Neste caso, temos as seguintes propriedades:

- Norma³:

$$\int_{-\infty}^{+\infty} f(t) dt = \mathcal{L}_f(0)$$

e dizemos que f é normalizada quando

$$\int_{-\infty}^{+\infty} f(t) dt = \mathcal{L}_f(0) = 1$$

- Média:

$$\mu = \int_{-\infty}^{+\infty} t f(t) dt = -\mathcal{L}'_f(0)$$

é o valor médio de f .

- Variância para funções normalizadas:

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{+\infty} (t - \mu)^2 f(t) dt = \\ &= \int_{-\infty}^{+\infty} t^2 f(t) dt - \mu^2 = \\ &= \mathcal{L}''_f(0) - \mu^2\end{aligned}$$

é a variância de f quando f é normalizada.

³Esta norma não tem relação direta com a norma L_2 apresentada anteriormente, isto é, $(f(t), f(t))^{\frac{1}{2}} = \left[\int_{-\infty}^{+\infty} |f(t)|^2 dt \right]^{\frac{1}{2}}$. Este tipo de "norma" é mais utilizada no contexto de probabilidades; no nosso caso, utilizaremos-la quando tratarmos de núcleos de convolução. Para um núcleo $K(t)$, freqüentemente pede-se que $\int_{-\infty}^{+\infty} K(t) dt = 1$.

- Convolução:

$$\mathcal{L}_{f_1 * f_2} = \mathcal{L}_{f_1} \mathcal{L}_{f_2}$$

isto é, a Transformada de Laplace leva convoluções em produtos (isto não deve ser uma surpresa, já que a Transformada de Fourier tinha a mesma propriedade). Portanto, dadas duas funções normalizadas f_1 e f_2 com transformadas \mathcal{L}_1 e \mathcal{L}_2 , médias μ_1 e μ_2 e variâncias σ_1^2 e σ_2^2 , a média e a variância da convolução serão

$$\begin{aligned} \mu &= -\mathcal{L}'(0) = -\mathcal{L}'_1(0) \mathcal{L}_2(0) - \mathcal{L}_1(0) \mathcal{L}'_2(0) = \mu_1 + \mu_2 \\ \sigma^2 &= \mathcal{L}''(0) - \mu^2 = \mathcal{L}''_1(0) \mathcal{L}_2(0) + 2\mathcal{L}'_1(0) \mathcal{L}'_2(0) + \mathcal{L}_1(0) \mathcal{L}''_2(0) - \mu^2 = \\ &= (\sigma_1^2 + \mu_1^2) + 2\mu_1\mu_2 + (\sigma_2^2 + \mu_2^2) - \mu^2 = \sigma_1^2 + \sigma_2^2 \end{aligned}$$

onde usamos $\mathcal{L} = \mathcal{L}_{f_1 * f_2}$.

- Derivada:

$$\mathcal{L}_{\frac{df}{dx}}(s) = s \mathcal{L}_f(s)$$

já que $\mathcal{L}_{\delta'} = s$ e $\frac{df}{dx} = \delta' * f$.

- Reflexão:

$$\begin{aligned} f_2(t) = f_1(-t) &\Rightarrow \\ \Rightarrow \mathcal{L}_2(s) = \mathcal{L}_1(-s); \mu_2 = -\mu_1; \sigma_2^2 = \sigma_1^2 \end{aligned}$$

Comentário 1.19 Tanto a Transformada de Fourier como a Transformada de Laplace são ferramentas bastante utilizadas para resolver equações diferenciais, já que ambas transformam derivadas em multiplicações. A título de exemplo, considere a equação diferencial parcial com condição inicial

$$\begin{aligned} \frac{\partial F(x, t)}{\partial t} &= \frac{\partial F(x, t)}{\partial x} \\ F(x, 0) &= f_0(x) \end{aligned}$$

Podemos obter a solução dela aplicando a Transformada de Fourier à variável x . Assim, no domínio da frequência

$$\begin{aligned} \frac{\partial \hat{F}(w, t)}{\partial t} &= 2\pi i w \hat{F}(w, t) \\ \hat{F}(w, 0) &= \hat{f}_0(w) \end{aligned}$$

cuja solução é claramente

$$\hat{F}(w, t) = e^{2\pi i w t} \hat{f}_0(w)$$

Dada a solução no domínio da frequência, a Transformada Inversa de Fourier nos dá a solução no domínio do espaço; neste caso, usando que $\delta_u = e^{-2\pi i w u}$

$$F(x, t) = (\delta_{-t} * f_0)(x) = f_0(x + t)$$

1.4 Transformadas Bidimensionais

A teoria discutida até aqui é facilmente generalizada para o caso multidimensional. Nesta seção, indicamos os principais resultados associados a convoluções e Transformadas de Fourier bidimensionais; a extensão para maiores dimensões é imediata.

Definição 1.20 *A função delta de Dirac bidimensional é a função com a propriedade de amostragem*

$$\int_{\mathbb{R}^2} f(x, y) \delta(x, y) dA = f(0, 0) \text{ para qualquer } f : \mathbb{R}^2 \rightarrow \mathbb{R}$$

que também satisfaz

$$\delta(x, y) = 0 \text{ para qualquer } (x, y) \neq (0, 0)$$

Afirmção 1.21 *Defina $\delta_{x_0, y_0} = \delta(x - x_0, y - y_0)$. Então*

$$\int_{\mathbb{R}^2} f(x, y) \delta_{x_0, y_0}(x, y) dA = f(x_0, y_0)$$

Afirmção 1.22 *Toda transformação linear L (que leva funções de duas variáveis em funções de duas variáveis) invariante por translações é dada por uma convolução*

$$Lf(a, b) = (f * h)(a, b) \doteq \int_{\mathbb{R}^2} f(x, y) h(a - x, b - y) dA$$

onde $h = L\delta$ é a resposta de L ao impulso unitário bidimensional.

A convolução bidimensional satisfaz as mesmas propriedades que a convolução unidimensional. Em particular, a propriedade da derivada da convolução pode ser aplicada a qualquer derivada parcial:

$$\begin{aligned} \frac{\partial (f * g)}{\partial x} &= \frac{\partial f}{\partial x} * g = f * \frac{\partial g}{\partial x} \\ \frac{\partial (f * g)}{\partial y} &= \frac{\partial f}{\partial y} * g = f * \frac{\partial g}{\partial y} \end{aligned}$$

Comentário 1.23 *Pode-se realizar a convolução em apenas uma das várias variáveis de duas funções; por exemplo, a partir de $f(x, y)$ e $g(x, y)$ podemos realizar uma convolução somente na primeira variável, mantendo y fixo*

$$h(x, y) = f *_x g = \int_{-\infty}^{+\infty} f(u, y) g(x - u, y) du$$

Quando isto ocorrer, usaremos a notação $*_x$ para indicar a variável que foi convoluída, como indicado acima (se tal variável não for óbvia do contexto). É

importante distinguir esta convolução da convolução verdadeiramente bidimensional! Por exemplo, não é surpresa que

$$\frac{\partial (f *_x g)}{\partial x} = \frac{\partial f}{\partial x} *_x g = f *_x \frac{\partial g}{\partial x}$$

mas para a derivação em y precisamos de uma regra tipo produto

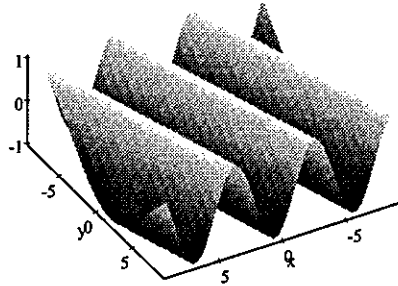
$$\frac{\partial (f *_x g)}{\partial y} = \frac{\partial f}{\partial y} *_x g + f *_x \frac{\partial g}{\partial y}$$

Esta última forma deve ser aplicada sempre que calcularmos derivadas parciais com respeito a variáveis que não estão na convolução!

Os autovetores da convolução são agora dados por autovetores da forma

$$e_{w_1, w_2}(x, y) = e^{2\pi i(w_1 x + w_2 y)}$$

que representam de certa forma frequências puras na direção (w_1, w_2) . A título de ilustração, segue abaixo o gráfico da parte real de e_{w_1, w_2} , a saber, $\cos(w_1 x + w_2 y)$



$$w_1 = 1; w_2 = 0.3$$

Definição 1.24 *Dada uma função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definimos a sua Transformada de Fourier como*

$$\hat{f}(w_1, w_2) = \int_{\mathbb{R}^2} f(x, y) e^{-2\pi i w_1 x} e^{-2\pi i w_2 y} dA$$

Note que a sua inversa é

$$f(x, y) = \int_{\mathbb{R}^2} \hat{f}(w_1, w_2) e^{2\pi i w_1 x} e^{2\pi i w_2 y} dw_1 dw_2$$

Exemplo 1.25 Se h somente depende de uma variável, digamos, $h(x, y) = g(x)$, então

$$\begin{aligned}\hat{h}(w_1, w_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x) e^{-2\pi i w_1 x} e^{-2\pi i w_2 y} dx dy = \\ &= \int_{-\infty}^{\infty} g(x) e^{-2\pi i w_1 x} \left(\int_{-\infty}^{\infty} e^{-2\pi i w_2 y} dy \right) dx = \\ &= \hat{g}(w_1) \lim_{a \rightarrow \infty} \frac{\sin 2\pi w_2 a}{\pi w_2}\end{aligned}$$

não é exatamente a Transformada de Fourier de g (e provavelmente não convergirá). O problema é que o sinal $h(x, y) = g(x)$ tem energia demais. Mesmo que tentemos consertar este problema tomando

$$h(x, y) = \begin{cases} g(x), & -a \leq y \leq a \\ 0, & \text{caso contrário} \end{cases}$$

ainda temos

$$\hat{h}(w_1, w_2) = \hat{g}(w_1) \frac{\sin 2\pi w_2 a}{\pi w_2}$$

e a transformada de h ainda não é exatamente a transformada de g . Estas observações são casos especiais da seguinte propriedade.

Proposição 1.26 Dadas duas funções $f, g: \mathbb{R} \rightarrow \mathbb{R}$, construa o produto tensorial

$$h(x, y) = f(x)g(y)$$

Então a Transformada de Fourier de h é

$$\hat{h}(w_1, w_2) = \hat{f}(w_1)\hat{g}(w_2)$$

Demonstração. De fato

$$\begin{aligned}\hat{h}(w_1, w_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) e^{-2\pi i w_1 x} e^{-2\pi i w_2 y} dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)g(y) e^{-2\pi i w_1 x} e^{-2\pi i w_2 y} dx dy = \\ &= \left(\int_{-\infty}^{\infty} f(x) e^{-2\pi i w_1 x} dx \right) \left(\int_{-\infty}^{\infty} g(y) e^{-2\pi i w_2 y} dy \right) = \\ &= \hat{f}(w_1)\hat{g}(w_2)\end{aligned}$$

onde assumimos que todas as integrais acima convergem. ■

Corolário 1.27 A função $h(x, y) = \delta(x)g(y)$ tem Transformada de Fourier $\hat{h}(w_1, w_2) = \hat{g}(w_2)$.

Afirmção 1.28 As seguintes propriedades são válidas (f e h são sinais bidimensionais quaisquer; $\mathbb{I}(x, y) \equiv 1$)

$$\hat{\delta} = \mathbb{I}(w_1, w_2) \text{ (e portanto } \widehat{\mathbb{I}} = \delta(w_1, w_2) \text{)}$$

$$\widehat{f * h} = \hat{f} \cdot \hat{h} \text{ (e portanto } \widehat{f h} = \hat{f} * \hat{h} \text{)}$$

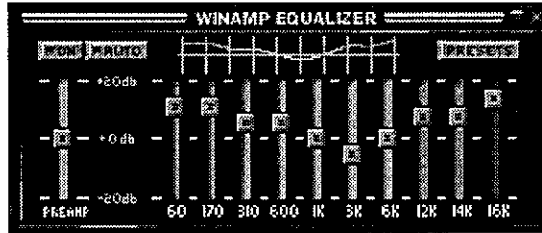
$$g(x, y) = f(x - a, y - b) \Rightarrow \hat{g}(w_1, w_2) = \hat{f}(w_1, w_2) e^{-2\pi i(w_1 a + w_2 b)} \text{ (} |\hat{g}| = |\hat{f}| \text{)}$$

$$g(x, y) = f(\alpha x, \alpha y) \Rightarrow \hat{g}(w_1, w_2) = \frac{1}{\alpha} \hat{f}\left(\frac{w_1}{\alpha}, \frac{w_2}{\alpha}\right)$$

$$g(x, y) = \frac{\partial f}{\partial x} = \frac{\partial \delta}{\partial x} * f \Rightarrow \hat{g}(w_1, w_2) = \left(\frac{\partial \delta}{\partial x}\right) \cdot \hat{f} = 2\pi i w_1 \hat{f}(w_1, w_2)$$

1.5 Exercícios

1) Você provavelmente já viu alguma espécie de “equalizador gráfico de áudio”. Por exemplo, as versões mais recentes do conhecido programa “free-ware” Winamp™ têm um equalizador embutido:



Use a linguagem deste capítulo para explicar o que um “equalizador” faz com um sinal de áudio (idealmente). Alguns pontos que devem estar presentes em sua resposta:

- Se a amplitude dos sinais que entram e saem do equalizador são $f_{in}(t)$ e $f_{out}(t)$ (em função do tempo), qual a relação entre f_{out} e f_{in} ?
- Qual a função de cada slider no equalizador? O que são “60”, “170”, ..., “16K”?
- A figura acima mostra uma espécie de gráfico que se assemelha à posição dos sliders. Isto seria o gráfico de quê?
- Se todos os sliders estiverem na posição central (0 db), o que acontece? E se todos estiverem no topo (+20 db)? Qual a diferença sonora entre esses dois casos?

2) Um sinal de áudio $f_{in}(t)$ é processado por um filtro linear gerando uma saída $f_{out}(t)$. Um modelo para este processamento é

$$f_{out}(t) = (f_{in} * h)(t)$$

onde h é um núcleo de convolução e t representa tempo. Tal expressão parece indicar que $f_{out}(t_0)$ (para um t_0 fixo qualquer) depende dos valores de $f_{in}(t)$

em toda a reta real. No entanto, filtros de áudio vão “cuspindo” a resposta f_{out} à medida que f_{in} entra (isto é, o filtro ainda não sabe os pedaços de f_{in} “estão por entrar”). Como explicar esta aparente contradição? Tendo em vista este fato, que tipo de núcleo h esperamos encontrar? Mais ainda, se o filtro tem um tempo de processamento Δt (isto é, mudanças no sinal f_{in} em $t = t_0$ só podem ser sentidas em f_{out} em $t = t_0 + \Delta t$), como fica a condição em h ?

3) Demonstre as propriedades básicas da convolução (Proposição 1.5) e suas versões bidimensionais.

4) Demonstre as propriedades da Transformada de Fourier (Afirmações 1.13 e 1.28).

5) Mostre que se $f(x)$ é real então $\hat{f}(w) = \overline{\hat{f}(-w)}$. Conclua que $\hat{f}(0)$ é real. Generalize para a Transformada de Fourier bidimensional.

6) Considere $f(x, y) = \cos(2\pi w_1 x + 2\pi w_2 y) = \operatorname{Re} e^{-2\pi i(w_1 x + w_2 y)}$. Vamos encontrar o comprimento de onda e a direção de f em função de w_1 e w_2 .

(a) O que são as curvas de nível de f ?

(b) Dê um vetor perpendicular às curvas de nível de f . Que ângulo este vetor faz com o eixo x ?

(c) Qual a distância entre as curvas sucessivas que correspondem a $f = 1$ (comprimento da onda)?

7) Mostre que se $f(x) > 0$ para todo x , então $|\hat{f}(w)| \leq \hat{f}(0)$ para qualquer w ; em outras palavras, no domínio da frequência, a componente mais forte de um sinal positivo é a componente contínua. Generalize para o caso bidimensional. Quando ocorre a igualdade?

8) Uma longa barra numa imagem parece produzir “linhas de frequências” na sua Transformada de Fourier. Vamos descobrir porque isto acontece:

(a) Calcule a Transformada de Fourier da função $f(x, y) = \delta(x)$ (versão idealizada de uma barra fina e longa). Há tais linhas?

(b) Calcule a Transformada de Fourier da função indicatriz de um retângulo $2m \times 2n$ centrado na origem com lados paralelos aos eixos (isto é, uma função que vale 1 dentro do retângulo e 0 fora dele).

(c) Suponha que $m \gg n$, por exemplo. Há tais linhas (ou segmentos)? Onde?

9) a) Utilize uma Transformada de Sinais (Fourier ou Laplace) para mostrar que a solução geral $f(x) : \mathbb{R} \rightarrow \mathbb{C}$ da Equação Diferencial Ordinária

$$f''(x) = -K f(x)$$

onde K é uma constante positiva, é dada por

$$f(x) = K_1 e^{\sqrt{K}ix} + K_2 e^{-\sqrt{K}ix}$$

b) Utilize uma Transformada de Sinais (Fourier ou Laplace) para encontrar todas as funções $F(x, t)$ que satisfaçam a Equação Diferencial Parcial abaixo

$$\begin{aligned}\frac{\partial^2 F}{\partial t^2} &= \frac{\partial^2 F}{\partial x^2} \\ F(x, 0) &= f_0(x) \\ \frac{\partial F}{\partial t}(x, 0) &= 0\end{aligned}$$

10) Se uma rotação de coordenadas de ângulo θ é aplicada a uma função $f(x, y)$, como muda a sua Transformada de Fourier?

Capítulo 2

Transformadas de Sinais Discretos

2.1 Convoluções

Como no capítulo anterior, considere um operador L cuja entrada é agora um sinal **discreto** unidimensional $f[n] : \mathbb{Z} \rightarrow \mathbb{R}$ e cuja saída será um outro sinal discreto $Lf[n] : \mathbb{Z} \rightarrow \mathbb{R}$. Note o uso de colchetes ao invés de parênteses para indicar que o sinal é discreto, notação que adotaremos daqui por diante.

Nesta seção examinaremos brevemente a classe dos operadores L que sejam lineares e invariantes por translações (cuja definição é similar à do capítulo anterior). Desta vez, faremos uso da versão discreta da função delta de Dirac:

Definição 2.1 *A função delta de Dirac discreta (ou impulso unitário discreto) é definida por*

$$\begin{aligned}\delta[0] &= 1 \\ \delta[n] &= 0 \text{ se } n \neq 0\end{aligned}$$

Corolário 2.2 *A função delta de Dirac discreta é a única com a propriedade de amostragem*

$$\sum_{\alpha=-\infty}^{+\infty} f[\alpha] \delta[\alpha] = f[0] \text{ para qualquer } f : \mathbb{Z} \rightarrow \mathbb{R}$$

Note que

$$f[n] = \sum_{\alpha=-\infty}^{+\infty} f[\alpha] \delta[n - \alpha]$$

e as seqüências $\delta_\alpha[n] = \delta[n - \alpha]$ são uma base natural do espaço de seqüências.

Proposição 2.3 *Toda transformação linear invariante por translações é dada por uma convolução discreta*

$$Lf[n] = (f * h)[n] \doteq \sum_{\alpha=-\infty}^{+\infty} f[\alpha] h[n - \alpha]$$

onde $h = L\delta$ é a resposta de L ao impulso unitário.

Demonstração. Seja $h = L\delta$. Então como L é linear e invariante por translações

$$\begin{aligned} Lf[n] &= L \left\{ \sum_{\alpha=-\infty}^{+\infty} f[\alpha] \delta[n - \alpha] \right\} = \sum_{\alpha=-\infty}^{+\infty} L \{ f[\alpha] \delta[n - \alpha] \} = \\ &= \sum_{\alpha=-\infty}^{+\infty} f[\alpha] L \{ \delta[n - \alpha] \} = \sum_{\alpha=-\infty}^{+\infty} f[\alpha] h[n - \alpha] = f * h \end{aligned}$$

sempre que os somatórios acima converjam fortemente. ■

A função $h = L\delta$ é chamada de *função de espalhamento pontual discreta* de L ; como antes, usaremos $L_h f = f * h$.

Note que nem sempre a convolução de duas seqüências está bem definida (de fato, até mesmo a demonstração da proposição acima falha se os somatórios nela envolvidos não convergirem). No entanto, se nos limitarmos a seqüências de quadrado somável

$$\sum_{-\infty}^{+\infty} |f[n]|^2 < +\infty$$

então a convolução está bem definida e as seguintes propriedades são válidas.

Proposição 2.4 *Dadas seqüências f , g e h de quadrado somável*

$$\begin{aligned} f * g &= g * f \\ (f * g) * h &= f * (g * h) \\ f * \delta &= f \\ \Delta(f * g) &= \Delta f * g = f * \Delta g \end{aligned}$$

onde Δf é a seqüência definida por $\Delta f[n] = f[n + 1] - f[n]$.

Em particular, sinais finitos (isto é, com suporte finito: $n > M \Rightarrow f[n] = 0$ para algum M) satisfazem todas essas propriedades. Note também que a última propriedade vem das duas primeiras, já que $\Delta f = K * f$ onde $K = [1 \ -1]$ é centrado em -1 (isto é, $K[0] = -1$).

2.2 Transformada de Fourier de Tempo Discreto (DTFT)

Se L_h é um operador convolução qualquer definido por $L_h\{f\} = f * h$, será que podemos diagonalizá-lo? Se sim, que base o faz? Quais os autovetores do operador convolução?

Proposição 2.5 *A seqüência $e_w[n] = e^{2\pi i w n}$ é um autovetor do operador L_h (qualquer que seja h).*

Demonstração. De fato,

$$\begin{aligned} L_h e_w[n] &= h * e_w = \sum_{\alpha=-\infty}^{+\infty} h[\alpha] e^{2\pi i w (n-\alpha)} = \\ &= \left(\sum_{\alpha=-\infty}^{+\infty} h[\alpha] e^{-2\pi i w \alpha} \right) e^{2\pi i w n} = \hat{h}(w) e_w[n] \end{aligned}$$

mostra que e_w é um autovetor com autovalor dado pela expressão

$$\hat{h}(w) = \sum_{\alpha=-\infty}^{+\infty} h[\alpha] e^{-2\pi i w \alpha}$$

Novamente, toda convolução tem como autovetores os mesmos vetores e_w ; apenas os autovalores $\hat{h}(w)$ é que mudam de convolução para convolução. Note uma peculiaridade deste conjunto de vetores: como n é sempre um inteiro, $e_w[n] = e_{w+1}[n]$ para qualquer w . Em outras palavras, basta considerarmos os vetores e_w para $0 \leq w < 1$ – qualquer outro valor de w apenas repete vetores já considerados.

Como decompor uma seqüência $f[n]$ como combinação linear desses e_w (com $0 \leq w < 1$) para que possamos realizar a convolução via produtos? O caminho certo é tentar algo na linha do produto interno

$$\langle f, g \rangle = \sum_{-\infty}^{+\infty} f[\alpha] \overline{g[\alpha]}$$

que está definido no sentido clássico para todas as seqüências de quadrado somável. De certa forma, os vetores e_w formam uma base ortogonal (isto não faz sentido na análise clássica – os vetores e_w não são seqüências de quadrado somável). Este fato nos sugere o uso do produto interno para decompor uma seqüência $f[n]$ em componentes de diversas freqüências.

Definição 2.6 *A Transformada de Fourier de Tempo Discreto¹ (Discrete-Time*

¹O nome que se dá a esta transformada é bastante infeliz: gostaríamos de chamá-la de Transformada de Fourier Discreta, mas este nome é reservado para uma outra situação que será discutida no capítulo seguinte: o caso em que $f[n]$ é uma seqüência finita (isto é, um vetor em \mathbb{R}^m)

Fourier Transform, ou DTFT) de uma seqüência $f[n]$ é dada por

$$\hat{f}(w) = \langle f, e_w \rangle = \sum_{\alpha=-\infty}^{+\infty} f[\alpha] e^{-2\pi i w \alpha}$$

sempre que este somatório for válido.

Como $\hat{f}(w)$ é periódica de período 1, é comum imaginá-la como uma função $\hat{f}: [0, 1] \rightarrow \mathbb{R}$ ou $\hat{f}: [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{R}$. Note também que, felizmente, a notação \hat{h} acima é consistente com essa definição.

Estranhamente, a afirmação de que os vetores e_w são ortogonais nos indica o caminho certo mas **não é necessária** para demonstrar a fórmula da Transformada Inversa. Ao invés, o resultado que precisamos é um cálculo simples cuja demonstração deixamos ao leitor:

Lema 2.7

$$\int_0^1 e^{2\pi i n w} dw = \delta[n]$$

Proposição 2.8 (Transformada Fourier Inversa de Tempo Discreto) A seqüência $f[n]$ pode ser decomposta como uma “combinação linear” dos vetores e_w da seguinte forma

$$f[n] = \int_0^1 \hat{f}(w) e^{2\pi i w n} dw$$

Demonstração. Usando a definição de \hat{f} e o lema acima

$$\begin{aligned} \int_0^1 \hat{f}(w) e^{2\pi i w n} dw &= \int_0^1 \sum_{\alpha=-\infty}^{+\infty} f[\alpha] e^{-2\pi i w \alpha} e^{2\pi i w n} dw = \\ &= \sum_{\alpha=-\infty}^{+\infty} f[\alpha] \left(\int_0^1 e^{2\pi i (n-\alpha)w} dw \right) = \sum_{\alpha=-\infty}^{+\infty} f[\alpha] \delta[n-\alpha] = f[n] \end{aligned}$$

Definição 2.9 A DTFT Inversa de um sinal $g(w)$ é denotada por

$$\tilde{g}[n] = \int_0^1 g(w) e^{2\pi i w n} dw$$

Intuitivamente, $\hat{f}(w)$ mede o quanto de uma freqüência w está presente na seqüência $f[n]$; como freqüências que diferem por 1 são indistinguíveis numa seqüência discreta ($e_1[n] = e^{2\pi n} = 1 = e_0[n]$), basta considerar $0 \leq w < 1$. Diz-se que $\hat{f}(w)$ é a representação da seqüência $f[n]$ no *espaço da freqüência*.

As propriedades básicas da DTFT seguem abaixo:

2.2. TRANSFORMADA DE FOURIER DE TEMPO DISCRETO (DTFT) 23

Afirmção 2.10 As seguintes propriedades são válidas (f e h são seqüências quaisquer; $\mathbb{I}[n] \equiv 1$)

$$\widehat{\delta[n]} \equiv 1$$

$$\widehat{\mathbb{I}[n]} = \sum_{k=-\infty}^{\infty} \delta(w - k) \text{ (extensão periódica de } \delta(w) \text{)}$$

$$\widehat{f * h} = \hat{f} \cdot \hat{h}$$

$$\widehat{fh} = \hat{f} * \hat{h}$$

$$g[n] = f[n - \alpha] \Rightarrow \hat{g}(w) = \hat{f}(w) e^{2\pi i w \alpha} \text{ (e } |\hat{g}(w)| = |\hat{f}(w)| \text{)}$$

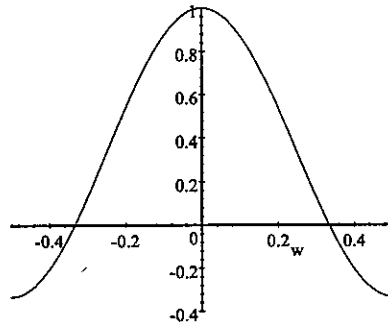
Exemplo 2.11 Seja $h[n] = \frac{1}{3}$ para $n \in \{-1, 0, 1\}$ ($h[n] = 0$ caso contrário). O filtro L_h é um filtro "box", que suaviza um sinal f . De fato, a convolução substitui $f[n]$ pela média de seus 3 valores vizinhos

$$L_h f[n] = \frac{f[n-1] + f[n] + f[n+1]}{3}$$

No espaço da freqüência, tem-se

$$\hat{h}(w) = \frac{1}{3} (e^{2\pi i w} + 1 + e^{-2\pi i w}) = \frac{1}{3} (2 \cos 2\pi w + 1)$$

cujo gráfico segue abaixo (usando $-0.5 \leq w \leq 0.5$)

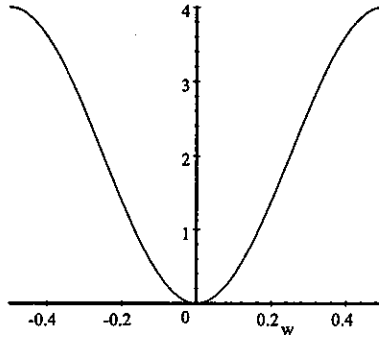


confirmando que h é um filtro razoável de passa-baixa (mas note que, por exemplo, $|\hat{h}(0.5)| > |\hat{h}(0.3)|$). Por exemplo, se f é uma discretização do degrau unitário, tem-se

$$f = \begin{matrix} \dots & 0 & 0 & 0 & 1 & 1 & 1 & \dots \end{matrix}$$

$$f * h = \begin{matrix} \dots & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 1 & 1 & \dots \end{matrix}$$

Exemplo 2.12 Por outro lado, se $h = [-1 \ 2 \ -1]$, sua DTFT é dada por



$$\hat{h}(w) = 2 - 2 \cos 2\pi w$$

e L_h é um ótimo filtro de passa-alta. Por exemplo

$$f = \begin{array}{cccccccccc} \dots & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \dots \\ f * h = & \dots & 0 & 0 & 0 & -1 & 1 & 0 & 0 & \dots \end{array}$$

Comentário 2.13 Note que a DTFT corresponde à Transformada de Fourier da amostragem pontual de um sinal, isto é,

$$g(x) = \sum_{n=-\infty}^{+\infty} f[n] \delta[x - n] \Rightarrow \hat{g}(w) = \hat{f}(w)$$

onde \hat{g} corresponde à Transformada de Fourier “tradicional” (que, neste caso, é naturalmente periódica) enquanto \hat{f} é a DTFT de f . Assim, se utilizarmos a prática (comum em Processamento de Sinais) de identificar um sinal discreto com uma soma de impulsos

$$f[n] \approx g(x) = \sum_{k=-\infty}^{\infty} f[k] \delta_k(x)$$

a Transformada de Fourier passa a ser uma única transformada.

2.3 Transformada Z

Definição 2.14 Dada uma seqüência $f[n]$ definida para n inteiro, definimos a sua Transformada Z (também denominada função geradora) como a série de potências

$$\varphi_f(z) = \sum_{n=-\infty}^{\infty} f[n] z^n$$

A Transformada Z é basicamente a DTFT onde usamos z ao invés de $e^{-2\pi i w}$. Muitas vezes esta substituição nos permite trabalhar de maneira bem mais

simples. Em geral, uma série de potências desse tipo converge apenas para $r_1 < |z| < r_2$ (para algum r_1, r_2); em particular, se tivermos uma expressão para a Transformada Z de uma seqüência válida para $|z| = 1$, podemos calcular a sua Transformada de Fourier via

$$\hat{f}(w) = \varphi_f(e^{-2\pi i w})$$

Caso a Transformada Z seja bem definida (e diferenciável em volta de $z = 1$), temos as seguintes propriedades:

- Normalização

$$\sum_{n=-\infty}^{\infty} f[n] = \varphi_f(1)$$

e diz-se que f é normalizada quando

$$\sum_{n=-\infty}^{\infty} f[n] = \varphi_f(1) = 1$$

- Média

$$\mu = \sum_{n=-\infty}^{\infty} n f[n] = \varphi'_f(1)$$

- Variância

$$\begin{aligned} \sigma^2 &= \sum_{n=-\infty}^{\infty} (n - \mu)^2 f[n] = \\ &= \sum_{n=-\infty}^{\infty} n(n-1) f[n] + \sum_{n=-\infty}^{\infty} n(1-2\mu) f[n] + \mu^2 \sum_{n=-\infty}^{\infty} f[n] = \\ &= \varphi''_f(1) + \mu - \mu^2 \end{aligned}$$

se f é normalizada.

- Convolução

$$\varphi_{f_1 * f_2} = \varphi_{f_1} \varphi_{f_2}$$

Em particular, se f_1 e f_2 são normalizadas, com Transformadas Z dadas por φ_1 e φ_2 , médias μ_1 e μ_2 e variâncias σ_1^2 e σ_2^2 , a convolução $f_1 * f_2$ terá média $\mu = \mu_1 + \mu_2$ e variância $\sigma^2 = \sigma_1^2 + \sigma_2^2$. De fato

$$\begin{aligned} \varphi &= \varphi_{f_1 * f_2} = \varphi_1 \varphi_2 \Rightarrow \\ \Rightarrow &\begin{cases} \mu = \varphi'(1) = \varphi'_1(1)\varphi_2(1) + \varphi_1(1)\varphi'_2(1) = \mu_1 + \mu_2 \\ \sigma^2 = \varphi''(1) + \mu - \mu^2 = \\ = \varphi''_1(1)\varphi_2(1) + 2\varphi'_1(1)\varphi'_2(1) + \varphi_1(1)\varphi''_2(1) + \mu - \mu^2 = \\ = (\sigma_1^2 - \mu_1 + \mu_1^2) + 2\mu_1\mu_2 + (\sigma_2^2 - \mu_2 + \mu_2^2) + \mu - \mu^2 = \\ = \sigma_1^2 + \sigma_2^2 \end{cases} \end{aligned}$$

• Reflexão

$$f_2[n] = f_1[-n] \Rightarrow \\ \Rightarrow \varphi_2(z) = \varphi_1(z^{-1}); \mu_2 = -\mu_1; \sigma_2^2 = \sigma_1^2$$

Exemplo 2.15 *Do mesmo modo que as Transformada de Fourier e Laplace são ferramentas utilizadas na resolução de equações diferenciais, a Transformada Z (ou a DTFT) pode ser usada para resolução de equações diferenciais discretas. Por exemplo, se quisermos $F : \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R}$ tal que*

$$\frac{\partial F_t[n]}{\partial t} = F_t[n] - F_t[n-1] \\ F_0[n] \equiv f_0[n]$$

podemos notar que $F_t[n] - F_t[n-1] = (K * F_t)[n]$ (onde $K = [-1 \ 1]$ é centrado em -1) e aplicar a Transformada Z

$$\frac{\partial \varphi_{F_t}}{\partial t} = (z-1)\varphi_{F_t} \\ \varphi_{F_0} = \varphi_{f_0}$$

então

$$\varphi_{F_t} = e^{(z-1)t} \varphi_{f_0}$$

Mas note que

$$e^{(z-1)t} = e^{-t} (e^{zt}) = e^{-t} \left(1 + tz + \frac{t^2}{2} z^2 + \frac{t^3}{3!} z^3 + \dots \right)$$

vale para $|z| = 1$. Portanto

$$F_t = e^{-t} \left[1 + t \frac{t^2}{2} \frac{t^3}{3!} \frac{t^4}{4!} \dots \right] * f_0$$

Note que a DTFT poderia ser utilizada com o mesmo propósito, mas as integrais envolvidas são difíceis de resolver (tente!); a Transformada Z providencia automaticamente uma substituição agradável ($e^{-2\pi i w}$ ao invés de z).

Proposição 2.16 *A solução da equação diferencial*

$$\frac{\partial F_t}{\partial t} = A * F_t$$

é dada por

$$F_t = T_t * F_0$$

onde T_t é dada por

$$T_t = \sum_{n=0}^{\infty} \frac{t^n}{n!} (A)^{*n} = \delta + tA + \frac{t^2}{2} A * A + \frac{t^3}{3!} A * A * A + \dots$$

Demonstração. A demonstração é análoga à resolução do exemplo acima; basta tomar a Transformada Z da equação diferencial e resolvê-la

$$\varphi_{F_t} = e^{t\varphi_A} \varphi_{F_0}$$

portanto, a solução é da forma $F_t = T_t * F_0$ onde

$$\varphi_{T_t} = e^{t\varphi_A} = 1 + t\varphi_A + \frac{t^2}{2}\varphi_A^2 + \frac{t^3}{3!}\varphi_A^3 + \dots$$

e aplicando a Transformada Z inversa² terminamos a demonstração. ■

Definição 2.17 Com a notação acima, o núcleo A é chamado gerador infinitesimal do semi-grupo T_t . Uma maneira mais sucinta de exprimir a relação entre A e T_t é

$$\frac{\partial T_t}{\partial t} = \lim_{h \rightarrow 0} \frac{T_{t+h} - T_t}{h} = A * T_t$$

Note que a condição de que T_t seja um semi-grupo (isto é, $T_{t_1} * T_{t_2} = T_{t_1+t_2}$) é necessária para que o gerador infinitesimal exista independente de t ; então podemos ser ainda mais sucintos:

$$A = \lim_{h \rightarrow 0} \frac{T_h - \delta}{h}$$

Comentário 2.18 A série

$$\varphi_f(z) = \sum_{n=-\infty}^{\infty} f[n] z^n$$

é conhecida em *Variáveis Complexas* como série de Laurent da função $\varphi_f(z)$. No exemplo acima, a expansão de e^{tz} como série de potências de z era simples. No caso geral, a resposta pode ser obtida usando a IDTFT

$$f[n] = \int_0^1 \hat{f}(w) e^{2\pi i w n} dw$$

Para escrevê-la a partir da Transformada Z, use a mudança de variáveis $z = e^{-2\pi i w}$

$$\hat{f}(w) = \varphi_f(z); dz = -2\pi i e^{-2\pi i w} dw = -2\pi i z dw$$

Assim, se C é a circunferência de centro 0 e raio 1 no plano complexo

$$f[n] = \int_{-C} \varphi_f(z) z^{-n} \frac{dz}{-2\pi i z} = \frac{1}{2\pi i} \int_C \frac{\varphi_f(z)}{z^{n+1}} dz$$

²E reconhecidamente varrendo possíveis problemas de convergência para debaixo do tapete...

e esta seria a fórmula da "Inversa da Transformada Z". Usando a teoria de Variáveis Complexas, pode-se mostrar que:

i) Se $\varphi_f(z)$ é analítica no anel $r_1 < |z| < r_2$, então a série de Laurent cujos coeficientes $f[n]$ são obtidos pela fórmula acima converge de fato para $\varphi_f(z)$ para $r_1 < |z| < r_2$;

ii) Se uma série de Laurent converge para uma função analítica $\varphi_f(z)$ num anel $r_1 < |z| < r_2$, então seus coeficientes têm de ser obtidos pela fórmula acima.

2.4 Transformadas Bidimensionais

Mais uma vez, não é difícil generalizar a teoria acima para o caso multidimensional. A título de referência, apresentamos os principais resultados associados a convoluções e Transformadas de Fourier bidimensionais, deixando a extensão multi-dimensional ao leitor.

Definição 2.19 A função delta de Dirac discreta é definida por

$$\delta[m, n] = \begin{cases} 1, & \text{se } m = n = 0 \\ 0, & \text{caso contrário} \end{cases}$$

Afirmção 2.20 Defina $\delta_{j,k}[m, n] = \delta[m - j, n - k]$. Então

$$f[m, n] = \sum_{j,k=-\infty}^{\infty} f[j, k] \delta_{j,k}[m, n]$$

para qualquer sinal bidimensional discreto $f: \mathbb{Z}^2 \rightarrow \mathbb{C}$.

Proposição 2.21 Toda transformação linear invariante por translações (que leva $f: \mathbb{Z}^2 \rightarrow \mathbb{C}$ em $Lf: \mathbb{Z}^2 \rightarrow \mathbb{C}$) é dada por uma convolução discreta

$$Lf[m, n] = (f * h)[m, n] \triangleq \sum_{j,k=-\infty}^{+\infty} f[j, k] h[m - j, n - k]$$

onde $h = L\delta$ é a resposta de L ao impulso unitário.

A convolução bidimensional discreta satisfaz as mesmas propriedades que a unidimensional discreta. Seus autovetores têm a forma

$$e_{w_1, w_2}[m, n] = e^{2\pi i(w_1 m + w_2 n)} \quad \text{para } 0 \leq w_1, w_2 < 1$$

Definição 2.22 Dada uma seqüência $f: \mathbb{Z}^2 \rightarrow \mathbb{C}$, definimos a sua DTFT como

$$\hat{f}(w_1, w_2) = \sum_{j,k=-\infty}^{\infty} f[j, k] e^{-2\pi i w_1 j} e^{-2\pi i w_2 k} \quad 0 \leq w_1, w_2 \leq 1$$

Sua inversa é

$$f[m, n] = \int_0^1 \int_0^1 \hat{f}(w_1, w_2) e^{2\pi i w_1 m} e^{2\pi i w_2 n} dw_1 dw_2$$

Afirmção 2.23 *As seguintes propriedades são válidas (f e h são seqüências bidimensionais quaisquer; $\mathbb{I}[m, n] \equiv 1$)*

$$\begin{aligned} \widehat{\delta[m, n]} &\equiv 1 \\ \widehat{\mathbb{I}[m, n]} &= \sum_{j, k=-\infty}^{\infty} \delta(w_1 - j, w_2 - k) \text{ (trem bidimensional de pulsos)} \\ \widehat{f * h} &= \widehat{f} \cdot \widehat{h} \\ \widehat{fh} &= \widehat{f} * \widehat{h} \\ g[m, n] &= f[m - a, n - b] \Rightarrow \\ &\Rightarrow \widehat{g}(w_1, w_2) = \widehat{f}(w_1, w_2) e^{-2\pi i(w_1 a + w_2 b)}; |\widehat{g}(w)| = |\widehat{f}(w)| \end{aligned}$$

Vale a pena também ressaltar a versão bidimensional da Transformada Z.

Definição 2.24 *Dada uma seqüência $f: \mathbb{Z}^2 \rightarrow \mathbb{R}$ definimos a sua Transformada Z como*

$$\varphi_f(z_1, z_2) = \sum_{m, n=-\infty}^{\infty} f[m, n] z_1^m z_2^n$$

Afirmção 2.25 *Em geral*

$$\widehat{f}(w_1, w_2) = \varphi_f(e^{-2\pi i w_1}, e^{-2\pi i w_2})$$

Afirmção 2.26 *Se f é separável em f_1 e f_2 , isto é, $f[m, n] = f_1[m] f_2[n]$, então*

$$\begin{aligned} \varphi_f(z_1, z_2) &= \varphi_{f_1}(z_1) \varphi_{f_2}(z_2) \\ \widehat{f}(w_1, w_2) &= \widehat{f}_1(w_1) \widehat{f}_2(w_2) \end{aligned}$$

Em particular, usando esta propriedade podemos construir filtros bidimensionais passa-baixa a partir de dois filtros unidimensionais passa-baixa.

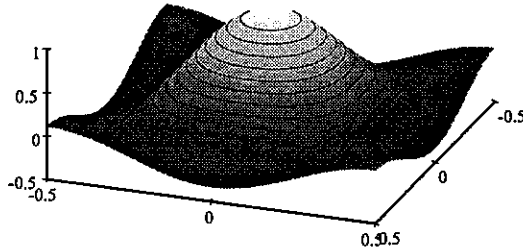
Exemplo 2.27 *A Transformada Z do filtro "box" bidimensional centrado em $(0, 0)$*

$$B = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

é

$$\begin{aligned} \varphi_B(x, y) &= \frac{1}{9} (xy + xy^{-1} + x^{-1}y + x^{-1}y^{-1} + x + y + x^{-1} + y^{-1} + 1) = \\ &= \left(\frac{1 + x + x^{-1}}{3} \right) \left(\frac{1 + y + y^{-1}}{3} \right) \end{aligned}$$

já que B é separável em dois filtros box unidimensionais. O gráfico de $\widehat{B}(w_1, w_2)$ é



$$\hat{B} = \frac{1}{9} (2 \cos 2\pi w_1 + 1) (2 \cos 2\pi w_2 + 1)$$

2.5 Exercícios

(1) Demonstre as propriedades básicas da convolução discreta (Proposição 2.4) e suas versões bidimensionais.

(2) Demonstre as propriedades da DTFT (Afirmações 2.10 e 2.23).

(3) Mostre que se f é periódica então $f * g$ (se existir) será também periódica.

(4) Mostre que

$$\sum_{j=-\infty}^{\infty} (f * g)[j] = \left(\sum_{j=-\infty}^{\infty} f[j] \right) \left(\sum_{j=-\infty}^{\infty} g[j] \right)$$

(assumindo que os somatórios duplos que aparecem convergem em qualquer ordem; isto é verdade, por exemplo, se f e g forem sinais finitos).

(5) Considere o filtro $b_1 = \frac{1}{2} \boxed{1 \quad 1}$, isto é,

$$b_1[n] = \begin{cases} \frac{1}{2}, & n = 0, 1 \\ 0, & \text{caso contrário} \end{cases}$$

e defina

$$b_n \doteq b_1 * b_1 * \dots * b_1 = b_{n-1} * b_1$$

(a) Mostre que $b_n = \frac{1}{2^n} \boxed{1 \quad n \quad \binom{n}{2} \quad \binom{n}{3} \quad \dots \quad n \quad 1}$, isto é,

$$b_n[k] = \frac{1}{2^n} \binom{n}{k}$$

(b) Calcule $\hat{b}_1(w)$ e use-o para determinar $\hat{b}_n(w)$. Faça o gráfico de \hat{b}_n . Conclua que b_n é um filtro de passa-baixa.

(6) Considere $f = C$

...	0	0	1	β	β^2	β^3	...
-----	---	---	---	---------	-----------	-----------	-----

 (centrado em 1, isto é, $f[0] = C$), onde $\beta < 1$. Encontre uma expressão compacta para $\varphi_f(z)$ (ela deve ser válida para $|z| < \frac{1}{\beta}$) e para a DTFT de f . Use-as para normalizar f , e então calcule a sua média e variância.

(7) Encontre todas as famílias de seqüências $f_t[n]$ que satisfazem

$$\frac{\partial f_t}{\partial t}[n] = f_t[n] + f_t[n-1]$$

Escreva f_t em função de f_0 usando convoluções.

(8) Uma longa barra numa imagem parece produzir “segmentos de frequências” na sua DTFT. Vamos descobrir porque isto acontece:

(a) Calcule a DTFT da função $f[x, y] = \delta[x]$ (uma barra fina e longa discreta). Há tais linhas?

(b) Calcule a DTFT da função indicatriz de um retângulo $(2m+1) \times (2n+1)$ com lados paralelos aos eixos, centrado na origem.

(c) Suponha que $m \gg n$, por exemplo. Há tais linhas (ou segmentos)? Onde?

(9) Dado um sinal discreto $f[n]$, considere a suavização $L_1 f = f * h_1$, onde

$$h_1[m] = 1, m = 0, 1$$

Nesta questão procuramos restaurar um sinal que foi suavizado pela aplicação de L_1 .

(a) Calcule $\varphi_1(z)$, a Transformada Z de h_1 . Usando a série formal

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots$$

escreva a Transformada Z Inversa de $\frac{1}{\varphi_1(z)}$ como

$$h_2[m] = (-1)^m, m \geq 0$$

(b) Mostre **diretamente** que h_2 é um inverso de convolução de h_1 , isto é, que $h_1 * h_2 = \delta$.

(c) Considere o sinal $f_0[n] = (-1)^n$ ($n \in \mathbb{Z}$). Calcule **diretamente** $f_0 * h_1$ e $(f_0 * h_1) * h_2$.

(d) Por uma simples inspeção do sinal f_0 , vemos que $w_0 = \frac{1}{2}$ é uma de suas frequências predominantes. Qual o valor de $\hat{h}_1(w_0) = \varphi_1(e^{-2\pi i w_0})$? Qual o ganho que a convolução com h_2 produz na frequência w_0 ?

(e) O que você conclui a respeito do uso de h_2 como filtro inverso de h_1 ?

[Nota: algumas das propriedades básicas da convolução não estão funcionando aqui; o problema ocorre porque f_0 e h_2 não decaem suficientemente rápido em $\pm\infty$, e a Transformada Z utilizada não converge para $|z| = 1$; dados estes h_1 e

h_2 , para que $(f * h_1) * h_2 = f$ é **necessário** que $f[n] \rightarrow 0$ quando $n \rightarrow -\infty$; surpreendentemente, isto não é suficiente!]

(10) Poderíamos driblar os problemas de convergência no problema anterior ao calcularmos a Transformada Z Inversa de

$$\frac{\varphi_1(z)}{\varphi_1^2(z) + B^2}$$

para alguma constante B .

(a) Considere o filtro

$$h_3[n] = (-1)^n (\cos \alpha)^{n+1} \cos((n+1)\alpha), \quad n \geq 0, \quad \cos \alpha \neq \pm 1$$

Mostre que $h_3[n] \rightarrow 0$ quando $n \rightarrow \pm\infty$. Mostre que

$$\lim_{\alpha \rightarrow 0} h_3[n] = h_2[n]$$

pontualmente. Isto é, h_3 é um filtro que aproxima h_2 , mas h_3 é mais “bem-comportado”.

(b) Usando que $2 \cos(n\alpha) = e^{in\alpha} + e^{-in\alpha}$, mostre que a Transformada Z de h_3 é

$$\frac{\varphi_1(z)}{\varphi_1^2(z) + \tan^2 \alpha}$$

(que vale para $|z| = 1$). Portanto, a convolução com h_3 é uma restauração parcial razoável de $f * h_1$.

Capítulo 3

Transformadas de Sinais Finitos

3.1 Convoluções Circulares

Desta vez, considere uma transformação linear $L : \mathbb{R}^N \rightarrow \mathbb{R}^N$ (ou $\mathbb{C}^N \rightarrow \mathbb{C}^N$). Estamos interessados nos operadores L que satisfazem a seguinte propriedade

$$\begin{aligned} L(a_0, a_1, \dots, a_{N-1}) &= (b_0, b_1, \dots, b_{N-1}) \Rightarrow \\ &\Rightarrow L(a_{N-1}, a_0, a_1, \dots, a_{N-2}) = (b_{N-1}, b_0, b_1, \dots, b_{N-2}) \end{aligned} \quad (3.1)$$

que é semelhante à invariância por translação dos capítulos anteriores.

Há uma maneira mais compacta de enunciar tal propriedade: seja \mathfrak{S} a transformação linear dada por $\mathfrak{S}(a_0, a_1, \dots, a_{N-1}) = (a_{N-1}, a_0, a_1, \dots, a_{N-2})$, isto é

$$\mathfrak{S} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

Então a propriedade 3.1 diz que $L\mathfrak{S} = \mathfrak{S}L$.

Denotemos a base canônica de \mathbb{R}^N por

$$\begin{aligned} \delta_0 &= (1, 0, 0, \dots, 0) \\ \delta_1 &= (0, 1, 0, \dots, 0) \\ &\dots \\ \delta_{N-1} &= (0, 0, 0, \dots, 1) \end{aligned}$$

Note que $L\delta_0$ é suficiente para determinar todos os outros $L\delta_k$. De fato, se

$$L\delta_0 = (h_0, h_1, h_2, \dots, h_N)$$

então usando a propriedade 3.1 seguidas vezes obtemos

$$\begin{aligned} L\delta_1 &= (h_{N-1}, h_0, h_1, \dots, h_{N-2}) \\ L\delta_2 &= (h_{N-2}, h_{N-1}, h_0, \dots, h_{N-3}) \\ &\dots \\ L\delta_{N-1} &= (h_1, h_2, \dots, h_{N-1}, h_0) \end{aligned}$$

e nesta base L pode ser representada pela matriz

$$\begin{bmatrix} h_0 & h_{N-1} & \dots & h_1 \\ h_1 & h_0 & \dots & h_2 \\ h_2 & h_1 & \dots & h_3 \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \dots & h_0 \end{bmatrix}$$

Corolário 3.1 *Toda transformação linear L satisfazendo 3.1 é dada por uma convolução circular*

$$(Lf)_j = (f *_c h)_j \doteq \sum_{k=0}^j f_k h_{j-k} + \sum_{k=j+1}^{N-1} f_k h_{N+j-k} = \sum_{k=0}^{N-1} f_k h_{(j-k) \bmod N}$$

para qualquer vetor $f \in \mathbb{R}^N$ (ou \mathbb{C}^N), onde $h = L\delta_0$.

O vetor δ_0 faz o papel da função delta de Dirac dos capítulos anteriores. A propriedade de amostragem torna-se

$$\langle f, \delta_0 \rangle = f_0$$

para qualquer vetor $f = (f_0, f_1, \dots, f_{N-1})$. Novamente, como $h = L\delta_0$ determina completamente a transformação L , usaremos a notação $L_h f = f *_c h$.

Proposição 3.2 *Dados vetores f, g e h em \mathbb{C}^N*

$$\begin{aligned} f *_c g &= g *_c f \\ (f *_c g) *_c h &= f *_c (g *_c h) \\ f *_c \delta_0 &= f \end{aligned}$$

3.2 Transformada de Fourier Discreta (DFT)

Não é difícil diagonalizar um operador do tipo $L_h : \mathbb{C}^N \rightarrow \mathbb{C}^N$.

Proposição 3.3 *Se α é uma raiz N -ésima da unidade ($\alpha = e^{\frac{2\pi i}{N}w}$), o vetor $e = (1, \alpha, \alpha^2, \dots, \alpha^{N-1})$ é um autovetor de L_h .*

Demonstração. A j -ésima componente de $L_h e = h * e$ é dada por

$$\sum_{k=0}^{N-1} h_k e_{(j-k) \bmod N} = \sum_{k=0}^{N-1} h_k \alpha^{j-k}$$

já que $\alpha^N = 1 \Rightarrow \alpha^{N+j-k} = \alpha^{j-k}$. Então

$$(L_h e)_j = \left(\sum_{k=0}^{N-1} h_k \alpha^{-k} \right) \alpha^j$$

Como o somatório não depende de j , concluímos que e é um autovetor

$$L_h e = \lambda e$$

com autovalor dado pela expressão

$$\lambda = \sum_{k=0}^{N-1} h_k e^{-\frac{2\pi i k}{N} w}$$

■

Note que se α percorre todos os valores possíveis para as N -ésimas raízes da unidade, os vetores e correspondentes formam uma base ortogonal de \mathbb{C}^N . Não é difícil renormalizá-los para obter uma base ortonormal.

Lema 3.4 Os vetores $e_w = \frac{1}{\sqrt{N}} (1, \alpha, \alpha^2, \dots, \alpha^{N-1})$ onde $\alpha = e^{\frac{2\pi i}{N} w}$ (para $w = 0, 1, \dots, N-1$) são ortonormais, isto é,

$$\langle e_{w_1}, e_{w_2} \rangle = \delta_{w_1, w_2} = \begin{cases} 1, & \text{se } w_1 = w_2 \\ 0, & \text{se } w_1 \neq w_2 \end{cases}$$

Agora é simples decompor um vetor qualquer $f = (f_0, f_1, \dots, f_{N-1})$ na base $\{e_w\}_{w=0,1,\dots,N-1}$. Primeiro, vamos à expressão de cada um dos coeficientes em e_w .

Definição 3.5 A Transformada de Fourier Discreta (Discrete Fourier Transform, ou DFT) de um vetor $f = (f_0, f_1, \dots, f_{N-1})$ é dada por

$$\hat{f}_w = \langle f, e_w \rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f_k e^{-\frac{2\pi i m}{N} k}$$

Para evitar (ou criar) confusão, devemos mencionar que há 2 outras definições ligeiramente diferentes da DFT na literatura, trocando a constante na frente do somatório por $\frac{1}{N}$ ou por 1. Também, o nome DFT é um pouco infeliz mas consagrado pela literatura (gostaríamos de chamá-la de Transformada de Fourier Finita). Finalmente, note que com esta definição o autovalor λ correspondente ao autovetor e_w de L_h não é exatamente \hat{h}_w ; ao invés, temos

$$L_h e_w = \sqrt{N} \hat{h}_w e_w$$

A recomposição do vetor f é dada pela Inversa da DFT:

Corolário 3.6 (Transformada Inversa de Fourier Discreta) Um vetor f em \mathbb{C}^N pode ser recomposto como

$$f = \sum_{w=0}^{N-1} \hat{f}_w e_w$$

ou seja, sua j -ésima componente (na base canônica) é dada por

$$f_j = \frac{1}{\sqrt{N}} \sum_{w=0}^{N-1} \hat{f}_w e^{\frac{2\pi i}{N} jw}$$

A nossa escolha mantém uma certa simetria entre as definições da DFT e da IDFT. Se escolhêssemos a constante $\frac{1}{N}$ na DFT, teríamos que utilizar 1 na IDFT, e vice-versa. Preferimos a nossa definição pois, além de simétrica, vem naturalmente de uma base ortonormal.

Definição 3.7 A DFT Inversa de uma seqüência finita $g = \{g_0, g_1, \dots, g_{N-1}\}$ é denotada por

$$\check{g}_j = \frac{1}{\sqrt{N}} \sum_{w=0}^{N-1} g_w e^{\frac{2\pi i}{N} jw}$$

Intuitivamente \hat{f}_w mede o quanto da frequência $\frac{w}{N}$ está presente na seqüência finita de valores $\{f_0, f_1, \dots, f_{N-1}\}$. A título de ilustração, no caso em que N é par, note que a oscilação mais rápida que pode ser bem representada por uma tal seqüência f corresponde a $e_{\frac{N}{2}} = \frac{1}{\sqrt{N}} (1, -1, 1, -1, \dots, 1, -1)$ (frequência $\frac{N/2}{N} = \frac{1}{2}$).

Com esta definição da DFT, temos:

Afirmção 3.8 As seguintes propriedades são válidas (f e h são vetores quaisquer em \mathbb{C}^N ; fh representa a multiplicação componente a componente)

$$\begin{aligned} \hat{\delta}_0 &= \frac{1}{\sqrt{N}} (1, 1, \dots, 1) \\ \widehat{(1, 1, \dots, 1)} &= \frac{1}{\sqrt{N}} \delta_0 \\ \widehat{f *_{\text{c}} h} &= \sqrt{N} \hat{f} \cdot \hat{h} \\ \widehat{fh} &= \sqrt{N} \hat{f} *_{\text{c}} \hat{h} \\ g &= \mathfrak{S}^\alpha f \Rightarrow \hat{g}_w = e^{-\frac{2\pi i \alpha}{N} w} \hat{f}_w \text{ (em particular, } |\hat{g}_w| = |\hat{f}_w| \text{)} \end{aligned}$$

Comentário 3.9 FFT (Fast Fourier Transform) e VFFT (Very Fast Fourier Transform) não são outras transformadas de sinais – tais siglas se referem a algoritmos rápidos que calculam a DFT aqui apresentada, usualmente quando

N é potência de 2. Tais algoritmos são tão eficientes que freqüentemente é preferível calcular

$$f *_c h = \sqrt{N} \text{IFFT}(\text{FFT}(f) \cdot \text{FFT}(h))$$

através da expressão do lado direito do que a partir da definição da convolução.

Comentário 3.10 Também vale a pena notar que é possível calcular $f * h$ para sinais discretos e finitos através de uma convolução circular – basta escolher N maior do que a soma dos tamanhos dos sinais f e h , que são completados com zeros:

$$f = (f_0 \ f_1 \ \dots \ f_{L-1}) \Rightarrow \tilde{f} = (f_0 \ f_1 \ \dots \ f_{L-1} \ 0 \ 0 \ \dots \ 0) \quad (\text{tamanho } N \geq L + M - 1)$$

$$h = (h_0 \ h_1 \ \dots \ h_{M-1}) \Rightarrow \tilde{h} = (h_0 \ h_1 \ \dots \ h_{M-1} \ 0 \ 0 \ \dots \ 0) \quad (\text{idem})$$

$$\Rightarrow \tilde{f} *_c \tilde{h} = \begin{pmatrix} \underbrace{f_0 h_0}_{\text{posição } 0} & \underbrace{f_0 h_1 + f_1 h_0}_{\text{posição } 1} & \dots & \underbrace{f_{L-1} h_{M-1}}_{\text{posição } L+M-2} & 0 & 0 & \dots & 0 \end{pmatrix} \approx f * h$$

Usando um algoritmo como o FFT (em geral, escolhe-se N como uma potência de 2)

$$f * h \approx \tilde{f} *_c \tilde{h} = \sqrt{N} \text{IFFT}(\text{FFT}(\tilde{f}) \text{FFT}(\tilde{h}))$$

e a convolução (não-circular!) do lado esquerdo é calculada através de FFT's (fato que aumenta a confusão entre as diversas Transformadas de Fourier existentes).

Além do comentário acima, há uma outra relação entre a DFT e a DTFT. Partindo de um vetor $f \in \mathbb{C}^N$, monte a seqüência g estendendo f periodicamente, isto é

$$g[n] = f_{n \bmod N}$$

Então

$$\hat{g}(w) = \frac{1}{\sqrt{N}} \sum_{\alpha=0}^{N-1} \hat{f}_\alpha \delta\left(w - \frac{\alpha}{N}\right) \quad \text{para } w \in [0, 1)$$

(o lado esquerdo \hat{g} é uma DTFT; no lado direito, \hat{f}_α é uma DFT). Isto é, \hat{g} (em seu período $[0, 1)$) é a soma de impulsos localizados nas freqüências $0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}$ cujas amplitudes são $\hat{f}_0, \hat{f}_1, \hat{f}_2, \dots, \hat{f}_{N-1}$, respectivamente. Sucintamente, \hat{f} é uma amostragem de \hat{g} !

Este fato simples permite o cálculo exato da convolução não-finita de sinais discretos $g[n] = f_{n \bmod N}$ (periódico) e $h[n]$ (não necessaria-

mente periódico) via DFTs. De fato, tomando a DTFT de $g * h$:

$$\begin{aligned} \widehat{(g * h)}(w) &= \hat{g}(w) \hat{h}(w) = \left(\frac{1}{\sqrt{N}} \sum_{\alpha=0}^{N-1} \hat{f}_\alpha \delta\left(w - \frac{\alpha}{N}\right) \right) \hat{h}(w) = \\ &= \frac{1}{\sqrt{N}} \sum_{\alpha=0}^{N-1} \hat{f}_\alpha \hat{h}\left(\frac{\alpha}{N}\right) \delta\left(w - \frac{\alpha}{N}\right) \end{aligned}$$

Seja agora $v \in \mathbb{C}^N$ a IDFT (finita!) de $\hat{f}_\alpha \hat{h}\left(\frac{\alpha}{N}\right)$, isto é

$$\hat{v}_\alpha = \hat{f}_\alpha \hat{h}\left(\frac{\alpha}{N}\right) \quad \alpha = 0, 1, \dots, N-1$$

(novamente: \hat{v} e \hat{f} são DFTs enquanto \hat{h} é uma DTFT). Comparando a DFT de v com a DTFT de $g * h$ e usando novamente a propriedade supracitada

$$(g * h)[n] = v_{n \bmod N}$$

Resumindo:

Proposição 3.11 *Se g é um sinal discreto periódico (extensão periódica de f cujo tamanho é N) e h é um sinal discreto qualquer, então*

$$g * h = \text{extensão periódica do sinal finito } v$$

$$\text{onde } (DFT(v))_w = (DFT(f))_w \hat{h}\left(\frac{w}{N}\right), \quad w = 0, 1, \dots, N-1$$

sempre que $g * h$ existir.

Este processo será fundamental para o cálculo computacional de convoluções entre dois núcleos infinitos discretos – veja o apêndice sobre convoluções (pense bem: que outra maneira você teria de calculá-las?).

3.3 Transformadas Bidimensionais

A generalização da DFT para o caso bidimensional parte da escolha das dimensões $M \times N$ dos sinais a serem utilizados. A função delta de Dirac e sua propriedade de amostragem são idênticas ao caso discreto. A título de notação, usaremos $I_n = \{0, 1, \dots, n-1\}$.

Proposição 3.12 *Toda transformação linear (que leva $f : I_M \times I_N \rightarrow \mathbb{R}$ em $Lf : I_M \times I_N \rightarrow \mathbb{R}$) invariante por translações circulares em ambas as direções (de I_M e I_N) é dada por uma convolução circular*

$$Lf_{m,n} = (f *_c h)_{m,n} \stackrel{\circ}{=} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f_{j,k} h_{(m-j) \bmod M, (n-k) \bmod N}$$

onde $h = L\delta$ é a resposta de L ao impulso unitário.

Os autovetores desta convolução circular têm a forma

$$(e_{w_1, w_2})_{m, n} = \frac{1}{\sqrt{MN}} e^{2\pi i (\frac{w_1}{M} m + \frac{w_2}{N} n)} \text{ para } (w_1, w_2) \in I_M \times I_N$$

Definição 3.13 Dada uma matriz (uma imagem monocromática) $f : I_M \times I_N \rightarrow \mathbb{R}$, definimos a sua DFT como

$$\hat{f}_{w_1, w_2} = \frac{1}{\sqrt{MN}} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f_{j, k} e^{-2\pi i (\frac{w_1}{M} j + \frac{w_2}{N} k)} \text{ para } (w_1, w_2) \in I_M \times I_N$$

Sua inversa é

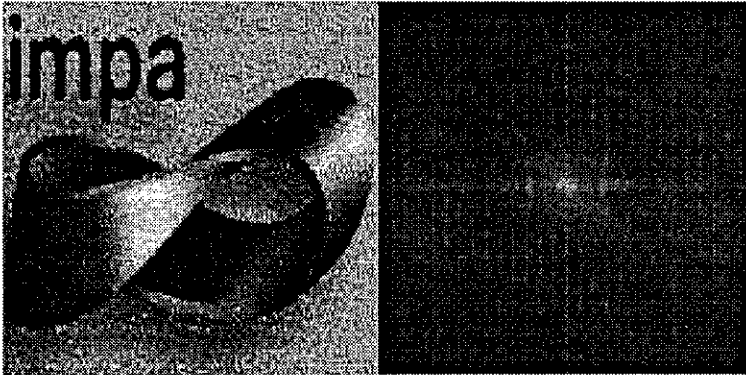
$$f_{m, n} = \frac{1}{\sqrt{MN}} \sum_{w_1=0}^{M-1} \sum_{w_2=0}^{N-1} \hat{f}_{w_1, w_2} e^{2\pi i (\frac{w_1}{M} m + \frac{w_2}{N} n)}$$

Afirmiação 3.14 Valem as propriedades (f e $h : I_M \times I_N \rightarrow \mathbb{R}; \mathbb{I}_{m, n} \equiv 1$)

$$\begin{aligned} \hat{\delta} &\equiv \frac{1}{\sqrt{MN}} \mathbb{I}; \hat{\mathbb{I}} = \frac{1}{\sqrt{MN}} \delta \\ \widehat{f *_{\text{c}} h} &= \frac{1}{\sqrt{MN}} \hat{f} \cdot \hat{h} \\ \widehat{f h} &= \frac{1}{\sqrt{MN}} \hat{f} *_{\text{c}} \hat{h} \end{aligned}$$

$$g_{m, n} = f_{(m-\alpha) \bmod M, (n-\beta) \bmod N} \Rightarrow \hat{g}_{w_1, w_2} = \hat{f}_{w_1, w_2} e^{2\pi i (w_1 \frac{\alpha}{M} + w_2 \frac{\beta}{N})}$$

Exemplo 3.15 Exibimos aqui uma imagem monocromática e a sua Transformada de Fourier. A imagem à esquerda foi renormalizada para que preto= 0 e branco= 1. O brilho da imagem à direita foi alterado para maior visibilidade e sua coordenada (0, 0) está no centro da figura.



Novamente, a seguinte afirmação é essencial para o cálculo de convoluções com imagens que são estendidas periodicamente ou por espelhamento.

Afirmção 3.16 *Seja g a extensão periódica de uma imagem f*

$$g[m, n] = f_{m \bmod M, n \bmod N}$$

então

$$(g * h)[m, n] = v_{m \bmod M, n \bmod N}$$

$$\text{onde } \hat{v}_{w_1, w_2} = \hat{f}_{w_1, w_2} \hat{h}\left(\frac{w_1}{M}, \frac{w_2}{N}\right) \quad \forall (w_1, w_2) \in I_M \times I_N$$

3.4 Exercícios

(1) Demonstre as propriedades básicas da convolução circular (Proposição 3.2) e suas versões bidimensionais.

(2) Mostre que os vetores e_w são, de fato, ortonormais (Lema 3.4).

(3) Demonstre as propriedades da DFT (Afirmções 3.8 e 3.14).

(4) Mostre a relação entre a DFT e a DTFT para sinais periódicos

$$g[n] = f_{n \bmod N} = \sum_{k=0}^{N-1} \sum_{j=-\infty}^{\infty} f_k \delta[n - (Nj + k)] \Rightarrow$$

$$\Rightarrow \hat{g}(w) = \frac{1}{\sqrt{N}} \sum_{\alpha=0}^{N-1} \hat{f}_\alpha \delta\left(w - \frac{\alpha}{N}\right) \quad \text{para } w \in [0, 1].$$

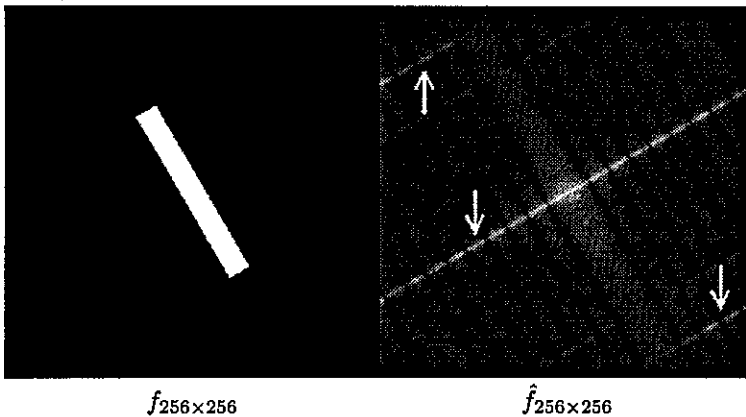
Dica: o valor da DTFT de $\mathbb{I}[n]$ indica que

$$\sum_{j=-\infty}^{\infty} e^{-2\pi i a j} = \sum_{k=-\infty}^{\infty} \delta(k - a) = \sum_{k=-\infty}^{\infty} \delta(-k - a)$$

para qualquer a ; você vai querer também usar o escalamento da função delta

$$\delta(\alpha x) = \frac{1}{|\alpha|} \delta(x)$$

(5) Considere a imagem $f_{j,k}$ a seguir, de tamanho 256×256 (branco = 1, preto = 0). A figura da direita representa a magnitude da DFT de f , cujo brilho foi alterado para maior visibilidade. A frequência $(0, 0)$ corresponde ao centro desta figura. As linhas brancas mais fortes são as indicadas pelas setas (de fato, estas eram as únicas razoavelmente visíveis antes da alteração em brilho; mais exatamente, tais “linhas” são compostas por vários segmentos concatenados):



- (a) Calcule a DFT da imagem $M \times N$ dada por $f_{j,k} = \delta[j]$. Há tais linhas?
- (b) Calcule a magnitude da DFT da imagem de um retângulo $(2m + 1) \times (2n + 1)$ com lados paralelos aos eixos, centrado na origem.
- (c) Suponha que $m \gg n$, por exemplo. Há tais linhas? Onde estão as mais fortes? Mostre que tais linhas diminuem de intensidade à medida que nos afastamos da origem do domínio da frequência.
- (d) As figuras acima parecem contradizer (c): do centro para as bordas, vemos uma linha forte (indicada pela seta central), algumas linhas fracas paralelas à central e então duas linhas mais fortes (indicadas pelas outras duas setas). Como explicar esta aparente contradição?

(6) Este exercício requer o auxílio de um programa que calcule a DFT de uma imagem bidimensional e realize operações matemáticas; qualquer bom pacote matemático terá tal capacidade.

- (a) Utilize uma imagem teste à sua escolha, digamos, $f[x, y]$. Calcule a sua DFT $\hat{f}[w_1, w_2] = r[w_1, w_2] e^{i\theta[w_1, w_2]}$ e extraia a magnitude $|\hat{f}[w_1, w_2]| = r[w_1, w_2]$ e a fase $e^{i\theta[w_1, w_2]}$. Tente reconstruir a imagem f fazendo a IDFT da magnitude, $\tilde{r}[x_1, x_2]$; você reconhece a imagem original? Tente várias formas de visualização – veja a magnitude, a fase, a parte real e a parte imaginária de \tilde{r} (que é complexa). Faça o mesmo reconstruindo a imagem como a IDFT da fase $e^{i\theta[w_1, w_2]}$. Qual das reconstruções reteve mais características da imagem original?
- (b) Agora tome duas imagens f e g de mesmo tamanho. Repita o processo do item anterior obtendo as imagens r_f , $e^{i\theta_f}$, r_g e $e^{i\theta_g}$. Agora multiplique-as para obter $r_f e^{i\theta_g}$ e $r_g e^{i\theta_f}$ e então aplique a IDFT a ambas estas imagens (não se esqueça de visualizá-las em várias formas). Você consegue reconhecer alguma das imagens originais?
- (c) De acordo com a análise acima, qual das duas componentes da Transformada de Fourier Discreta parece ser mais importante no reconhecimento de estruturas de uma imagem: a magnitude (amplitude) ou a fase?

Capítulo 4

A Função Gaussiana e a Equação do Calor

4.1 A Função Gaussiana

Definição 4.1 A função Gaussiana (também conhecida como distribuição normal) de variância σ^2 é a função dada por

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}x^2}$$

O parâmetro σ é costumeiramente chamado *desvio-padrão* de $G(x)$. Como $G(x)$ é uma função par ($G(x) = G(-x)$), é imediato verificar que sua média é $\mu = 0$. Antes de prosseguir, devemos mostrar que $G(x)$ é uma função normalizada.

Proposição 4.2 Se $G(x)$ é dada pela fórmula acima, então

$$\int_{-\infty}^{+\infty} G(x) dx = 1$$

Demonstração. Seja A o valor da integral acima. Um truque interessante é passar o problema para o cálculo de duas variáveis

$$\begin{aligned} A^2 &= \left(\int_{-\infty}^{+\infty} G(x) dx \right) \left(\int_{-\infty}^{+\infty} G(y) dy \right) = \\ &= \int_{\mathbb{R}^2} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x^2+y^2)} dA \end{aligned}$$

e então usar coordenadas polares

$$\begin{aligned} A^2 &= \int_{\mathbb{R}^2} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}r^2} r \, dr \, d\theta = \\ &= \frac{1}{2\pi\sigma^2} \int_0^{2\pi} \int_0^{+\infty} r e^{-\frac{1}{2\sigma^2}r^2} \, dr \, d\theta = \\ &= \frac{1}{2\pi\sigma^2} (2\pi) \left[-\sigma^2 e^{-\frac{1}{2\sigma^2}r^2} \right]_{r=0}^{\infty} = 1 \end{aligned}$$

Como claramente $A > 0$, temos $A = 1$. ■

A partir de uma função Gaussiana, cria-se um filtro Gaussiano $L_G f = f * G$. Filtros Gaussianos são muito utilizados em processamento de sinais pois eles nos dão filtros de passa-baixa que não geram oscilações, como veremos a seguir.

Lema 4.3 *Dados b e c reais, $c > 0$*

$$\int_{-\infty+bi}^{\infty+bi} e^{-cz^2} \, dz = \int_{-\infty}^{+\infty} e^{-cx^2} \, dx$$

Demonstração. Este fato pode ser demonstrado usando fatos de cálculo de duas variáveis (Teorema de Green); preferimos uma demonstração análoga mas mais sucinta por utilizar o jargão de Variáveis Complexas: considere no plano complexo o retângulo R com vértices $-a$, a , $a+bi$, $-a+bi$. Como e^{-cx^2} é analítica dentro deste retângulo, tem-se

$$\int_R e^{-cx^2} \, dx = 0$$

Mas

$$\begin{aligned} \left| \int_a^{a+bi} e^{-cz^2} \, dz \right| &= \left| \int_0^b e^{-c(a+ti)^2} \, dt \right| \leq \int_0^b \left| e^{-c(a+ti)^2} \right| \, dt = \\ &= \int_0^b e^{c(t^2-a^2)} \, dt = e^{-ca^2} \int_0^b e^{ct^2} \, dt \leq e^{-ca^2} (be^c) \end{aligned}$$

O mesmo pode ser feito para o outro lado vertical do retângulo; assim

$$\left| \int_{-a}^a e^{-cx^2} \, dx - \int_{-a+bi}^{a+bi} e^{-cz^2} \, dz \right| \leq 2be^{-ca^2} e^c$$

Tome $a \rightarrow +\infty$ e o resultado está demonstrado. ■

Proposição 4.4 *A Transformada de Fourier de um filtro Gaussiano de variância σ^2 é dada por*

$$\hat{G}(w) = e^{-2\sigma^2\pi^2 w^2}$$

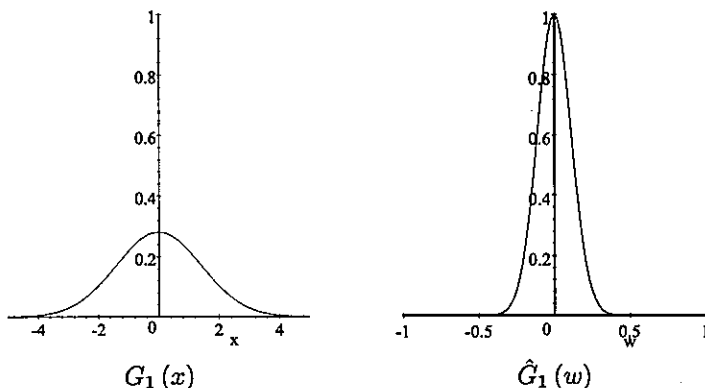
Demonstração. Basta completar os quadrados no expoente para obter

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}x^2} e^{-2\pi iwx} dx &= \\ &= \left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x+2\sigma^2\pi iw)^2} dx \right) e^{2\sigma^2 \cdot (-\pi^2 w^2)} = \\ &= \left(\int_{-\infty+4\pi iw\sigma^2}^{+\infty+4\pi iw\sigma^2} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}x^2} dx \right) \cdot e^{-2\pi^2\sigma^2 w^2} \end{aligned}$$

Pelo lema acima, a integral à esquerda pode ser calculada de $-\infty$ a ∞ ; portanto, pela proposição 4.2, ela vale 1; então

$$\hat{G}(w) = e^{-2\pi^2\sigma^2 w^2}$$

Note que \hat{G} é, a menos de uma constante de proporcionalidade, a Gaussiana de variância $(\frac{1}{2\pi\sigma})^2$. Seguem abaixo os gráficos de $G(x)$ e $\hat{G}(w)$ para $\sigma^2 = 2$ (note que as escalas dos gráficos são diferentes):



Corolário 4.5 A Transformada bilateral de Laplace de $G(x)$ é

$$\mathcal{L}_G(s) = e^{\frac{1}{2}\sigma^2 s^2}$$

Corolário 4.6 A variância de G é, de fato, σ^2 , isto é

$$\mathcal{L}_G''(0) = \sigma^2$$

Por comodidade, daqui por diante utilizaremos $\sigma^2 = 2t$ para obter a notação

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{1}{4t}x^2}$$

e portanto

$$\begin{aligned}\hat{G}_t(w) &= e^{-4\pi^2 w^2 t} \\ \mathcal{L}_{G_t}(s) &= e^{s^2 t}\end{aligned}$$

Vale notar que o núcleo G_t é conhecido em Equações Diferenciais Parciais como *núcleo de Poisson* ou *núcleo do calor*. Repetindo: G_t nada mais é do que um núcleo Gaussiano de variância $2t$. O motivo desta escolha de representação ficará claro em breve.

Proposição 4.7 (Semi-grupo) *A convolução de dois núcleos Gaussianos de variâncias σ_1^2 e σ_2^2 produz um núcleo Gaussiano de variância $\sigma_1^2 + \sigma_2^2$, isto é*

$$G_{t_1} * G_{t_2} = G_{t_1+t_2}$$

Demonstração. Basta notar que

$$\mathcal{L}_{G_{t_1}}(s)\mathcal{L}_{G_{t_2}}(s) = e^{s^2(t_1+t_2)} = \mathcal{L}_{G_{t_1+t_2}}(s)$$

▪

Finalmente, cálculos que deixamos ao leitor mostram que:

Afirmção 4.8 *O núcleo Gaussiano $G_t(x)$ satisfaz*

$$\begin{aligned}\frac{\partial G}{\partial t} &= \frac{\partial^2 G}{\partial x^2} \\ \lim_{t \rightarrow 0} G_t(x) &= \delta(x)\end{aligned}$$

4.2 A Equação do Calor

Apesar da nomenclatura ser consagrada, a equação do calor descreve não só a maneira como calor se difunde em um meio homogêneo, mas serve para a difusão de qualquer substância que obedeça a Lei de Difusão de Fick:

Afirmção 4.9 (Lei de Difusão de Fick) *A difusão de uma substância A em uma substância B se propaga da região de maior densidade para as regiões de menor densidade; a taxa de difusão é proporcional ao gradiente da concentração de A em B.*

Em outras palavras, se v é a velocidade de propagação da substância A e ρ é a densidade de massa de A em B, então a Lei de Fick diz que

$$v = -K \nabla \rho$$

para alguma constante positiva K . No caso unidimensional, isto se resume a

$$v(x) = -K \frac{d\rho}{dx}$$

Queremos obter uma equação diferencial que, a partir desta lei, nos diga a variação da densidade ρ com o tempo. Para tanto, seja $M(x, t)$ a quantidade de massa no intervalo $I = [x_0, x]$ (para algum x_0 fixo) no tempo t . Uma simples lei de conservação de massa relaciona v com esta quantidade

$$\frac{\partial M}{\partial t} = -v(x) + v(x_0)$$

Mas a densidade é definida como $\rho(x, t) = \frac{\partial M}{\partial x}(x, t)$, portanto

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2 M}{\partial x \partial t} = \frac{\partial (-v(x) + v(x_0))}{\partial x} = -\frac{\partial v}{\partial x}$$

Esta equação é simplesmente a lei de conservação de massa. Se agora utilizarmos a Lei de Fick, obtemos a seguinte equação para ρ

$$\frac{\partial \rho}{\partial t}(x, t) = K \frac{\partial^2 \rho}{\partial x^2}(x, t)$$

Se soubermos a distribuição inicial de massa $\rho(x, 0)$, é possível então resolver esta equação para descobrir como ρ varia com o tempo.

De fato, a equação diferencial acima foi estabelecida para a condução de calor em meios homogêneos por J. Fourier antes da sua interpretação como um processo de difusão explicitada acima (a analogia entre os processos de transporte de calor e transporte de matéria é que foi estabelecida por A. Fick, em 1855).

Definição 4.10 A equação do calor *unidimensional* (com condição inicial) é a seguinte equação diferencial parcial parabólica

$$\begin{aligned} \frac{\partial F(x, t)}{\partial t} &= \frac{\partial^2 F(x, t)}{\partial x^2} \\ F(x, 0) &= f_0(x) \end{aligned}$$

Tal EDP também é conhecida como *equação de difusão*. Note que tomamos $K = 1$ (o que fisicamente significa apenas uma mudança de unidades, e não afetará o comportamento da solução).

Proposição 4.11 A solução da equação do calor unidimensional

$$\begin{aligned} \frac{\partial F}{\partial t} &= \frac{\partial^2 F}{\partial x^2} \\ F(x, 0) &= f_0(x) \end{aligned}$$

é dada pela convolução

$$F(x, t) = G_t(x) * f_0(x)$$

Demonstração. Há várias maneiras equivalentes de resolver a equação do calor: pode-se simplesmente verificar a solução acima (o que é imediato, usando a afirmação 4.8); outra opção é observar que a transformação que leva $F(x, 0)$ em $F(x, t)$ tem de ser linear e invariante por translações, e portanto dada por uma convolução a ser determinada; aqui, preferimos mostrar como encontrar a solução aplicando a Transformada de Fourier (em x) a ambas equações para obter uma equação em $\hat{F}(w, t)$

$$\frac{\partial \hat{F}}{\partial t} = (2\pi iw)^2 \hat{F}$$

$$\hat{F}(w, 0) = \hat{f}_0(w)$$

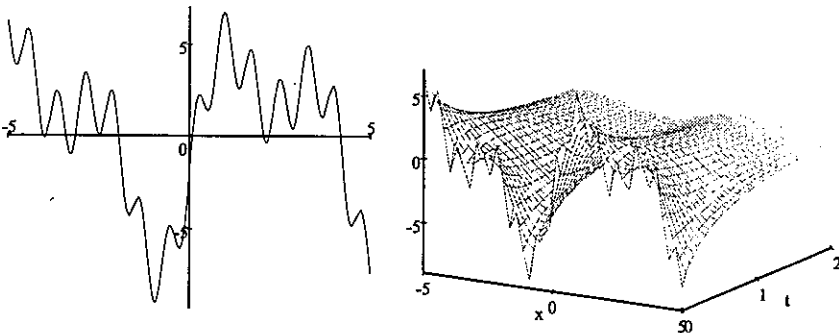
então

$$\hat{F}(w, t) = \hat{f}_0(w) e^{-4\pi^2 w^2 t}$$

Assim, vê-se que $F(x, t)$ é dada por uma convolução de f_0 com a Transformada de Fourier Inversa de $e^{-4\pi^2 w^2 t}$; esta poderia ser calculada pela fórmula da IFT, mas já sabemos que é $G_t(x)$; então

$$F(x, t) = f_0(x) * G_t(x)$$

A figura a seguir mostra o gráfico de um sinal $f_0(x)$ e da correspondente solução da equação do calor $F(x, t) = f_0(x) * G_t(x)$:



A condição inicial $f_0(x)$...

...e $F(x, t) = G_t(x) * f_0(x)$.

Note um dos efeitos da aplicação da equação de difusão a um sinal inicial f_0 ; a difusão parece fazer com que F fique cada vez mais simples à medida que t aumenta, tendendo a uma situação estacionária (“temperatura constante”); assim, $F(x, t)$ parece perder estrutura à medida que t aumenta.

4.3 Caso Bidimensional

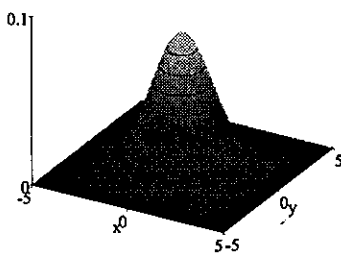
A extensão da Gaussiana ao caso bidimensional é dada pela fórmula

$$G_t(x, y) = G_t(x) G_t(y) = \frac{1}{4\pi t} e^{-\frac{1}{4t}(x^2+y^2)}$$

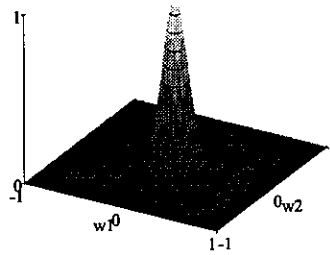
Note que a Gaussiana assim definida é uma função separável e rotacionalmente simétrica (de fato, pode-se mostrar que a Gaussiana é praticamente a única função com tais propriedades; vide exercícios). Assim, é fácil calcular a sua Transformada de Fourier.

Corolário 4.12

$$\hat{G}_t(w_1, w_2) = \hat{G}_t(w_1) \hat{G}_t(w_2) = e^{-4\pi^2(w_1^2+w_2^2)t}$$

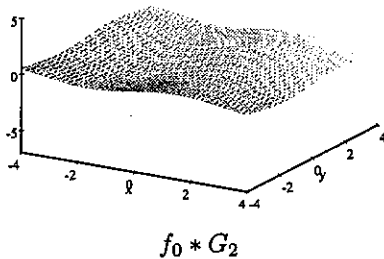
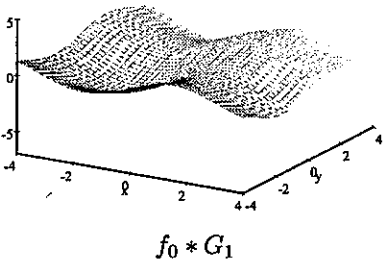
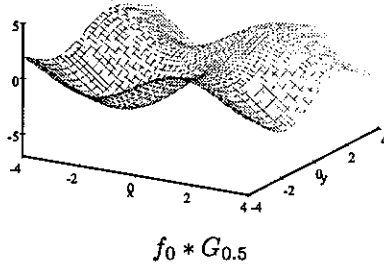
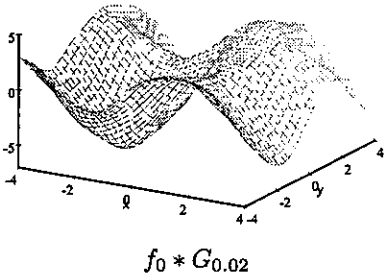
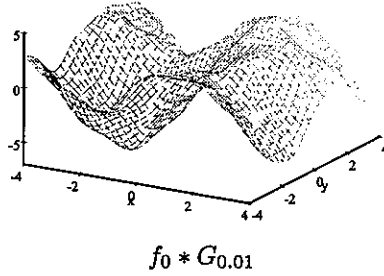
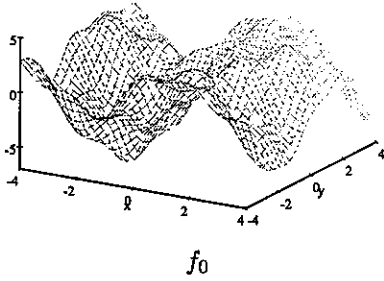


$G_1(x, y)$



$\hat{G}_1(w_1, w_2)$

Por estes motivos, o filtro Gaussiano $L_{G_t} : f_0 \rightarrow f_0 * G_t$ é um dos mais utilizados como filtro passa-baixa em processamento de imagens (em geral com desvio-padrão de no máximo alguns pixels – o valor exato de t é frequentemente obtido por experimentação, sem uma razão mais forte). A convolução com a Gaussiana tende a eliminar rapidamente ruídos de alta frequência, como evidenciado nas figuras abaixo (lembre-se que $t = 2\sigma^2$ não é exatamente o desvio-padrão).



Corolário 4.13 *O núcleo Gaussiano bidimensional também tem a propriedade de semi-grupo*

$$G_{t_1} * G_{t_2} = G_{t_1+t_2}$$

Afirmção 4.14 *O núcleo Gaussiano $G_t(x, y)$ satisfaz*

$$\frac{\partial G}{\partial t} = \nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2}$$

Afirmção 4.15 A solução $F(x, y; t)$ da equação do calor bidimensional

$$\frac{\partial F}{\partial t} = \nabla^2 F$$

$$F(x, y; 0) = f_0(x, y)$$

(onde o Laplaciano é calculado apenas nas variáveis x e y) é dada pela convolução

$$F(x, y; t) = G_t(x, y) * f_0(x, y)$$

Assim, um exemplo da equação do calor em ação é dado pelas figuras acima – basta interpretar os índices dos filtros Gaussianos como tempo de difusão. Portanto, as figuras mostram $F(x, y; t)$ para $t = 0, 0.01, 0.02, 0.5, 1$ e 2 .

Finalmente, a extensão ao caso multidimensional é imediata; basta usar

$$G_t(\mathbf{x}) = \frac{1}{(4\pi t)^{\frac{N}{2}}} e^{-\frac{1}{4t}|\mathbf{x}|^2}$$

onde \mathbf{x} é um vetor em \mathbb{R}^N . Então, a equação do calor multidimensional

$$\frac{\partial F}{\partial t} = \nabla^2 F$$

$$F(\mathbf{x}; 0) = f_0(\mathbf{x})$$

tem solução dada por

$$F(\mathbf{x}; t) = G_t(\mathbf{x}) * f_0(\mathbf{x})$$

4.4 Exercícios

(1) Vamos mostrar que as funções da forma $Ce^{\alpha(x^2+y^2)}$ são as únicas funções rotacionalmente simétricas que podem ser decompostas como um produto tensorial de funções de uma variável. De fato:

(a) Se escrevermos f em coordenadas polares $f(x, y) = g(r, \theta)$, então f é rotacionalmente simétrica $\frac{\partial g}{\partial \theta} = 0$; conclua que toda função rotacionalmente simétrica tem que satisfazer

$$\frac{1}{x} \frac{\partial f}{\partial x} = \frac{1}{y} \frac{\partial f}{\partial y}$$

(b) Descubra que se $f(x, y) = f_1(x) f_2(y)$ então

$$\frac{df_1}{dx} = cx f_1(x) \text{ e } \frac{df_2}{dy} = cy f_2(y)$$

para a mesma constante c .

(c) Resolva estas equações diferenciais e mostre que f é da forma $Ce^{\alpha(x^2+y^2)}$.

(2) Neste capítulo, encontramos a Transformada de Fourier da Gaussiana

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{1}{4t}x^2} \Leftrightarrow \hat{G}_t(w) = e^{-4\pi^2 w^2 t}$$

usando argumentos de variáveis complexas. Vamos aqui descobrir duas outras maneiras de chegar ao mesmo resultado.

(a) Mostre diretamente que a Gaussiana G_t satisfaz a equação do calor

$$\frac{\partial G_t}{\partial t} = \frac{\partial^2 G_t}{\partial x^2}$$

aplique a Transformada de Fourier a ambos os lados da equação e use que $G_0 = \delta$ para obter \hat{G}_t .

(b) Encontre todas as soluções de

$$\begin{aligned} f'(x) &= K_1 x f(x) \\ f(0) &= K_2 \end{aligned}$$

e conclua que a Gaussiana G_t é a única solução da equação

$$\frac{\partial G_t(x)}{\partial x} = -\frac{x}{2t} G_t(x)$$

a menos de uma constante de multiplicidade. Aplique a Transformada de Fourier (ou Laplace) na variável x a ambos os lados da EDO acima e então obtenha \hat{G}_t (você vai precisar de $\hat{G}_t(0) = ?$).

Parte II

Espaço de Escala Gaussiano

Capítulo 5

Espaço de Escala Gaussiano

Como vimos no capítulo 4, a convolução de um sinal com o núcleo Gaussiano tende a eliminar estruturas de escala menor, dependendo da variância escolhida. Uma pergunta comum em processamento de sinais é a seguinte: que variância escolher para a Gaussiana a ser utilizada numa determinada suavização de um sinal? O quanto devemos suavizar nosso sinal (ou imagem) de maneira a eliminar estruturas pequenas e pouco importantes, sejam elas ruído ou textura ou mesmo objetos de pequeno tamanho que desejamos ignorar?

Como discutido no capítulo introdutório, um sistema genérico de visão computacional não tem informação alguma *a priori* sobre as escalas importantes presentes na imagem de entrada. A solução natural é realizar o processamento de baixo nível *simultaneamente em todas as escalas*, deixando a decisão sobre as escalas a serem ignoradas para os módulos de alto-nível (mais “inteligentes”) do nosso sistema de visão.

Assim, nos parece natural definir como o espaço de escala de um sinal a composição deste sinal convoluído com Gaussianas de todas as variâncias possíveis. Esse processo aumenta a dimensão da nossa estrutura de dados em 1: um sinal unidimensional gera um espaço de escala bidimensional, uma imagem gera um espaço de escala tridimensional, e assim por diante. A dimensão extra é a escala.

Devemos notar que o objetivo da criação de um espaço de escala não é somente o de suavizar um sinal em diversas quantidades. Queremos também, se possível, aproveitar as estruturas do sinal vistas *através de escalas distintas*. Por exemplo, uma aresta detectada numa escala maior será trazida para escalas menores para que se encontre sua verdadeira localização na imagem original; uma bolha multi-escala (a ser definida no capítulo 14) terá importância determinada por seu volume através de várias escalas.

Mais tarde, veremos que é possível definir espaços de escala usando outros núcleos de convolução ligeiramente diferentes do Gaussiano (capítulo 10); mais ainda, é possível criar estruturas semelhantes a espaços de escala que não vêm de convoluções. Por este motivo, apresentaremos uma definição para o espaço de escala Gaussiano que envolve a equação do calor ao invés da convolução com a Gaussiana; isto pode parecer um pouco menos intuitivo, mas veremos que é

convenientemente generalizável.

Por simplicidade, começaremos nosso estudo do espaço de escala Gaussiano com sinais unidimensionais. Procuraremos então descrever matematicamente algumas das boas propriedades desse espaço de escala, criando princípios que nos guiarão através do livro. Só então estenderemos a teoria para imagens; a partir daí, a extensão para sinais multidimensionais é imediata.

5.1 Caso Unidimensional

Definição 5.1 Dado um sinal $f_0 : \mathbb{R} \rightarrow \mathbb{R}$, definimos o espaço de escala Gaussiano de f como a função $F : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ (representada por $F(x, t) = F_t(x)$) que é solução da equação do calor

$$\begin{aligned} \frac{\partial F_t(x)}{\partial t} &= \frac{\partial^2 F_t(x)}{\partial x^2} \\ F_0(x) &= f_0(x) \end{aligned} \tag{5.1}$$

Corolário 5.2 O espaço de escala de f pode ser obtido através de convoluções com Gaussianas

$$F_t(x) = G_t(x) * f_0(x)$$

Veja a figura da página 48 para um exemplo de um espaço de escala de um sinal unidimensional. O espaço de escala Gaussiano tem as seguintes propriedades básicas:

- **Linearidade:** a transformação que leva o sinal original $f_0(x)$ ao espaço de escala $F_t(x)$ é linear;
- **Invariância por translações:** se \mathfrak{S} é uma translação qualquer, o espaço de escala de $\mathfrak{S}f_0$ é $\mathfrak{S}F$, pois

$$K(x) * \mathfrak{S}f_0(x) = \mathfrak{S}(K(x) * f_0(x));$$

para qualquer núcleo de convolução K (no caso, $K = G_t$);

- **Causalidade:** O sinal f_0 é simplificado com a escala.

As duas primeiras propriedades vêm do fato que a equação do calor é linear e todos os seus termos são invariantes por translações¹. Como vimos no capítulo 1, linearidade e invariância por translações equivalem a exigir que F_t seja gerado por uma convolução de um núcleo K_t (que pode depender da escala t) com o sinal original f_0 .

Já a terceira propriedade, uma característica essencial do espaço de escala Gaussiano, não está ainda bem definida. Para entendê-la melhor (e também para classificar outros tipos de espaços de escala quando eles aparecerem), devemos formalizá-la.

¹De fato, esta última propriedade vale para qualquer equação diferencial parcial envolvendo apenas constantes, F e suas derivadas parciais; veja capítulo 10.

Definição 5.3 Um cruzamento de zero (CZ) de uma função contínua $f(x)$ é um intervalo fechado $[a, b]$ (possivelmente com $a = b$) tal que

$$f([a, b]) = 0 \\ \lim_{x \rightarrow a_-} \text{sinal } f(x) = - \lim_{x \rightarrow b_+} \text{sinal } f(x) \neq 0$$

A definição acima pode parecer um tanto críptica mas sua explicação é simples: conta-se um cruzamento de zero para $f(x)$ cada vez que o gráfico de f cruze o eixo Ox . Tal cruzamento pode ocorrer num ponto (se $f(x_0) = 0$ e $f'(x_0) \neq 0$ isto definitivamente acontece; tome $a = b = x_0$) ou num intervalo (pode ser que f seja não-nula por algum tempo, alcance 0 em a e permaneça ali até b , só então mudando de sinal; conta-se um CZ também neste caso). Por exemplo, $f(x) = x^2$ não tem CZ's enquanto $f(x) = x^3$ tem um CZ em $x = 0$. Se $f(x) \geq 0$ em \mathbb{R} , f não pode ter cruzamentos de zero (mesmo que f tenha zeros).

Proposição 5.4 (Causalidade CZ) Dada $f_0(x)$ contínua qualquer, considere o seu espaço de escala Gaussiano $F_t(x)$; o número de cruzamentos de zero (CZ) de $F_t(x)$ não aumenta à medida que t cresce.

Demonstração. Veja capítulo 10 para uma discussão mais detalhada desta propriedade e referências. ■

Corolário 5.5 Se f_0 é diferenciável, então o número de máximos e mínimos de $F_t(x)$ não aumenta à medida que t cresce.

Demonstração. De fato, máximos e mínimos (com relação a x) de $F_t(x) = G_t(x) * f_0(x)$ são CZ de

$$\frac{\partial F_t}{\partial x} = G_t(x) * \frac{\partial f_0}{\partial x}$$

que é o espaço de escala de $f_0'(x)$. ■

Em particular, se os detalhes que procuramos forem dados por CZ de expressões lineares envolvendo as derivadas espaciais de F_t , esta propriedade nos garante que o número de detalhes diminui (ou, pelo menos, não aumenta) à medida que aumentamos a escala! Esta idéia será explorada no capítulo 8.

No entanto, as duas propriedades acima são muito restritivas; em breve, veremos que o espaço de escala Gaussiano não as satisfaz no caso multidimensional. Uma propriedade mais fraca que também reflete a simplificação do sinal é dada pelo princípio do máximo.

Proposição 5.6 (Princípio do Máximo) Se $F_t(x)$ tem um máximo local em $x = x_0$ (para t fixo), então

$$\frac{\partial F_t(x_0)}{\partial t} \leq 0$$

Se x_0 é um mínimo local

$$\frac{\partial F_t(x_0)}{\partial t} \geq 0$$

Demonstração. A condição de máximo (mínimo) local claramente indica que $\frac{\partial^2 F_t}{\partial x^2}(x_0) \leq 0$ (ou ≥ 0). Como F_t satisfaz a equação do calor, a propriedade está demonstrada. ■

De fato, o princípio do máximo clássico diz que $F_t(x_0)$ desce sempre que x_0 for estritamente um máximo local ou fizer parte de um platô máximo local (mesmo que $\frac{\partial F_t}{\partial t}(x_0) = 0$ – o leitor que quiser uma abordagem rigorosa pode consultar [38]); assim, máximos locais “descem” e mínimos locais “sobem” à medida que a escala cresce².

Da figura da página 48, é possível vislumbrar duas outras propriedades do espaço de escala Gaussiano: a suavização quase que instantânea do sinal original e a propriedade de semi-grupo, descritas a seguir.

Proposição 5.7 $F_t(x)$ é uma função suave (infinitamente diferenciável) para qualquer $t > 0$ fixo.

Demonstração. De fato, note que

$$\frac{\partial^n F_t}{\partial x^n} = \frac{\partial^n}{\partial x^n} (G_t(x) * f_0(x)) = \frac{\partial^n G_t(x)}{\partial x^n} * f_0(x)$$

que existe pois a função Gaussiana $G_t(x)$ é suave para qualquer $t > 0$. ■

Proposição 5.8 A partir de um sinal $f_0(x)$ montamos o espaço de escala Gaussiano $F_t(x)$. Então o espaço de escala de $h_0(x) = F_{t_0}(x)$ é dado por $H_t(x) = F_{t+t_0}(x)$.

Demonstração. Basta utilizar a propriedade de semi-grupo da função Gaussiana

$$\begin{aligned} H_t(x) &= G_t(x) * h_0(x) = G_t(x) * (G_{t_0}(x) * f_0(x)) = \\ &= G_{t+t_0}(x) * f_0(x) = F_{t+t_0}(x) \end{aligned}$$

²Eis aqui o princípio do máximo unidimensional em sua generalidade: suponha que $F(x, t)$ satisfaz

$$\frac{\partial F}{\partial t} \leq a(x, t) \frac{\partial^2 F}{\partial x^2} + b(x, t) \frac{\partial F}{\partial x}$$

em um domínio E do plano x, t onde $a(x, t)$ e $b(x, t)$ são limitadas e $a(x, t) \geq \mu > 0$ para alguma constante μ . Suponha que o máximo de F em E é atingido num ponto interior $P(x_0, t_0)$. Se Q é um ponto de E que pode ser conectado a P por um caminho de segmentos horizontais ou verticais subindo (isto é, com t crescente), então $F(Q) = F(P)$. Em particular, se $E(t_0)$ é a reta ou segmento horizontal ($t = t_0$) dentro de E que contém P , $F(E(t_0)) = F(P)$. Estes resultados são válidos também se P estiver na componente $E(t_0)$ da fronteira no topo de E (isto é, tal que $(x_0, t_0 - \varepsilon) \in E$ para ε pequeno positivo), desde que F , $\frac{\partial F}{\partial x}$, $\frac{\partial^2 F}{\partial x^2}$ e $\frac{\partial F}{\partial t}$ sejam contínuos em $E \cup E(t_0)$.

A demonstraco acima   um tanto desnecess ria se voltarmos   equaco do calor: a propriedade simplesmente diz que “jogar $f_0(x)$ na m quina do espaço de escala, esperar um tempo t_0 e ent o esperar mais t_1   o mesmo que esperar $t_0 + t_1$ desde o começo do processo”. De fato, qualquer processo que seja expresso por uma equaco diferencial do tipo $\frac{\partial F_t}{\partial t} = \mathcal{F} \left\{ F, \frac{\partial F}{\partial x}, \frac{\partial^2 F}{\partial x^2}, \dots \right\}$ ter  esta propriedade.

5.2 Caso Bidimensional

Definiç o 5.9 Dado uma imagem cont nua $f_0 : \mathbb{R}^2 \rightarrow \mathbb{R}$, definimos o espaço de escala Gaussiano de f_0 como a funç o $F : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$ (denotada $F_t(x, y)$ onde $t \in \mathbb{R}_+$) que   soluç o da equaco do calor bidimensional

$$\frac{\partial F_t(x, y)}{\partial t} = \nabla^2 F_t(x, y) = \frac{\partial^2 F_t(x, y)}{\partial x^2} + \frac{\partial^2 F_t(x, y)}{\partial y^2} \quad (5.2)$$

$$F_0(x, y) = f_0(x, y)$$

De maneira similar ao caso unidimensional, sabemos que

Afirmaç o 5.10 A soluç o de 5.2 pode ser expressa como uma convoluç o com Gaussianas bidimensionais

$$F_t(x, y) = G_t(x, y) * f_0(x, y)$$

Coment rio 5.11 (Separabilidade) A convoluç o acima pode ser calculada atrav s de duas convoluç es com Gaussianas unidimensionais

$$F_t(x, y) = G_t(y) *_y (G_t(x) *_x f_0(x, y))$$

onde cada convoluç o   feita apenas na vari vel indicada. Esta propriedade facilitar  a computaç o dos espaços de escala Gaussianos para imagens.

A figura 5.1 na p gina 60 apresenta o espaço de escala F_t de uma imagem para nove valores positivos de t : 0.25, 0.5, 1, 2, 4, 8, 16 e 32.

Al m da linearidade e da invari ncia por translaç es, o espaço de escala Gaussiano bidimensional tem uma outra vantagem: a invari ncia por rotaç es do dom nio.

Proposiç o 5.12 Seja f_0 uma imagem qualquer e $g_0 = R_\theta f_0$ correspondente   rotaç o de f_0 de  ngulo θ . Ent o o espaço de escala G_t (de g_0)   a rotaç o de  ngulo θ de F_t (espaço de escala de f_0), isto  

$$G_t = R_\theta F_t$$

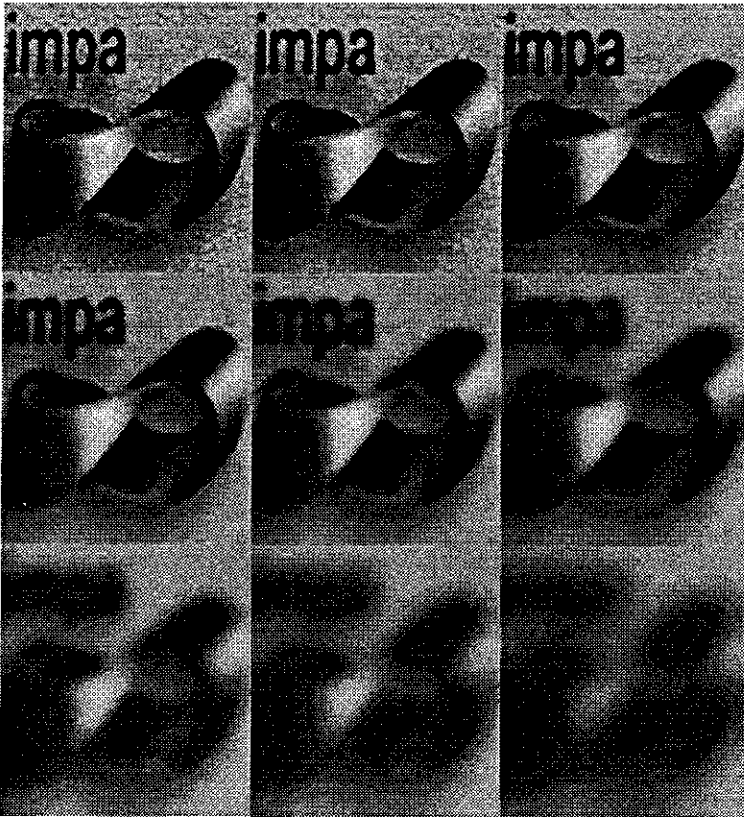


Figura 5.1: Espaço de escala do logotipo do IMPA.

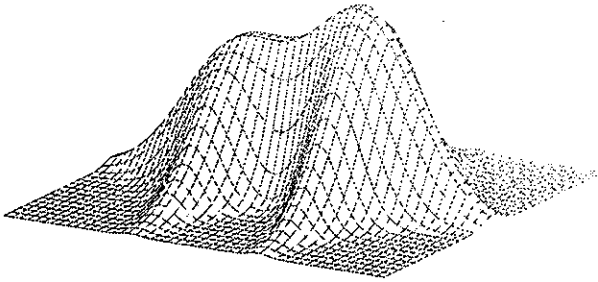
Demonstração. De fato, como convoluções são invariantes por rotações

$$\begin{aligned}
 R_\theta(F_t) &= R_\theta(G_t(x, y) * f_0(x, y)) = \\
 &= R_\theta(G_t(x, y)) * R_\theta(f_0(x, y)) = \\
 &= R_\theta(G_t(x, y)) * g_0(x, y)
 \end{aligned}$$

e a proposição segue imediatamente já que a Gaussiana bidimensional é rotacionalmente simétrica ($R_\theta G_t = G_t$). ■

Tentemos agora entender o que será o princípio da causalidade em 2D. Note que não faz sentido falar em número de cruzamentos de zero de uma imagem, já que em geral os cruzamentos de zero de uma imagem formam um conjunto de curvas, não um conjunto discreto de pontos.

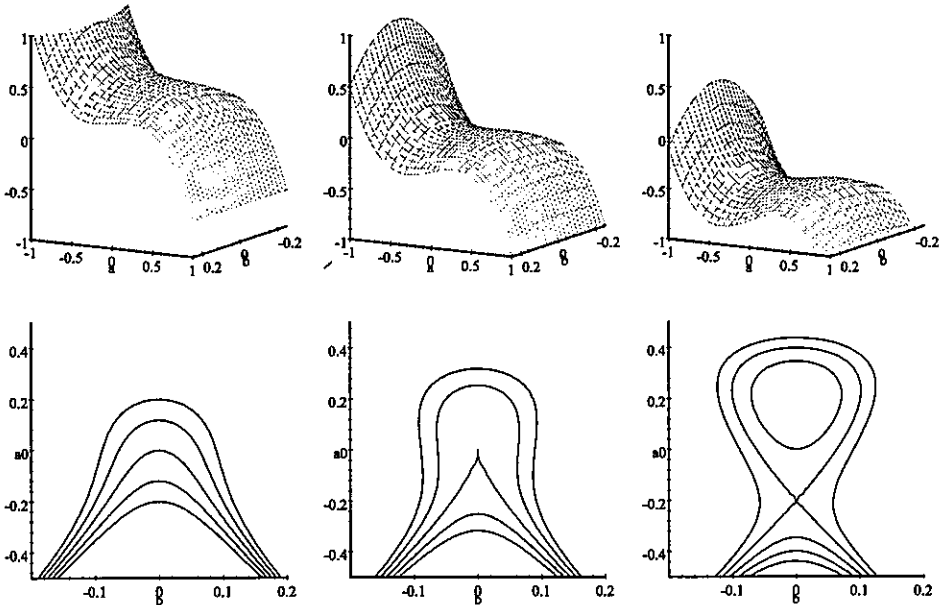
Por outro lado, pode-se falar do número de máximos e mínimos locais de uma imagem genérica (ou de um sinal N -dimensional). No entanto, *não é verdade* que o número de pontos críticos diminua com a escala no espaço de escala de uma imagem qualquer.



Exemplo 5.13 Tome uma imagem cujo gráfico consiste de dois morros, um mais alto do que o outro, conectados por uma fina crista (veja figura acima) – a crista evita o que seria um segundo ponto de máximo no morro mais baixo. Aplicando a equação do calor a esta imagem, a crista desce muito mais rápido que os morros, e um segundo pico se forma no morro mais baixo. Mais explicitamente, uma função que exhibe este comportamento é

$$F_t(x, y) = -x^3 + 4xy^2 - 4y^2 - 8t + 2xt$$

O leitor pode verificar que F_t satisfaz a equação do calor e tem pontos críticos em $(\pm\sqrt{\frac{2t}{3}}, 0)$ e $(1, \pm\sqrt{\frac{3-2t}{2}})$, sendo que apenas $(\sqrt{\frac{2t}{3}}, 0)$ é máximo local (os outros são selas). No entanto, note que este ponto de máximo não existe para $t < 0$ e portanto é criado com o aumento de escala. Abaixo estão os gráficos de F_t perto da origem para $t = -0.06$, $t = 0$ e $t = 0.06$



Assim, para sinais N -dimensionais, somos praticamente forçados a perseguir apenas a versão mais fraca do princípio da causalidade: o princípio do máximo.

Proposição 5.14 *Se $F_t(x, y)$ tem um mínimo (máximo) local em (x_0, y_0) (para t fixo), então*

$$\frac{\partial F_t}{\partial t}(x_0, t_0) \geq 0 \quad (\leq 0)$$

Demonstração. A condição de mínimo local indica que a matriz das segundas derivadas

$$\begin{pmatrix} \frac{\partial^2 F_t}{\partial x^2} & \frac{\partial^2 F_t}{\partial x \partial y} \\ \frac{\partial^2 F_t}{\partial y \partial x} & \frac{\partial^2 F_t}{\partial y^2} \end{pmatrix}$$

é positiva definida. Em particular, todas segundas derivadas contendo apenas uma variável

$$\frac{\partial^2 F}{\partial x^2}, \frac{\partial^2 F}{\partial y^2}$$

são positivas. Como $\nabla^2 F$ é exatamente a soma dessas derivadas, $\nabla^2 F \geq 0$. Se F satisfaz a equação do calor, a propriedade está demonstrada. ■

Vale a pena notar que a propriedade de semi-grupo ainda vale; pode-se deduzi-la a partir da equação diferencial, ou a partir da propriedade de semi-grupo das Gaussianas.

Finalmente, o caso multidimensional é uma extensão natural dos resultados obtidos nesta seção. Novamente, o único princípio de causalidade que pode ser perseguido é o princípio do máximo – os outros falham (ou não fazem sentido) no caso de espaços de escala Gaussianos.

5.3 Exercícios

(1) Encontre os espaços de escala das seguintes funções (sugestão: pense nos autovetores da convolução com G_t ; quais são os autovalores correspondentes?)

(a) $f_0(x) = \cos ax$

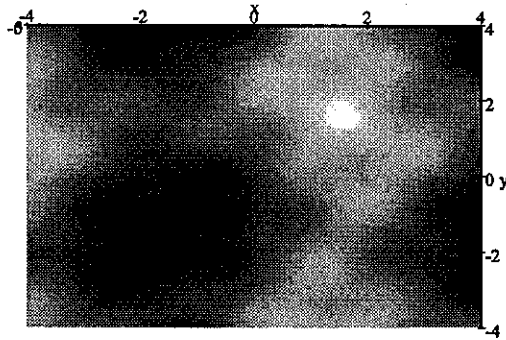
(b) $f_0(x) = \sin ax$

(c) $f_0(x) = 4 \sin x + 3.1 \sin 2.1x + 1.8 \sin 8.2x - \cos x$

(d) $f_0(x, y) = \cos(ax + by)$

(e) $f_0(x, y) = \sin(ax + by)$

(f) $f_0(x, y) = 3 \sin x + 2 \sin y + 0.5 \sin(2x + 3y) + 0.6 \sin(3x - 2y) - \cos(x - 1.1y)$



A imagem da questão 2(f) é a imagem do exemplo na página 49.

(2) Encontre os espaços de escala das seguintes funções (sugestão: pense nos autovetores da convolução com G_t ; quais são os autovalores correspondentes?)

- (a) $f_0(x) = x$
- (b) $f_0(x) = x^2$
- (c) $f_0(x) = x^3$
- (d) $f_0(x) = x^4$
- (e) $f_0(x) = 6x^4 + 3x^2 - 2$
- (f) $f_0(x, y) = x^3 - y^2 + \sin(2x)$

(3) Escreva o operador Laplaciano bidimensional $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ em coordenadas polares. Conclua que o Laplaciano é de fato rotacionalmente simétrico. Sugestão: obtenha primeiro

$$\begin{aligned} r \frac{\partial}{\partial r} &= x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \\ \frac{\partial}{\partial \theta} &= -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \end{aligned}$$

e a partir daí

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} &= y^2 \frac{\partial^2}{\partial x^2} + x^2 \frac{\partial^2}{\partial y^2} - 2xy \frac{\partial^2}{\partial x \partial y} - y \frac{\partial}{\partial y} - x \frac{\partial}{\partial x} \\ r \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) &= x^2 \frac{\partial^2}{\partial x^2} + y^2 \frac{\partial^2}{\partial y^2} + 2xy \frac{\partial^2}{\partial x \partial y} + y \frac{\partial}{\partial y} + x \frac{\partial}{\partial x} \end{aligned}$$

concluindo que

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2}$$

Capítulo 6

Discretização Espacial

Vimos que o espaço de escala Gaussiano, definido a partir da equação do calor

$$\begin{aligned}\frac{\partial F_t}{\partial t}(x) &= \frac{\partial^2 F_t}{\partial x^2}(x) \\ F_0(x) &= f_0(x)\end{aligned}$$

ou, equivalentemente, a partir de convoluções com Gaussianas

$$\begin{aligned}F_t(x) &= G_t(x) * f_0(x) \\ G_t(x) &= \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}\end{aligned}$$

tem boas propriedades de linearidade, invariância por translações e causalidade. No entanto, para aplicações computacionais, precisamos responder à seguinte pergunta: como discretizar o espaço de escala Gaussiano? Neste capítulo, vamos nos preocupar com a discretização no espaço (x) mantendo a escala (t) contínua.

É viável não discretizar a escala t ? A resposta é positiva: o espaço de escala não precisa ser **calculado** em todas as escalas simultaneamente; para encontrar estruturas multi-escala, é suficiente calculá-lo em diferentes escalas **sob encomenda**. Assim, a teoria dos espaços de escala com escala contínua pode ser utilizada computacionalmente.

Então, como discretizar o espaço de escala na variável espaço? Inicialmente, apresentamos 3 alternativas:

1. Fazer **amostragens** de Gaussianas de diferentes variâncias e usá-las como um núcleo de convolução discreto. O problema deste método é que os núcleos encontrados não satisfazem exatamente as propriedades de causalidade e de semi-grupo – portanto, não há garantias de que o sinal se simplifique à medida que a escala aumenta.
2. Construir uma teoria de espaço de escala discreto a partir dos axiomas de convolução e causalidade; podemos descobrir todos os núcleos de convolução *discretos* que tenham tais propriedades e então verificar quais (se

algum) se aproximam do núcleo Gaussiano. Este método será discutido no capítulo 10; para o leitor interessado, este é o método utilizado em [28].

3. Discretizar a equação do calor; é este o método que analisamos na seção seguinte; ele é mais simples e direto do que o método axiomático, e acaba apresentando os mesmos resultados.

Infelizmente, a computação das convoluções que encontraremos é um pouco trabalhosa. O apêndice 17 descreve com mais detalhes como implementar convoluções “infinitas”; em particular, para aplicações que precisam de uma implementação eficiente da convolução com a Gaussiana mais do que das propriedades de causalidade e semi-grupo, apresentamos em 17.0.2 uma quarta alternativa: o método recursivo de Deriche.

6.1 Caso Unidimensional

Discretizando a equação do calor unidimensional apenas no espaço, encontramos

$$\begin{aligned} \frac{\partial F_t[n]}{\partial t} &= F_t[n+1] - 2F_t[n] + F_t[n-1] = (1, -2, 1) * F_t[n] \\ F_0[n] &= f_0[n] \end{aligned} \quad (6.1)$$

onde lembramos que $n \in \mathbb{Z}$ enquanto $t \in \mathbb{R}_+$.

Proposição 6.1 *A solução de 6.1 pode ser expressa através da convolução*

$$F_t[n] = P_t[n] * f_0[n]$$

onde $P_t[n]$ é o núcleo de Poisson simétrico

$$P_t[n] = e^{-2t} I_n(2t)$$

e $I_n(2t)$ é a função modificada de Bessel¹ de ordem n aplicada no ponto $2t$.

Demonstração. Para verificar a solução, basta usar as propriedades das funções de Bessel

$$\begin{aligned} I_n'(t) &= \frac{I_{n-1}(t) + I_{n+1}(t)}{2} \\ I_0(t) &= \delta(t) \end{aligned}$$

em 6.1. Ao invés de fazê-lo, preferimos mostrar como utilizar a Transformada Z para encontrar a solução. Primeiro, escreva 6.1 como uma convolução

$$\frac{\partial F}{\partial t} = (1, -2, 1) * F$$

¹Para a definição e propriedades das funções modificadas de Bessel, vide Apêndice.

A Transformada Z de F deve então satisfazer

$$\begin{aligned}\frac{\partial \varphi_F}{\partial t} &= \left(\frac{1}{z} - 2 + z \right) \varphi_F \Rightarrow \\ &\Rightarrow \varphi_F = \varphi_{F_0} e^{(z + \frac{1}{z} - 2)t}\end{aligned}$$

e F pode ser escrita como uma convolução

$$F[n] = f_0[n] * P_t[n]$$

onde o núcleo $P_t[n]$ vem do desenvolvimento da expressão de φ_{P_t}

$$e^{(z + \frac{1}{z} - 2)t} = \sum_{n=-\infty}^{\infty} P_t[n] z^n$$

Comparando tal expressão com a definição da função modificada de Bessel de ordem inteira, vê-se que

$$P_t = e^{-2t} (\dots I_{-2}(2t), I_{-1}(2t), I_0(2t), I_1(2t), I_2(2t), \dots)$$

As funções modificadas de Bessel fazem parte de várias bibliotecas matemáticas de programação. Em geral, elas são calculadas por relações recursivas ou como soluções de uma equação diferencial (vide Apêndice).

Há uma outra maneira de encarar os núcleos de Poisson simétricos e as funções de Bessel; para tanto escreva

$$\varphi_{P_t} = e^{(z-1)t} e^{(z^{-1}-1)t}$$

e observe que os coeficientes do desenvolvimento da primeira exponencial em série de potências de z , a saber,

$$e^{(z-1)t} = e^{-t} \sum_{n=0}^{\infty} \frac{t^n z^n}{n!}$$

formam a distribuição de Poisson de média $\mu = t$

$$\text{PoissonDen}(n; \mu) = \frac{\mu^n e^{-\mu}}{n!}$$

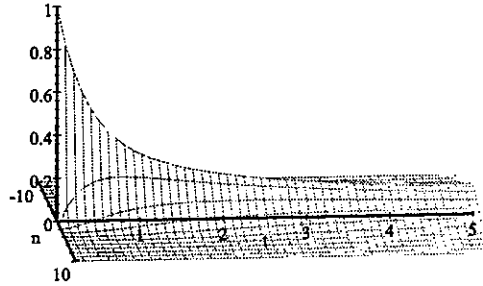
A segunda exponencial terá os mesmos coeficientes em ordem inversa (troque z por z^{-1} na Transformada Z). Isto é, P_t pode ser expresso explicitamente como uma convolução de duas distribuições de Poisson de média $\mu = t$

$$P_t = e^{-t} \left(\dots 0, 1, t, \frac{t^2}{2!}, \frac{t^3}{3!}, \frac{t^4}{4!}, \dots \right) * e^{-t} \left(\dots \frac{t^4}{4!}, \frac{t^3}{3!}, \frac{t^2}{2!}, t, 1, 0, \dots \right)$$

ou, usando a definição da convolução, chegamos à seguinte expressão explícita

$$P_t[n] = e^{-2t} t^{|n|} \sum_{k=0}^{\infty} \frac{t^{2k}}{k! (k + |n|)!}$$

A seguir, mostramos o gráfico de $P_t[n]$ para uma gama de valores de t e n ; note como $P_0[n] = \delta[n]$ e como para outros valores de t o núcleo de Poisson simétrico se assemelha à função Gaussiana G_t .



$$P_t[n] = e^{-2t} I_n(2t)$$

Pode-se mostrar (vide capítulo 10) que o número de cruzamentos de zero de $F_t[n] = P_t[n] * f_0[n]$ diminui à medida que t aumenta para qualquer f_0 ; como F_t é gerado por uma convolução, isto significa que o número de máximos (mínimos) locais também diminui com a escala.

Uma das grandes vantagens do núcleo de Poisson sobre outras discretizações da Gaussiana é que ele possui a propriedade de semi-grupo! De fato,

$$\begin{aligned} \varphi_{P_\alpha * P_\beta} &= \varphi_{P_\alpha} \varphi_{P_\beta} = e^{(\alpha+\beta)(z+\frac{1}{z}-2)} = \varphi_{P_{\alpha+\beta}} \Rightarrow \\ &\Rightarrow P_\alpha * P_\beta = P_{\alpha+\beta} \end{aligned}$$

Assim, o cálculo de um espaço de escala que utilize núcleos de Poisson pode ser feito “incrementalmente”, usando pequenos passos de tempo (escala). Esta propriedade já era de se esperar, já que a definição do espaço de escala associado a esses núcleos vem de uma equação diferencial.

Há vários outros paralelos entre o comportamento das Gaussianas (no caso contínuo) e do núcleo de Poisson simétrico (no caso discreto). Por exemplo, como acontece com $G_t(x)$, a variância de $P_t[n]$ é $2t$:

$$\begin{aligned} \varphi_{P_t} &= e^{(z+z^{-1}-2)t} \Rightarrow \varphi''_{P_t}(z) = \left(\frac{2t}{z^3} + t^2 \left(1 - \frac{1}{z^2} \right)^2 \right) e^{(z+z^{-1}-2)t} \Rightarrow \\ &\Rightarrow \text{Variância}_{P_t} = \varphi''_{P_t}(1) + \mu_{P_t} - \mu_{P_t}^2 = 2t \end{aligned}$$

A Transformada de Fourier de $P_t[n]$ é obtida colocando $z = e^{2\pi i w}$ na expressão de φ_{P_t} :

$$\hat{P}_t(w) = e^{-4t \sin^2 \pi w} \approx e^{-4\pi^2 w^2 t} = \hat{G}_t(x) \text{ para } w \approx 0$$

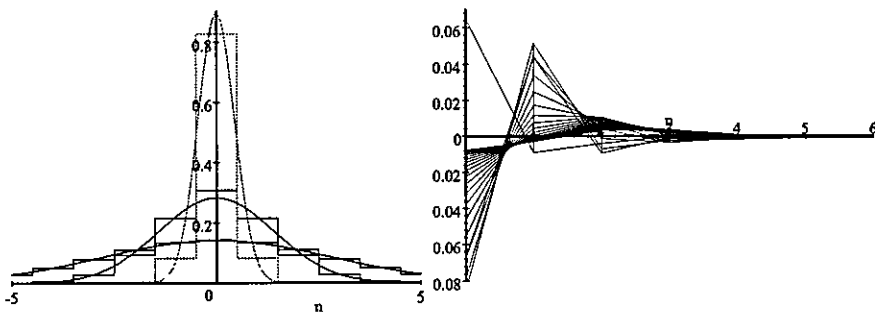
Também, com a escolha mais natural das derivadas discretas² D_x e D_{xx}

$$\begin{aligned}\partial_x G_t(x) &= \partial_x \left(\frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \right) = -\frac{x}{2t} G_t(x) \\ D_x P_t[n] &= \frac{1}{2} (P_t[n+1] - P_t[n-1]) = \frac{e^{-2t}}{2} (I_{n+1}(2t) - I_{n-1}(2t)) = \\ &= \frac{e^{-2t}}{2} \left(-\frac{2n}{2t} I_n(2t) \right) = -\frac{n}{2t} P_t[n]\end{aligned}$$

e, é claro

$$\begin{aligned}\partial_{xx} G_t(x) &= \partial_t G_t(x) \\ D_{xx} P_t[n] &= P_t[n+1] - 2P_t[n] + P_t[n-1] = \partial_t P_t[n]\end{aligned}$$

Como P_t tem tais propriedades semelhantes às da Gaussiana, dizemos que o núcleo de Poisson simétrico $P_t[n]$ é a versão discreta da Gaussiana $G_t(x)$. Os gráficos abaixo mostram como este núcleo discreto aproxima bem uma Gaussiana de variância $\sigma^2 = 2t$, especialmente para altas variâncias. À esquerda mostramos os gráficos de $P_t[n]$ e $G_t(x)$ superpostos para $t = 0.1$, $t = 1$ e $t = 4$. À direita está o gráfico de $G_t(n) - P_t[n]$ para valores de t que vão de 0.1 a 2.0 (as melhores aproximações correspondem a valores maiores de t).



G_t e P_t para $\sigma^2 = 2t = 0.2, 2$ e 8 $G_t(n) - P_t[n]$ para $t = 0.1, 0.2, \dots, 2.0$

6.2 Caso Bidimensional

Como estender nossa análise para espaços de escala de imagens? Basta discretizar a equação do calor

$$\frac{\partial F_t}{\partial t} = L * F$$

²O leitor que quiser se convencer que estas são escolhas naturais para as derivadas discretas deve se dirigir ao capítulo 9, enquanto as propriedades de $I_n(t)$ estão no apêndice.

onde L será uma discretização do Laplaciano bidimensional atuando em F . Por exemplo, a maneira mais simples de discretizar tal equação é usar derivadas discretas semelhantes às apresentadas na seção anterior em cada direção e somá-las

$$\begin{aligned} \frac{\partial F_t [m, n]}{\partial t} &= \\ &= F [m + 1, n] + F [m - 1, n] + F [m, n + 1] + F [m, n - 1] - 4F [m, n] \end{aligned}$$

Tal equação corresponde a

$$\frac{\partial F}{\partial t} = L * F$$

onde L é escolhido como o núcleo de convolução

$$L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Esta não é a única escolha possível para a discretização do Laplaciano; a discretização mais geral do Laplaciano é dada pela convolução com

$$L_c = \begin{bmatrix} c & 1 - 2c & c \\ 1 - 2c & -4 + 4c & 1 - 2c \\ c & 1 - 2c & c \end{bmatrix}$$

onde, na literatura, costuma-se usar apenas $c \leq \frac{1}{2}$. Nossa preferência é $c = \frac{1}{6}$, por gerar o Laplaciano mais rotacionalmente simétrico possível (vide capítulo 9).

Proposição 6.2 *A solução de*

$$\begin{aligned} \frac{\partial F_t [m, n]}{\partial t} &= L_c * F_t [m, n] \\ F_0 [m, n] &= f_0 [m, n] \end{aligned}$$

pode ser descrita pela convolução

$$F_t [m, n] = P_{(1-2c)t}^x * P_{(1-2c)t}^y * P_{ct}^{x+y} * P_{ct}^{x-y} * f_0 [m, n]$$

onde os núcleos de convolução bidimensionais são dados por

$$\begin{aligned} P_\alpha^x [m, n] &= P_\alpha [m] \delta [n]; & P_\alpha^y [m, n] &= P_\alpha [n] \delta [m] \\ P_\alpha^{x+y} [m, n] &= P_\alpha [m] \delta [m - n]; & P_\alpha^{x-y} [m, n] &= P_\alpha [m] \delta [m + n] \end{aligned}$$

Demonstração. Para verificar a solução; use $\alpha = (1 - 2c)t$ e $\beta = ct$; lembre que t não é uma das variáveis da convolução, portanto

$$\begin{aligned} \frac{\partial F_t}{\partial t} &= \\ &= (1 - 2c) \left\{ \frac{\partial P_\alpha^x}{\partial \alpha} * P_\alpha^y + P_\alpha^x * \frac{\partial P_\alpha^y}{\partial \alpha} \right\} * P_\beta^{x+y} * P_\beta^{x-y} * f_0 + \\ &+ c P_\alpha^x * P_\alpha^y * \left\{ \frac{\partial P_\beta^{x+y}}{\partial \beta} * P_\beta^{x-y} + P_\beta^{x+y} * \frac{\partial P_\beta^{x-y}}{\partial \beta} \right\} * f_0 \end{aligned}$$

A partir daí, basta usar repetidas vezes a propriedade do núcleo de Poisson simétrico

$$\frac{\partial P_t}{\partial t} = (1 - 2c) * P_t$$

nas respectivas direções, para terminar os cálculos. No entanto, estamos novamente mais interessados em mostrar como encontrar a solução usando a Transformada Z a duas variáveis (aqui denominadas x e y). Temos

$$\begin{aligned} \frac{\partial \varphi_F}{\partial t} &= \varphi_L \varphi_F \Rightarrow \varphi_F = \varphi_{f_0} e^{\varphi_L t} \Rightarrow \\ &\Rightarrow F = K * f_0 \end{aligned}$$

onde $K [m, n]$ é o coeficiente em $x^m y^n$ da série

$$\begin{aligned} \varphi_K &= e^{\varphi_L t} = e^{t\{c(xy + xy^{-1} + x^{-1}y + x^{-1}y^{-1}) + (1-2c)(x + x^{-1} + y + y^{-1}) - 4 + 4c\}} = \\ &= e^{(1-2c)t(x + \frac{1}{x} - 2)} e^{(1-2c)t(y + \frac{1}{y} - 2)} e^{ct(xy + \frac{1}{xy} - 2)} e^{ct(\frac{x}{y} + \frac{y}{x} - 2)} = \\ &= \varphi_{P_{(1-2c)t}^x} \varphi_{P_{(1-2c)t}^y} \varphi_{P_{ct}^{x+y}} \varphi_{P_{ct}^{x-y}} \end{aligned}$$

isto é,

$$K = P_{(1-2c)t}^x * P_{(1-2c)t}^y * P_{ct}^{x+y} * P_{ct}^{x-y}$$

Comentário 6.3 O leitor deve notar que os 4 núcleos presentes na convolução acima são de fato bastante simples: P^x é um núcleo de Poisson simétrico disposto somente no "eixo x ", ou direção 0° , de sua estrutura bidimensional (todos os outros coeficientes são 0); uma convolução de uma imagem com P^x pode ser feita convolvendo separadamente cada linha da imagem com o núcleo de Poisson simétrico. Analogamente, P^y , P^{x+y} e P^{x-y} somente têm coeficientes nas direções 90° , 45° e -45° , respectivamente. Convoluções com estes núcleos correspondem a convoluir separadamente cada coluna (diagonal principal, diagonal secundária, respectivamente) com um núcleo de Poisson simétrico.

Exemplo 6.4 Considere o caso do espaço de escala do Laplaciano mais simples ($c = 0$); pode-se calcular $F_t[m, n]$ a partir da imagem inicial $f_0[m, n]$ através de duas convoluções com núcleos de Poisson dispostos nas direções das linhas e das colunas da imagem

$$F_t = P_t^x * P_t^y * f_0$$

A generalização deste método para maiores dimensões é trabalhosa mas completamente análoga. Por exemplo, em 3 dimensões, o Laplaciano é comumente discretizado como uma matriz $3 \times 3 \times 3$. O caso mais simples apresentará uma solução do tipo

$$F_t[m, n, k] = P_t^x * P_t^y * P_t^z * f_0[m, n, k]$$

mas o uso de Laplacianos mais genéricos (possivelmente para aumentar a simetria rotacional) pode envolver até 13 convoluções.

Em geral, não só o Laplaciano, mas qualquer matriz N -dimensional $3 \times 3 \times 3 \times \dots \times 3$ simétrica com relação ao seu centro permite o uso do método acima; a solução será dada por $\frac{3^N - 1}{2}$ convoluções com núcleos de Poisson diversos em várias direções.

Comentário 6.5 Podemos também explicar agora a utilização dos filtros “diferença de Gaussianas” (filtros DOG) para o cálculo do Laplaciano de um espaço de escala, comumente encontrado na literatura. Como

$$\nabla^2 G_t(x) = \partial_t G_t(x) \approx \frac{G_{t+\Delta t}(x) - G_t(x)}{\Delta t}$$

vale a aproximação

$$\begin{aligned} \nabla^2 F_t &= \nabla^2 (G_t * f_0) = \nabla^2 G_t * f_0 \approx \\ &\approx \frac{1}{\Delta t} (G_{t+\Delta t} * f_0 - G_t * f_0) = \frac{1}{\Delta t} (F_{t+\Delta t} - F_t) \propto F_{t+\Delta t} - F_t \end{aligned}$$

Note que essa aproximação também é válida se utilizarmos as discretizações com núcleos de Poisson e substituirmos o Laplaciano por sua discretização correspondente.

6.3 Exercícios

1) O objetivo desta questão é mostrar que $P_t[n]$ é unimodal e positivo no domínio do espaço (o roteiro abaixo mostra como chegar ao resultado sem utilizar causalidade CZ). Suponha $n, t \geq 0$

a) Complete as lacunas corretamente:

$$P_t[n+1] - P_t[n-1] = ???$$

$$\frac{\partial P_t[n]}{\partial t} = ((???) * P_t)[n]$$

$$P_t[n] = \sum_{k=0}^{\infty} ???$$

b) Mostre que o núcleo de Poisson simétrico é positivo, isto é

$$P_t [n] \geq 0$$

c) Mostre que

$$P_t [n - 1] \geq P_t [n + 1]$$

d) Mostre que, para $0 < t < n + 1$, tem-se

$$P_t [n] > P_t [n + 1]$$

Em particular, isto garante que P_t é unimodal para $0 \leq t < 1$.

e) Use a função

$$f(t) = e^{4t} (P_t [n] - P_t [n + 1])$$

para mostrar que P_t é, de fato, unimodal para qualquer t .

Capítulo 7

Discretização Temporal

No capítulo anterior discutimos como discretizar a Gaussiana no espaço; isto será viável se calcularmos o espaço de escala em cada escala sob encomenda; em outras palavras, isto funcionará se nunca precisarmos calcular o espaço de escala como um todo para todas as escalas.

Caso contrário, se quisermos uma representação computacional do espaço de escala como um todo, precisaremos também discretizar a variável tempo (escala); como fazê-lo?

Imagine que as escalas a serem consideradas serão da forma $n\Delta t$ para $n \in \mathbb{N}$ e Δt escolhido a priori. Uma opção de construção de espaço de escala é simplesmente usar os núcleos simétricos de Poisson calculados somente em tais escalas discretas

$$F_{n\Delta t} = (P_{\Delta t})^{*n} * f_0 = P_{n\Delta t} * f_0$$

onde $P_{\Delta t} = e^{-2\Delta t} (\dots I_{-2}(2\Delta t), I_{-1}(2\Delta t), I_0(2\Delta t), I_1(2\Delta t), I_2(2\Delta t), \dots)$

É importante notar que esta é uma solução viável, que certamente satisfará as propriedades de homogeneidade, semi-grupo e causalidade. O motivo de não darmos mais atenção a esta solução neste capítulo é um só: já a discutimos bastante no capítulo anterior.

Mas será que há outras opções? Afinal, já que decidimos discretizar a escala, não nos parece tão importante satisfazer uma equação diferencial que contenha o termo $\frac{\partial F}{\partial t}$ (e foi de uma tal equação que surgiram os núcleos de Poisson).

Uma idéia é buscar todos os espaços de escala discretos na escala e no tempo que satisfaçam os axiomas de linearidade, homogeneidade (invariância por translação) e causalidade. Esta abordagem axiomática será discutida no capítulo 10. No entanto, preferimos aqui obter resultados semelhantes seguindo uma linha de raciocínio mais direta: vamos discretizar diretamente a equação diferencial.

7.1 Caso Unidimensional

7.1.1 Equação Explícita

Uma maneira de discretizar a equação do calor é

$$\frac{F_{t+\Delta t}[n] - F_t[n]}{\Delta t} = F_t[n+1] - 2F_t[n] + F_t[n-1] \quad (7.1)$$

$$F_0[n] = f_0[n]$$

Proposição 7.1 *A solução da equação 7.1 é dada por*

$$F_{j\Delta t} = (B_{\Delta t})^{*j} * f_0[n]$$

onde $B_{\Delta t} = (\Delta t \ 1 - 2\Delta t \ \Delta t)$

A equação e sua solução são estáveis se e somente se $0 \leq \Delta t \leq \frac{1}{2}$.

Demonstração. Basta re-escrever 7.1 como

$$F_{t+\Delta t}[n] = (\Delta t \ 1 - 2\Delta t \ \Delta t) * F_t[n]$$

e a solução é imediata. Para verificar sua estabilidade, calcule o ganho máximo da DTFT de $B_{\Delta t}$ (usando que $-1 \leq \cos 2\pi w \leq 1$)

$$\hat{B}_{\Delta t} = 2\Delta t (\cos 2\pi w - 1) + 1 \Rightarrow$$

$$\Rightarrow -4\Delta t + 1 \leq \hat{B}_{\Delta t} \leq 1$$

Para que a solução seja estável, é necessário que **nenhuma frequência seja amplificada** quando da convolução com $B_{\Delta t}$ (caso contrário, as sucessivas convoluções fariam com que uma tal frequência “explodisse” em $F_{t+j\Delta t}$), isto é

$$|\hat{B}_{\Delta t}| \leq 1 \Rightarrow \Delta t \leq \frac{1}{2}$$

Transformada Z e DTFT

A Transformada Z de $B_{\Delta t}$ é facilmente obtida

$$\varphi_{B_{\Delta t}} = 1 + \Delta t (z + z^{-1} - 2)$$

Colocando $z = e^{-2\pi i w}$, obtemos a DTFT

$$\hat{B}_{\Delta t}(w) = 1 - 4\Delta t \sin^2 \pi w$$

Note que $B_{\Delta t}$ é um filtro de passa-baixa, e sua DTFT é unimodal somente para $0 \leq \Delta t \leq \frac{1}{4}$

Causalidade

Que princípios de causalidade são satisfeitos por um espaço de escala assim determinado? O princípio do máximo é facilmente deduzido:

$$\begin{aligned} F_t[n] &\geq F_t[n-1], F_t[n+1] \Rightarrow \\ &\Rightarrow \frac{F_{t+\Delta t}[n] - F_t[n]}{\Delta t} = F_t[n+1] - 2F_t[n] + F_t[n-1] \leq 0 \Rightarrow \\ &\Rightarrow F_{t+\Delta t}[n] \leq F_t[n] \end{aligned}$$

Já o princípio de causalidade por cruzamentos de zero, a saber,

$$j > k \Rightarrow \#CZ(F_{t+j\Delta t}) \leq \#CZ(F_{t+k\Delta t}) \quad (7.2)$$

somente vale para $0 \leq \Delta t \leq \frac{1}{4}$. A demonstração não é muito difícil, mas preferimos adiar sua discussão para o capítulo 10.

Aqui vale a pena esclarecer (ou criar) uma confusão: se utilizarmos passos de tempo diferentes Δt_1 e Δt_2 , não podemos garantir o princípio da causalidade! Em outras palavras,

$$j\Delta t_1 > k\Delta t_2 \not\Rightarrow \#CZ\left((B_{\Delta t_1})^{*j} * f_0\right) \leq \#CZ\left((B_{\Delta t_2})^{*k} * f_0\right).$$

Repetimos: a única afirmação válida feita a respeito da diminuição do número de cruzamento de zeros é 7.2, que vale para Δt fixo e determinado a priori, menor do que $\frac{1}{4}$.

Semi-grupo

Claramente, a família de núcleos $K_j = (B_{\Delta t})^{*j}$ satisfaz a propriedade de semi-grupo para Δt fixo, escolhido a priori, isto é

$$K_{(m+n)\Delta t} = (B_{\Delta t})^{*(m+n)} = (B_{\Delta t})^{*m} * (B_{\Delta t})^{*n} = K_m * K_n$$

No entanto, a mudança na escolha do passo de tempo, além de quebrar a propriedade de causalidade, também destrói a propriedade de semi-grupo. Em outras palavras, em geral

$$B_{\Delta t_1 + \Delta t_2} \neq B_{\Delta t_1} * B_{\Delta t_2}.$$

Isto quer dizer que o valor de F_t não é independente do passo de tempo Δt escolhido! Por exemplo, se escolhermos $\Delta t = \frac{1}{4}$, temos $F_2 = (B_{1/4})^{*8} * f_0$; se escolhermos $\Delta t = \frac{1}{8}$, temos $F_2 = (B_{1/8})^{*16} * f_0$. Infelizmente, estas duas expressões para F_2 dão resultados diferentes!

Assim, suponha que escolhermos um certo valor para Δt e calculamos todo o espaço de escala $F_{k\Delta t}$ ($k \in \mathbb{N}$) baseado neste Δt . Se a aplicação em questão pedir então $F_{\frac{\Delta t}{2}}$, não há saída: temos de diminuir o passo de tempo para $\frac{\Delta t}{2}$ – e, se quisermos manter a coerência de nossa discretização (e a propriedade

de causalidade dos cruzamentos de zero), todo o espaço de escala tem de ser recalculado com este novo passo de tempo! Este é um dos motivos pelos quais deve-se considerar com carinho o espaço de escala com escala contínua discutido no capítulo anterior (que satisfaz a propriedade de semi-grupo para qualquer escala não-negativa).

Caso limite: $\Delta t \rightarrow 0$

Como o cálculo de F_t depende do passo de tempo Δt escolhido, uma idéia natural é tomar $\Delta t = \frac{t}{n}$ e deixar $n \rightarrow \infty$, isto é, olhar para

$$\lim_{n \rightarrow \infty} \left(B_{\frac{t}{n}} \right)^{*n} * f_0$$

e talvez tomar este limite como o valor “verdadeiro” de F_t . Qual é esse limite?

Não é difícil adivinhar a resposta observando o que acontece com a equação do calor discreta à medida que $\Delta t \rightarrow 0$: o espaço de escala F_t se aproxima ao espaço de escala gerado pelo núcleo de Poisson apresentado no capítulo anterior. Em outras palavras:

Afirmção 7.2

$$\lim_{n \rightarrow \infty} \left(B_{\frac{t}{n}} \right)^{*n} = P_t$$

Demonstração. Analise a Transformada Z de $K_{n,t} = \left(B_{\frac{t}{n}} \right)^{*n}$:

$$\varphi_{K_{n,t}} = \left(\varphi_{B_{\frac{t}{n}}} \right)^n = \left(\frac{t}{n} z^{-1} + \left(1 - 2\frac{t}{n} \right) + \frac{t}{n} z \right)^n = \left(1 + \frac{t(z^{-1} - 2 + z)}{n} \right)^n$$

Tomando $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \varphi_{K_{n,t}} = e^{t\left(\frac{1}{z} - 2 + z\right)} = \varphi_{P_t}$$

7.1.2 Equação Implícita

A equação explícita da subseção anterior não é a única maneira coerente de discretizar a equação do calor – também podemos usar uma discretização com aproximação implícita, isto é, tomar a derivada espacial no tempo $t + \Delta t$:

$$\frac{F_{t+\Delta t}[n] - F_t[n]}{\Delta t} = F_{t+\Delta t}[n+1] - 2F_{t+\Delta t}[n] + F_{t+\Delta t}[n-1] \quad (7.3)$$

$$F_0[n] = f_0[n]$$

Proposição 7.3 *A solução da equação 7.3 é dada por*

$$F_{j\Delta t} = (A_{\Delta t})^{*j} * f_0 [n]$$

$$\text{onde } A_{\Delta t} = \frac{1-\alpha}{1+\alpha} (\dots\alpha^3 \alpha^2 \alpha 1 \alpha \alpha^2 \alpha^3 \dots)$$

$$\text{e } \alpha = \frac{\sqrt{1+4\Delta t} - 1}{\sqrt{1+4\Delta t} + 1}$$

A equação e sua solução são estáveis para qualquer $\Delta t > 0$.

Demonstração. Re-escreva a equação na forma

$$(-\Delta t 1 + 2\Delta t - \Delta t) * F_{t+\Delta t} [n] = F_t [n]$$

Como resolver esta “deconvolução”? Aplique a Transformada Z

$$(-\Delta t z^{-1} + (1 + 2\Delta t) - \Delta t z) \varphi_{F_{t+\Delta t}} = \varphi_{F_t} \Rightarrow$$

$$\Rightarrow F_{t+\Delta t} = A_{\Delta t} * F_t$$

onde

$$\varphi_{A_{\Delta t}} = \frac{1}{-\Delta t z^{-1} + (1 + 2\Delta t) - \Delta t z} = \frac{z}{z - \Delta t (1 - z)^2}$$

Para encontrar $A_{\Delta t}$, devemos expandir esta expressão como uma série de potências em z . Abrindo $\varphi_{A_{\Delta t}}$ em frações parciais, obtém-se

$$\varphi_{A_{\Delta t}} = \frac{1-\alpha}{1+\alpha} \left(\frac{1}{1-\alpha z} + \frac{\alpha}{z-\alpha} \right)$$

onde $0 < \alpha = \frac{\sqrt{4\Delta t+1}-1}{\sqrt{4\Delta t+1}+1} < 1$ é raiz de $(1-\alpha)^2 \Delta t = \alpha$. Usando agora as expansões¹

$$\frac{1}{1-\alpha z} = 1 + \alpha z + \alpha^2 z^2 + \dots$$

$$\frac{\alpha}{z-\alpha} = \frac{\alpha z^{-1}}{1-\alpha z^{-1}} = \alpha z^{-1} + \alpha^2 z^{-2} + \alpha^3 z^{-3} + \dots$$

tem-se

$$\varphi_{A_{\Delta t}} = \frac{1-\alpha}{1+\alpha} \sum_{n=-\infty}^{+\infty} \alpha^{|n|} z^n$$

e portanto podemos escrever o núcleo $A_{\Delta t}$ explicitamente

$$A_{\Delta t} = \frac{1-\alpha}{1+\alpha} (\dots\alpha^3 \alpha^2 \alpha 1 \alpha \alpha^2 \alpha^3 \dots)$$

¹Note que a primeira expansão é válida para $|\alpha z| < 1$ e a segunda é válida para $|\frac{\alpha}{z}| < 1$; como $|\alpha| < 1$, ambas são válidas para $|z| = 1$ - fato crucial para a correção de nosso método.

onde $A[0] = \frac{1-\alpha}{1+\alpha} = \frac{1}{\sqrt{4\Delta t+1}}$ é o elemento central.

Finalmente, note que a magnitude da DTFT de $A_{\Delta t}$ é obtida usando $z + z^{-1} = 2 \cos 2\pi w$ na expressão de $\varphi_{A_{\Delta t}}$

$$\begin{aligned}\hat{A}_{\Delta t}(w) &= \frac{1}{1 + \Delta t (2 - 2 \cos 2\pi w)} \Rightarrow \\ &\Rightarrow \frac{1}{1 + 4\Delta t} \leq \hat{A}_{\Delta t}(w) \leq 1\end{aligned}$$

e portanto nenhuma frequência é amplificada na convolução com $\hat{A}_{\Delta t}$; portanto, a solução é estável para qualquer $\Delta t > 0$. ■

Comentário 7.4 Note que, estendendo a definição de A e B para índices negativos, tem-se

$$A_{-\Delta t} * B_{\Delta t} = \delta$$

para $\Delta t < \frac{1}{4}$ ($\hat{A}_{-\Delta t}$ e $A_{-\Delta t}$ não estão bem definidos para $\Delta t \geq \frac{1}{4}$).

Comentário 7.5 Computacionalmente, a convolução com o núcleo $A_{\Delta t}$ expresso na forma acima pode ser muito custosa. Uma outra expressão para este espaço de escala é sugerida pela fatoração

$$\varphi_{A_{\Delta t}} = (1 - \alpha)^2 \frac{1}{1 - \alpha z} \frac{1}{1 - \alpha z^{-1}}$$

que mostra que $A_{\Delta t}$ é a convolução de dois núcleos geométricos, a saber,

$$A_{\Delta t} = (1 - \alpha) (\dots 0 0 0 0 1 \alpha \alpha^2 \alpha^3 \dots) * (1 - \alpha) (\dots \alpha^3 \alpha^2 \alpha 1 0 0 0 0 \dots)$$

Esta expressão é útil pois a convolução com um núcleo geométrico unilateral pode ser feita de maneira simples (veja o apêndice, seção 17.0.2). De fato, se

$$T_{\alpha} = (1 - \alpha) (\dots 0 0 0 0 1 \alpha \alpha^2 \alpha^3 \dots)$$

então calcular a convolução com T_{α}

$$L_2 = L_1 * T_{\alpha}$$

corresponde à interpolação recursiva

$$L_2[n+1] = \alpha L_2[n] + (1 - \alpha) L_1[n]$$

e pode ser feita “da esquerda para a direita”. A convolução com o núcleo refletido equivale à recursão oposta

$$L_3[n-1] = \alpha L_3[n] + (1 - \alpha) L_2[n]$$

As duas recursões realizadas consecutivamente equivalem à convolução com $A_{\Delta t}$.

Comentário 7.6 *Analogamente ao comentário anterior, podemos resolver a equação 7.3 esquecendo completamente a convolução e voltando à equação diferencial discreta*

$$F_{t+\Delta t}[n+1] = \frac{F_{t+\Delta t}[n] - F_t[n]}{\Delta t} + 2F_{t+\Delta t}[n] - F_{t+\Delta t}[n-1]$$

Basta usar esta relação recursiva diretamente para montar o espaço de escala da esquerda para a direita.

Transformada Z e DTFT

Como visto na demonstração cima, a Transformada Z do núcleo $A_{\Delta t}$ é

$$\varphi_{A_{\Delta t}} = \frac{1}{1 - \Delta t(z + z^{-1} - 2)}$$

e a sua DTFT, obtida tomando $z = e^{-2\pi i w}$, é

$$\hat{A}_{\Delta t}(w) = \frac{1}{1 + 4\Delta t \sin^2 \pi w}$$

Desta expressão, é fácil ver que $A_{\Delta t}$ é um filtro de passa-baixa, unimodal para qualquer $\Delta t > 0$.

Causalidade

Pode-se mostrar (vide capítulo 10) que o número de Cruzamentos de Zero de F_t diminui com t . Isto é, se $\Delta t > 0$ é fixo, então para qualquer f_0

$$j > k \Rightarrow \#CZ \left((A_{\Delta t})^{*j} * f_0 \right) \leq \#CZ \left((A_{\Delta t})^{*k} * f_0 \right)$$

Já o princípio do máximo, se escrito na forma da última seção, **não é válido**. De fato, tome por exemplo $f_0 = (0 \ 1 \ 0 \ K)$, centrado em $f_0[0] = 1$. Se usarmos f_0 para montar o espaço de escala, o valor de

$$F_{\Delta t}[0] = (A_{\Delta t} * f_0)[0] = \frac{1 - \alpha}{1 + \alpha} (1 + K\alpha^2)$$

pode ser feito arbitrariamente grande simplesmente tomando-se $K \gg 0$; assim, o máximo local $F_0[0] = 1$ **sobe** ainda mais para $F_{\Delta t}[0] > 1$.

Podemos salvar o princípio do máximo da seguinte forma: se $n = n_0$ nos dá um máximo local de $F_t[n]$, então $F_t[n_0]$ **desceu** na iteração passada, isto é

$$F_{t-\Delta t}[n] \geq F_t[n]$$

A demonstração é imediata e deixamo-la para o leitor.

Finalmente, vale a pena citar novamente que o uso de diferentes passos de tempo Δt quebra o princípio da causalidade (como na seção anterior). Em outras palavras

$$j\Delta t_1 > k\Delta t_2 \not\Rightarrow \#CZ \left((A_{\Delta t_1})^{*j} * f_0 \right) \leq \#CZ \left((A_{\Delta t_2})^{*k} * f_0 \right).$$

Semi-grupo

A mesma discussão da seção anterior vale para o espaço de escala gerado pela discretização implícita. Enquanto a família de núcleos $K_j = (A_{\Delta t})^{*j}$ satisfaz a propriedade de semi-grupo (discretamente) para Δt fixo, a mistura de passos de tempo traz problemas (já que infelizmente $A_{\Delta t} * A_{\Delta t} \neq A_{2\Delta t}$). Assim, uma vez escolhida a discretização em Δt , tem-se

$$F_{n\Delta t}(x) = (A_{\Delta t})^{*n} * f_0(x) \neq A_{n\Delta t} * f_0(x)$$

Uma vez calculado o espaço de escala com passo Δt , se precisarmos calcular $F_{\frac{t}{2}}$, devemos recalcular todo o espaço de escala (ou abandonar as propriedades de causalidade e semi-grupo).

Caso limite: $\Delta t \rightarrow 0$

Novamente, como o cálculo de F_t depende do passo de tempo Δt escolhido, tome $\Delta t = \frac{t}{n}$ e analise

$$\lim_{n \rightarrow \infty} \left(A_{\frac{t}{n}}\right)^{*n} * f_0$$

Observando o que acontece com a discretização implícita à medida que $\Delta t \rightarrow 0$, vislumbramos este limite: o espaço de escala F_t deve se aproximar novamente ao espaço de escala gerado por núcleos de Poisson.

Afirmção 7.7

$$\lim_{n \rightarrow \infty} \left(A_{\frac{t}{n}}\right)^{*n} = P_t$$

Demonstração. Analise a Transformada Z de $T_n = \left(A_{\frac{t}{n}}\right)^{*n}$

$$\varphi_{T_n} = \left(\varphi_{A_{\frac{t}{n}}}\right)^n = \left(-\frac{t}{n}z^{-1} + \left(1 + 2\frac{t}{n}\right) - \frac{t}{n}z\right)^{-n} = \left(1 - \frac{z^{-1} - 2 + z}{n}t\right)^{-n}$$

e, tome $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \varphi_{T_n} = e^{t(\frac{1}{z} - 2 + z)} = \varphi_{P_t}$$

■

7.2 Caso Bidimensional

Agora, além da escolha entre uma discretização implícita ou explícita, há várias escolhas possíveis para a discretização do Laplaciano na equação do calor. Note-se que o caso multidimensional é uma extensão imediata dos métodos a seguir (mas as contas a se resolver ficam cada vez mais trabalhosas).

7.2.1 Equação Explícita

Como mencionamos no capítulo anterior, a forma genérica para a discretização do Laplaciano de F é $L_c * F$, onde o núcleo L_c é dado por

$$L_c = \begin{bmatrix} c & 1-2c & c \\ 1-2c & -4+4c & 1-2c \\ c & 1-2c & c \end{bmatrix}$$

Assim, a discretização explícita da equação do calor é

$$\frac{F_{t+\Delta t}[m, n] - F_t[m, n]}{\Delta t} = (L_c * F_t)[m, n]$$

$$F_0[m, n] = f_0[m, n]$$

Proposição 7.8 *A solução da equação acima é dada por*

$$F_{j\Delta t} = (B_{\Delta t})^{*j} * f_0[m, n]$$

onde $B_{\Delta t} = \delta + \Delta t L_c$

A equação e sua solução são estáveis se e somente se as seguintes condições são satisfeitas: $c \leq \frac{1}{2}$ e $0 \leq \Delta t \leq \min\left(\frac{1}{2}, \frac{1}{4(1-2c)}\right)$.

Demonstração. A solução é imediata a partir de

$$F_{t+\Delta t}[m, n] = (\delta + \Delta t L_c) * F_t[m, n]$$

Para a estabilidade, calcule a DTF de $B_{\Delta t} = (\delta + \Delta t L_c)$:

$$\begin{aligned} \hat{B}_{\Delta t} &= \\ &= 1 + 2\Delta t (\cos 2\pi w_1 + \cos 2\pi w_2 - 2) + 4c\Delta t (\cos 2\pi w_1 - 1)(\cos 2\pi w_2 - 1) = \\ &= 1 - 4\Delta t (\sin^2 \pi w_1 + \sin^2 \pi w_2) + 16c\Delta t \sin^2 \pi w_1 \sin^2 \pi w_2 = \\ &= 1 + \Delta t \frac{(4c \sin^2 \pi w_1 - 1)(4c \sin^2 \pi w_2 - 1) - 1}{c} \end{aligned}$$

Para que a solução seja estável, é necessário que $|\hat{B}_{\Delta t}| < 1$ para quaisquer w_1, w_2 . Uma análise cuidadosa revela apenas cinco pontos críticos de \hat{B} , independentes de Δt , correspondendo a $(\sin^2 \pi w_1, \sin^2 \pi w_2) = (0, 0), (1, 0), (0, 1), (1, 1)$ e $(\frac{1}{4c}, \frac{1}{4c})$ (este último é sempre um ponto de sela e só existe se $c \geq \frac{1}{4}$). Assim, os possíveis máximos e mínimos de \hat{B} são, respectivamente, $1, 1 - 4\Delta t, 1 - 4\Delta t, 1 - 8\Delta t + 16c\Delta t$ (o quinto ponto não precisa ser considerado). Então

$$\begin{aligned} -1 \leq 1 - 4\Delta t \leq 1 &\Rightarrow 0 \leq \Delta t \leq \frac{1}{2} \\ -1 \leq 1 - 8\Delta t + 16c\Delta t \leq 1 &\Rightarrow c \leq \frac{1}{2} \text{ e } \Delta t \leq \frac{1}{4(1-2c)} \end{aligned}$$

■

Transformada Z e DTFT

A Transformada Z (aqui, com variáveis x e y) de $B_{\Delta t}$ pode ser expressa como

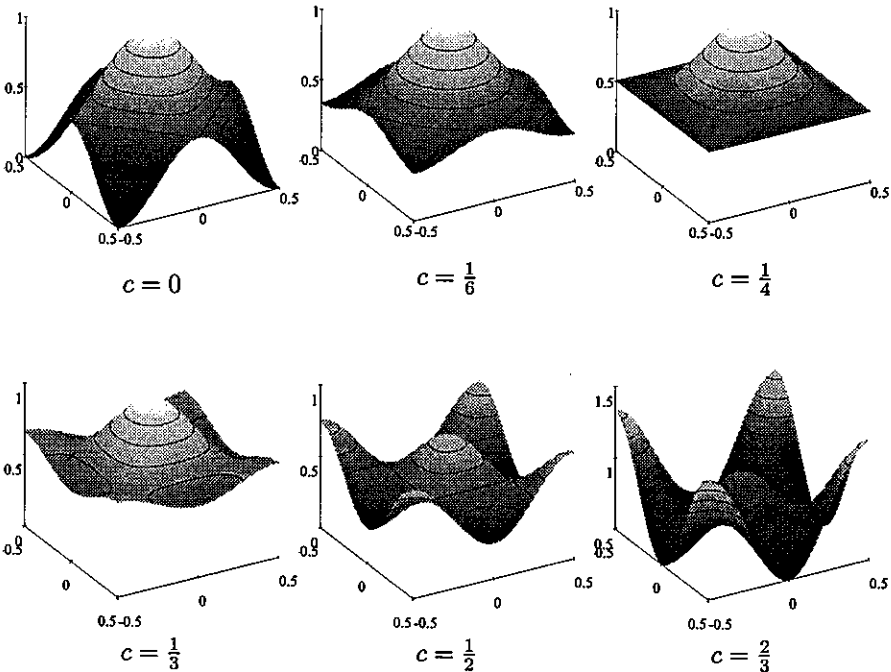
$$\varphi_{B_{\Delta t}}(x, y) = 1 + \Delta t \frac{(2c(x + x^{-1} - 1) + 1)(2c(y + y^{-1} - 1) + 1) - 1}{c}$$

e, como visto acima, sua DTFT é

$$\hat{B}_{\Delta t}(w_1, w_2) = 1 + \Delta t \frac{(4c \sin^2 \pi w_1 - 1)(4c \sin^2 \pi w_2 - 1) - 1}{c}$$

que será unimodal para $0 \leq c \leq \frac{1}{4}$.

A seguir, mostramos o gráfico de $\hat{B}_{\frac{1}{8}}$ ($\Delta t = \frac{1}{8}$ fixo) para diversos valores de c : 0 , $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{2}$ e $\frac{2}{3}$. Note como o ponto de sela aparece após $c = \frac{1}{4}$ e a solução torna-se instável para $c \geq \frac{1}{2}$ (pois há frequências onde $\hat{B} > 1$).



Vale a pena mencionar que, perto da origem, $c = \frac{1}{6}$ gera as curvas de nível mais "circulares" dentre os gráficos acima (correspondendo ao Laplaciano isotrópico - veja capítulo 9).

Causalidade

Como é difícil falar em número de cruzamentos de zero de sinais bidimensionais, perseguiremos apenas o princípio do máximo. Se definirmos um máximo local de $F[m, n]$ usando vizinhanças de 8 pontos, isto é

$$F[m_0, n_0] \text{ é máx. local} \Leftrightarrow F[m_0, n_0] \geq F[m_0 + \mu, n_0 + \nu] \text{ para } |\mu|, |\nu| \leq 1$$

então é fácil verificar que o princípio do máximo vale para qualquer discretização L_c com $0 \leq c \leq \frac{1}{2}$, já que num máximo local

$$\begin{aligned} (L_c * F)[m_0, n_0] &= (1 - 2c) \sum_{|j|+|k|=1} (F[m_0 + j, n_0 + k] - F[m_0, n_0]) \\ &\quad + c \sum_{|j|=|k|=1} (F[m_0 + j, n_0 + k] - F[m_0, n_0]) \leq 0 \end{aligned}$$

Por outro lado, se usarmos vizinhanças de 4 pontos, isto é,

$$F[m_0, n_0] \text{ é máx. local} \Leftrightarrow F[m_0, n_0] \geq F[m_0 + \mu, n_0 + \nu] \text{ para } |\mu| + |\nu| \leq 1$$

o mesmo argumento mostra que o princípio do máximo só vale para $c = 0$ (que corresponde à discretização separável do Laplaciano).

Semi-grupo

Novamente, a família de núcleos $T_{j\Delta t} = (B_{\Delta t})^{*j}$ satisfaz a propriedade de semi-grupo para Δt fixo, escolhido a priori, enquanto o uso de diversos passos de tempo quebra a propriedade de semi-grupo, isto é, em geral

$$B_{\Delta t_1 + \Delta t_2} \neq B_{\Delta t_1} * B_{\Delta t_2}.$$

7.2.2 Equação Implícita

A discretização implícita da equação do calor bidimensional

$$\begin{aligned} \frac{F_{t+\Delta t} - F_t}{\Delta t} &= L_c * F_{t+\Delta t} \\ F_0[m, n] &= f_0[m, n] \end{aligned}$$

tem propriedades semelhantes à sua versão unidimensional.

Proposição 7.9 *A solução da equação acima é dada por*

$$\begin{aligned} F_{j\Delta t} &= (A_{\Delta t})^{*j} * f_0[m, n] \\ \text{onde } A_{\Delta t} &= \text{IDTFFT} \left(\frac{c}{c + \Delta t - \Delta t (4c \sin^2 \pi w_1 - 1) (4c \sin^2 \pi w_2 - 1)} \right) \end{aligned}$$

A equação e sua solução são estáveis para $c \leq \frac{1}{2}$ e qualquer $\Delta t \geq 0$.

Demonstração. Re-escreva a equação implícita na forma

$$(\delta - \Delta t L_c) * F_{t+\Delta t} = F_t$$

e aplique a Transformada Z

$$\begin{aligned} (1 - \Delta t \varphi_{L_c}) \varphi_{F_{t+\Delta t}} &= \varphi_{F_t} \Rightarrow \\ \Rightarrow F_{t+\Delta t} &= A_{\Delta t} * F_t \end{aligned}$$

onde

$$\begin{aligned} \varphi_{A_{\Delta t}} &= \frac{1}{1 - \Delta t \varphi_{L_c}} = \\ &= \frac{1}{1 + \Delta t (z_1 + z_1^{-1} + z_2 + z_2^{-1} - 4) - \Delta t c (z_1 + z_1^{-1} - 2) (z_2 + z_2^{-1} - 2)} = \\ &= \frac{c}{c + \Delta t - \Delta t (c (z_1 + z_1^{-1} - 2) - 1) (c (z_2 + z_2^{-1} - 2) - 1)} \end{aligned}$$

Para encontrar $A_{\Delta t}$, deveríamos expandir esta expressão como uma série de potências em z_1 e z_2 ; no entanto, a expressão é complicada e preferimos simplesmente expressá-la através da DTFT

$$\begin{aligned} \hat{A}_{\Delta t} &= \frac{1}{1 + 4\Delta t (\sin^2 \pi w_1 + \sin^2 \pi w_2) - 16c\Delta t \sin^2 \pi w_1 \sin^2 \pi w_2} = \\ &= \frac{c}{c + \Delta t - \Delta t (4c \sin^2 \pi w_1 - 1) (4c \sin^2 \pi w_2 - 1)} \end{aligned}$$

Quanto à estabilidade, note que $\hat{A}_{\Delta t}(0, 0) = 1$ e $\hat{A}_{\Delta t}(w_1, w_2) \neq 0$ para quaisquer (w_1, w_2) . Assim, $|\hat{A}_{\Delta t}| \leq 1$, é equivalente a

$$\frac{1}{\hat{A}_{\Delta t}} = 1 + \frac{\Delta t}{c} - \frac{\Delta t}{c} (4c \sin^2 \pi w_1 - 1) (4c \sin^2 \pi w_2 - 1) \geq 1$$

Já vimos que os pontos críticos desta expressão (que é $\hat{B}_{-\Delta t}$) independem de Δt e correspondem a $(\sin^2 \pi w_1, \sin^2 \pi w_2) = (0, 0), (1, 0), (0, 1), (1, 1)$ e $(\frac{1}{4c}, \frac{1}{4c})$, sendo este último um ponto de sela; os valores dela nas primeiras quatro opções são $1, 1 + 4\Delta t, 1 + 4\Delta t, 1 + 8\Delta t - 16c\Delta t$. Assim, a estabilidade da solução pede

$$1 + 4\Delta t \geq 1 \Rightarrow \Delta t \geq 0$$

$$1 + 8\Delta t - 16c\Delta t \geq 1 \Rightarrow c \leq \frac{1}{2}$$

Comentário 7.10 Novamente, note que se estendermos as definições dos núcleos bidimensionais A e B para índices negativos, então

$$A_{\Delta t} * B_{-\Delta t} = \delta$$

para $\Delta t \geq -\frac{1}{4}$ e $\Delta t(1 - 2c) \geq -\frac{1}{8}$. Fora destes intervalos $\hat{A}_{\Delta t}$ não está bem definido.

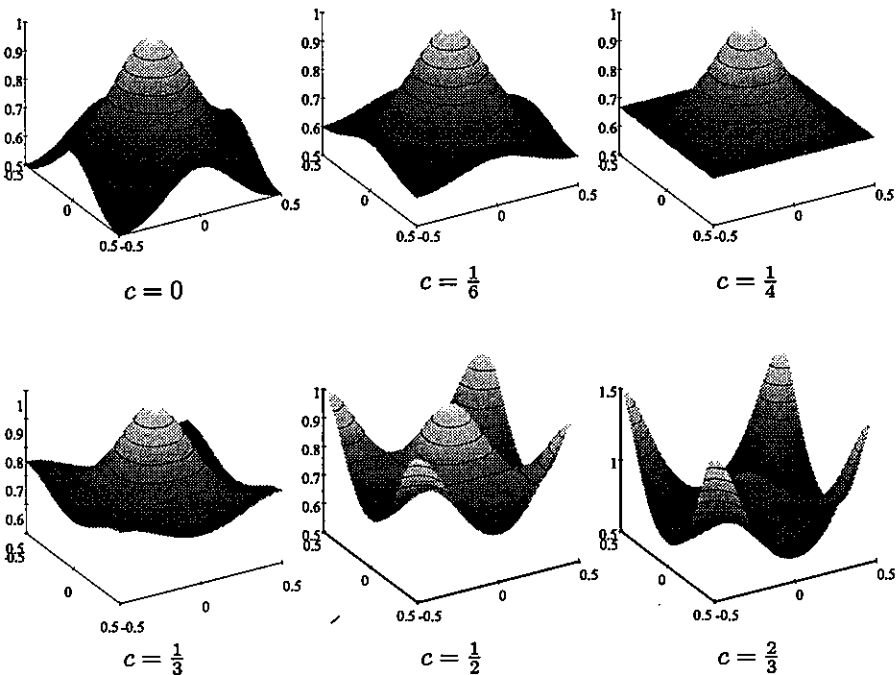
Transformada Z e DTFT

Desta vez, esta informação é o que temos a respeito de $A_{\Delta t}$. Resumindo o que se viu acima, a Transformada Z e a DTFT de $A_{\Delta t}$ podem ser escritas como

$$\varphi_{A_{\Delta t}}(x, y) = \frac{1}{1 - \Delta t \frac{(2c(x+x^{-1}-1)+1)(2c(y+y^{-1}-1)+1)-1}{c}}$$

$$\hat{A}_{\Delta t}(w_1, w_2) = \frac{1}{1 - \Delta t \frac{(4c \sin^2 \pi w_1 - 1)(4c \sin^2 \pi w_2 - 1) - 1}{c}}$$

Ilustramos a seguir o gráfico de $\hat{A}_{\frac{1}{3}}$ ($\Delta t = \frac{1}{8}$ fixo) para $c = 0, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}$. Novamente, note como a solução torna-se rapidamente instável para $c \geq \frac{1}{2}$ (pois há frequências onde $\hat{A} > 1$).



Causalidade

O princípio do máximo não vale na sua forma habitual: máximos locais de $F_{j\Delta t}$ podem subir na convolução com $A_{\Delta t}$ atingindo um valor maior em $F_{(j+1)\Delta t}$. No entanto, se falarmos de máximos locais (usando 8-vizinhanças) de $F_{(j+1)\Delta t}$ e suas posições relativas a $F_{j\Delta t}$, então podemos salvar um princípio de causalidade: tais máximos locais desceram na iteração passada. A verificação deste fato é simples e deixada ao leitor.

No caso de 4-vizinhanças, não só é necessário usar esta segunda forma do princípio do máximo mas também devemos exigir $c = 0$ na discretização do laplaciano.

Semi-grupo

Claramente a família de núcleos $T_{j\Delta t} = (A_{\Delta t})^{*j}$ satisfaz a propriedade de semi-grupo para Δt fixo. O uso de diversos passos de tempo quebra a propriedade de semi-grupo, isto é, em geral

$$A_{\Delta t_1 + \Delta t_2} \neq A_{\Delta t_1} * A_{\Delta t_2}.$$

7.3 Discretização Somente no Tempo*

A propósito, é possível discretizar a equação do calor apenas no tempo (apesar de não termos em mente quaisquer aplicações específicas deste caso).

7.3.1 Equação Explícita

Obtém-se

$$\frac{F_{t+\Delta t}(x) - F_t(x)}{\Delta t} = F_t''(x)$$

Pode-se escrever isto como uma convolução para cada avanço Δt no tempo

$$F_{t+\Delta t}(x) = (\delta + \Delta t \delta'') * F_t(x)$$

onde δ é a Função Delta de Dirac e δ'' é a sua segunda derivada. Isto gera um espaço de escala do tipo

$$\begin{aligned} F_{n\Delta t} &= (\delta + \Delta t \delta'')^{*n} * f_0 = \sum_{k=0}^n \binom{n}{k} (\Delta t)^k \frac{\partial^{2k} f_0}{\partial x^{2k}} = \\ &= f_0 + n\Delta t f_0'' + \binom{n}{2} (\Delta t)^2 f_0^{iv} + \dots + (\Delta t)^n f_0^{(2n)} \end{aligned}$$

Nó caso multidimensional, o mesmo raciocínio nos dá

$$F_{t+\Delta t}(\mathbf{x}) = (\delta + \Delta t \cdot \nabla^2 \delta)(\mathbf{x}) * F_t(\mathbf{x})$$

onde $\nabla^2 \delta$ é o Laplaciano da função δ multidimensional.

O espaço de escala obtido não satisfaz as propriedades de causalidade; até mesmo a estabilidade falha, pois a Transformada de Fourier do núcleo base

$$(\delta + \Delta t \delta'')(\omega) = 1 - 4\pi^2 \omega^2 \Delta t$$

indica que a convolução ampliará frequências grandes ($\omega \geq \frac{\sqrt{2}}{2\pi\sqrt{\Delta t}}$ - note que há também mudança de fase). Observe que, assim mesmo, o princípio do máximo

claramente é satisfeito. A propriedade de semi-grupo mais uma vez vale se Δt for fixado a priori, mas rui se vários passos de escala forem misturados.

Finalmente, vale a pena observar o que acontece tomando $\Delta t = \frac{t}{n}$ e $n \rightarrow \infty$ com t fixo. Tem-se

$$F_t = \left(\delta + \frac{t}{n} \delta'' \right)^{*n} * F_0$$

onde

$$\lim_{n \rightarrow \infty} \left(\widehat{\delta + \frac{t}{n} \delta''} \right)^{*n} = \lim_{n \rightarrow \infty} \left(1 - \frac{4\pi^2 w^2 t}{n} \right)^n = e^{-4\pi^2 w^2 t} = \hat{G}_t(w)$$

como era de se esperar.

7.3.2 Equação Implícita

Neste caso

$$\frac{F_{t+\Delta t}(x) - F_t(x)}{\Delta t} = F''_{t+\Delta t}(x)$$

Proposição 7.11 *A solução de*

$$\frac{F_{t+\Delta t}(x) - F_t(x)}{\Delta t} = F''_{t+\Delta t}(x)$$

é

$$F_{t+\Delta t}(x) = (E_{\Delta t} * F_t)(x)$$

onde E_t é a exponencial simétrica

$$E_t(x) = \frac{\lambda}{2} e^{-\lambda|x|}; \lambda = \frac{1}{\sqrt{t}}$$

Demonstração. Não é difícil verificar a solução diretamente, usando que $\left(\frac{d|x|}{dx}\right)^2 = 1$ q.t.p. e que $\frac{d^2|x|}{dx^2} = \delta(x)$. Para encontrar a solução, todavia, escreva a equação como uma convolução

$$(\delta - \Delta t \delta'') * F_{t+\Delta t} = F_t$$

e aplique a Transformada de Fourier na variável x a ambos os lados

$$\begin{aligned} (1 + \Delta t 4\pi^2 w^2) \hat{F}_{t+\Delta t}(w) &= \hat{F}_t(w) \\ \hat{F}_{t+\Delta t}(w) &= \frac{1}{1 + \Delta t 4\pi^2 w^2} \hat{F}_t(w) \end{aligned}$$

Portanto, $F_{t+\Delta t}$ é dado por uma convolução de F_t com um núcleo $E_{\Delta t}$ que satisfaz

$$\hat{E}_t(w) = \frac{1}{1 + 4\pi^2 t w^2} = \frac{1}{4\pi i \sqrt{t}} \left(\frac{1}{\frac{1}{2\pi i \sqrt{t}} + w} + \frac{1}{\frac{1}{2\pi i \sqrt{t}} - w} \right)$$

Um raciocínio não muito rigoroso é separar $\hat{E}_t(w)$ em frações como acima, e aplicar as regras conhecidas às translações em w ; assim, bastaria calcular $H(x)$ tal que $\hat{H} = \frac{1}{w}$, assim:

$$\begin{aligned}\hat{H}(w) &= \frac{1}{w} \Rightarrow \widehat{H'}(w) = 2\pi i \Rightarrow H'(x) = 2\pi i \delta \Rightarrow \\ &\Rightarrow H(x) = 2\pi i u(x)\end{aligned}$$

onde $u(x)$ é a função degrau unitário. Portanto

$$E_t(x) = \frac{1}{4\pi i \sqrt{t}} \left(2\pi i u(x) e^{-\frac{2\pi i x}{2\pi i \sqrt{t}}} + 2\pi i u(-x) e^{\frac{2\pi i x}{2\pi i \sqrt{t}}} \right) = \frac{\lambda}{2} e^{-\lambda|x|}$$

Repetimos que este raciocínio não é muito rigoroso; o melhor seria ir direto à IFT de $\hat{E}_t(w)$ usando $z = 2\pi i w$:

$$E_t(x) = \int_{-\infty}^{\infty} \frac{e^{2\pi i w x}}{1 + 4\pi^2 t w^2} dw = \frac{1}{2\pi i} \int_{-\infty i}^{\infty i} \frac{e^{xz}}{1 - z^2 t} dz$$

Se $x > 0$, use o quadrado R de vértices $-ai$, ai , $-2a + ai$, $-2a - ai$ (com a grande o suficiente) para estimar a integral acima. Em primeiro lugar, note que

$$\frac{1}{2\pi i} \int_R \frac{e^{xz}}{1 - z^2 t} dz = \frac{1}{2\pi i} \int_R \frac{\lambda}{\lambda - z} \frac{\lambda}{\lambda + z} e^{xz} dz = \frac{e^{-\lambda x}}{2\lambda}$$

devido à presença do pólo em $z = -\lambda = -\frac{1}{\sqrt{t}}$. À medida que $a \rightarrow \infty$, esperamos que a integral em R seja a integral de $-ai$ a ai . De fato, em cada um dos outros três lados de R temos

$$\left| \int \frac{e^{xz}}{1 - z^2 t} dz \right| \leq \left| \int \frac{|e^{xz}|}{|z^2 t - 1|} dz \right| \leq \left| \int \frac{1}{a^2 t - 1} dz \right| = \frac{2a}{a^2 t - 1}$$

Assim, quando $a \rightarrow \infty$, a contribuição destes três lados vai para 0, e então

$$E_t(x) = \frac{1}{2\pi i} \lim_{a \rightarrow \infty} \int_R \frac{e^{xz}}{1 - z^2 t} dz = \frac{e^{-\lambda x}}{2\lambda}$$

Finalmente, note que o raciocínio tem de ser revertido se $x < 0$; para que as estimativas funcionem, o quadrado a ser usado terá vértices $-ai$, ai , $2a + ai$, $2a - ai$, e então

$$E_t(x) = \frac{e^{\lambda x}}{2\lambda}$$

■

Assim, temos o espaço de escala gerado por

$$F_{n\Delta t}(x) = \left(\frac{e^{\lambda|x|}}{2\lambda} \right)^{*n} * F_0(x), \quad \lambda = \frac{1}{\sqrt{\Delta t}}$$

Este espaço de escala satisfaz a causalidade de cruzamentos de zero, e portanto, a causalidade de número de extremos (veja capítulo 10), desde que Δt seja fixo a priori. A Transformada de Fourier do núcleo base

$$\left| \hat{E}_{\Delta t}(w) \right| = \frac{1}{1 + 4\pi^2 w^2 \Delta t} \leq 1$$

indica que este espaço de escala é estável para qualquer $\Delta t > 0$. Também é imediato verificar que o princípio do máximo vale em sua forma modificada (um máximo local de $F_{t+\Delta t}$ acaba de descer, etc.). Mais uma vez, a propriedade de semi-grupo claramente vale se Δt é fixo a priori, mas não funciona para diferentes passos de escala.

Mais uma vez, note o que acontece fixando $t = n\Delta t$ e tomando $n \rightarrow \infty$. Tem-se mais uma vez

$$\lim_{n \rightarrow \infty} \left(\hat{E}_{\Delta t}(w) \right)^n = \lim_{n \rightarrow \infty} \left(\frac{1}{1 + 4\pi^2 w^2 \frac{t}{n}} \right)^n = e^{-4\pi^2 w^2 t} = \hat{G}_t(w)$$

Finalmente, o caso N -dimensional é bem mais complicado e gera

$$F_{n\Delta t} = (E_{\Delta t})^{*n} * F_0$$

onde

$$E_t(\mathbf{x}) = IFT \left(\frac{1}{1 + 4\pi^2 |\mathbf{w}|^2 t} \right)$$

tem de ser calculado.

7.4 Exercícios

(1) Considere uma discretização mista da equação do calor do tipo

$$\frac{F_{t+\Delta t} - F_t}{\Delta t} = \frac{1}{2} [-1 \ 2 \ -1] * F_t + \frac{1}{2} [-1 \ 2 \ -1] * F_{t+\Delta t}$$

Usando os métodos apresentados neste capítulo, mostre que ela é equivalente a

$$F_{t+\Delta t} = T_{\Delta t} * F_t$$

onde $T_{\Delta t} = 2A_{\frac{\Delta t}{2}} - \delta$

(2) O que ocorre se na discretização mista acima as constantes a e b (com $a + b = 1$) forem utilizadas ao invés de $\frac{1}{2}$ e $\frac{1}{2}$? Mostre que, neste caso

$$F_{t+\Delta t} = T_{\Delta t} * F_t$$

onde $T_{\Delta t} = \frac{1}{a} (A_{a\Delta t} - b\delta)$

Em particular, usa as expressões explícitas de $A_{\Delta t}$ e $B_{\Delta t}$, a saber

$$A_{\Delta t} = \frac{1-\alpha}{1+\alpha} (\dots\alpha^3 \alpha^2 \alpha 1 \alpha \alpha^2 \alpha^3 \dots) \text{ onde } 0 < (1-\alpha)^2 \Delta t = \alpha < 1$$

$$B_{\Delta t} = (\Delta t 1 - 2\Delta t \Delta t)$$

para mostrar que

$$\lim_{n \rightarrow \infty} n \left(A_{\frac{\Delta t}{n}} - \delta \right) = B_{\Delta t} - \delta$$

Por que isso é desejável?

Capítulo 8

Invariantes geométricos

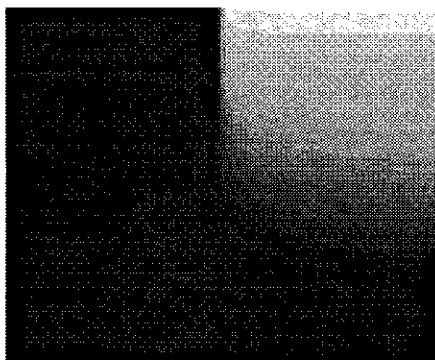
8.1 Motivação: detecção de arestas

Um passo importante do processamento de imagens é a detecção de arestas, isto é, dos pontos da imagem onde há variação brusca de intensidade de cinza (estas arestas tipicamente geram um “cartoon” da imagem, e tipicamente contêm muito mais informação do que o resto da imagem). Como fazê-lo?

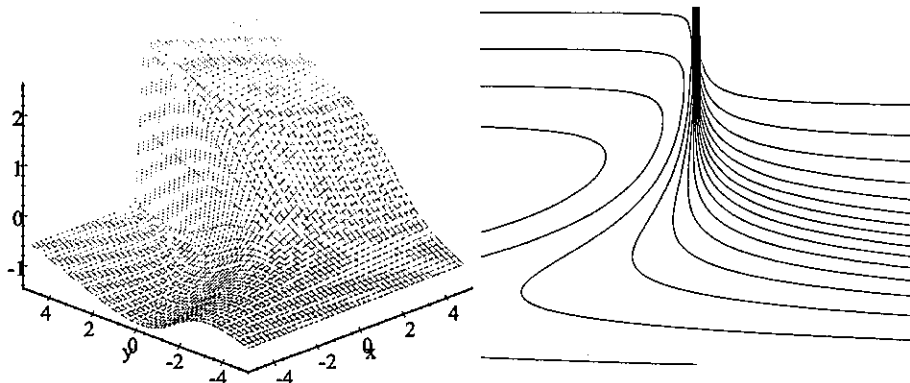
A título de motivação, convidamos o leitor a pensar primeiramente no caso 1D – o que seriam pontos de variação “brusca” de uma função $f : \mathbb{R} \rightarrow \mathbb{R}$? A resposta mais natural é procurar por pontos onde a derivada de f é grande. Mais especificamente, se quisermos eleger um ponto para cada “aresta”, a resposta natural é escolher pontos de máximo local de f_x – ou seja, pontos onde $f_{xx} = 0$ (para ser um pouco mais preciso, procuraríamos máximos locais de $|f_x|$ – que são pontos onde $f_{xx} = 0$ e $f_x f_{xxx} < 0$).

Como estender esta noção para o caso bidimensional mantendo a idéia intuitiva de “aresta”? Dada uma imagem $I : \mathbb{R}^2 \rightarrow \mathbb{R}$, infelizmente, não podemos tentar máximos locais de $|\nabla I|$ – tipicamente, esse máximos locais consistiriam de um conjunto de pontos isolados – mas queremos *curvas* no conceito de aresta! Enquanto esses pontos isolados devem fazer parte do conjunto de pontos de aresta, devemos de conectá-los de alguma forma através de curvas.

Aliás, uma das grandes dificuldades da formalização matemática do conceito de aresta é que pessoas diferentes têm idéias intuitivas distintas sobre o que deve ser uma aresta nos casos mais ambíguos. Convidamos o leitor a dizer onde ele acha que deveriam estar as arestas da imagem a seguir – e depois observar seu gráfico e suas curvas de nível, perceber que ela não é simétrica e, possivelmente, mudar ligeiramente a localização de suas arestas... Ao final do capítulo, vale a pena também examinar as imagens do exercício 1.



Onde estão as arestas?



Isso tudo dito, uma maneira razoavelmente natural de definir as arestas de I é considerar o conjunto de pontos que são máximos de $|\nabla I|$ na direção desse mesmo gradiente. Em outras palavras, escolhemos um “corte” da imagem numa direção perpendicular a suas curvas de nível; dentre os pontos deste corte, escolhemos aquele onde $|\nabla I|$ é máximo para ser o ponto de aresta. Fazendo isso para todos os cortes possíveis, teremos uma curva de pontos de aresta como desejado. Não encontraremos apenas pontos isolados pois não descartamos pontos onde $|\nabla I|$ é grande só porque há outros próximos onde $|\nabla I|$ é ainda maior – a comparação é restrita a cada um dos cortes que fizemos.

Como expressar essa condição através da geometria diferencial da função I ? Para facilitar os cálculos, utilizaremos o sistema gradiente de coordenadas locais (u, v) – isto é, dado um ponto P na imagem, olhamos para um sistema de

coordenadas com origem em $O \equiv P$ e eixos $\vec{O}u$ (tangente à curva de nível de I passando por P) e $\vec{O}v$ (na direção do gradiente ∇I passando por P). Note que, indicando derivadas parciais por índices, $I_u = 0$ e $I_v = |\nabla I|$. Queremos então pontos onde I_v é máximo na direção de v , isto é, pontos onde

$$\begin{aligned} I_{vv} &= 0 \\ I_{vvv} &< 0 \end{aligned}$$

Em coordenadas cartesianas, temos que calcular a direção do gradiente

$$\vec{e}_v = \frac{I_x \vec{i} + I_y \vec{j}}{\sqrt{I_x^2 + I_y^2}}$$

para obter

$$\begin{aligned} I_{vv} &= \partial_v |\nabla I| = \nabla \left(\sqrt{I_x^2 + I_y^2} \right) \cdot \vec{e}_v = \\ &= \left(\frac{(I_{xx}I_x + I_y I_{xy}) \vec{i} + (I_{xy}I_x + I_{yy}I_y) \vec{j}}{\sqrt{I_x^2 + I_y^2}} \right) \cdot \left(\frac{I_x \vec{i} + I_y \vec{j}}{\sqrt{I_x^2 + I_y^2}} \right) \Rightarrow \\ \Rightarrow I_{vv} &= \frac{I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2}{I_x^2 + I_y^2} \end{aligned}$$

e, de maneira semelhante

$$I_{vvv} = \nabla (I_{vv}) \cdot \vec{e}_v = \frac{I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3}{(I_x^2 + I_y^2)^{3/2}}$$

Como só interessam os sinais, podemos então olhar somente para as condições envolvendo os numeradores

$$\begin{aligned} \tilde{I}_{vv} &= I_v^2 I_{vv} = I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 = 0 \\ \tilde{I}_{vvv} &= I_v^3 I_{vvv} = I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 < 0 \end{aligned}$$

Em outras palavras, as arestas de I são os CZ de \tilde{I}_{vv} onde \tilde{I}_{vvv} é negativo.

Imediatamente, vemos no raciocínio acima uma necessidade básica do processamento de imagens: se quisermos calcular tais quantidades computacionalmente, é essencial que saibamos discretizar todas as derivadas parciais de uma imagem I . Em particular, se trabalhamos com um espaço de escala de I , podemos querer calcular as arestas desta imagem em várias escalas. Este problema será abordado no capítulo seguinte.

Mais ainda, o processo acima de detecção de arestas tem outras propriedades interessantes. Sendo apresentado como o conjunto de zeros de uma função (I_{vv}), as arestas de I podem ser calculadas de maneira simples com precisão sub-pixel; automaticamente, tais arestas serão curvas (já conectadas corretamente; compare com processos que criam vários pontos a serem conectados mais tarde).

A condição extra $I_{vvv} < 0$ simplesmente elimina partes dessas arestas, mantendo a conectividade proveniente diretamente do modelo.

Finalmente, uma análise mais cuidadosa desse modelo revela outras propriedades altamente desejáveis. Por exemplo, essa definição de arestas é invariante por rotações e translações de coordenadas. Isto pode ser visto de maneira geométrica (é claro que \vec{e}_u e \vec{e}_v acompanham tais rotações e translações) ou de maneira algébrica (a expressão que usamos pode ser escrita como derivadas de I com relação a u e v , no caso, como I_{vv}). Outras propriedades dessa representação são resumidas no quadro abaixo:

Invariância da representação	Razão geométrica	Razão algébrica
Translação de coordenadas	\vec{e}_u e \vec{e}_v acompanham translações	Derivadas de I invariantes por translações
Rotação de coordenadas	\vec{e}_u e \vec{e}_v acompanham rotações	Derivadas de I com relação a u, v invariantes por rotações
Rescalamento de coordenadas	Máximo de I_v acompanha rescalamentos	Expressões I_{vv} e I_{vvv} homogêneas em ordem de derivação; sinais não mudam
Transformação afim da imagem de I	∇I não muda direção; Máx de I_v acompanha transf. afins	Expressões I_{vv} e I_{vvv} homogêneas em grau; sinais não mudam

8.2 Invariantes diferenciais

Em geral, imagine que procuramos algum *detalhe* (característica, “feature”) da imagem I . Queremos descobrir descritores (expressões envolvendo derivadas de I) que expressem propriedades da imagem (tais detalhes serão, digamos, zeros destas expressões). No entanto, não é qualquer expressão que nos dá um bom descritor. Por exemplo, a expressão I_x depende demais do sistema de coordenadas escolhido. Queremos uma certa invariância de nossos descritores com relação a transformações básicas na imagem. As expressões que apresentam essas invariâncias são chamadas **invariantes diferenciais**.

Invariância por translações de coordenadas

Para que um descritor seja invariante por translações, é suficiente utilizar expressões que dependam somente de derivadas de I .

Invariância por rotações de coordenadas

Neste caso, basta que o descritor dependa somente de derivadas de I com relação às coordenadas locais u e v . Por exemplo, $I_u = 0$ (não muito interessante) e $I_v = |\nabla I|$ são invariantes por rotações. Em geral, qualquer expressão envolvendo $I_u, I_v, I_{uu}, I_{uv}, I_{vv}, I_{vvv}$, etc. será invariante por rotações de coordenadas.

Invariância por transformações monótonas da imagem

Estamos interessados aqui em descritores que não mudem caso recodifiquemos os níveis de cinza de uma imagem I . Matematicamente, queremos que o descritor seja o mesmo para as expressões

$$I \quad f \circ I$$

qualquer que seja a função monótona $f: \mathbf{R} \rightarrow \mathbf{R}$.

Por que isto é desejável? Imagine que mudamos o brilho da imagem (ou talvez o nível de iluminação da sala onde ela foi obtida, como um todo); os valores dos níveis de cinza mudam por uma transformação monótona de intensidade, isto é, pixels com níveis iguais de cinza continuam com os mesmos valores entre si e pixels mais escuros continuam mais escuros (apesar dos valores de cada um ter mudado possivelmente de maneira não linear). Seria interessante que nossos detalhes fossem encontrados nos mesmos locais, isto é, que a nossa função descritora não variasse à medida que alteramos o brilho da imagem original.

Intuitivamente, uma transformação monótona de níveis de cinza mantém as curvas de nível de I intactas, simplesmente mudando seus rótulos (reempilhando-as em novas alturas). Assim, qualquer expressão que dependa somente das formas das curvas de nível permanece invariante. Em particular, temos a curvatura¹ de suas curvas de nível

$$\kappa_I = \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{(I_x^2 + I_y^2)^{\frac{3}{2}}} = \frac{I_{uu}}{I_v}$$

e a curvatura das curvas integrais de seu gradiente

$$\mu_I = \frac{(I_x^2 - I_y^2)I_{xy} + I_xI_y(I_{yy} - I_{xx})}{(I_x^2 + I_y^2)^{\frac{3}{2}}} = -\frac{I_{uv}}{I_v}$$

Pode-se mostrar que κ e μ são os únicos invariantes básicos de uma imagem I utilizando derivadas de I de ordem até 2. Em outras palavras,

$$\kappa_{f \circ I} = \kappa_I$$

$$\mu_{f \circ I} = \mu_I$$

¹Para uma revisão da definição de curvatura e das expressões abaixo, veja apêndice ao final deste livro.

para qualquer função monótona $f : \mathbb{R} \rightarrow \mathbb{R}$.

Resumindo, se quisermos criar um descritor que seja invariante por mudanças de brilho, basta usarmos qualquer função de κ e μ .

Invariância por transformações afins da imagem

Se exigirmos somente invariância por transformações afins (imagine na discussão anterior que f seja da forma $f(x) = ax + b$), outros invariantes aparecem: de fato, $\partial_{x^\alpha y^\beta} (f \circ I) = a^{\alpha+\beta} (\partial_{x^\alpha y^\beta} I)$, portanto funções como

$$\frac{I_x}{I_y} \quad \frac{I_{xx}}{I_{xy}} \quad \frac{I_{xx}}{I_{yy}} \quad \frac{I_{xx}}{I_x I_y}$$

são todas invariantes por funções afins.

Note também que as soluções de equações homogêneas em grau de diferenciação, como por exemplo

$$I_x I_y I_{xx} + I_{xx} I_{yy} + I_{xxx} + I_{xyy} I_y = 0$$

(note: cada “monômio” tem 4 derivadas parciais) serão invariantes se aplicarmos uma transformação afim a I . Em outras palavras, se um detalhe é o conjunto de zeros de uma função homogênea em grau de diferenciação, tal detalhe permanecerá invariante ao aplicarmos uma função afim ao conjunto de níveis de cinza de I .

8.3 Detecção de detalhes

Como discutido acima, nossa abordagem para detecção de detalhes é encontrar expressões (descritores) dependendo de I cujos zeros sejam os detalhes procurados (alguns zeros podem ser eliminados através de seletores, isto é, inequações que envolvam outras expressões diferenciais). Se o detalhe deve apresentar certas invariâncias, espera-se que o descritores e seletores apresentem as mesmas invariâncias.

8.3.1 Arestas

Como visto acima, defina uma aresta como um conjunto de pontos onde $|\nabla I|$ é máximo na direção da curva integral do gradiente. Utilizando coordenadas u e v , procuramos pontos onde I_v é um máximo local na direção v :

$$\begin{aligned} I_{vv} &= 0 \\ I_{vvv} &< 0 \end{aligned}$$

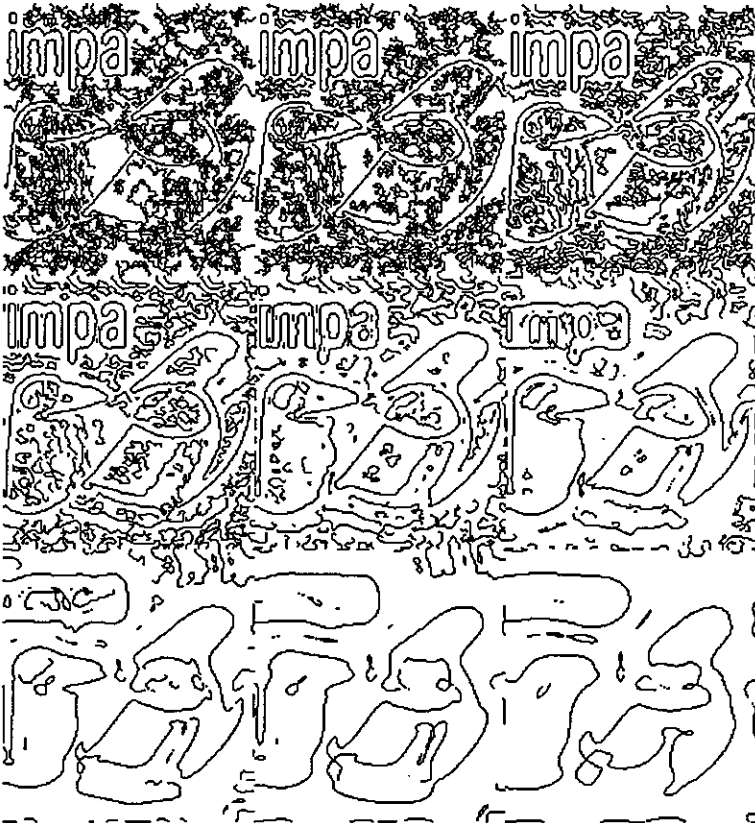
Em coordenadas cartesianas, nosso descritor I_{vv} e nosso seletor I_{vvv} são expressos como

$$I_{vv} = \frac{I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2}{I_x^2 + I_y^2} = 0$$

$$I_{vvv} = \frac{I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3}{(I_x^2 + I_y^2)^{3/2}} < 0$$

Arestas assim detectadas são invariantes por translações e rotações de coordenadas de I , assim como por transformações afins do brilho da imagem. Este detector de arestas é comumente chamado *detector de Haralick*.

A figura a seguir mostra a detecção de arestas usando este método aplicada ao espaço de escala do logotipo do Impa (que foi exibido na figura 5.1, página 60). Note como as texturas são proeminentes em escalas menores, escalas intermediárias são boas para registrar as letras do logotipo e escalas maiores tendem a detectar a estrutura da faixa de Möebius como um todo.



8.3.2 Junções

Junções são caracterizadas por pontos de alta curvatura e de alto gradiente (por proximidade a arestas). Isto sugere o uso de algum tipo de produto entre as expressões abaixo

$$\kappa = \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{(I_x^2 + I_y^2)^{3/2}} = \frac{I_{uu}}{I_v}$$

$$|\nabla I| = (I_x^2 + I_y^2)^{\frac{1}{2}} = I_v$$

enquanto $\kappa|\nabla I|$ é uma expressão razoável, seria interessante eliminarmos os problemas que surgem perto de pontos onde o denominador envolvido é pequeno. De fato, um bom descritor é

$$\tilde{\kappa} = \kappa|\nabla I|^3 = I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2 = I_{uu}I_v^2$$

pela simplicidade computacional da expressão. Para escrevermos tais detalhes como zeros de expressões, note que os máximos locais de $\tilde{\kappa}$ correspondem a

$$\begin{aligned}\tilde{\kappa}_u &= 0 \\ \tilde{\kappa}_v &= 0 \\ \tilde{\kappa}_{uu}\tilde{\kappa}_{vv} - \tilde{\kappa}_{uv}^2 &> 0 \\ \tilde{\kappa}_{uu} &< 0\end{aligned}$$

Quando assim apresentadas, junções são invariantes por translações e rotações de coordenadas da imagem, assim como por transformações afins do brilho.

8.3.3 Arestas por Laplaciano

Um descritor comum em processamento de imagens é o Laplaciano. Pontos onde o Laplaciano é zero são comumente usados para detecção de arestas, estéreo e detecção de bolhas:

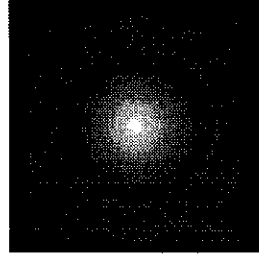
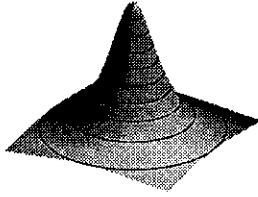
$$\Delta^2 F = F_{xx} + F_{yy} = F_{uu} + F_{vv} = 0$$

O descritor Laplaciano é invariante por rotações e translações de coordenadas; seus zeros são invariantes por mudanças de brilho. Mais ainda, o Laplaciano é um operador linear! Apesar destas propriedades desejáveis e da simplicidade de sua computação, gostaríamos de ressaltar que o uso da expressão acima providencia uma localização muito pobre de arestas que tenham alta curvatura (vide exercício 1).

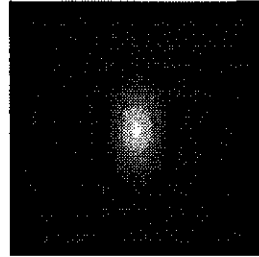
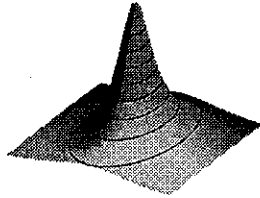
8.4 Exercícios

Laplaciano e Detecção de Arestas

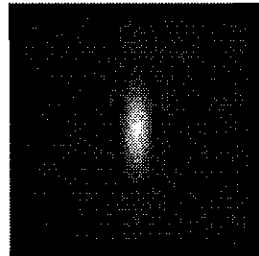
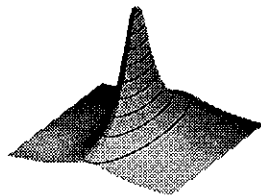
(1) Apresentamos a seguir os gráficos e as imagens correspondentes à função $f(x, y) = \frac{1}{1+x^2+cy^2}$ para 3 valores distintos de c .



(i) $c = 1$



(ii) $c = 3$



(iii) $c = 9$

(a) Marque nas imagens à direita as arestas que você detecta (use seu olho apenas; use os gráficos à esquerda apenas para driblar possíveis ruídos devidos à impressão).

(b) Utilize o descritor Laplaciano para detectar as arestas de $f(x, y)$ para qualquer $c \geq 1$. Classifique a forma das arestas de acordo com os possíveis valores de c . Em que casos a sua resposta se assemelha à obtida em (a)?

(c) O que este exercício revela a respeito do uso do Laplaciano em detecção de arestas?

(d) Numericamente, utilize o detector de Haralick para obter as mesmas arestas e compare os resultados com (a).

2) Neste capítulo, apresentamos o *sistema de coordenadas gradiente*; resumidamente, dada uma imagem (diferenciável) $I(x, y)$ e um ponto P nesta imagem, este é o par de eixos perpendiculares $u-v$ tal que **no ponto P**

$$I_u = 0 \text{ e } I_v = |\nabla I|$$

Vimos também como escrever alguns invariantes geométricos básicos neste sistema de coordenadas

$$\text{Detector de Haralick: } I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 = I_{vv}$$

$$\text{Laplaciano: } I_{xx} + I_{yy} = I_{uu} + I_{vv}$$

$$\text{Curvatura de curvas de nível: } \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{(I_x^2 + I_y^2)^{\frac{3}{2}}} = \frac{I_{uu}}{I_v}$$

as igualdades sendo válidas no ponto P . Nesta questão vamos definir o *sistema de coordenadas $p-q$* .

a) Dada uma imagem $I(x, y)$ e um ponto P nesta imagem, prove que é possível encontrar um par de eixos perpendiculares $p-q$ tais que $I_{pq} = 0$ **no ponto P** . [Sugestão: pense no valor de I_{xy} no ponto P à medida que giramos os eixos; o que acontece depois de um giro de 90 graus?]

b) É simples escrever os três invariantes citados acima neste sistema de coordenadas no ponto P . Faça-o. Exatamente que **propriedade dos invariantes** acima faz com que este processo seja, de fato, simples?

3) A expressão do detector de Haralick

$$H = F_{xx}F_x^2 + 2F_{xy}F_xF_y + F_{yy}F_y^2 = F_{vv}$$

é invariante por rotações (onde v é a direção do gradiente de F no ponto onde I vai ser calculado). No exercício acima, você viu que, no sistema de coordenadas pq

$$H = F_{pp}F_p^2 + F_{qq}F_q^2$$

pois $F_{pq} = 0$ em qualquer ponto do plano. No entanto, é importante notar que

$$\partial_p H \neq \partial_p (F_{pp}F_p^2 + F_{qq}F_q^2) = F_{ppp}F_p^2 + 2F_{pp}^2F_p + F_{pqq}F_q^2 + 2F_{qq}F_{pq}F_q$$

Como explicar esta aparente contradição?

4) A expressão

$$\tilde{\kappa} = F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2 = F_{uu}$$

é invariante por rotações (onde u é a direção perpendicular ao gradiente de F no ponto onde $\tilde{\kappa}$ vai ser calculado).

a) O sistema de coordenadas pq é tal que $F_{pq} = 0$ em A ; qual a expressão mais simples para $\tilde{\kappa}$ que pode ser escrita no ponto A ?

b) Calcule $\partial_p \tilde{\kappa}$ no ponto A . Escreva-o em função das derivadas parciais de F nas direções p e q no ponto A . [Cuidado!]

Capítulo 9

Derivadas Discretas

9.1 Introdução

Se um espaço de escala é gerado a partir de uma convolução, digamos, $F_t(\mathbf{x}) = K_t(\mathbf{x}) * f_0(\mathbf{x})$, podemos obter qualquer uma de suas derivadas parciais espaciais usando uma das seguintes expressões

$$\partial_{x^\alpha} F_t(\mathbf{x}) = \partial_{x^\alpha} (K_t(\mathbf{x}) * f_0(\mathbf{x})) = \partial_{x^\alpha} K_t(\mathbf{x}) * f_0(\mathbf{x}) = K_t(\mathbf{x}) * \partial_{x^\alpha} f_0(\mathbf{x})$$

onde ∂_{x^α} pode ser qualquer derivada parcial (possivelmente mista). Note que, mesmo que a função original f_0 não seja diferenciável o bastante para o cálculo de $\partial_{x^\alpha} f_0$, podemos calcular $\partial_{x^\alpha} (K_t(\mathbf{x}) * f_0(\mathbf{x}))$ desde que o núcleo de convolução K_t o seja. Esta expressão também mostra que $\partial_{x^\alpha} F$ é um espaço de escala (correspondendo ao sinal inicial $\partial_{x^\alpha} f_0$). No caso da difusão linear, podemos escrever

$$\partial_t (\partial_{x^\alpha} F_t) = \nabla^2 (\partial_{x^\alpha} F_t)$$

e então $\partial_{x^\alpha} F_t(\mathbf{x})$ é um espaço de escala.

Nesta seção, examinaremos como obter tais derivadas quando o espaço de escala é discreto espacialmente.

9.2 Discretização

Gostaríamos de definir um operador D_{x^α} discreto (correspondente à ∂_{x^α} no caso contínuo). Desejamos que este operador seja linear e invariante por translações, portanto podemos escrevê-lo como uma convolução:

$$D_{x^\alpha} F_t = D_{x^\alpha} * (K_t * f_0) = (D_{x^\alpha} * K_t) * f_0$$

Note que se K_t é uma família de núcleos de espaço de escala (com a propriedade de semi-grupo, como ocorre no caso N -dimensional) que comute com D_{x^α} ,

$$D_{x^\alpha} F_{t_2} = D_{x^\alpha} * K_{t_2} * f_0 = K_{t_2-t_1} * (D_{x^\alpha} * K_{t_1} * f_0) = K_{t_2-t_1} * D_{x^\alpha} F_{t_1}$$

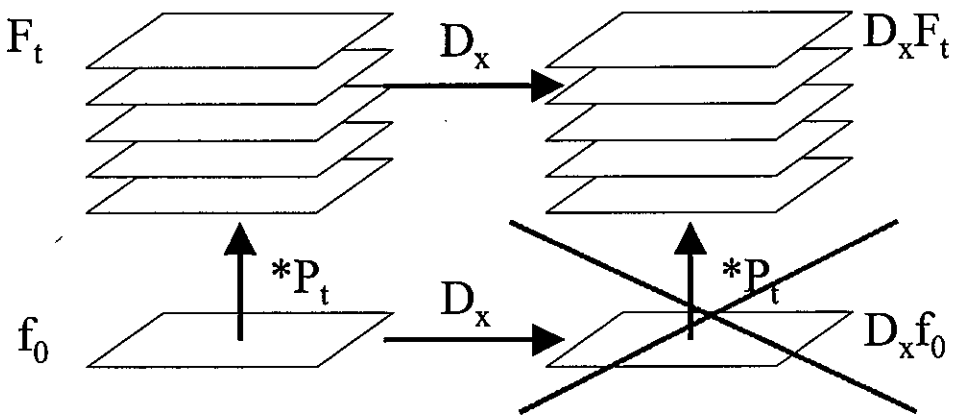
e $D_{x^\alpha} F$ também será um espaço de escala no caso discreto. De fato, quase qualquer operador/núcleo D_{x^α} satisfaz essa propriedade; nossa grande preocupação é encontrar um que seja coerente, isto é, um que aproxime a derivada contínua à medida que diminuimos o “passo espacial”.

A expressão $D_{x^\alpha} * (K_t * f_0) = K_t * (D_{x^\alpha} * f_0)$ mostra então que há duas maneiras equivalentes de obter o espaço de escala da derivada $D_{x^\alpha} * f_0$:

- Construindo diretamente o espaço de escala a partir de $D_{x^\alpha} * f_0$ por convolução com os núcleos de espaço de escala K_t ;
- Diferenciando o espaço de escala $F_t = K_t * f_0$;

Veremos abaixo que a diferenciação é, em geral, uma convolução com um núcleo de pequeno suporte (por exemplo, para imagens, há boas aproximações das derivadas de primeira e segunda ordem que correspondem a núcleos 3×3); portanto, é mais simples utilizar o segundo método caso o espaço de escala F_t esteja (ou precise ser) calculado!

Este princípio é provavelmente o mais importante deste capítulo, portanto vamos repeti-lo: caso você já tenha calculado o espaço de escala de f_0 , não calcule o espaço de escala Gaussiano de $D_{x^\alpha} f_0$ através de convoluções deste com a Gaussiana (discreta)! Ao invés, simplesmente calcule a derivada D_{x^α} do espaço de escala de f_0 . Aliás, mesmo que o espaço de escala de f_0 não tenha sido ainda calculado, o segundo método é quase tão custoso quanto o primeiro e provavelmente gerará melhores resultados; mais ainda, o segundo método é a única opção se f_0 não for diferenciável!



Calcular $F_t = f_0 * P_t$ e então $D_x F_t$ é mais simples.

Só nos resta agora encontrar tais operadores derivada. Acreditamos que a maneira correta de encontrar tais operadores discretos coerentes vem da análise de séries de Taylor próximas do ponto em questão; o resultado é o resumo abaixo, onde apenas apresentamos as *máscaras* correspondentes a cada derivada

discreta (uma “máscara” é um núcleo de convolução já refletido para facilitar a sua visualização; na nossa notação, a origem de cada máscara corresponde ao número central ou ao número destacado). O leitor interessado no processo de obtenção de tais núcleos por Fórmulas de Taylor o encontrará na seção a seguir (9.3).

9.2.1 Máscaras para derivadas 1D

Sugerimos o uso das seguintes aproximações $o(h^2)$ (daqui por diante, h representará o espaçamento do grid usado para discretização)

$$D_x = \frac{1}{2h} [-1 \ 0 \ 1]; \quad D_{xx} = \frac{1}{h^2} [1 \ -2 \ 1]$$

para gerar todas as outras derivadas através da relação recursiva

$$D_{x^n} = D_{x^{n-2}} * D_{xx}$$

obtendo uma família de aproximações (todas de ordem $o(h^2)$) cujas máscaras tem tamanho $n+1$ (se n é par) ou $n+2$ (se n é ímpar). Note que, infelizmente, $D_{xx} \neq D_x * D_x$. De fato, D_{xx} é uma versão amostrada de $D_x * D_x$.

Enquanto a opção acima é a nossa favorita, abaixo apresentamos outras opções de máscaras coerentes no caso unidimensional. A lista mostra todas as máscaras ótimas para cada tamanho apresentado.

- Tamanho 2

$$D_{x+} = \frac{1}{h} [-\dot{1} \ 1] \quad \text{ou} \quad D_{x-} = \frac{1}{h} [-1 \ \dot{1}] \quad o(h)$$

- Tamanho 3

$$D_x = \frac{1}{2h} [-1 \ 0 \ 1] \quad o(h^2)$$

$$D_{xx} = \frac{1}{h^2} [1 \ -2 \ 1] = D_{x+} * D_{x-} \quad o(h^2)$$

- Tamanho 4

$${}^4D_{x+} = \frac{1}{6h} [-2 \ -\dot{3} \ 6 \ -1] = D_{x+} * \frac{1}{6} [2 \ 5 \ -1] \quad o(h^3)$$

$${}^4D_{x-} = \frac{1}{6h} [1 \ -6 \ \dot{3} \ 2] = D_{x-} * \frac{1}{6} [2 \ 5 \ -1] \quad o(h^3)$$

$$D_{xx} = \frac{1}{h^2} [1 \ -\dot{2} \ 1 \ 0] = \frac{1}{h^2} [0 \ 1 \ -\dot{2} \ 1] \quad o(h^2)$$

$$D_{xxx+} = \frac{1}{h^3} [-1 \ \dot{3} \ -3 \ 1] = D_{xx} * D_{x+} \quad o(h)$$

$$D_{xxx-} = \frac{1}{h^3} [-1 \ 3 \ -\dot{3} \ 1] = D_{xx} * D_{x-} \quad o(h)$$

- Tamanho 5

$$\begin{aligned} {}_5D_x &= \frac{1}{12h} [1 \ -8 \ 0 \ 8 \ -1] = D_x * \frac{1}{6} [-1 \ 8 \ -1] \quad o(h^4) \\ {}_5D_{xx} &= \frac{1}{12h^2} [-1 \ 16 \ -30 \ 16 \ -1] = D_{xx} * \frac{1}{12} [-1 \ 14 \ -1] \quad o(h^4) \\ D_{xxx} &= \frac{1}{2h^3} [-1 \ 2 \ 0 \ -2 \ 1] = D_{xx} * D_x \quad o(h^2) \\ D_{xxxx} &= \frac{1}{h^4} [1 \ -4 \ 6 \ -4 \ 1] = D_{xx} * D_{xx} \quad o(h^2) \end{aligned}$$

9.2.2 Máscaras para derivadas 2D

Sugerimos o uso das seguintes máscaras 3×3 que nos dão aproximações $o(h^2)$ de suas respectivas derivadas

$$\begin{aligned} D_x &= \frac{1}{12h} \begin{bmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -1 & 0 & 1 \end{bmatrix}; \quad D_y = \frac{1}{12h} \begin{bmatrix} 1 & 4 & 1 \\ 0 & 0 & 0 \\ -1 & -4 & -1 \end{bmatrix} \\ D_{xx} &= \frac{1}{12h^2} \begin{bmatrix} 1 & -2 & 1 \\ 10 & -20 & 10 \\ 1 & -2 & 1 \end{bmatrix}; \quad D_{yy} = \frac{1}{12h^2} \begin{bmatrix} 1 & 10 & 1 \\ -2 & -20 & -2 \\ 1 & 10 & 1 \end{bmatrix} \\ D_{xy} &= \frac{1}{4h^2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \quad D_{xxyy} = \frac{1}{h^4} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \end{aligned}$$

pois estas apresentam “simetria rotacional” (num sentido que ficará mais claro adiante) de ordem $o(h^4)$. Para derivadas de maior ordem, sugerimos o uso das relações recursivas (para $m, n \geq 0$)

$$\begin{aligned} D_{x^{m+2}y^{n+2}} &= D_{xxyy} * D_{x^m y^n} \\ D_{x^{m+2}y} &= D_{xx} * D_{x^m y} \\ D_{xy^{m+2}} &= D_{yy} * D_{xy^m} \\ D_{x^{m+2}} &= D_{xx} * D_{x^m} \\ D_{y^{m+2}} &= D_{yy} * D_{y^m} \end{aligned}$$

que definem todas as derivadas parciais de maneira única.

Em geral, mostraremos que toda máscara 3×3 coerente de ordem $o(h^2)$ para D_x tem a forma

$$D_x = \frac{1}{2h} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{A}{h} M_{xyy} + \frac{B}{h} M_{xxy} + \frac{C}{h} M_{xxyy}$$

onde

$$M_{xyy} = \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix}; M_{xxy} = \frac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix};$$

$$M_{xxyy} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Para determinar os parâmetros A , B e C , vale notar algumas restrições desejáveis. Por exemplo, para que D_x seja simétrica na direção y , precisamos ter $B = 0$ (o que equivale a forçar $D_x(f(x, y)) = D_x(f(x, -y))$ para qualquer f discreta). Na literatura, é comum usar também $C = 0$ e escolher A arbitrariamente no intervalo $[0, \frac{1}{2}]$ sem uma explicação convincente; acreditamos que $C = 0$ e $A = \frac{1}{6}$ geram a aproximação mais “corretamente assimétrica”, como mostraremos na seção seguinte.

De maneira análoga, a expressão $o(h^2)$ mais genérica para D_{xx} é

$$D_{xx} = \frac{1}{h^2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{A}{h^2} M_{xyy} + \frac{B}{h^2} M_{xxy} + \frac{C}{h^2} M_{xxyy}$$

mas por motivos de simetria rotacional a serem explicitados mais tarde preferimos $A = B = 0$ e $C = \frac{1}{12}$ (que geram o conhecido “Laplaciano simétrico” encontrado na literatura).

Para D_{xy} temos a expressão $o(h^2)$ mais geral como

$$D_{xy} = \frac{1}{4h^2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} + \frac{A}{h^2} M_{xyy} + \frac{B}{h^2} M_{xxy} + \frac{C}{h^2} M_{xxyy}$$

mas preferimos $A = B = C = 0$.

Finalmente, mostraremos que a única aproximação coerente de D_{xxyy} em tamanho 3×3 é a apresentada acima.

9.3 Séries de Taylor e Derivadas Discretas*

9.3.1 Caso 1D

Máscaras de tamanho 2

Para obter máscaras de tamanho 2 que sejam candidatos para D_x , usamos uma das seguintes séries de Taylor de f

$$f(x \pm h) = f(x) \pm hf'(x) + \frac{h^2}{2} f''(x) \pm \frac{h^3}{3!} f'''(x) + o(h^4)$$

Então

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2} f''(x) + o(h^2)$$

$$\frac{f(x) - f(x-h)}{h} = f'(x) - \frac{h}{2} f''(x) + o(h^2)$$

nos dão aproximações $o(h)$ para $f'(x)$ cujos erros são comandados pela segunda derivada de f . As máscaras correspondentes são

$$D_{x+} = \frac{1}{h} [-1 \ 1] \text{ ou } D_{x-} = \frac{1}{h} [-1 \ 1]$$

Máscaras de tamanho 3

Se usarmos ambas as séries de Taylor ao mesmo tempo, temos

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{h^2}{3} f'''(x) + o(h^3)$$

que gera uma aproximação $o(h^2)$ de $f'(x)$, comandada pela terceira derivada de f , cuja máscara é a conhecida aproximação central da derivada

$$D_x = \frac{1}{2h} [-1 \ 0 \ 1]$$

Note que $D_x = D_{x+} * \frac{1}{2} [1 \ 1] = D_{x-} * \frac{1}{2} [1 \ 1]$; portanto, D_x pode ser vista como uma versão suavizada de D_{x+} ou D_{x-} . A série de Taylor explicita porque D_x é uma aproximação melhor do que D_{x+} ou D_{x-} : ganhamos uma ordem de aproximação (de h a h^2). O preço a pagar é uma máscara maior.

Podemos também usar uma combinação linear de $f(x+h)$, $f(x-h)$ e $f(x)$ para zerar os termos em f e f' :

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f''(x) + \frac{h^2}{4!} f^{(4)}(x) + o(h^3)$$

obtendo uma aproximação $o(h^2)$ de $f''(x)$ comandada pela quarta derivada de f e cuja máscara é

$$D_{xx} = \frac{1}{h^2} [1 \ -2 \ 1]$$

Máscaras de tamanho 4

Para tamanho 4, escrevemos as séries em ordem $o(h^5)$

$$f(x+2h) = f(x) + 2hf'(x) + \frac{4h^2}{2!} f''(x) + \frac{8h^3}{3!} f'''(x) + \frac{16h^4}{4!} f^{(4)}(x) + o(h^5)$$

$$f(x \pm h) = f(x) \pm hf'(x) + \frac{h^2}{2} f''(x) \pm \frac{h^3}{3!} f'''(x) + \frac{h^4}{4!} f^{(4)}(x) + o(h^5)$$

Para simplificar o trabalho, escreva os coeficientes destas séries numa matriz 4×4 como abaixo

$$\begin{array}{cccc}
 & f & hf' & \frac{h^2}{2}f'' & \frac{h^3}{3}f''' \\
 f(x-h) & 1 & -1 & 1 & -1 \\
 f(x) & 1 & 0 & 0 & 0 \\
 f(x+h) & 1 & 1 & 1 & 1 \\
 f(x+2h) & 1 & 2 & 4 & 8
 \end{array}$$

A matriz acima mostra como escrever $f(x+ah)$ ($a = -1, 0, 1, 2$) como combinações lineares das derivadas de f no ponto x (de acordo com as séries de Taylor de ordem $o(h^4)$). Queremos exatamente o contrário: escrever as derivadas de f como combinações lineares das expressões do tipo $f(x+ah)$. Para tanto, tudo o que temos de fazer é inverter a matriz 4×4 acima! Assim, como

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{1}{3} & -\frac{1}{2} & 1 & -\frac{1}{6} \\ \frac{1}{2} & -1 & \frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{6} \end{bmatrix}$$

então, lendo as linhas

$$\begin{aligned}
 f(x) &= 0f(x-h) + f(x) + 0f(x+h) + 0f(x+2h) + o(h^4) \\
 hf'(x) &= -\frac{1}{3}f(x-h) - \frac{1}{2}f(x) + f(x+h) - \frac{1}{6}f(x+2h) + o(h^4) \\
 \frac{h^2}{2}f''(x) &= \frac{1}{2}f(x-h) - f(x) + \frac{1}{2}f(x+h) + 0f(x+2h) + o(h^4) \\
 \frac{h^3}{3!}f'''(x) &= -\frac{1}{6}f(x-h) + \frac{1}{2}f(x) - \frac{1}{2}f(x+h) + \frac{1}{6}f(x+2h) + o(h^4)
 \end{aligned}$$

Enquanto a primeira expressão não é útil, as outras três nos dão máscaras de tamanho 4 que aproximam as derivadas do lado esquerdo. As máscaras são

$$\begin{aligned}
 {}_4D_{x+} &= \frac{1}{6h} [-2 \quad -\overset{\circ}{3} \quad 6 \quad -1] \\
 D_{xx} &= \frac{1}{h^2} [1 \quad -\overset{\circ}{2} \quad 1 \quad 0] \\
 D_{xxx+} &= \frac{1}{h^3} [-1 \quad \overset{\circ}{3} \quad -3 \quad 1]
 \end{aligned}$$

que correspondem a aproximações pelo menos $o(h^3)$, $o(h^2)$ e $o(h)$, respectivamente. Note que D_{xx} é a máscara de tamanho 3 encontrada anteriormente. Analisemos as outras máscaras individualmente.

Primeiro, note que ${}_4D_{x+} = D_{x+} * \frac{1}{6} [2 \ 5 \ -1]$ pode ser considerada um tipo de suavização de D_{x+} . A combinação correspondente

$$\frac{-2f(x-h) - 3f(x) + 6f(x+h) - f(x+2h)}{6h} = f'(x) + \frac{h^3}{12}f^{iv}(x) + o(h^4)$$

nos dá uma aproximação $o(h^3)$ cujo erro é comandado pela quarta derivada.

Para D_{xxx+} , obtemos a combinação

$$\frac{-f(x-h) + 3f(x) - 3f(x+h) + f(x+2h)}{h^3} = f'''(x) + \frac{h}{2}f^{iv}(x) + o(h^2)$$

que nos dá uma aproximação $o(h)$ cujo erro também é comandado pela quarta derivada. Note que $D_{xxx+} = D_{xx} * D_{x+}$.

Outras possibilidades surgem se utilizarmos a série de $f(x-2h)$ ao invés de $f(x+2h)$. Enquanto a máscara para D_{xx} não muda, temos

$$\begin{aligned} {}_4D_{x-} &= \frac{1}{6h} [1 \ -6 \ 3 \ 2] = D_{x-} * \frac{1}{6} [-1 \ 5 \ 2] \\ D_{xxx-} &= \frac{1}{h^3} [-1 \ 3 \ 3 \ 1] = D_{xx} * D_{x-} \end{aligned}$$

Máscaras de tamanho 5

As séries

$$\begin{aligned} f(x \pm 2h) &= \\ &= f(x) \pm 2hf'(x) + \frac{4h^2}{2!}f''(x) \pm \frac{8h^3}{3!}f'''(x) + \frac{16h^4}{4!}f^{iv}(x) \pm \frac{32h^5}{5!}f^v(x) + o(h^6) \end{aligned}$$

$$\begin{aligned} f(x \pm h) &= \\ &= f(x) \pm hf'(x) + \frac{h^2}{2}f''(x) \pm \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f^{iv}(x) \pm \frac{h^5}{5!}f^v(x) + o(h^6) \end{aligned}$$

dão origem à matriz (em $o(h^5)$)

	f	hf'	$\frac{h^2}{2}f''$	$\frac{h^3}{3!}f'''$	$\frac{h^4}{4!}f^{iv}$
$f(x-2h)$	1	-2	4	-8	16
$f(x-h)$	1	-1	1	-1	1
$f(x)$	1	0	0	0	0
$f(x+h)$	1	1	1	1	1
$f(x+2h)$	1	2	4	8	16

cuja inversa é

	$f(x-2h)$	$f(x-h)$	$f(x)$	$f(x+h)$	$f(x+2h)$
f	0	0	1	0	0
hf'	$\frac{1}{12}$	$-\frac{2}{3}$	0	$\frac{2}{3}$	$-\frac{1}{12}$
$\frac{h^2}{2}f''$	$-\frac{1}{24}$	$\frac{2}{3}$	$-\frac{5}{4}$	$\frac{2}{3}$	$-\frac{1}{24}$
$\frac{h^3}{3!}f'''$	$-\frac{1}{12}$	$\frac{1}{6}$	0	$-\frac{1}{6}$	$\frac{1}{12}$
$\frac{h^4}{4!}f^{iv}$	$\frac{1}{24}$	$-\frac{1}{6}$	$\frac{1}{4}$	$-\frac{1}{6}$	$\frac{1}{24}$

As linhas nos dão as seguintes máscaras

$$\begin{aligned} {}_5D_x &= \frac{1}{12h} [1 \ -8 \ 0 \ 8 \ -1] = D_x * \frac{1}{6} [-1 \ 8 \ -1] \\ {}_5D_{xx} &= \frac{1}{12h^2} [-1 \ 16 \ -30 \ 16 \ -1] = D_{xx} * \frac{1}{12} [-1 \ 14 \ -1] \\ D_{xxx} &= \frac{1}{2h^3} [-1 \ 2 \ 0 \ -2 \ 1] = D_x * D_{xx} \\ D_{xxxx} &= \frac{1}{h^4} [1 \ -4 \ 6 \ -4 \ 1] = D_{xx} * D_{xx} \end{aligned}$$

que serão aproximações pelo menos $o(h^4)$, $o(h^3)$, $o(h^2)$ e $o(h)$ de suas respectivas derivadas.

De fato, ${}_5D_x$ corresponde à expressão

$$\begin{aligned} \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h} &= \\ &= f'(x) - \frac{2h^4}{5} f^{iv}(x) + o(h^5) \end{aligned}$$

que nos dá uma aproximação $o(h^4)$ cujo erro é comandado pela quinta derivada de f .

Para ${}_5D_{xx}$, obtemos um “cancelamento extra” e

$$\begin{aligned} \frac{-f(x-2h) + 16f(x-h) - 30f(x) + 16f(x+h) - f(x+2h)}{12h^2} &= \\ &= f''(x) - \frac{4h^4}{15} f^{vi}(x) + o(h^5) \end{aligned}$$

é uma aproximação $o(h^4)$ com erro comandado pela sexta derivada¹.

Para D_{xxx} , temos uma aproximação $o(h^2)$ comandada pela quinta derivada

$$\begin{aligned} \frac{-f(x-2h) + 2f(x-h) - 2f(x+h) + f(x+2h)}{2h^3} &= \\ &= f'''(x) - 3h^2 f^{iv}(x) + o(h^3) \end{aligned}$$

Finalmente, para D_{xxxx} , obtemos uma ordem extra: $o(h^2)$ na sexta derivada

$$\begin{aligned} \frac{f(x-2h) - 4f(x-h) + 6f(x) - 4f(x+h) + f(x+2h)}{h^4} &= \\ &= f^{iv}(x) + 4h^2 f^{vi}(x) + o(h^3) \end{aligned}$$

¹Note que esta aproximação é melhor do que a gerada por $D_x * D_x = \frac{1}{4h^2} [1 \ 0 \ -2 \ 0 \ 1]$. De fato, esta expressão corresponde a

$$\frac{f(x-2h) - 2f(x) + f(x+2h)}{4h^2} = f''(x) + \frac{h^2}{3} f^{iv}(x) + o(h^3)$$

que geraria uma aproximação apenas $o(h^2)$ (precisão já obtida com a máscara de tamanho 3).

Conclusão

Note como as melhores máscaras de tamanho 5 para a terceira e quarta derivadas são $D_{xxx} = D_x * D_{xx}$ e $D_{xxxx} = D_{xx} * D_{xx}$ (ambas de ordem $o(h^2)$). Em geral, pode-se mostrar que a relação recursiva

$$D_{x^n} = D_{x^{n-2}} * D_{xx}$$

nos dá aproximações $o(h^2)$ cujas máscaras tem tamanho $n + 1$ (se n é par) ou $n + 2$ (se n é ímpar). Tais derivadas discretas são ótimas para seus tamanhos e simples de computar, justificando essa nossa escolha.

O leitor familiar com Álgebra Linear poderá entender o problema de obtenção de máscaras da seguinte forma:

- O espaço das máscaras de tamanho n é identificado com \mathbb{R}^n ;
- Uma base é dada pelas máscaras elementares $e_1 = [1\ 0\ 0 \dots 0]$, $e_2 = [0\ 1\ 0 \dots 0]$, ... e $e_n = [0\ 0\ 0 \dots 1]$, cada uma delas representando uma expressão do tipo $f(x \pm ah)$;
- As séries de Taylor de ordem $o(h^n)$ nos mostram como escrever a base $\{e_i\}$ em termos de uma outra base, dada pelos vetores $v_0 = f(x)$, $v_1 = hf'(x)$, $v_2 = \frac{h^2}{2} f''(x)$, ..., $v_n = \frac{h^n}{n!} f^{(n)}(x)$;
- As matrizes montadas a partir dos coeficientes da série de Taylor são nada mais que matrizes de mudança de base!
- Queremos as derivadas escritas como máscaras, isto é, queremos escrever os vetores v_j em função dos vetores e_i ; portanto, tudo que há a fazer é obter a matriz de mudança de base inversa.
- Note que as matrizes que aparecem segundo este método serão matrizes de Vandermonde **não-singulares**, justificando o porquê de $\{v_j\}$ ser de fato uma base; assim, o método funciona para qualquer tamanho de máscara desejado.

9.3.2 Caso 2D

O método acima pode ser estendido para o caso bidimensional se utilizarmos séries de Taylor a duas variáveis. Na primeira subseção abaixo, abordamos o problema de determinar boas máscaras 3×3 na forma da resolução de um sistema linear de equações; veremos que sobram 3 “graus de liberdade” na determinação das derivadas de primeira e segunda ordem, e restringimos tal liberdade procurando outras propriedades desejáveis de nossas aproximações (além da coerência quando $h \rightarrow 0$).

A presença de tais graus de liberdade não deve surpreender os leitores mais familiares com Álgebra Linear: a escolha da “base de derivadas” (a base $\{v_j\}$ na notação anterior) não é tão imediata quanto era no caso 1D. No exemplo acima, as máscaras 3×3 correspondem a um espaço de dimensão 9. Se escrevermos

as séries de Taylor em ordem $o(h^3)$, obtemos as derivadas $f, f_x, f_y, f_{xx}, f_{xy}$ e f_{yy} - 6 candidatos a vetores. Assim, haverá uma ambigüidade na escolha de máscaras, explicando a presença dos 3 graus de liberdade. Por outro lado, a utilização de ordem $o(h^4)$ leva a 10 candidatos a vetores; pior ainda, não só eles são linearmente dependentes (como teriam de ser), mas pode-se ver que estes 10 candidatos varrem um espaço de dimensão apenas 8. Mesmo que joguemos fora um deles (digamos, f_{xxx}), ainda não podemos resolver o problema de mudança de base! Na segunda subseção abaixo mostramos como escolher uma base de derivadas corretamente.

Máscaras 3×3 por sistemas lineares

Como no caso 1D, utilizamos combinações lineares de fórmulas de Taylor; a duas variáveis e em ordem $o(h^4)$, temos

$$\begin{aligned} f(x+ah, y+bh) = & f(x, y) + (ah) f_x(x, y) + (bh) f_y(x, y) + \\ & + \frac{(ah)^2}{2} f_{xx}(x, y) + (ah)(bh) f_{xy}(x, y) + \frac{(bh)^2}{2} f_{yy}(x, y) + \\ & + \frac{(ah)^3}{3!} f_{xxx}(x, y) + \frac{3(ah)^2(bh)}{3!} f_{xxy}(x, y) + \\ & + \frac{3(ah)(bh)^2}{3!} f_{xyy}(x, y) + \frac{(bh)^3}{3!} f_{yyy}(x, y) + o(h^4) \end{aligned}$$

Portanto uma máscara 3×3 dada por

$$\begin{bmatrix} \alpha_6 & \alpha_7 & \alpha_8 \\ \alpha_3 & \alpha_4 & \alpha_5 \\ \alpha_0 & \alpha_1 & \alpha_2 \end{bmatrix}$$

corresponde a $\alpha_0 f(x-h, y-h) + \alpha_1 f(x, y-h) + \dots + \alpha_8 f(x+h, y+h)$, que terá os seguintes termos de ordem até h^2

$$\begin{aligned} \text{em } f &: \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 + \alpha_7 + \alpha_8 \\ \text{em } hf_x &: -\alpha_0 + \alpha_2 - \alpha_3 + \alpha_5 - \alpha_6 + \alpha_8 \\ \text{em } hf_y &: -\alpha_0 - \alpha_1 - \alpha_2 + \alpha_5 + \alpha_7 + \alpha_8 \\ \text{em } \frac{h^2}{2} f_{xx} &: \alpha_0 + \alpha_2 + \alpha_3 + \alpha_5 + \alpha_6 + \alpha_8 \\ \text{em } \frac{h^2}{2} f_{yy} &: \alpha_0 + \alpha_1 + \alpha_2 + \alpha_6 + \alpha_7 + \alpha_8 \\ \text{em } h^2 f_{xy} &: \alpha_0 - \alpha_2 - \alpha_6 + \alpha_8 \end{aligned}$$

Agora escolha a sua derivada discreta favorita e resolva o sistema linear correspondente! Por exemplo, para encontrar uma aproximação para f_x , force os termos acima a serem $0, \frac{1}{h}, 0, 0, 0$ e 0 respectivamente – encontrar-se-á uma aproximação **pelo menos** $o(h^2)$ para tal derivada (após a inevitável divisão por h); para encontrar f_{xx} , tente $0, 0, 0, \frac{2}{h^2}, 0$ e 0 (aproximação **pelo menos** $o(h)$ de f_{xx}); para f_{xy} , tente $0, 0, 0, 0, \frac{2}{h^2}$; e assim por diante.

Vale a pena notar que há 6 equações lineares para 9 incógnitas; portanto, nossas soluções terão 3 graus de liberdade. Aliás, como os sistemas de equações lineares para f_x, f_{xx}, f_{xy} , etc. são **exatamente os mesmos a menos dos termos independentes**, vale a pena encontrar primeiro a solução homogênea do sistema linear (que nos dá os tais 3 graus de liberdade). Igualando os 6 termos acima a 0, encontramos a seguinte solução homogênea

$$S_0 : \begin{bmatrix} \alpha_6 & \alpha_7 & \alpha_8 \\ \alpha_3 & \alpha_4 & \alpha_5 \\ \alpha_0 & \alpha_1 & \alpha_2 \end{bmatrix} = AM_{xyy} + BM_{xxy} + CM_{xxyy}$$

Agora, basta encontrar uma solução particular para cada sistema mencionado acima. Por exemplo, o sistema linear para f_x tem como solução particular

$$\alpha_0 = \alpha_1 = \alpha_2 = \alpha_4 = \alpha_7 = \alpha_8 = \alpha_9 = 0; \quad \alpha_3 = -\frac{1}{2h}; \quad \alpha_5 = \frac{1}{2h}$$

chegando à solução geral

$$D_x = \frac{1}{2h} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{S_0}{h}$$

(separamos h dos coeficientes A, B e C dentro de S_0 por conveniência).

Similarmente, as soluções particulares para D_{xx} e D_{xy} dão as expressões gerais

$$D_{xx} = \frac{1}{h^2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{S_0}{h^2}$$

$$D_{xy} = \frac{1}{4h^2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} + \frac{S_0}{h^2}$$

As expressões gerais de D_y e D_{yy} são análogas às de D_x e D_{xx} .

Agora, qual o significado de cada um desses graus de liberdade em S_0 ? Como escolher os seus parâmetros? Para responder a esta pergunta, temos de voltar às Fórmulas de Taylor e usar ordem $o(h^6)$. Os termos até h^2 já haviam sido determinados; os termos com h^3 e h^4 são

$$\begin{aligned} \text{em } \frac{h^3}{3!} f_{xxx} &: -\alpha_0 + \alpha_2 - \alpha_3 + \alpha_5 - \alpha_6 + \alpha_8 \text{ (igual ao termo em } hf_x) \\ \text{em } \frac{3h^3}{3!} f_{xxy} &: -\alpha_0 - \alpha_2 + \alpha_6 + \alpha_8 \\ \text{em } \frac{3h^3}{3!} f_{xyy} &: -\alpha_0 + \alpha_2 - \alpha_6 + \alpha_8 \\ \text{em } \frac{h^3}{3!} f_{yyy} &: -\alpha_0 - \alpha_1 - \alpha_2 + \alpha_6 + \alpha_7 + \alpha_8 \text{ (igual ao termo em } hf_y) \\ \text{em } \frac{h^4}{4!} f_{xxxx} &: \alpha_0 + \alpha_2 + \alpha_3 + \alpha_5 + \alpha_6 + \alpha_8 \text{ (igual ao termo em } \frac{h^2}{2} f_{xx}) \\ \text{em } \frac{4h^4}{4!} f_{xxyy} &: \alpha_0 - \alpha_2 - \alpha_6 + \alpha_8 \text{ (igual ao termo em } h^2 f_{xy}) \\ \text{em } \frac{6h^4}{4!} f_{xyyy} &: \alpha_0 + \alpha_2 + \alpha_6 + \alpha_8 \\ \text{em } \frac{4h^4}{4!} f_{xyyy} &: \alpha_0 - \alpha_2 - \alpha_6 + \alpha_8 \text{ (igual ao termo em } h^2 f_{xy}) \\ \text{em } \frac{h^4}{4!} f_{yyyy} &: \alpha_0 + \alpha_1 + \alpha_2 + \alpha_6 + \alpha_7 + \alpha_8 \text{ (igual ao termo em } \frac{h^2}{2} f_{yy}) \end{aligned}$$

Assim, estabelecemos as seguintes 9 correspondências (aqui com termos em

h^5 também; a última foi calculada em $o(h^8)$)

$$\begin{aligned}
 & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot f = f \\
 \frac{1}{2} & \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot f = hf_x + \frac{h^3}{6}f_{xxx} + \frac{h^5}{120}f_{xxxxx} + o(h^6) \\
 \frac{1}{2} & \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \cdot f = hf_y + \frac{h^3}{6}f_{yyy} + \frac{h^5}{120}f_{yyyyy} + o(h^6) \\
 & \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot f = h^2f_{xx} + \frac{h^4}{12}f_{xxxx} + o(h^6) \\
 & \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot f = h^2f_{yy} + \frac{h^4}{12}f_{yyyy} + o(h^6) \\
 \frac{1}{4} & \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \cdot f = h^2f_{xy} + \frac{h^4}{6}f_{xxy} + \frac{h^4}{6}f_{xyy} + o(h^6) \\
 \frac{1}{2} & \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix} \cdot f = h^3f_{xyy} + \frac{20h^5}{5!}f_{xxyy} + \frac{10h^5}{5!}f_{xyyy} + o(h^6) \\
 \frac{1}{2} & \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix} \cdot f = h^3f_{xxy} + \frac{10h^5}{5!}f_{xxxy} + \frac{20h^5}{5!}f_{xxyy} + o(h^6) \\
 & \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \cdot f = h^4f_{xxyy} + \frac{h^6}{12}(f_{xxxxyy} + f_{xxyyyy}) + o(h^8)
 \end{aligned}$$

que explicitam uma base do espaço vetorial de todas as máscaras 3×3 que é adequado à descrição das suas correspondentes aproximações de Taylor. Estas relações são fundamentais no entendimento da relação entre máscaras e aproximações discretas de derivadas! Por exemplo, note que as 3 últimas expressões (correspondentes aos tais graus de liberdade) são de ordem h^3 e h^4 , e portanto não farão diferença alguma se procuramos combinações lineares que são aproximações $o(h^2)$ de algo! Denotaremos as 9 máscaras acima respectivamente por M_0 , M_x , M_y , M_{xx} , M_{yy} , M_{xy} , M_{xyy} , M_{xxy} e M_{xxyy} (de acordo com o primeiro termo de suas Fórmulas de Taylor).

Muitos argumentariam a esta altura que a melhor escolha para D_x seria então $\frac{M_x}{h}$. Afinal, por que acrescentar M_{xyy} , M_{xxy} ou M_{xxyy} ? No entanto, note que a série de Taylor associada a M_x já vem com um termo em f_{xxx} do qual não podemos nos livrar (pelo menos não dentro do tamanho 3×3 imposto). Talvez a presença dessas outras derivadas de ordem mais alta possa “abrandar” a

intrusão do termo em f_{xxx} . A escolha dos termos a serem adicionados dependem do objetivo da aproximação.

Um possível critério de escolha é o seguinte: no caso contínuo, o operador ∂_x não é rotacionalmente simétrico, mas sua integral com relação a x é rotacionalmente simétrica. Em outras palavras, queremos manter a propriedade de que f_x é a derivada com relação a x de um objeto rotacionalmente simétrico (a saber, f). Enquanto é impossível manter tal propriedade no caso discreto (já que a introdução de um grid dificulta até mesmo a definição do conceito "rotacionalmente simétrico"), podemos pelo menos manter tal propriedade na aproximação de Taylor correspondente. Afinal, a aproximação geral de D_x com parâmetros A , B e C corresponde a

$$\begin{aligned} D_x \cdot f &= \frac{1}{h} (M_x + AM_{xyy} + BM_{xxy} + CM_{xxyy}) \cdot f = \\ &= f_x + \frac{h^2}{6} f_{xxx} + Ah^2 f_{xyy} + Bh^2 f_{xxy} + Ch^3 f_{xxyy} + o(h^4) = \\ &= \left(f + h^2 \left(\frac{f_{xx}}{6} + Af_{yy} \right) + Bh^2 f_{xy} + Ch^3 f_{xxy} \right)_x + o(h^4) \end{aligned}$$

que será a derivada com relação a x de uma expressão rotacionalmente simétrica (em ordem $o(h^4)$) se e somente se $A = \frac{1}{6}$ e $B = C = 0$. Acrescentamos ainda mais uma ordem para ver que essa escolha de parâmetros nos dá

$$D_x \cdot f = \frac{\partial}{\partial x} \left(f + \frac{h^2}{6} \nabla^2 f \right) + \frac{h^4}{5!} \left(f_{xxxx} + \frac{10}{3} f_{xxxyy} + \frac{5}{3} f_{xyyyy} \right) + o(h^5)$$

(note que é impossível conseguir a simetria rotacional em ordem $o(h^5)$) e a máscara correspondente é, explicitamente

$$D_x = \frac{1}{12h} \begin{bmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -1 & 0 & 1 \end{bmatrix}$$

(a máscara D_y é obtida de modo análogo).

Uma vantagem adicional de tal escolha é que o cálculo da norma do gradiente usando tais máscaras será automaticamente rotacionalmente simétrico em ordem $o(h^4)$. De fato, escrevendo $g = f + \frac{h^2}{6} \nabla^2 f$, vê-se que

$$\begin{aligned} D_x \cdot f &= g_x + o(h^4); \quad D_y \cdot f = g_y + o(h^4) \Rightarrow \\ &\Rightarrow (D_x \cdot f)^2 + (D_y \cdot f)^2 = |\nabla g|^2 + o(h^4) \end{aligned}$$

note-se que nenhuma outra máscara 3×3 apresenta tal simetria no cálculo da norma do gradiente até esta ordem.

Similarmente, partimos da expressão geral

$$\begin{aligned} \frac{1}{h^2} (M_{xx} + AM_{xyy} + BM_{xxy} + CM_{xxyy}) \cdot f &= \\ &= f_{xx} + \frac{h^2}{12} f_{xxx} + Ahf_{xyy} + Bhf_{xxy} + Ch^2 f_{xxyy} + o(h^3) = \\ &= \frac{\partial^2}{\partial x^2} \left(f + Bhf_y + \frac{h^2}{12} f_{xx} + Ch^2 f_{yy} \right) + Ahf_{xxy} + o(h^3) \end{aligned}$$

e vemos que D_{xx} será a derivada segunda com relação a x de algo rotacionalmente simétrico até a maior ordem possível quando $A = B = 0$ e $C = \frac{1}{12}$.

Usando esses parâmetros (e uma ordem adicional)

$$D_{xx} \cdot f = \frac{\partial^2}{\partial x^2} \left(f + \frac{h^2}{12} \nabla^2 f \right) + o(h^4)$$

e obtemos uma “simetria rotacional” de ordem $o(h^4)$! A expressão correspondente de D_{xx} é

$$D_{xx} = \frac{1}{12h^2} \begin{bmatrix} 1 & -2 & 1 \\ 10 & -20 & 10 \\ 1 & -2 & 1 \end{bmatrix}$$

A vantagem adicional desta escolha (e da escolha análoga para D_{yy}) é que o cálculo do Laplaciano usando esta máscara é rotacionalmente simétrico em ordem $o(h^4)$. De fato, tome $g = f + \frac{h^2}{12} \nabla^2 f$ e temos

$$\begin{aligned} D_{xx} \cdot f &= \frac{\partial^2 g}{\partial x^2} + o(h^4); \quad D_{yy} \cdot f = \frac{\partial^2 g}{\partial y^2} + o(h^4) \Rightarrow \\ &\Rightarrow (D_{xx} + D_{yy}) \cdot f = \nabla^2 g + o(h^4) \end{aligned}$$

e note-se novamente que **nenhuma outra escolha de máscaras 3×3 apresenta tal simetria rotacional até esta ordem.**

Para D_{xy}

$$\begin{aligned} \frac{1}{h^2} (M_{xy} + AM_{xyy} + BM_{xxy} + CM_{xxyy}) &= \\ &= f_{xy} + \frac{h^2}{6} f_{xxy} + \frac{h^2}{6} f_{xyy} + Ahf_{xyy} + Bhf_{xxy} + Ch^2 f_{xxyy} + o(h^3) = \\ &= \frac{\partial^2}{\partial x \partial y} \left(f + \frac{h^2}{6} \nabla^2 f + Ahf_y + Bhf_x + Ch^2 f_{xy} \right) + o(h^3) \end{aligned}$$

e a escolha natural é de fato $A = B = C = 0$.

Usando uma ordem adicional nota-se que

$$D_{xy} \cdot f = \frac{\partial^2}{\partial x \partial y} \left(f + \frac{h^2}{6} \nabla^2 f \right) + o(h^4)$$

e obtemos novamente $o(h^4)$ na “simetria” desejada. A expressão final para D_{xy} é

$$D_{xy} = \frac{1}{4h^2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

Finalmente, note que a única combinação linear das 9 máscaras acima que dá uma aproximação de D_{xxyy} corresponde a tomarmos apenas a última, a saber

$$D_{xxyy} = \frac{1}{h^2} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

então

$$\begin{aligned} D_{xxyy} \cdot f &= f_{xxyy} + \frac{h^2}{12} (f_{xxxxyy} + f_{xxyyyy}) + o(h^4) = \\ &= \frac{\partial^4}{\partial x^2 \partial y^2} \left(f + \frac{h^2}{12} \nabla^2 f \right) + o(h^4) \end{aligned}$$

é novamente “rotacionalmente simétrica” em $o(h^4)$.

Método geral

No caso geral, se procurarmos máscaras $m \times n$, a base de derivadas correta a se usar é $\frac{h^{i+j}}{i!j!} f_{x^i y^j}$ onde $i \in \{0, 1, \dots, m-1\}$ e $j \in \{0, 1, \dots, n-1\}$. Com esta base podemos estender o método do caso 1D com sucesso.

- Caso 2×2 : usamos a seguinte matriz 4×4

Coef.de	→ em	↓	f	hf_x	hf_y	$h^2 f_{xy}$	
			$f(x, y)$	1	0	0	0
			$f(x+h, y)$	1	1	0	0
			$f(x, y+h)$	1	0	1	0
			$f(x+h, y+h)$	1	1	1	1

e sua inversa

	$f(x, y)$	$f(x+h, y)$	$f(x, y+h)$	$f(x+h, y+h)$
f	1	0	0	0
hf_x	-1	1	0	0
hf_y	-1	0	1	0
$h^2 f_{xy}$	1	-1	-1	1

cujas linhas (quando reformatadas em 2×2) nos dão

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} \cdot f &= hf_x + \frac{h^2}{2}f_{xx} + \frac{h^3}{3!}f_{xxx} + o(h^4) \\ \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \cdot f &= hf_y + \frac{h^2}{2}f_{yy} + \frac{h^3}{3!}f_{yyy} + o(h^4) \\ \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \cdot f &= h^2f_{xy} + \frac{h^3}{2}f_{xxy} + \frac{h^3}{2}f_{xyy} + o(h^4) \end{aligned}$$

(note como temos de verificar para onde vão os termos f_{xx} e f_{yy} que não apareciam na nossa base). Uma aproximação de $\frac{\partial}{\partial x}$ tem um grau de liberdade (assim como $\frac{\partial}{\partial y}$); é possível também aproximar $\frac{\partial^2}{\partial x \partial y}$ com máscaras 2×2 :

$$\begin{aligned} D_{x+} &= \frac{1}{h} \left(\begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} + C \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \right) \cdot \frac{\partial}{\partial x} \quad o(h) \\ D_{x+y+} &= \frac{1}{h^2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \frac{\partial^2}{\partial x \partial y} \quad o(h) \end{aligned}$$

- Caso 3×3 : usamos a seguinte matriz 9×9

Ponto	f	hf_x	$\frac{h^2}{2}f_{xx}$	hf_y	h^2f_{xy}	$\frac{h^3}{2}f_{xxy}$	$\frac{h^2}{2}f_{yy}$	$\frac{h^3}{2}f_{xyy}$	$\frac{h^4}{4}f_{xyyy}$
$-h, -h$	1	-1	1	-1	1	-1	1	-1	1
$0, -h$	1	0	0	-1	0	0	1	0	0
$h, -h$	1	1	1	-1	-1	-1	1	1	1
$-h, 0$	1	-1	1	0	0	0	0	0	0
$0, 0$	1	0	0	0	0	0	0	0	0
$h, 0$	1	1	1	0	0	0	0	0	0
$-h, h$	1	-1	1	1	-1	1	1	-1	1
$0, h$	1	0	0	1	0	0	1	0	0
h, h	1	1	1	1	1	1	1	1	1

cuja inversa é

$$\begin{bmatrix} f & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ hf_x & 0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{h^2}{2}f_{xx} & 0 & 0 & 0 & \frac{1}{2} & -1 & \frac{1}{2} & 0 & 0 & 0 \\ hf_y & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ h^2f_{xy} & \frac{1}{4} & 0 & -\frac{1}{4} & 0 & 0 & 0 & -\frac{1}{4} & 0 & \frac{1}{4} \\ \frac{h^3}{2}f_{xxy} & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ \frac{h^2}{2}f_{yy} & 0 & \frac{1}{2} & 0 & 0 & -1 & 0 & 0 & \frac{1}{2} & 0 \\ \frac{h^3}{2}f_{xyy} & -\frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{4} & 0 & \frac{1}{4} \\ \frac{h^4}{4}f_{xyyy} & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} & -\frac{1}{2} & 1 & -\frac{1}{2} & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{bmatrix};$$

cada linha desta matriz (depois de uma normalização e colocação em formato 3×3) nos dá uma das máscaras $M_0, M_x, M_y, M_{xx}, M_{yy}, M_{xy}, M_{xyy}$,

M_{xxy} e M_{xyy} . Como f_{xxx} , f_{yyy} , f_{xxx} , f_{xxx} , f_{xxy} , f_{xyy} e f_{yyy} não aparecem diretamente na tabela acima, os seus coeficientes devem ser calculados a posteriori utilizando as máscaras encontradas para que determinemos corretamente todas as expressões em $o(h^5)$ (como feito na subseção anterior).

9.4 Exercícios

1) Para calcular invariantes geométricos (que venham de convoluções) perto das bordas de um sinal, a solução mais comum é simplesmente estender a imagem além de suas bordas de uma forma pré-determinada: com zeros, periodicamente ou por espelhamento. Uma outra idéia é simplesmente assumir que a imagem é desconhecida além de suas bordas e adaptar o **processamento** a esta condição. No sentido desta última idéia, dado um sinal $f[n]$ definido apenas para $n \geq 0$, encontre as melhores aproximações discretas da primeira e segunda derivadas de f em $n = 0$ usando máscaras de tamanho 3 (para clarificar: sua resposta dependerá de $f[0]$, $f[1]$ e $f[2]$ mas **não pode depender de $f[-1]$** ; você pode supor que f é a amostragem de um sinal contínuo onde um espaçamento de grid h foi utilizado). Qual a ordem das aproximações encontradas?

Capítulo 10

Abordagens Axiomáticas

O espaço de escala Gaussiano revelou-se útil por suas propriedades de linearidade, invariância por translações, semi-grupo e causalidade (cuja forma mais fraca é o princípio do máximo). Será que há outros espaços de escala com tais características? A resposta é, sucintamente, não: neste capítulo, mostraremos que os únicos espaços de escala satisfazendo propriedades de linearidade, invariância, semi-grupo e causalidade (e isotropia para o caso multidimensional contínuo, simetria para o caso multidimensional discreto) são os apresentados previamente neste livro.

Há duas abordagens diferentes para a análise deste assunto (que, basicamente, levam aos mesmos resultados). A primeira é partir de uma equação diferencial parcial genérica e então pedir que cada uma das propriedades sejam verificadas. No final, esperamos ficar com uma família de equações que nos dão possíveis espaços de escala. Esta é a abordagem da próxima seção.

Uma segunda abordagem (que é a linha principal de [28]) é supor que o espaço de escala será gerado por uma convolução do tipo $K_t * f_0$ para uma família de núcleos de convolução K_t que satisfazem uma condição de causalidade de cruzamentos de zero. Esta abordagem já assume linearidade e invariância por translações (via convolução), mas não força a condição de semi-grupo a priori. Infelizmente, por tratar de cruzamentos de zero, esta teoria só funciona bem para o caso unidimensional. A segunda seção deste capítulo descreve brevemente esta abordagem.

10.1 Via Equações Diferenciais Parciais

10.1.1 Caso contínuo

Caso Unidimensional

Esta subseção é completamente dedicada a mostrar e esclarecer a seguinte proposição de unicidade.

Proposição 10.1 *A equação do calor define o único (a menos de translações proporcionais à escala) espaço de escala não-trivial linear, invariante por translações que satisfaz a propriedade de semi-grupo e o princípio do máximo.*

Se supusermos que o espaço de escala F_t vem de uma equação diferencial, a propriedade de semi-grupo pede uma equação parabólica do tipo

$$\frac{\partial F_t}{\partial t} = \mathcal{F} \left\{ F_t, \frac{\partial F_t}{\partial x}, \frac{\partial^2 F_t}{\partial x^2}, \dots \right\}$$

$$F_0(x) = f_0(x)$$

Se nos restringirmos a equações lineares, somente teremos equações do tipo

$$\frac{\partial F_t}{\partial t} = C_0(x) F_t + C_1(x) \frac{\partial F_t}{\partial x} + C_2(x) \frac{\partial^2 F_t}{\partial x^2} + \dots$$

onde $C_0(x)$, $C_1(x)$, etc. são funções quaisquer.

Agora, se pedirmos também invariância por translações, somos forçados a concluir que C_0 , C_1 , etc. têm de ser invariantes por translações, e, portanto, constantes. De fato, queremos que tanto $F_t(x)$ como $G_t(x) = F_t(x-a)$ satisfaçam a equação acima, isto é

$$\sum_k C_k(x) \frac{\partial^k G(x)}{\partial x^k} = \frac{\partial G_t(x)}{\partial t} = \frac{\partial F_t(x-a)}{\partial t} = \sum_k C_k(x-a) \frac{\partial^k F_t(x-a)}{\partial x^k}$$

Finalmente, o princípio do máximo pede que $\frac{\partial F_t}{\partial t} \leq 0$ sempre que F_t tenha um mínimo local; em particular, devemos ter $\frac{\partial F}{\partial t} \leq 0$ sempre que $\frac{\partial F}{\partial x} = 0$ e $\frac{\partial^2 F}{\partial x^2} < 0$, não importando quais os valores das outras derivadas de $F(x)$ (que pode ser tomado como condição inicial f_0)!. Isto indica que a única equação diferencial satisfazendo linearidade, invariância por translações, semi-grupo e causalidade quaisquer que sejam as condições iniciais *tem de ser da forma*¹

$$\frac{\partial F_t}{\partial t} = \alpha \frac{\partial F_t}{\partial x} + \beta \frac{\partial^2 F_t}{\partial x^2}$$

$$F_0(x) = f_0(x)$$

com α e β constantes e $\beta \geq 0$. Há somente quatro casos a considerar:

- Se $\alpha = \beta = 0$, a equação claramente tem solução constante em t

$$F_t(x) = f_0(x)$$

que não tem graça.

¹Para o leitor mais cético, tome $f_0(x) = x^2 - \lambda x^n$ como condição inicial onde $n > 2$. Note que $x = 0$ é um mínimo de $f_0(x)$ e portanto deve subir; mas $\frac{\partial F_t}{\partial t} = 2C_2 - \lambda n(n-1)C_n$ em $t = 0$. A única maneira disto ser positivo para qualquer λ é tomando $C_n = 0$. O mesmo argumento funciona com $f_0(x) = \lambda + x^2$ para mostrar que $C_0 = 0$.

- Se $\alpha \neq 0$ mas $\beta = 0$, a solução é simplesmente uma translação do sinal original com a escala (veja exemplo 1.19)

$$F_t(x) = f_0(x + \alpha t) = (\delta_{-\alpha t} * f_0)(x)$$

que também não tem graça.

- Se $\alpha = 0$ e $\beta \neq 0$, temos a equação do calor com a solução dada por convoluções com Gaussianas, possivelmente acelerada pelo fator β , vistas no capítulo 5

$$F_t(x) = G_{\beta t}(x) * f_0(x)$$

- Finalmente, se $\alpha, \beta \neq 0$ a situação é simplesmente uma mistura das duas anteriores; a solução é uma translação (progressiva) do espaço de escala Gaussiano

$$F_t(x) = (\delta_{-\alpha t} * G_{\beta t} * f_0)(x) = G_{\beta t}(x + \alpha t) * f_0(x)$$

Caso bidimensional

O raciocínio da seção anterior se estende ao caso bidimensional e multidimensional; a única possível surpresa é que precisamos pedir uma condição de isotropia para chegar a um único espaço de escala. Esta subseção se dedica a mostrar e discutir:

Proposição 10.2 *A equação do calor define o único (a menos de translações proporcionais à escala) espaço de escala multidimensional não-trivial linear, invariante por translações e rotações que satisfaz a propriedade de semi-grupo e o princípio do máximo.*

Vamos nos restringir ao caso bidimensional (o caso multidimensional é análogo). As propriedades de **semi-grupo** e **linearidade** levam a

$$\frac{\partial F_t}{\partial t} = \sum_{j,k} C_{j,k} \frac{\partial^{j+k} F_t}{\partial x^j \partial y^k}$$

e a **invariância por translações** força os $C_{j,k}$ a serem constantes a serem determinadas.

Agora, o **princípio do máximo** vai exigir que $\frac{\partial F_t}{\partial t} \leq 0$ sempre que houver um máximo local; em particular, isto deve valer sempre que

$$\frac{\partial F_t}{\partial x} = \frac{\partial F_t}{\partial y} = 0, \quad \frac{\partial^2 F_t}{\partial x^2}, \frac{\partial^2 F_t}{\partial y^2} < 0 \quad \text{e} \quad \left| \begin{array}{cc} \frac{\partial^2 F_t}{\partial x^2} & \frac{\partial^2 F_t}{\partial x \partial y} \\ \frac{\partial^2 F_t}{\partial x \partial y} & \frac{\partial^2 F_t}{\partial y^2} \end{array} \right| > 0$$

independentemente dos valores das outras derivadas parciais de F_t . Assim, a equação tem de ser da forma

$$\frac{\partial F_t}{\partial t} = A \frac{\partial F_t}{\partial x} + B \frac{\partial F_t}{\partial y} + C^2 \frac{\partial^2 F_t}{\partial x^2} + 2D \frac{\partial^2 F_t}{\partial x \partial y} + E^2 \frac{\partial^2 F_t}{\partial y^2}$$

$$F_0(x, y) = f_0(x, y)$$

com $|D| \leq CE$. Uma rotação de coordenadas pode ser feita para desaparecer com a derivada mista; por abuso de notação escreveremos

$$\frac{\partial F_t}{\partial t} = A \frac{\partial F_t}{\partial x} + B \frac{\partial F_t}{\partial y} + C^2 \frac{\partial^2 F_t}{\partial x^2} + E^2 \frac{\partial^2 F_t}{\partial y^2}$$

$$F_0(x, y) = f_0(x, y)$$

A solução é uma Gaussiana anisotrópica (dependendo da razão entre C e E), composta com uma translação (dependendo de A e B)

$$F_t(x, y) = G_{CEt} \left(\sqrt{\frac{E}{C}}(x + At), \sqrt{\frac{C}{E}}(y + Bt) \right) * f_0(x, y)$$

Vale a pena notar que tal equação pode ser utilizada para o tratamento de imagens que tenham direções preferenciais conhecidas a priori. Em particular, há forte evidência que indica a presença de células em nosso sistema visual que realizam operações de convolução semelhantes à convolução com Gaussianas anisotrópicas como a mostrada acima!

Quando adicionamos a condição de que nosso espaço de escala tem de ser isotrópico, isto é, **invariante por rotações**, é que somente nos resta a opção de tomarmos $C = E$; então, voltamos à solução já conhecida da equação do calor (composta com uma translação, re-escalada):

$$F_t(x, y) = G_{C^2t}(x + At, y + Bt) * f_0(x, y)$$

10.1.2 Caso Discreto

Caso Unidimensional

Argumentos semelhantes se aplicam ao caso discreto. Agora, temos:

Proposição 10.3 *A equação do calor discreta define o único espaço de escala discreto no espaço, não-trivial linear, invariante por translações e simetrias que satisfaz a propriedade de semi-grupo e o princípio do máximo.*

Suponha que $F_t[n]$ é um espaço de escala discreto no espaço. O fato de que $F_t[n]$ satisfaz a condição de **semi-grupo** em t exige uma equação diferencial parabólica do tipo

$$\frac{\partial F_t}{\partial t}[n] = \mathcal{F}\{F_t, n\}$$

Se o espaço de escala é **linear** em $f_0[n]$, não nos resta opção exceto aceitar que $\mathcal{F}\{F_t, n\}$ é linear em F_t , isto é

$$\frac{\partial F_t}{\partial t}[n] = \sum_{j=-\infty}^{\infty} A[j, n] F_t[j]$$

onde os coeficientes $A[j, n]$ podem muito bem ser diferentes para cada n .

A **invariância por translações** diz que o lado direito é expresso por uma convolução. De fato, considere o espaço de escala $F[n]$ gerado por $f_0[n] = \delta[n]$ e o espaço de escala $G_t[n]$ gerado por $g_0[n] = f_0[n - a]$

$$\frac{\partial G_t}{\partial t}[n] = \sum_{j=-\infty}^{\infty} A[j, n] G_t[j] = \sum_{j=-\infty}^{\infty} A[j, n] F_t[j - a]$$

A invariância diz que $G_t[n] = F_t[n - a]$, e portanto,

$$\frac{\partial G_t}{\partial t}[n] = \frac{\partial F_t}{\partial t}[n - a] = \sum_{j=-\infty}^{\infty} A[j, n - a] F_t[j]$$

Tomando $t = 0$, onde $F_t = \delta$, ficamos apenas com

$$A[0, n - a] = A[a, n]$$

e a equação original torna-se

$$\frac{\partial F_t}{\partial t}[n] = \sum_{j=-\infty}^{\infty} A[0, n - j] F_t[j] = A * F_t$$

onde o núcleo de convolução A é definido por $A[n] = A[0, n]$.

Se A for "bem comportado", ele será o gerador infinitesimal de um semi-grupo T_t e a solução da equação do espaço de escala será simplesmente (veja seção 2.3):

$$F_t = T_t * f_0$$

$$T_t = \delta + tA + \frac{t^2}{2!} A * A + \dots$$

Agora adicione uma pitada de **causalidade** à nossa mistura; para uma idéia da teoria em volta da causalidade de cruzamentos de zero, veja a próxima seção. Aqui, analisaremos o princípio do máximo: queremos que $\frac{\partial F_t}{\partial t}[n] \leq 0$ sempre que $F_t[n]$ tiver um máximo local, isto é, sempre que $F_t[n] \geq F_t[n - 1], F_t[n + 1]$. Em particular, use as seguintes funções teste para tirar conclusões a respeito de A :

- Tome $f_0[n] = \pm(\delta[n] + \delta[n + 1] + \delta[n - 1])$, que tem um máximo local em $n = 0$; então

$$\pm(A * (\delta + \delta_1 + \delta_{-1}))[0] \leq 0 \Rightarrow A[0] + A[1] + A[-1] = 0$$

- Tome $f_0[n] = -\delta[n \pm 1]$, que tem um máximo local em $n = 0$; então

$$-(A * \delta_{\pm 1})[0] \leq 0 \Rightarrow A[\pm 1] \geq 0$$

- Para $|b| \geq 2$; tome $f_0[n] = \delta[n] + \lambda\delta[n-b]$, que tem um máximo local em $n = 0$; então

$$\forall \lambda, (A * (\delta - \lambda\delta_b)) [0] = A[0] - \lambda A[b] \leq 0 \Rightarrow A[b] = 0$$

Juntando as informações acima, concluímos que o gerador infinitesimal A é da forma

$$A = \begin{pmatrix} a & & & \\ & -a-b & & \\ & & b & \\ & & & \ddots \end{pmatrix}$$

onde $a, b \geq 0$. Não é difícil ver que isto é suficiente para satisfazer o princípio do máximo; afinal

$$(A * f_0)[n] = a(f_0[n+1] - f_0[n]) + b(f_0[n-1] - f_0[n]) \leq 0$$

sempre que $f_0[n]$ for um máximo local.

A equação resultante

$$\begin{aligned} \frac{\partial F_t}{\partial t}[n] &= aF_t[n+1] - (a+b)F_t[n] + bF_t[n-1] \\ F_0[n] &= f_0[n] \end{aligned} \quad (10.1)$$

pode ser resolvida pelos métodos da seção 2.3 (o melhor exemplo sendo o núcleo simétrico de Poisson do capítulo 6). Mas vamos colocar um último toque na nossa receita: se pedirmos uma simples condição natural de **invariância por reflexão**, isto é,

$$\begin{aligned} F_t[n] \text{ é o espaço de escala de } f_0[n] &\Rightarrow \\ \Rightarrow F_t[-n] \text{ é o espaço de escala de } f_0[-n] \end{aligned}$$

então é fácil ver que A deve ser simétrico, isto é, $a = b$, e a equação torna-se

$$\begin{aligned} \frac{\partial F_t}{\partial t} &= a(1 - 2\mathbf{1}) * F_t \\ F_0[n] &= f_0[n] \end{aligned}$$

cuja solução é exatamente a encontrada no capítulo 6 (exceto pelo possível re-escalamento dado por a)

$$\begin{aligned} F_t &= P_{at} * f_0 \\ P_t[n] &= \text{núcleo de Poisson simétrico} \end{aligned}$$

Caso Bidimensional

A aplicação dos argumentos da seção anterior ao caso multidimensional discreto leva a:

Proposição 10.4 *A equação do calor discreta define o único espaço de escala discreto no espaço multidimensional, simétrico, não-trivial, linear, invariante por translações que satisfaz a propriedade de semi-grupo e o princípio do máximo.*

Exibimos o caso bidimensional. As propriedades de semi-grupo e linearidade levam a

$$\frac{\partial F_t}{\partial t} [m, n] = \sum_{j, k=-\infty}^{\infty} A [j, k, m, n] F_t [j, k]$$

e a invariância por translações faz com que

$$A [j, k, m, n] = A [0, 0, m - j, n - k]$$

Portanto, definindo um núcleo bidimensional por $A [m, n] = A [0, 0, m, n]$ chegamos à forma da equação diferencial com uma convolução

$$\begin{aligned} \frac{\partial F_t}{\partial t} [m, n] &= (A * F_t) [m, n] \\ F_0 [m, n] &= f_0 [m, n] \end{aligned}$$

Cuja solução é novamente

$$\begin{aligned} F_t [m, n] &= T_t * f_0 \\ T_t &= \delta + tA + \frac{t^2}{2!} A * A + \dots \end{aligned}$$

Sabemos que é difícil falar em causalidade por cruzamentos de zero neste caso, e temos poucas esperanças com causalidade de número de máximos e mínimos. Assim, nosso toque de causalidade fica por conta do **princípio do máximo**, que vai exigir que $\frac{\partial F_t}{\partial t} \leq 0$ sempre que houver um máximo local em F_t . Usando uma vizinhança de 8 pontos para a definição de máximo local e usando funções teste semelhantes às da subseção anterior, chegamos a

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{pmatrix} \quad (10.2)$$

$$\sum_{j=1}^9 a_j = 0$$

$$a_1, a_2, a_3, a_4, a_6, a_7, a_8, a_9 \geq 0$$

(uma vizinhança de 4 pontos torna $a_1 = a_3 = a_7 = a_9 = 0$). Isto gera uma família de espaços de escala assimétricos (alguns dos quais podem muito bem ser aplicados).

Se desejarmos, condições de **invariância por simetria** impostas ao espaço de escala se traduzem em condições de simetria em A . Por exemplo, se exigirmos invariância por simetria em relação à origem

$$F_t[m, n] \text{ é o espaço de escala de } f_0[m, n] \Rightarrow \\ \Rightarrow F_t[-m, -n] \text{ é o espaço de escala de } f_0[-m, -n]$$

então A é simétrica com relação a seu centro, isto é

$$a_1 = a_9; a_2 = a_8; a_3 = a_7; a_4 = a_6$$

Se também pedirmos simetria com relação ao eixo horizontal (e vertical), então só nos resta

$$A = \begin{pmatrix} a & b & a \\ b & -4a - 4b & b \\ a & b & a \end{pmatrix} = (2a + b) \begin{pmatrix} c & 1 - 2c & c \\ 1 - 2c & -4 + 4c & 1 - 2c \\ c & 1 - 2c & c \end{pmatrix} \\ c = \frac{a}{2a + b} \Rightarrow 0 \leq c \leq \frac{1}{2}$$

que é um múltiplo do Laplaciano mais geral L_c já apresentado nos capítulos 6 e 9 (para vizinhanças de 4 pontos, $c = 0$ é a única opção).

10.2 Via Convoluções

Aqui, abandonamos a propriedade de semi-grupo, retendo apenas as condições de linearidade e invariância por translações (leia-se, convoluções), mais a causalidade de cruzamentos de zero (que só funciona bem no caso unidimensional); esta é a abordagem em [28]; vamos descrevê-la brevemente nesta subseção.

10.2.1 Caso contínuo

Definição 10.5 $K : \mathbb{R} \rightarrow \mathbb{R}$ é um núcleo de espaço de escala se para qualquer sinal f_0

$$\#_{CZ}(K * f_0) \leq \#_{CZ}(f_0)$$

Afirmção 10.6 K é um núcleo de espaço de escala se e somente se sua Transformada Bilateral de Laplace for da forma

$$\mathcal{L}_K(s) = Ce^{-\mu s + t s^2} \prod_{i=1}^{\infty} \frac{\lambda_i}{\lambda_i + s}$$

para uma região do tipo $|\operatorname{Re}(s)| \leq k$, ou, equivalentemente, se e somente se

$$\begin{vmatrix} K(x_1 - y_1) & K(x_2 - y_1) & \dots & K(x_n - y_1) \\ K(x_1 - y_2) & K(x_2 - y_2) & \dots & K(x_n - y_2) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_1 - y_n) & K(x_2 - y_n) & \dots & K(x_n - y_n) \end{vmatrix} \geq 0$$

para quaisquer $x_1 < x_2 < \dots < x_n$ e $y_1 < y_2 < \dots < y_n$ em \mathbb{R} .

Demonstração. Veja [42], [18] e [20]. ■

Como a convolução de núcleos corresponde ao produto de suas transformadas de Laplace, a primeira parte da afirmação acima indica que **todo núcleo de espaço de escala é uma convolução de núcleos básicos que pertencem às seguintes categorias:**

- Amplificações

$$\mathcal{L}_K(s) = C \iff K_C = C\delta$$

- Translações

$$\mathcal{L}_K(s) = e^{-\mu s} \iff K_\mu = \delta_\mu$$

- Gaussianas

$$\mathcal{L}_K(s) = e^{ts^2} \iff K_t = G_t$$

- Exponenciais truncadas:

$$\mathcal{L}_K(s) = \frac{\lambda}{\lambda + s} \iff K_\lambda = \begin{cases} |\lambda| e^{-|\lambda|x} & \text{se } \text{sign}(\lambda)x \geq 0 \\ 0 & \text{se } \text{sign}(\lambda)x < 0 \end{cases}$$

O único núcleo que pode parecer um pouco surpreendente é o último. No entanto, note que sua “versão simétrica” é o núcleo exponencial simétrico apresentado na seção 7.3.2, já que

$$\begin{aligned} \mathcal{L}_{K_\lambda}(s) \mathcal{L}_{K_\lambda}(-s) &= \frac{\lambda}{\lambda + s} \frac{\lambda}{\lambda - s} = \frac{1}{2} \left(\frac{\lambda}{\lambda + s} - \frac{\lambda}{s - \lambda} \right) = \\ &= \frac{1}{2} (K_\lambda + K_{-\lambda}) = \mathcal{L}_{E_t} \text{ para } \lambda = \frac{1}{\sqrt{t}}. \end{aligned}$$

A partir daí, a idéia é criar famílias de núcleos K_t usando os núcleos acima, e gerar um espaço de escala como $F_t = K_t * f_0$; note que não é qualquer família que gera um bom espaço de escala – afinal, se escolhermos K_t aleatoriamente, somente sabemos que $\#_{CZ}(F_t) \leq \#_{CZ}(f_0)$, mas nada se sabe sobre $\#_{CZ}(F_{t_1})$ comparado a $\#_{CZ}(F_{t_2})$ para $t_1 \geq t_2$ – a causalidade CZ em geral só é satisfeita se os núcleos escolhidos formarem um semi-grupo! Assim, a única boa opção para t contínuo é a família de Gaussianas; para passos de tempo discretos, escolha qualquer núcleo de espaço de escala e convolua-o repetidas vezes com o sinal original.

10.2.2 Caso Discreto

Definição 10.7 Um núcleo $K : \mathbb{Z} \rightarrow \mathbb{R}$ é um núcleo de espaço de escala se para qualquer seqüência f_0

$$\#_{CZ}(K * f_0) \leq \#_{CZ}(f_0)$$

Afirmção 10.8 K é um núcleo de espaço de escala se e somente se sua Transformada Z tiver a forma

$$\varphi_K(z) = Cz^k e^{\alpha z + \beta z^{-1}} \prod_i \frac{(1 + \alpha_i z)}{(1 - \beta_i z)(1 - \gamma_i z^{-1})}$$

com $C > 0$, $k \in \mathbb{Z}$, $\alpha_i, \beta_i, \gamma_i \geq 0$, $\beta_i, \gamma_i < 1$ e $\sum \alpha_i + \beta_i + \gamma_i < \infty$. Equivalentemente, K é um núcleo de espaço de escala se e somente se todos os determinantes menores de

$$\begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & K([0]) & K[-1] & K[-2] & \dots \\ \dots & K([1]) & K[0] & K[-1] & \dots \\ \dots & K([2]) & K[1] & K[0] & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

forem não negativos ($K[n]$ é também chamada uma seqüência de freqüência de Pólya).

Demonstração. Veja [41], [43] e [20]. ■

Novamente, a primeira parte da afirmação acima indica que todo núcleo discreto de espaço de escala é uma convolução de núcleos básicos que pertencem às seguintes categorias:

- Amplificações

$$\varphi_K(z) = C \iff K_C = C\delta$$

- Translações

$$\varphi_K(z) = z^k \iff K_k = \delta_k$$

- Poisson (e reflexões)

$$\varphi_K(z) = e^{\alpha z} \iff K_\alpha = \left(0 \ 0 \ 0 \ 1 \ \alpha \ \frac{\alpha^2}{2} \ \frac{\alpha^3}{3!} \ \dots \right)$$

$$\varphi_K(z) = e^{\beta z^{-1}} \iff K_\beta = \left(\dots \ \frac{\beta^3}{3!} \ \frac{\beta^2}{2} \ \beta \ 1 \ 0 \ 0 \ 0 \dots \right)$$

- Binomiais

$$\varphi_K(z) = 1 + \alpha_i z \iff K_{\alpha_i} = (1 \ \alpha_i)$$

- Geométricos ou binomiais recursivos

$$\varphi_K(z) = \frac{1}{1 - \beta_i z} \iff K_{\beta_i} = \{(1 \ \beta_i \ \beta_i^2 \ \beta_i^3 \ \dots)\} \text{ se } 0 < \beta_i < 1$$

$$\varphi_K(z) = \frac{1}{1 - \gamma_i z^{-1}} \iff K_{\gamma_i} = \{(\gamma_i^3 \ \gamma_i^2 \ \gamma_i^1 \ 1)\} \text{ se } 0 < \gamma_i < 1$$

Todos estes núcleos já foram vislumbrados em capítulos anteriores, especialmente suas versões simétricas que foram encontradas nos capítulos 6 e 7.

A partir daí, escolhe-se uma família de núcleos K_t usando os núcleos acima para gerar um espaço de escala $F_t = K_t * f_0$; para que t seja contínuo, precisamos da propriedade de semi-grupo e as distribuições de Poisson são as únicas aceitáveis. Caso contrário, qualquer núcleo dos citados acima pode ser utilizado como base de um espaço de escala.

Vale a pena notar que a amostragem do núcleo Gaussiano dada por

$$K_t[n] = \frac{1}{\sqrt{4\pi t}} e^{-\frac{n^2}{4t}}$$

é um núcleo de espaço de escala como definido acima (baseado em uma “multitude” de núcleos binomiais – calcule a sua Transformada Z ou veja exercícios). No entanto, a família K_t não tem a propriedade de semi-grupo; portanto, dados $t_1 \geq t_2 > 0$, existe um sinal $f_0[n]$ para o qual

$$\#_{CZ}(K_{t_1} * f_0) \geq \#_{CZ}(K_{t_2} * f_0)$$

apesar de valer que, para qualquer sinal $f_0[n]$:

$$\#_{CZ}(K_{t_1} * f_0) \leq \#_{CZ}(f_0)$$

10.3 Exercícios

[Nota: a maioria dos exercícios deste capítulo são inspirados por exemplos de [28]].

(1) Escreva a solução da equação 10.1 como a convolução do sinal original f_0 com dois núcleos de Poisson (unilaterais) distintos.

(2) Que funções teste você usaria para provar que um gerador infinitesimal de um espaço de escala bidimensional discreto tem a forma mostrada em 10.2?

(3) Utilizando somente as definições 10.5 e 10.7, mostre que núcleos K e K_1 de espaço de escala têm as seguintes propriedades, tanto no caso contínuo como no caso discreto:

- K é positivo (ou K é negativo, mas K não troca de sinal);
- $K * f_0$ não tem mais pontos extremos que f_0 (assuma f_0 diferenciável no caso contínuo);

- c) K é unimodal.
 d) $K * K_1$ é também um núcleo de espaço de escala.

(4) Mostre que o seguinte núcleo contínuo $K(x)$ (definido através de sua Transformada de Fourier) não é de espaço de escala

$$\hat{K}(w) = \begin{cases} 1, & |w| \leq a \\ 0, & |w| > a \end{cases}$$

(5) Mostre que o seguinte núcleo (contínuo) não é de espaço de escala

$$K(x) = \begin{cases} 1, & |x| \leq a \\ 0, & |x| > a \end{cases}$$

(6) Seja K o núcleo discreto com três elementos (1 1 1). Encontre um sinal discreto $f_0[n]$ tal que

$$\#_{CZ}(K * f_0) \geq \#_{CZ}(f_0)$$

mostrando assim que K não é um núcleo de espaço de escala (apesar de ser positivo e unimodal).

(7) Explique (intuitivamente) porque você espera que um núcleo de espaço de escala K deva ser unimodal tanto no domínio do tempo como no domínio da frequência.

(8) Mostre que

$$K = [8 \ 36 \ 81 \ 36 \ 8]$$

é unimodal tanto no domínio do tempo como no domínio da frequência (para $w \in (-\frac{1}{2}, \frac{1}{2})$). No entanto, mostre que K não é um núcleo de espaço de escala. [Sugestão: tente sinais da forma $f_0 = [a \ -b \ c \ -d]$].

(9) Seja $K_t[n] = \frac{1}{\sqrt{4\pi t}} e^{-\frac{n^2}{4t}}$ (uma amostragem da Gaussiana). Mostre que

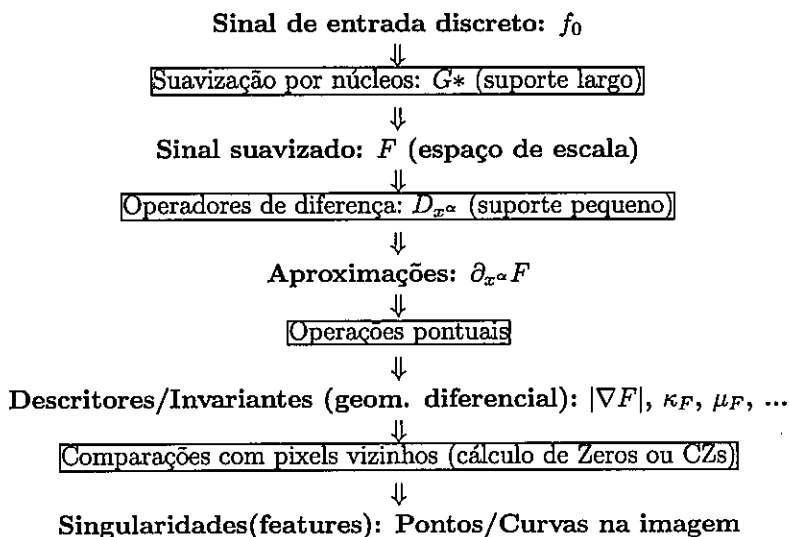
$$\varphi_K(z) = C_t \prod_{n=0}^{\infty} (1 + q_t^{2n+1} z) (1 + q_t^{2n+1} z^{-1})$$

para constantes C_t e q_t (dependendo de t). Use a caracterização em 10.2.2 para concluir que $K_t[n]$ é um núcleo de espaço de escala.

Capítulo 11

Conclusão da Parte II

A esta altura, podemos gerar um esboço do que seria um sistema de processamento de imagens e detecção de detalhes (características, features) baseado em espaço multi-escala através do esquema abaixo:



Normalmente, o primeiro passo desse processo é feito através de convoluções com Gaussianas (possivelmente discretas), mas há outras opções. Assim, é possível desistir das condições de linearidade (desistindo por completo da idéia da convolução) e adotar algum outro método de “suavização”. Esta tem sido uma área extremamente ativa em Visão Computacional recentemente – na próxima parte, apresentaremos algumas EDPs não-lineares que perfazem esta “suavização”.

Na parte seguinte do livro, discutiremos um possível descritor a ser usado com os espaços de escala que não segue exatamente o processo acima – as bolhas multi-escala, que geram uma espécie de esboço primitivo que acompanha uma imagem monocromática. Voltaremos à suavização linear pela Gaussiana, e provaremos alguns resultados interessantes sobre a formação e destruição de bolhas usando conhecimentos básicos de Teoria das Singularidades.

Parte III

EDPs Não-Lineares

Capítulo 12

Suavização Fora das Arestas

Um dos principais defeitos do Espaço de Escala Gaussiano é que ele pode modificar violentamente a estrutura das arestas de uma imagem, como por exemplo na figura 8.3.1. É claro que parte deste comportamento é desejável, já que as arestas de estruturas de pequena escala devem desaparecer na suavização, mas às vezes até mesmo arestas importantes são deslocadas durante a suavização, como comprovado pelo comentário 15.8.

Mas esta característica ocorre porque o Espaço de Escala Gaussiano “difunde” igualmente nas arestas ou fora delas; o modelo físico da equação do calor linear surge quando assumimos uma condutância de calor constante ao longo de toda a imagem (veja a seção 4.2). E se modificássemos essa condutância para que houvesse mais condução de calor (mais difusão) longe das arestas, e menos condutância perto delas? A equação pode ficar mais complicada, mas há esperança de que as arestas permaneçam nítidas por mais tempo à medida que a escala aumenta, talvez até mesmo eliminando a necessidade de um “back-tracking” rumo à escalas menores para a localização espacial da aresta. Esta é a idéia do modelo de Perona-Malik [36] e suas variantes, que apresentaremos a seguir.

No final deste capítulo, apresentaremos brevemente uma outra idéia para detectar arestas, que vem de um modelo probabilístico e acaba aparecendo como um problema de minimização para um funcional de energia. Ao fim, resolver o problema de minimização acaba se transformando em resolver uma equação diferencial, cujos algoritmos de resolução são muito semelhantes à criação de novos espaços de escala.

12.1 O modelo de Perona-Malik

Na seção 4.2, vimos que a Lei de Difusão de Fick leva a uma equação diferencial do tipo

$$\frac{\partial F(x, t)}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial F(x, t)}{\partial x} \right)$$

onde K era uma constante positiva que estabelecia quão rapidamente duas substâncias se misturavam. No caso da difusão do calor ao longo de uma barra não-homogênea, escrevemos

$$\frac{\partial F}{\partial t} = \frac{\partial}{\partial x} \left(C(x) \frac{\partial F}{\partial x} \right)$$

onde $F(x, t)$ é a temperatura da barra no ponto x no tempo t e $C(x)$ é a condutância térmica da barra no ponto x . A equação torna-se linear apenas se a condutância for constante (tomamo-la como 1 para obter a equação do calor clássica). Em geral, o modelo de difusão torna-se uma equação não-linear

$$\frac{\partial F}{\partial t} = \frac{\partial}{\partial x} \left(C(x) \frac{\partial F}{\partial x} \right) = \frac{\partial C}{\partial x} \frac{\partial F}{\partial x} + C \frac{\partial^2 F}{\partial x^2}$$

$$F(x, 0) = I(x)$$

onde $I(x)$ é novamente a distribuição inicial de calor na barra. Note que $C(x)$ é alto nas partes da barra onde o calor se espalha rapidamente, e pequeno nas áreas mais “isolantes” com respeito ao calor. Se $C(x) = 0$ numa parte da barra então $\frac{\partial F}{\partial t} = 0 \Rightarrow F(x, t) = I(x)$ e a distribuição de calor, por mais desigual que seja, não muda com o tempo t .

A versão bidimensional é semelhante

$$\frac{\partial F}{\partial t} = \nabla \cdot (CF) = \nabla C \cdot \nabla F + C \Delta F \quad (12.1)$$

$$F(x, y; 0) = I(x, y)$$

(o termo à direita mostra como a equação acima é rotacionalmente simétrica).

Aqui está o centro da idéia de Perona e Malik [36]: ao invés de tomarmos a “condutância térmica” como constante e chegarmos à difusão linear, por que não colocarmos uma condutância menor onde esperamos encontrar arestas, isto é, onde o gradiente de F é grande? Assim, a imagem F se suavizaria apenas nos pontos onde a suavização é necessária, mantendo as arestas nítidas para uma análise via um detector de arestas.

Idealmente, gostaríamos de fazer com que $C(x, y) = 0$ se há uma aresta passando pelo ponto (x, y) e $C(x, y) = 1$ caso contrário. Infelizmente, não há como fazer isto – se soubéssemos exatamente onde as arestas estão, não haveria necessidade de criar este processo todo para manter as arestas no lugar!

Ao invés disso, colocamos C com valores menores onde há mais chance de encontrar uma aresta, isto é, nos pontos onde $|\nabla F|$ é maior. Por exemplo, podemos tentar

$$C_1(x, y) = \frac{1}{1 + \frac{|\nabla F|^2}{k^2}}$$

ou

$$C_2(x, y) = e^{-\frac{|\nabla F|^2}{2k^2}}$$

(onde k é um parâmetro a ser estimado dependendo do espaço de escala desejado) ou outras funções $C(|\nabla F|)$ onde C é uma função decrescente de $|\nabla F|$.

Vamos proceder a uma rápida análise da equação diferencial 12.1 para ambas as condutâncias C_1 e C_2 apresentadas acima. Assim, suponha que a condutância C é uma função do gradiente de F , isto é, $C = C(|\nabla F|)$:

$$C_1(u) = \frac{1}{1 + \frac{u^2}{k^2}}$$

$$C_2(u) = e^{-\frac{u^2}{2k^2}}$$

Para simplificar a análise, vamos nos limitar ao caso unidimensional, que será então escrito como

$$\frac{\partial F}{\partial t} = \frac{\partial}{\partial x} (C(F_x) \cdot F_x) = C' F_x F_{xx} + C F_{xx} = (C' F_x + C) F_{xx} \quad (12.2)$$

(a derivada à direita é com relação ao argumento da função $C(u)$).

Em primeiro lugar, note que 12.2 é causal no sentido do princípio do máximo. De fato, num ponto crítico ($F_x = 0$) ela se resume a $F_t = C F_{xx}$ e portanto máximos locais não sobem e mínimos locais não descem (desde que a condutância satisfaça $C(0) > 0$, que é o caso tanto para C_1 como C_2). Assim, os valores máximos e mínimos da imagem $F(\cdot, t)$ serão limitados pelos valores máximo e mínimo de $F(\cdot, 0)$ – o que é um sinal de estabilidade desejável.

Por outro lado, as derivadas de F podem explodir à medida que a equação avança. Podemos verificar isto escrevendo $u = F_x$ e diferenciando 12.2 com relação a x para encontrar

$$\frac{\partial u}{\partial t} = (C'(u) \cdot u + C(u)) u_{xx} + (C''(u) u + 2C'(u)) u_x^2$$

Note que nos pontos críticos de u temos

$$\frac{\partial u}{\partial t} = (C' u + C) u_{xx}$$

e o princípio do máximo não valerá se houver pontos críticos de u tais que

$$C' F_x + C < 0$$

(aliás, diretamente de 12.2, vê-se que em pontos onde isto acontece, aquela EDP se assemelha a uma Equação do Calor Inversa). Montando a função $A(u) = uC(u)$, tem-se

$$A'(u) = \frac{d}{du} (uC(u)) = C' u + C$$

e portanto esta instabilidade sempre ocorrerá para valores de $u = F_x$ que estejam na parte decrescente da função A . Para as condutâncias sugeridas acima, vemos

que

$$A_1(u) = uC_1(u) = \frac{u}{1 + \frac{u^2}{k^2}} \Rightarrow A'_1(u) = \frac{\left(1 + \frac{u^2}{k^2}\right) - \frac{2u}{k^2} \cdot u}{\left(1 + \frac{u^2}{k^2}\right)^2} = \frac{1 - \frac{u^2}{k^2}}{\left(1 + \frac{u^2}{k^2}\right)^2}$$

$$A_2(u) = uC_2(u) = ue^{-\frac{u^2}{2k^2}} \Rightarrow A'_2(u) = \left(1 - \frac{u^2}{k^2}\right) e^{-\frac{u^2}{2k^2}}$$

e ambas as equações diferenciais resultantes serão instáveis em $u = F_x$ próximo a pontos onde $|F_x| > k$ - e tais pontos de alto gradiente estão próximos exatamente das arestas que queríamos manter intactas!

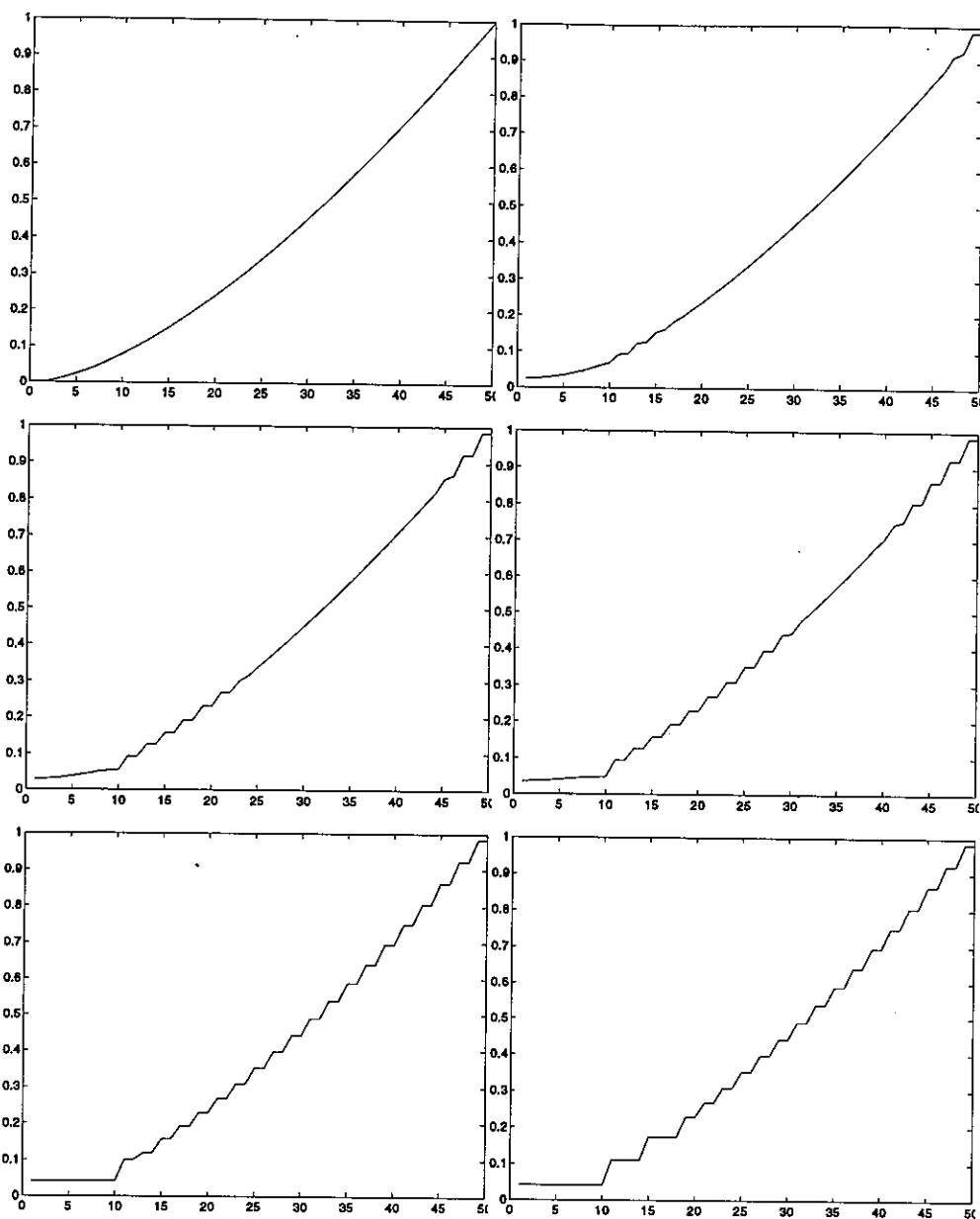
Em suma, enquanto a função F satisfaz o princípio do máximo, não há nenhum controle nos máximos e mínimos das derivadas espaciais de F . O efeito desta instabilidade na aproximação numérica é especialmente notado quando a imagem inicial I é uma faixa cujo gradiente de cinza é comparável a k . Por exemplo, se F gradativamente se estende de 0 (preto) a 1 (branco), a equação diferencial 12.1 gerará um “feito escada” em F , mantendo-a limitada entre 0 e 1 (F_x é que “estourará” gerando os vários degraus de F). Uma aresta gradual se divide em várias arestas menores, produzindo um efeito indesejável. Veja as figuras a seguir para um exemplo deste efeito.

Este efeito escada pode ser evitado através de uma regularização do cálculo do gradiente, através de uma suavização Gaussiana:

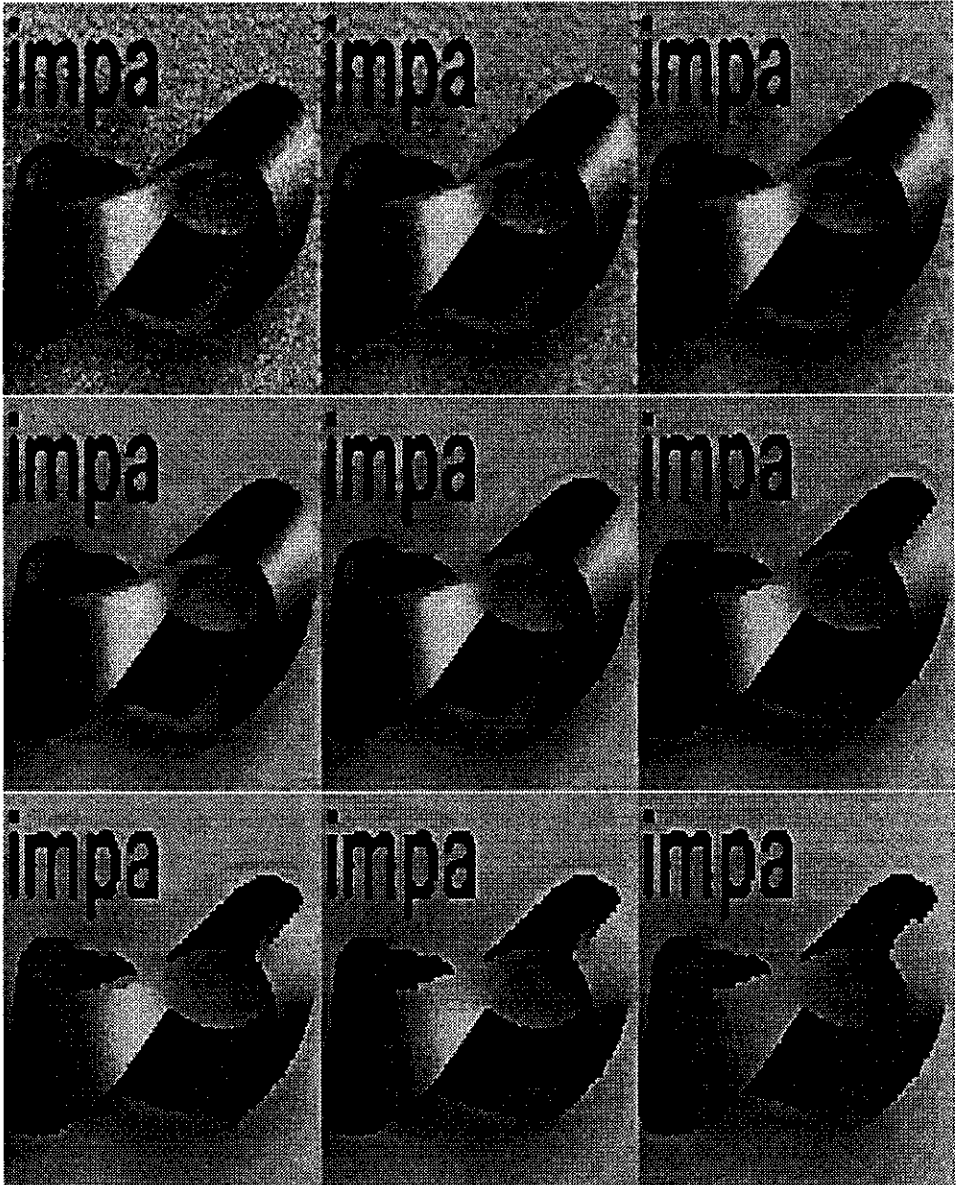
$$C(x, y) = \frac{1}{1 + \frac{1}{k^2} |\nabla (G_{\sigma^2} * F)|^2}$$

A discretização da equação diferencial é razoavelmente imediata (apenas precisamos dos operadores $\frac{\partial}{\partial x}$ e $\frac{\partial}{\partial y}$, e então precisamos decidir como discretizar $\frac{\partial}{\partial t}$) e oferece resultados interessantes. A presença da Gaussiana no denominador de $C(x, y)$ não deve assustar - afinal de contas, qualquer discretização de $|\nabla F|$ já envolveria uma suavização de qualquer forma.

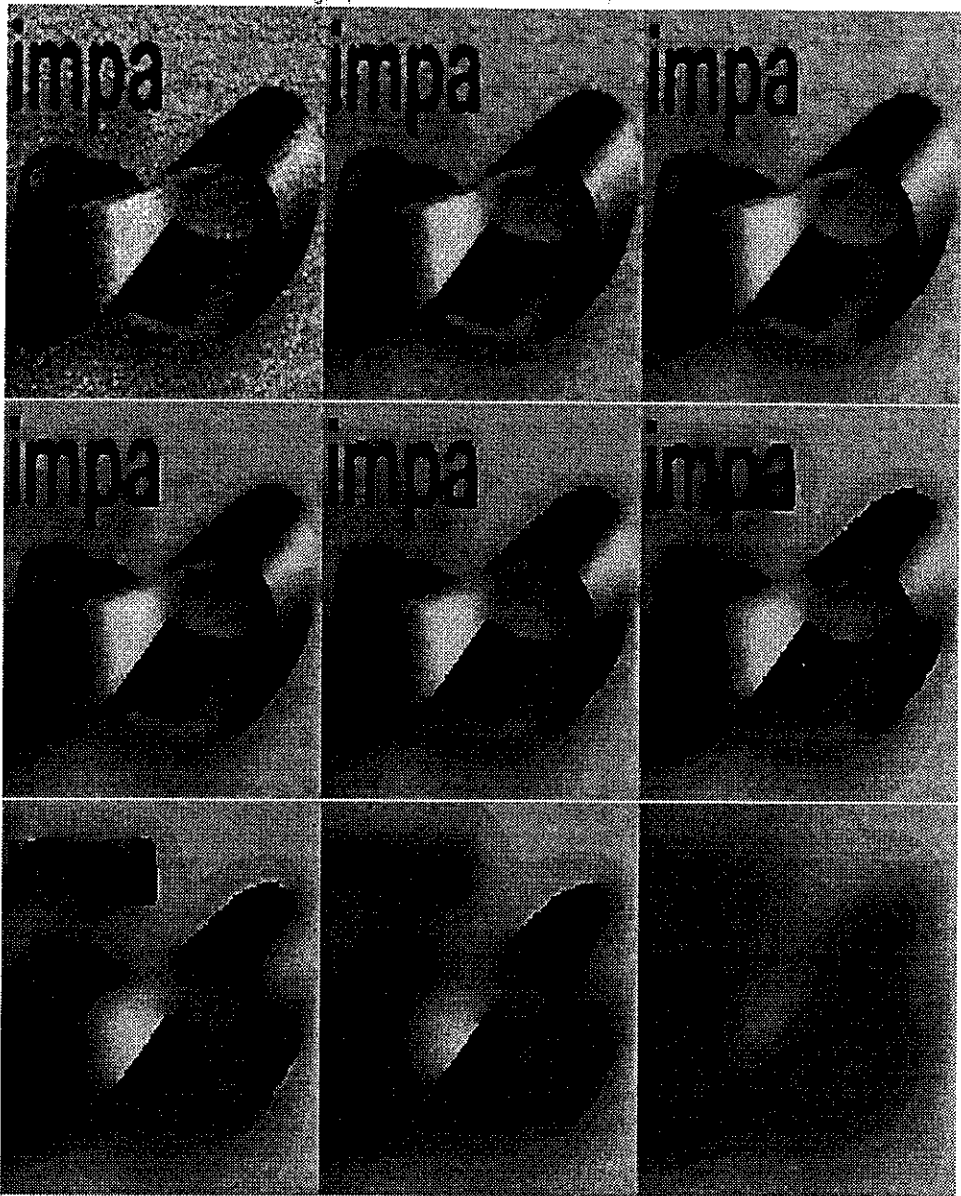
Há outras alternativas para a escolha de $C(x, y)$; Nitzberg e Shiota [33], por exemplo, propuseram uma equação diferencial que não só preserva arestas mas também procura enfatizá-las espalhando níveis de cinza até os lados das arestas em potencial. Infelizmente, uma análise mais profunda destes métodos escapa ao escopo deste livro.



O efeito escada: note como a versão 1D da EDP de Perona-Malik quebra uma aresta suave em várias arestas menores **sem ofender o princípio do máximo**; este efeito é devido à instabilidade da EDP em relação a F_x .



O logotipo do IMPA passa pela EDP de Perona-Malik; usamos a condutância C_2 sem suavização extra e $k = 10$; as imagens correspondem a 0, 5, 10, 20, 40, 100, 200, 400 e 1000 iterações. Note que a imagem atinge um estado praticamente estacionário, e que as arestas que na imagem original eram suavizadas (anti-aliased) no final aparecem pixeladas.



EDP de Perona-Malik com um valor maior para k (usamos $k = 20$ – isto é, apenas arestas onde o módulo do gradiente vale mais que 20 são respeitadas). Note os “vazamentos Gaussianos” perto das arestas mais fracas (que já podia ser observado para $k = 10$ em alguns locais). O efeito dos vazamentos acaba por borrar completamente a imagem para um grande número de iterações.

12.2 Detecção de Arestas via Funcionais de Energia

Há uma gama de métodos de detecção de arestas que se baseiam em minimização por funcionais de energia. Apresentamos aqui apenas uma brevíssima introdução a tais métodos (veja o artigo de David Mumford em [9] para outras referências). Inicialmente, considere o seguinte exemplo de suavização de sinais.

Exemplo 12.1 Dada uma função $f : [a, b] \rightarrow \mathbb{R}$ considere os funcionais

$$E_1 \{g\} = \int_a^b (g(x) - f(x))^2 dx$$

$$E_2 \{g\} = \int_a^b (g'(x))^2 dx$$

para uma função diferenciável qualquer $g : [a, b] \rightarrow \mathbb{R}$. Note que E_1 mede quanto g se distancia de f e E_2 é uma medida de quanto g é “bem comportada”. Para obter uma aproximação suave de f , podemos considerar a função g que minimiza

$$E \{g\} = (E_1 + \alpha E_2) \{g\}$$

onde o parâmetro α controla o que é mais importante: obter uma função suave ou ficar perto de f . Assim, se $\alpha = 0$, uma solução possível é $g = f$ (se f for diferenciável); por outro lado, à medida que $\alpha \rightarrow \infty$, $g(x) = K$ (constante) torna-se uma das melhores soluções. Para encontrar g matematicamente, precisamos utilizar o cálculo das variações e encontrar uma equação diferencial para g , neste caso

$$g - \alpha g'' = f$$

Uma idéia semelhante pode ser usada para segmentação de sinais unidimensionais – o chamado “Weak String Model”. A idéia é usar a suavização do exemplo acima, mas permitir a introdução de alguns pontos onde g pode se “quebrar” – mas não podem ser pontos demais, senão haverá uma tendência para quebrar g em todos os lugares!

Definição 12.2 Dada uma função $f : [a, b] \rightarrow \mathbb{R}$, o funcional de energia do weak string model é dado por

$$E \{g, S\} = \mu^2 \int_a^b (g(t) - f(t))^2 dt + \int_{[a,b]-S} (g'(t))^2 dt + \nu \cdot \#(S)$$

onde S é um conjunto de saltos $\{t_1, t_2, \dots, t_n\}$ tal que $a = t_0 < t_1 < t_2 < \dots < t_n < t_{n+1} = b$; g é uma função diferenciável em cada intervalo da forma $[t_i, t_{i+1}]$ (mas possivelmente descontínua em cada t_k); $\#(S) = n$ é simplesmente o número de elementos de S . Note que podemos ajustar a fidelidade ao sinal f (através do ajuste do parâmetro μ) e a quantidade de arestas desejados (através do ajuste de ν).

Queremos então encontrar g e S que minimizem $E\{g, S\}$; S será considerado o conjunto de arestas do sinal f original enquanto g será uma reconstrução suave-por-partes de f que atende a tal segmentação.

O caso bidimensional é semelhante, mas o funcional de energia (comumente denominado “funcional de Mumford-Shah”) se parece com

$$E\{g, S\} = \mu^2 \int_D (g - f)^2 dA + \int_{D-S} |\nabla g|^2 dA + \nu.e(S)$$

onde S é em geral um conjunto de curvas e $e(S)$ pode ser tomado como a soma de seus comprimentos (há outras opções para $e(S)$).

O caso discreto unidimensional é mais simples. O funcional se torna

$$E\{g, S\} = \mu^2 \sum_j (g_j - f_j)^2 + \sum_{j \notin S} (g_{j+1} - g_j)^2 + \nu.\#(S)$$

onde S é um conjunto de índices, cada um indicando a presença de um bordo entre f_j e f_{j+1} . Este funcional é normalmente re-escrito como

$$E\{g, l\} = \mu^2 \sum (g_j - f_j)^2 + \sum (g_{j+1} - g_j)^2 (1 - l_j) + \nu \sum l_j$$

onde l_j é a função característica de S , isto é, $l_j = 1$ se $j \in S$ e $l_j = 0$ quando $j \notin S$. O problema é então encontrar a seqüência g_j e a seqüência binária l_j que minimizem o funcional $E\{g, l\}$. Surpreendentemente, existem algoritmos de programação dinâmica que encontram g e l rapidamente neste caso discreto unidimensional (dada a seqüência inicial f).

Os casos contínuos e o caso discreto multidimensional são bem mais difíceis. Uma idéia é começar de $\{g, S\}$ quaisquer e mudá-los lentamente (localmente) para que a energia $E\{g, S\}$ caia. Tais métodos encontram mínimos locais do funcional E , mas não há garantia que os mínimos encontrados sejam próximos ao mínimo global. Resultados melhores são obtidos se mudanças bruscas (globais) são adicionadas esporadicamente, cada vez menos freqüentemente – a idéia básica do “simulated annealing”. Uma discussão mais profunda deste tópico foge ao escopo de nosso curso.

É interessante notar que métodos de interpretação de imagens com “funcionais de energia” têm uma interpretação probabilística forte. A idéia geral é que a “energia” de uma disposição de arestas refletiria a “probabilidade” daquela interpretação, algo como

$$\text{Pr}(\text{Arestas corretamente localizadas}) \propto e^{-K.\text{Energia}}$$

Ao leitor interessado, sugerimos novamente a consulta a [9] e referências ali contidas.

Capítulo 13

Suavização Morfológica

Neste capítulo abordaremos equações diferenciais que sejam *morfológicas*¹ (também chamados de *geométricas*), isto é, invariantes por mudança de brilho. Na seção 8.2 discutimos brevemente uma possível motivação para criar descritores que possuam tal invariância: mudanças de brilho podem ser causadas por mudanças na intensidade de iluminação de uma cena, durante o processo de captação de uma imagem ou mesmo durante sua reprodução em uma tela ou impressão; dentro de certos limites, estas mudanças não deveriam modificar muito a interpretação da imagem obtida. Naquela seção, discutimos como encontrar descritores que possuíssem tal invariância, mas o leitor atento deve ter notado uma certa futilidade em usá-los em um espaço de escala Gaussiano – afinal, a equação linear do calor não é invariante por contraste! Em suma, um processo de detecção de detalhes multi-escala que seja invariante por mudanças de brilho deveria começar por uma equação diferencial que também o fosse!

Há uma segunda motivação para analisar espaços de escala morfológicos: se quisermos criar estruturas multi-escala para formas (“shapes”, em geral uma região do plano limitada por uma curva fechada), faz sentido criar uma equação diferencial para a evolução desta forma (ou, equivalentemente, da curva que a limita) que dependa apenas de suas propriedades geométricas. Se esta forma for determinada por uma curva de nível de uma imagem, queremos evoluir a imagem independentemente em cada curva de nível. Mas que processo deformará cada uma das curvas de nível? Que propriedades desejamos para este processo?

Iniciaremos este capítulo apresentando um tal processo de deformação de curvas com belíssimas propriedades matemáticas: o *movimento por curvatura*.

¹A morfologia matemática em processamento de imagens é uma coleção de métodos que trata uma imagem como formada por suas curvas de nível, independentemente de seus rótulos. Assim, a morfologia matemática é a essência dos métodos invariantes por mudança de brilho, justificando o título deste capítulo – apesar de quase não tratarmos aqui dos métodos da Morfologia Matemática na maneira tradicional.

13.1 Movimento por Curvatura

Propagação de “frontes” determinadas por equações diferenciais que dependam apenas de propriedades locais são assuntos estudados há um bom tempo; algumas aplicações físicas incluem propagação de fogo ([26] e [44]) e crescimento de cristais ([4] e [27]), dentre outras². Nos últimos 20 anos, o assunto tem recebido bastante atenção, especialmente depois que uma série de artigos (a saber, [15], [13] e [16]) provou que o Movimento por Curvatura mais simples (para curvas imersas em \mathbf{R}^2) tem várias propriedades fascinantes (sem querer estragar a surpresa: qualquer curva imersa no plano permanece imersa, sem auto-interseções e converge a um “ponto redondo” – explicaremos estas propriedades com mais detalhes a seguir). Aproximações numéricas são abordadas em [35] e [45], e os problemas com o surgimento de singularidades são abordados em [3] e [10].

Não é nenhuma surpresa que uma evolução de curvas com tal apelo geométrico tenha encontrado aplicações em Visão Computacional. Por exemplo, em [23], [21] e [24], Kimia apresenta um método de classificação de formas bastante atraente (veja também [47] e [32] – uma noção dessas idéias será apresentada ainda neste capítulo). Aplicações do Movimento por Curvatura à suavização de imagens podem ser encontradas em [1] e [2].

13.1.1 Definição e Propriedades

Definição 13.1 *Seja $\mathcal{C}(t)$, $0 \leq t \leq T_{MAX}$, uma família de curvas \mathcal{C}^2 fechadas imersas em \mathbf{R}^2 parametrizadas por $Q(s, t)$. Dizemos que $\mathcal{C}(t)$ satisfaz o movimento por curvatura média³ (MCM) quando*

$$Q_t(s, t) = \kappa(s, t)N(s, t) \quad (13.1)$$

onde $\kappa(s, t)$ e $N(s, t)$ são a curvatura de $\mathcal{C}(t)$ e seu vetor normal unitário apontando para dentro no ponto $Q(s, t)$ (veja figura 13.1).

A solução mais simples do MCM é a família de círculos de raio $R(t) = \sqrt{R_0^2 - 2t}$ (deixamos ao leitor a verificação de que esta é de fato uma solução de 13.1). Intuitivamente, o MCM faz com que dedos finos da curva $\mathcal{C}(t)$ encolham rapidamente devido à alta curvatura de suas extremidades, partes retas permaneçam razoavelmente imóveis e partes côncavas se movam para fora e desapareçam, fazendo com que a curva fique rapidamente convexa. De fato, esperamos que a forma fique “arredondada” à medida que o tempo passa, e então encolha e desapareça.

²De fato, como Superfícies Mínicas são estacionárias no Movimento por Curvatura Média, o assunto possui raízes ainda mais antigas.

³A palavra *média* aparece porque pensamos neste movimento como a instância bidimensional de um movimento de superfícies mais genérico.

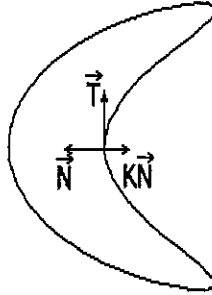


Figura 13.1: Convenção de orientações para o MCM

É interessante notar o que ocorre com o comprimento $L(t)$ e a área $A(t)$ da curva em questão. De fato, não é difícil mostrar que (veja exercícios)

$$L'(t) = - \int_{C(t)} \kappa^2 dL$$

$$A'(t) = -2\pi$$

Em particular, note que $A(t) = A_0 - 2\pi t$ e portanto a região em questão desaparecerá depois de exatamente $A_0/2\pi$ unidades de tempo!

Quanto à curvatura, alguns cálculos mostram que $\kappa(s, t)$ satisfaz

$$\frac{\partial \kappa}{\partial t} = \kappa'' + \kappa^3$$

desde que a segunda derivada do lado direito seja tomada com relação ao **parâmetro de comprimento de arco**. Assim, a curvatura satisfaz uma espécie de equação do calor (o que indica suavização instantânea da curva em questão mesmo que ela seja inicialmente não diferenciável). Em particular, usando esta equação é fácil notar que uma curva convexa (caracterizada por $\kappa > 0$ em todos seus pontos) tem de permanecer convexa (já que, se $\kappa > 0$, podemos garantir uma espécie de “princípio do mínimo” para $\kappa(s)$ – teremos $\kappa'' > 0$ e $\kappa^3 > 0$ em qualquer mínimo local, portanto mínimos locais de $\kappa(s)$ não “descem” e κ permanece positivo). Por outro lado, não há nenhum princípio do máximo para $\kappa(s)$ – de fato, o termo κ^3 faz com que a curvatura cresça rapidamente e vá para infinito **em tempo finito**, como sugerido pelo exemplo do círculo acima.

O leitor que queira ver as idéias acima apresentadas com mais rigor deve consultar [15], [13] e [16]; estes artigos mostram que

- Curvas inicialmente não diferenciáveis tornam-se diferenciáveis “instantaneamente”;

- O movimento por curvatura não cria singularidades ou auto-interseções até o momento final $t = \frac{An}{2\pi}$;
- Curvas não-convexas tornam-se convexas em tempo finito;
- Curvas convexas permanecem convexas e finalmente convergem para um “ponto redondo”.

A idéia de um “ponto redondo” merece uma clarificação: a idéia é olhar para a razão isoperimétrica $\frac{L^2}{A}$ da curva em questão. A conhecida desigualdade isoperimétrica afirma que $\frac{L^2}{A} \geq 4\pi$ para curvas \mathbf{C}^2 fechadas no plano, com igualdade ocorrendo somente para circunferências. Para o caso do MCM, pode-se mostrar facilmente que

$$\left(\frac{L^2}{A}\right)_t = -2\frac{L}{A} \left(\int_{C(t)} \kappa^2 dL - \pi \frac{L}{A} \right)$$

e [14] mostra que o lado direito é sempre negativo e que a razão $\frac{L^2}{A}$ decresce para 4π à medida que A se aproxima de 0 (exceto para circunferências, quando $\frac{L^2}{A} = 4\pi$ independentemente de t e o lado direito é sempre 0).

13.2 Formas Multi-Escala

Nesta seção, descreveremos brevemente algumas idéias por trás da teoria computacional de formas proposta por Kimia em [22], [23], [21], [24] e [47] que usa o MCM de maneira essencial.

Nós, humanos, conseguimos classificar formas quase sem pensar. Enquanto é verdade que a descrição de uma forma é bastante dependente do observador (e há evidências fortíssimas de que a relação “forma A é semelhante à forma B ” não é uma relação de equivalência, talvez nem mesmo uma relação no sentido matemático da palavra!), nós tendemos a concordar que, por exemplo, a forma do corpo humano é basicamente composta por cabeça, tronco, dois braços e duas pernas (pelo menos numa certa “escala”). Como chegamos a essa subdivisão? Como ensinar um computador a fazê-la (não estamos tão interessados em encontrar os **nomes**, mas em encontrar a subdivisão em si)? Mais ainda, associamos uma hierarquia a esses objetos (como formas, o tronco é a “raiz” à qual pertencem a cabeça, tronco e membros) e temos uma flexibilidade incrível em encontrá-los (veja um Dali ou um Picasso para notar isso).

Para o resto de nossa discussão, vamos ignorar toda a informação que vem de percepção de profundidade ou funcional, cor, movimento, textura etc. e vamos nos concentrar apenas na forma. Assim, imagine que temos uma certa curva no plano que descreva essa forma, e gostaríamos de extrair informações sobre essa forma, algo como “olha, são três manchas, duas das quais são conectadas por uma espécie de pescoço e uma terceira que está colada na primeira”. Como fazer isto?

A idéia de Kimia e usar movimentos por curvatura para fazer este processamento computacionalmente. Mais exatamente, apliquemos o seguinte movimento à curva $Q(s, t)$ parametrizada por s no tempo t

$$Q_t(s, t) = (\alpha + \beta\kappa(s, t))N(s, t)$$

onde α e β são constantes, κ é a curvatura e N é o vetor normal unitário como discutido anteriormente.

Se $\alpha = 0$, temos o MCM discutido na seção anterior (a presença de β simplesmente reescalará o tempo t , fazendo o processo mais rápido ou mais lento; podemos muito bem imaginar que $\beta = 1$). Como vimos, qualquer curva se tornará um ponto redondo permanecendo suave durante todo o processo. Uma figura humana com braços, pernas, pescoço e dedos simplesmente se tornará uma pequena mancha, nunca se subdividindo durante o processo, não importa quão finos sejam o pescoço ou os seu dedos. Este processo é chamado um *processo de difusão* aplicado à nossa forma inicial.

Por outro lado, se $\beta = 0$, temos $Q_t = \alpha N$, que é chamado um *processo de reação* aplicado à forma inicial⁴. A solução dessa equação pode ser escrita como

$$Q(s, t) = Q(s, 0) + \alpha t N(s, 0)$$

por algum tempo, mas não é imediato definir o que a solução deva ser em geral – a curva pode desenvolver rapidamente auto-interseções e singularidades onde N não está bem definido (por exemplo, assim que $t = \frac{1}{\alpha\kappa_{\max}}$, isto acontecerá)! Ainda assim, é possível definir uma espécie de “solução fraca” – basicamente, a solução será composta pelas curvas de nível da função distância à curva original. O leitor deve se convencer de que qualquer “pescoço” presente na forma inicial acabará quebrando a forma em partes, não importa quão pouco “pescoçudo” ele seja. Por exemplo, começando com uma figura humana, os dedos das mãos e dos pés rapidamente se separam do resto do corpo; a seguir, os braços e pernas e a cabeça se separam do tronco. Os dedos devem desaparecer primeiro, depois braços e pernas, a cabeça e finalmente o tronco some. Enfim, um processo ligeiramente desagradável para a cobaia utilizada.

E se utilizarmos outros valores para α e β ? Observe a razão α/β – quando esta for pequena, esperamos que o processo seja semelhante ao de difusão e apenas pescoços bem pescoçudos devam se quebrar durante o processo; por outro lado, quando a razão é grande o processo é de reação e quase todos os pescoços se quebram – a menos que eles sejam muito pouco proeminentes. Construindo um *espaço de reação-difusão* com todas as evoluções para todos os valores de α/β , podemos considerar o valor de α/β necessário para fazer um pescoço se quebrar como a “pescoçudez” daquele detalhe; assim, podemos determinar computacionalmente onde estão as divisões e determinar quão importante cada uma delas é (observando também quanto tempo aquela divisão demora para ocorrer dentro de um desses processos).

⁴Este é o momento deste capítulo em que chegamos mais perto dos operadores tradicionais da morfologia matemática tradicional. Afinal de contas, o processo a ser descrito neste parágrafo corresponde à versão contínua do clássico operador *erosão* ou *dilatação*.

Este é apenas um apanhado das idéias do trabalho de Kimia, e novamente encorajamos o leitor a consultar as referências apresentadas no início desta seção para mais detalhes no assunto.

13.3 Computando Movimentos por Curvatura

Uma questão surpreendentemente não trivial a respeito do MCM e outros movimentos por curvatura é: como computá-los?

Uma idéia imediata é colocar “marcadores” (isto é, uma lista de pontos) na curva a ser deformada. Para cada pequeno passo de tempo, obtemos aproximações numéricas da curvatura em cada um desses marcadores, da normal e movemos cada um de acordo com a lei da equação diferencial. Depois de mover todos os marcadores por um passo de tempo, repetimos o processo. Esta é a chamada *abordagem Lagrangiana*. Enquanto esta abordagem pode funcionar, ela tem problemas: à medida que o tempo passa, os marcadores se aproximam uns dos outros, e a aproximação numérica da curvatura torna-se extremamente instável (especialmente perto de “ondas de choque” onde a curva não é diferenciável, que não ocorrem para o MCM mas podem ocorrer para as variações do tipo $Q_t = (\alpha + \beta\kappa)N$, por exemplo). Também, é difícil decidir como reorganizar os marcadores quando há transições topológicas, isto é, quando a curva se quebra em pedaços ou pedaços se juntam para formar formas compostas (aliás, decidir se *houve* ou não uma quebra já é difícil com marcadores). Para mais detalhes sobre tais dificuldades, veja [35], [45] ou [46].

Já a *abordagem Euleriana* resolve quase todos esses problemas, mas paga o preço de adicionar uma dimensão ao problema. A idéia é a seguinte: imagine que a curva $Q(s, 0)$ corresponde à solução de $F(x, y) = 0$, isto é, $Q(s, 0)$ é a curva de nível 0 da função $F(x, y)$. Se fizermos a função F variar com o tempo t , é claro que tal curva de nível pode mudar. Mas lembremos que a curvatura de qualquer curva de nível é dada por⁵

$$\kappa = \operatorname{div} \left(\frac{\nabla F}{|\nabla F|} \right) = \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)^{3/2}}$$

Por outro lado, não é difícil ver que mover o gráfico de $F(x, y)$ num ponto para cima com velocidade v faz com que a curva de nível passando por aquele ponto mova-se com velocidade

$$-v \frac{\nabla F}{|\nabla F|^2} = \frac{v}{|\nabla F|} N$$

onde N é a normal unitária à curva de nível naquele ponto apontando na direção em que F **diminui**. Juntando tudo, se $F(x, y, t)$ satisfaz a equação

$$\frac{\partial F}{\partial t} = \kappa |\nabla F| = |\nabla F| \operatorname{div} \left(\frac{\nabla F}{|\nabla F|} \right) = \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_x^2 + F_y^2} \quad (13.2)$$

⁵Veja o apêndice para uma demonstração deste fato.

então sua curva de nível $Q(t) : F(x, y) = 0$ se moverá de acordo com MCM (aliás, todas as curvas de nível, não só a de rótulo 0, farão seus respectivos MCMs).

Em geral, se F satisfaz

$$\frac{\partial F}{\partial t} = |\nabla F| G \left(\operatorname{div} \left(\frac{\nabla F}{|\nabla F|} \right) \right) \quad (13.3)$$

então suas curvas de nível realizarão o seguinte movimento por curvatura

$$\frac{\partial Q}{\partial t} = G(\kappa)N$$

desde que você force $F < 0$ dentro da curva original e $F > 0$ fora.

Agora parece simples usar aproximações numéricas das derivadas de F para fazer uma implementação numérica que resolva 13.3; é verdade que introduzimos uma dimensão extra ao problema e que temos que calcular a curva Q a cada passo, mas temos as seguintes vantagens:

- Não há necessidade de fazer um gerenciamento complicado de marcadores;
- Transições topológicas são automaticamente calculadas;
- Temos uma precisão sub-pixel no cálculo de Q ;
- A equação 13.3 é estável (mas há que se tomar cuidado com a particular implementação – veja as referências do próximo parágrafo);

A equação 13.3 tem sido bastante estudada recentemente. Há problemas teóricos devido à presença de $|\nabla F|$ no denominador, que podem ser dirimidos via soluções por viscosidade ([3] e [10]). Há também preocupações com respeito à particular implementação a ser usada e, para diminuir o impacto da dimensão extra e reduzir o custo computacional, métodos do tipo “Banda Estreita” (Narrow Band) podem ser usados (veja [46]). No entanto, se estes cuidados forem tomados, a equação 13.3 acaba oferecendo a melhor maneira computacional de calcular o MCM.

13.4 A Equação do Calor Intrínseca

A abordagem Euleriana acima para o MCM sugere um espaço de escala não-linear $F_t(x, y)$ que seria gerado a partir de uma imagem original $I(x, y)$ pela EDP

$$\frac{\partial F_t}{\partial t} = \kappa |\nabla F| = \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_x^2 + F_y^2} \quad (13.4)$$

$$F_0(x, y) = I(x, y)$$

Note que não é fácil analisar esta equação diretamente a partir da fórmula acima. Por exemplo, não sabemos o que fazer com pontos onde $F_x = F_y = 0$ –

a expressão do lado direito é indeterminada em tal caso. Aliás, se escrevemos $F_x = \lambda F_y$, o lado direito é

$$A = \frac{F_{xx}F_y^2 - 2F_{xy}\lambda F_y^2 + F_{yy}\lambda^2 F_y^2}{\lambda^2 F_y^2 + F_y^2} = \frac{F_{xx} - 2\lambda F_{xy} + \lambda^2 F_{yy}}{\lambda^2 + 1} \quad (13.5)$$

mostrando que o limite da expressão $\kappa |\nabla F|$ quando $|\nabla F| \rightarrow 0$ depende exatamente de **como** este gradiente se aproxima de 0.

Ainda assim, há esperança de obter alguma espécie de princípio do máximo. Afinal, enquanto 13.5 depende de λ , pode-se mostrar que

$$\lambda_1 < \frac{F_{xx} - 2\lambda F_{xy} + \lambda^2 F_{yy}}{\lambda^2 + 1} < \lambda_2$$

onde λ_1 e λ_2 são os autovalores da Hessiana

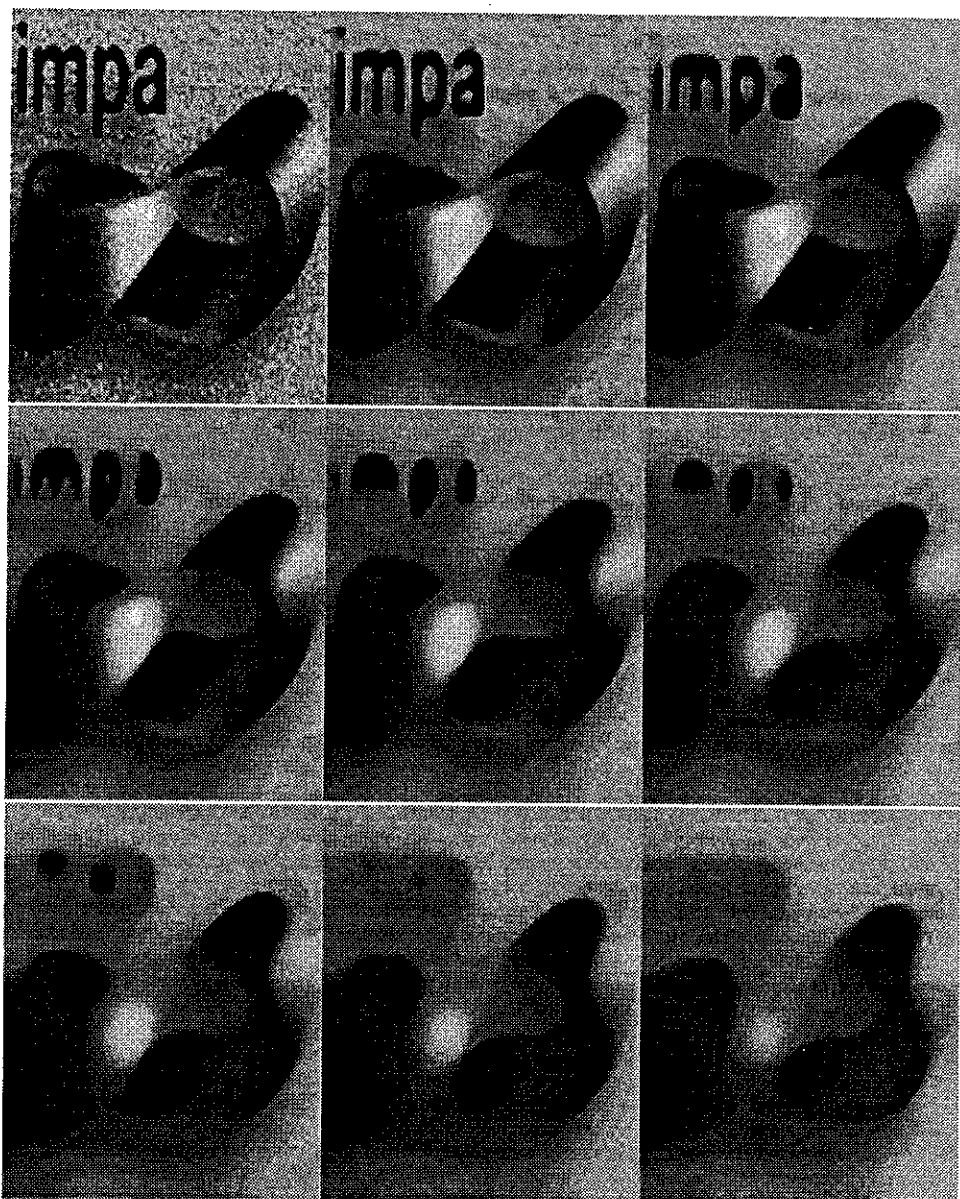
$$H = \begin{pmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{pmatrix}$$

Assim, num ponto de mínimo local de $F_t(x, y)$, onde $|\nabla F| = 0$ e ambos λ_1 e λ_2 são não-negativos, qualquer que seja a maneira de tomar o limite $|\nabla F| \rightarrow 0$ temos que o limite para a expressão A (se existir) está entre λ_1 e λ_2 e portanto é não-negativo; assim, este mínimo local não poderá descer. De maneira análoga, um máximo local não deve subir.

Em suma: nós não temos a mínima idéia de como determinar o que $\frac{\partial F_t}{\partial t}$ deveria ser num ponto crítico (a equação simplesmente não nos diz o que fazer ali), mas, se arrumarmos uma maneira minimamente coerente de estender 13.4 a tais pontos críticos, esta EDP deverá satisfazer o princípio do máximo.

Esta maneira “coerente” de estender esta EDP a pontos críticos existe, e é baseada na teoria de Soluções por Viscosidade (veja [6] para um ótimo resumo desta teoria; [46] é menos rigoroso mas dá uma idéia melhor do que “está acontecendo”; no caso específico desta EDP, veja [3] e [10]), que é uma maneira de definir soluções “fracas” para essa EDP que ainda satisfazem condições de unicidade de solução e estabilidade com relação a condições iniciais.

Porém, não queremos entrar em detalhes desta teoria – é suficiente dizer que as soluções geradas pela Teoria da Viscosidade equivalem às soluções geradas pela análise do MCM da seção anterior, isto é, cada curva de nível perfaz o seu MCM independentemente das demais. Pense que cada curva de nível do gráfico de sua função $F_t(x, y)$ passa pelo MCM e o autor aqui promete que a solução obtida será a mesma que a solução por viscosidade da EDP 13.4 correspondente. Como mencionamos na seção anterior, os métodos numéricos tem de ser implementados com certo cuidado (e são baseados muito mais na EDP do que na idéia do MCM), mas o problema está bem estudado (veja as referências da seção anterior).



O logotipo do IMPA derretendo sob o MCM. Note como os passo iniciais são ótimos para limpar ruídos e texturas, e como todas as “manchas” tornam-se convexas e se arredondam antes de desaparecerem (o pingo do “i” é uma das primeiras regiões a fazê-lo).

Com essa interpretação, podemos revisitar o princípio do máximo: como o MCM **encolhe** as curvas de nível, ele jamais “criará” novas curvas de nível; mas a criação de curvas de nível é equivalente a ter um máximo local que “sobe” ou um mínimo local que “desce”. Como esta criação não ocorre, aplicar o MCM às curvas de nível de $F_i(x, y)$ é um processo que respeita o princípio do máximo.

Assim, esta EDP (pense nela como um MCM) nos dá uma espaço de escala não-linear que satisfaz o princípio de **semi-grupo** (repetimos que praticamente qualquer EDP o satisfaz) e possui a **causalidade** no sentido do princípio do máximo (que nem mesmo o Espaço de Escala Gaussiano possuía!). Note também que ele é **invariante por translações e rotações** (já que estas não alteram a curvatura das curvas de nível). Finalmente, note que este espaço de escala é **invariante por mudanças de brilho** – de fato, o MCM atua em cada curva de nível independentemente das demais; como mudanças de brilho apenas reempilham as curvas de nível sem alterá-las, o MCM é invariante por mudanças monótonas de brilho; a palavra “monótona” é importante apenas para que a direção do vetor normal unitário usado no MCM não mude.

Por estes motivos, o MCM (também chamado de *Equação do Calor Intrínseca*) tem sido sugerido como uma boa maneira de suavizar imagens e criar espaços de escala. Ruídos pontuais, que têm escala pequena, correspondem a curvas de nível que englobam pequenas áreas, e portanto desaparecem rapidamente.

13.4.1 Filtro Mediano

A mediana de um número finito de valores é simplesmente o valor “do meio” quando tais valores são ordenados em maneira crescente. Assim, por exemplo, a mediana do conjunto com 7 números $\{-10, 1, 2, 2, 3, 3, 1000\}$ é o quarto valor, no caso, 2. No caso de haver um número par de amostras a considerar, costuma-se tomar a média aritmética dos dois valores “do meio”. A mediana é bastante insensível a amostras “isoladamente estranhas” e por isso é freqüentemente utilizada ao invés da média aritmética para conjuntos de amostras.

É interessante notar que, dado um conjunto de amostras $\{A_i\}_{i=1,2,\dots,n}$, a média aritmética pode ser definida como o valor μ que minimiza a expressão $\sum (A_i - \mu)^2$, enquanto a mediana m é um dos valores que minimizam a expressão $\sum |A_i - m|$.

Um dos filtros discretos mais conhecidos para eliminação de ruídos pontuais em imagens é o filtro mediano. No caso discreto, a idéia é simples: para cada pixel P de sua imagem, considere uma vizinhança $V(P)$ e substitua o valor de P pelo valor da mediana de todos os valores dos pixels em $V(P)$. É comum tomar $V(P)$ como um quadrado 3×3 , 5×5 ou 7×7 em volta de P , ou como o conjunto de pixels a distância menor que R de P para algum R fixo. Não é difícil ver porque tão filtro funciona bem: ruídos pontuais correspondem a amostras “isoladamente estranhas” dentro da vizinhança $V(P)$, e portanto tendem a ser eliminados por este filtro.

Também vale a pena notar que o uso de uma média aritmética ou de uma média ponderada para o novo valor do pixel P equivale simplesmente a fazer

uma convolução com uma máscara determinada pelos pesos da média tomada. Assim, por exemplo, se $V(P)$ é um quadrado 5×5 e tomamos a média aritmética dos 25 pixels como o novo valor do pixel P , estamos de fato fazendo uma convolução com a máscara

$$\frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

criando uma suavização linear que já foi bem estudada na parte II deste livro. Por outro lado, o filtro mediano é não-linear.

Qual seria o análogo contínuo do filtro mediano? Dada uma função $f(x, y)$, podemos considerar o seguinte processo: para cada ponto P considere uma vizinhança $V(P)$ – podemos simplificar e imaginar que $V(P)$ é simplesmente um disco de raio R em torno de P . Considere agora a função

$$\text{Mediana}_{V(P)} f(x, y) = g_R(x, y)$$

que é definida pela propriedade

$$\text{Área} \{(x, y) | f(x, y) \geq g_R(x, y)\} = \frac{\text{Área}(V(P))}{2}$$

Assim, $g_R(x, y)$ pode ser imaginado como um espaço de escala de uma imagem $f(x, y)$ onde $R = R(t)$ faz o papel de escala (note que $R \rightarrow 0$ faz com que $g \rightarrow f$ sempre que f for razoavelmente bem comportada).

A principal conexão entre o filtro mediano e o MCM é o seguinte resultado encontrado em [17]:

Teorema 13.2 *Com a notação acima*

$$g_R(x, y) - f(x, y) = \frac{1}{6} R^2 \kappa_f |\nabla f| + o(R^2)$$

Em particular, se definirmos um espaço de escala $F_t(x, y)$ a partir de uma imagem inicial $f(x, y)$ por filtros medianos

$$F_t(x, y) = g_{\sqrt{6t}}(x, y)$$

$$F_0(x, y) = f(x, y)$$

então

$$\frac{\partial F_0(x, y)}{\partial t} = \kappa_{F_0} |\nabla F_0|$$

Em outras palavras, assintoticamente para R pequeno, o filtro mediano perfaz um processo semelhante à deformação inicial feita pela equação do MCM.

É importante notar que esta semelhança somente vale para R e t pequenos (a última expressão só vale em $t = R = 0$), e em nenhum momento estamos afirmando que o espaço de escala gerado pelos filtros medianos é o mesmo do que o espaço gerado pelo MCM.

Aliás, [17] mostra que um filtro mediano é assintoticamente equivalente a um MCM qualquer que seja a vizinhança escolhida, mesmo que esta não seja um disco (usando um parâmetro R para escalar a vizinhança $V(P)$). A única diferença é a constante $\frac{1}{6}$ que pode ser diferente dependendo do tamanho da vizinhança escolhida.

Mais geralmente, [17] mostra como obter o comportamento assintótico de qualquer filtro morfológico criado via escalamento de elementos estruturais, mas preferimos deixar apenas dirigir o leitor àquele livro ao invés de repetir suas idéias.

13.5 Espaço de Escala Morfológico Afim

Baseado nas idéias do capítulo 8, o leitor deve ter notado que é fácil criar espaços de escala invariantes por mudança de brilho. De fato, no caso 2D, qualquer EDP da forma

$$\frac{\partial F_t(x, y)}{\partial t} = |\nabla F| \cdot G(\kappa, t)$$

possuirá tal invariância, mais a invariância por translações e rotações. Em particular $G = 1$ e $G = -1$ correspondem a processos de reação como descritos na seção 13.3 (e portanto, a versões contínuas dos processos de erosão e dilatação); $G(\kappa, t) = \kappa$ gera o MCM discutido anteriormente. Invente a sua função G e crie o seu espaço de escala morfológico!

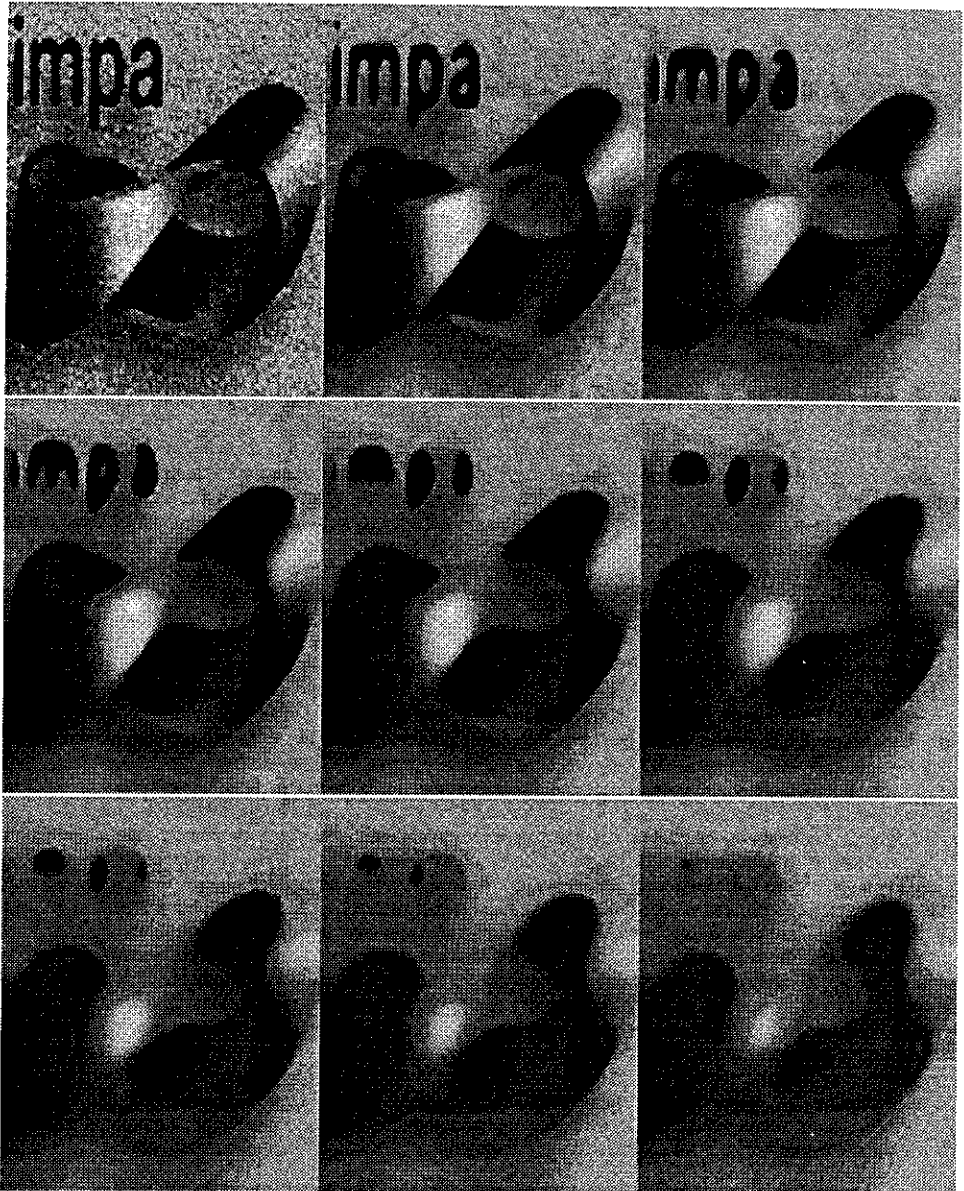
Em suma, ainda temos bastante liberdade para escolher nosso espaço de escala morfológico. Que outra condição podemos impor nele?

Uma opção notável é a sugerida pelo Espaço de Escala Morfológico Afim em [1] e [40]: exija também que o espaço de escala gerado seja invariante não somente por transformações Euclidianas mas também por transformações afins de seu domínio! Adicionando esta condição, chega-se a um único espaço de escala causal, regular⁶ e morfológico em duas dimensões

$$\frac{\partial F_t}{\partial t} = |\nabla F| (t, \kappa_F)^{1/3}$$

Ao leitor interessado recomendamos o artigo de Alvarez e Morel em [9] para uma análise mais detalhada desta equação e sua versão tridimensional (para filmes!).

⁶A noção de regularidade apresentada no artigo de Alvarez e Morel em [9] inclui o conceito de que a equação diferencial a ser obtida pode ser determinada a partir de suas derivadas de ordem até 2. Há outras opções de espaços de escala morfológicos usando derivadas de ordem mais alta que dependem de derivadas da curvatura.



O logotipo do IMPA derretendo sob o AMSS. As imagens são tomadas nos mesmos tempos que as imagens do MCM mostradas anteriormente. Note como as “manchas” da figura tornam-se convexas com o tempo, mas **não se arredondam** antes de desaparecer – o que é consistente com a invariância por Transformações Afins.

13.6 Exercícios

1) Suponha que uma família de circunferências de raio $R(t)$ satisfaz o MCM. Mostre que $R'(t) = -K(t) = -\frac{1}{R(t)}$ e resolva esta equação diferencial para mostrar que $R(t) = \sqrt{R_0^2 - 2t}$.

2) Se uma curva $\mathcal{C}(t)$ satisfazendo o MCM é parametrizada por $Q(s, t)$ então podemos escrever

$$\begin{aligned} Q_s &= gT \\ Q_t &= \kappa N \end{aligned}$$

onde $g(s, t)$ é a métrica da curva e T é o vetor tangente unitário à curva. Mostre que

$$\begin{aligned} g_t &= -g\kappa^2 \\ T_t &= \frac{\kappa_s}{g}N \\ N_t &= -\frac{\kappa_s}{g}T \end{aligned}$$

(Sugestão: calcule a derivada mista Q_{st} em ambas as ordens possíveis; compare as expressões obtidas lembrando que T e T_t são ortogonais pois T é unitário; enfim, note que a terceira equação é consequência imediata da segunda)

3) O comprimento da curva $\mathcal{C}(t)$ é dado por

$$L(t) = \int_{\mathcal{C}(t)} dL = \int_{\mathcal{C}(t)} g(s, t) ds$$

Use o problema (2) para mostrar que, no MCM,

$$L'(t) = - \int_{\mathcal{C}(t)} \kappa^2 dL$$

4) Pelo Teorema de Green, a área englobada por uma curva $\mathcal{C}(t)$ parametrizada por $Q(s, t) = (x(s, t), y(s, t))$ pode ser escrita como

$$A(t) = \frac{1}{2} \int_{\mathcal{C}(t)} xy_s - yx_s ds = \frac{1}{2} \int_{\mathcal{C}(t)} -\langle Q, N \rangle g(s, t) ds$$

Use (2) para mostrar que, no MCM,

$$A'(t) = -2\pi$$

(Sugestão: você deverá usar uma integração por partes e, ao final, a propriedade de que a integral de linha da curvatura ao longo de $\mathcal{C}(t)$ é independente de $\mathcal{C}(t)$ e vale 2π , isto é

$$\int_{\mathcal{C}(t)} \kappa dL = \int_{\mathcal{C}(t)} \kappa g ds = 2\pi$$

para qualquer curva fechada simples $\mathcal{C}(t)$ orientada no sentido anti-horário)

5) Mostre que, no MCM,

$$\kappa_t = \frac{g\kappa_{ss} - g_s\kappa_s + g^3\kappa^3}{g^2}$$

Em particular, se no tempo t a curva é parametrizada pelo comprimento de arco, então $g(s) \equiv 1$ e

$$\kappa_t = \kappa_{ss} + \kappa^3$$

(Sugestão: comece com $T_s = g\kappa N$ e derive com relação a t , use o problema (2)).

6) Mostre que a derivada da razão isoperimétrica no MCM é de fato

$$\left(\frac{L^2}{A}\right)_t = -2\frac{L}{A} \left(\int_{\mathcal{C}(t)} \kappa^2 dL - \pi\frac{L}{A}\right)$$

7) Considere a função

$$f(\lambda) = \frac{\lambda^2 c - 2\lambda b + a}{\lambda^2 + 1}$$

Mostre que ela se encontra de fato entre λ_1 e λ_2 onde estes são os autovalores da matriz

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

(Sugestão: se você gosta de álgebra linear, note que $f(\lambda) = \frac{v^T H v}{v^T v}$ onde $v = \begin{pmatrix} 1 \\ \lambda \end{pmatrix}$)

8) Dado uma seqüência de valores $\{A_i\}_{i=1,2,\dots,n}$ encontre o valor de μ que minimiza $\sum (A_i - \mu)^2$. Quais são os valores de m que minimizam $\sum |A_i - m|$? (Sugestão: separe este último caso em dois casos, para n ímpar e n par; mostre que m deve estar entre os pontos extremos da seqüência, e então elimine-os; repita o processo até encontrar a resposta)

9) Como um círculo evolui via AMSS? Determine o raio $R(t)$ do círculo em função do tempo dado $R_0 = R(0)$.

Parte IV

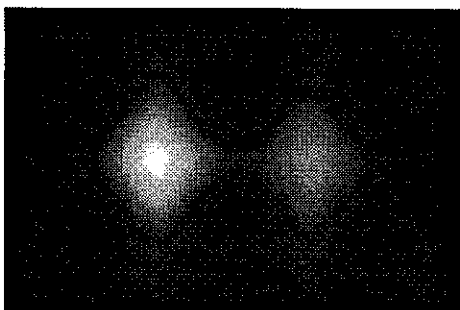
Esboço Primitivo

Capítulo 14

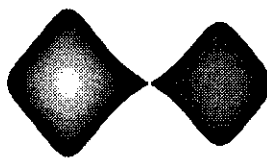
Esboço Primitivo

14.1 Bolhas

Neste capítulo, buscamos uma representação mais explícita das “estruturas” de uma imagem. Em se tratando de imagens preto-e-branco, tais estruturas correspondem em geral a áreas brancas em fundo preto ou vice-versa. Note também que uma divisão intuitiva de uma imagem em estruturas leva à geração de várias áreas conexas da imagem. Esta representação é uma possível versão computacional do “esboço primitivo” citado no capítulo introdutório.



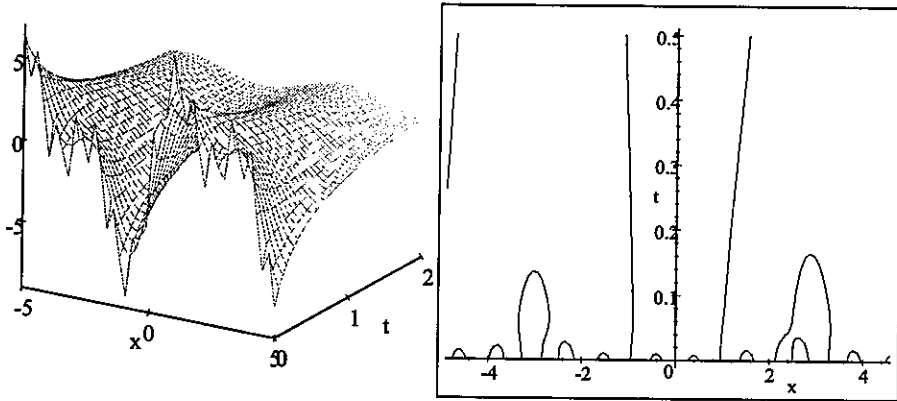
Uma imagem sintética...



...e uma possível divisão em duas “estruturas”.

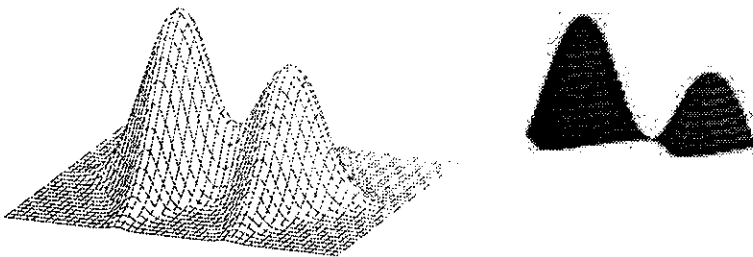
Uma outra motivação para a criação destas “regiões” é a seguinte: seguir curvas (como, por exemplo, arestas) ou pontos através das escalas não é sempre um processo simples. A figura abaixo mostra um espaço de escala de um sinal 1D e a trajetória de seus máximos e mínimos. Note que tais máximos e mínimos podem simplesmente colidir e desaparecer à medida que a escala aumenta; perto do momento da colisão, as velocidades “laterais” destes máximos e mínimos

pode ser alta, dificultando o seu rastreamento computacional; também, quando eles estão perto uns dos outros, pode ser difícil decidir que ramo pertence a que ponto crítico em escalas discretas sucessivas. Criando “bolhas” ao redor desses pontos críticos definimos “regiões de interesse” que se prestam melhor a ligações entre escalas.



Voltando ao caso 2D, associaremos então a cada máximo da imagem $F(x, y)$ uma certa região conexa. A questão é, se um máximo local é uma semente de uma bolha, quais são os limites dessa bolha? Imagine que uma bolha cresce a partir de um máximo – onde parar?

A definição a ser usada no resto deste capítulo crescerá uma bolha adicionando curvas de nível correspondendo a níveis de cinza cada vez mais baixos, parando quando esta bolha se junta a alguma outra bolha em formação. A idéia fica mais clara a partir das figuras a seguir:



Se este é o gráfico da função $F(x, y)$aqui estão suas duas bolhas.

Formalmente, temos:

Definição 14.1 *Seja $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma imagem com um máximo local em $P \in \mathbb{R}^2$. Para $z_0 < F(P)$, considere o conjunto*

$$X_{z_0} = \{(x, y, z) \in \mathbb{R}^2 \times \mathbb{R} \mid z_0 \leq z \leq F(x, y)\}$$

Sejam

$X_{z_0}^P =$ componente conexa de X_{z_0} que contém $(P, F(P))$

$G_{z_0}^P =$ caminhos $\gamma(t) = (x(t), y(t), z(t))$ em $X_{z_0}^P$
saindo de $(P, F(P))$ tais que $z'(t) < 0$

$H_{z_0}^P =$ caminhos $\gamma(t) = (x(t), y(t), z(t))$ em $X_{z_0}^P$
saindo de $(P, F(P))$ tais que $z'(t) \leq 0$

$$z_{base}(P) = \max_{z < F(P)} (\overline{G_z^P} \neq H_z^P)$$

então a bolha associada ao máximo local P é o conjunto tridimensional

$$B_P = \overline{G_{z_{base}(P)}^P}$$

e seu volume, suporte e contraste são definidos como¹

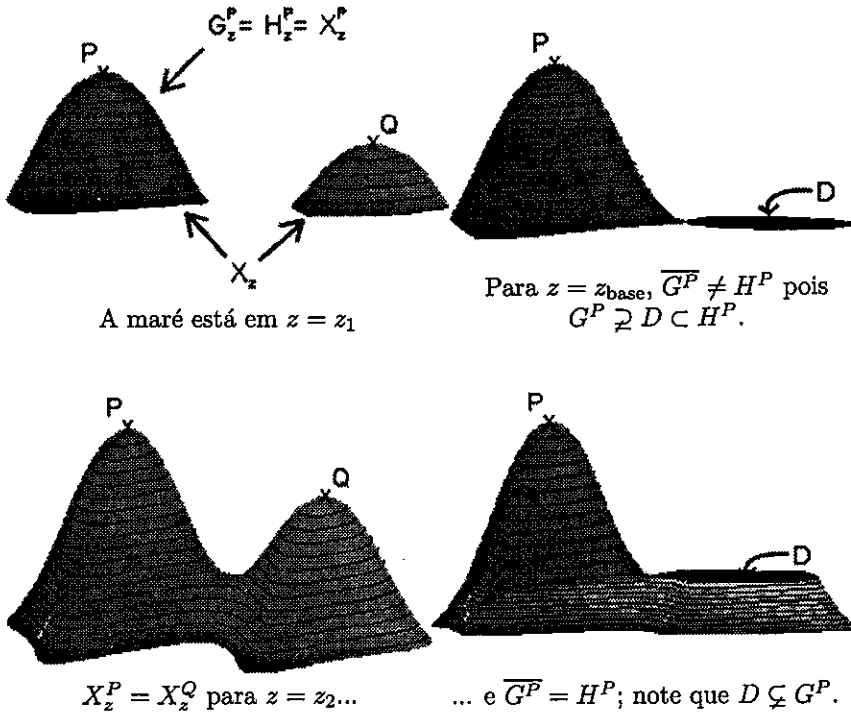
$$\text{Suporte} = \pi_{xy}(B_P)$$

$$\text{Volume} = V_{B_P} = \int_{B_P} dV = \int_{\text{Suporte}} (F(x, y) - z_{base}(P)) \, dx \, dy$$

$$\text{Contraste} = F(P) - z_{base}(P)$$

Finalmente, usaremos a notação $P = e(B)$ sempre que P for um ponto extremo (máximo ou mínimo) a partir do qual formamos a bolha B .

¹Denotamos por $\pi_{xy}(B_P)$ a projeção ortogonal de B_P no plano xy (isto é, em \mathbb{R}^2). Note que, estritamente falando, bolhas são objetos tridimensionais que "moram embaixo do gráfico de $F(x, y)$ ". No entanto, muitas vezes nos referimos a bolhas citando apenas o seu suporte. O leitor deve se convencer que, dada F , é simples obter um desses conjuntos a partir do outro.



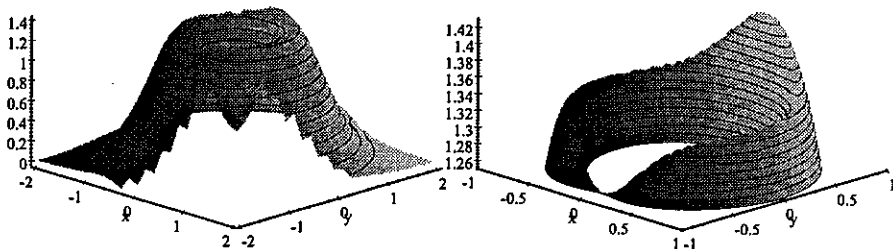
Comentário 14.2 Não se assuste com a aparência críptica dessa definição. Vamos decifrá-la voltando à nossa intuição original: z é o nível da maré. À medida que diminuímos z , um conjunto de ilhas aparece (X_z); a ilha que contém o ponto P é a ilha X_z^P . Na primeira figura, para $z = z_1$ ligeiramente abaixo de $F(P)$, os conjuntos G_z^P e H_z^P são em geral o mesmo que a ilha X_z^P . No entanto, se diminuirmos z bastante, note que X^P , G^P e H^P ganham grandes quantidades de terra de uma só vez. X^P incorpora ilhas inteiras de uma vez (terceira figura), enquanto G^P e H^P somente crescem “para baixo” (quarta figura). A diferença entre G^P e H^P será algo como o “disco” D indicado, mas, como D está no fecho de G^P , não há diferença entre $\overline{G^P}$ e H^P . Definimos $z_{\text{base}}(P)$ como o “primeiro” ou máximo valor de z abaixo de $F(P)$ que produz uma diferença entre os conjuntos $\overline{G^P}$ e H^P (a diferença é o disco D da segunda figura). Isto ocorrerá em geral quando a maré atingir um ponto de sela S .

Se altos níveis de cinza são usados para representar áreas brancas, esta definição nos dá bolhas brancas em fundo preto. É simples adaptar esta definição para bolhas pretas em fundo branco: use mínimos ao invés de máximos e suba a maré observando a parte das ilhas sob a linha d’água. Daqui por diante, nos fixaremos na discussão de bolhas brancas em fundo preto.

Note que, como esta definição é baseada somente em conceitos de comparação entre pixels vizinhos (as definições de G e H são baseadas em caminhos decrescentes em z), ela é facilmente discretizável: os conceitos de caminho e função decrescente num caminho estão bem definidos para espaços discretos. Também, o ponto S tal que $z_{\text{base}}(P) = f(S)$ onde bolhas vizinhas em crescimento colidem é em geral um ponto de sela (mas pode não sê-lo). Outras propriedades da definição de bolhas apresentada acima incluem:

- Bolhas são conexas;
- Em geral, uma bolha é determinada por um par máximo + sela;
- Bolhas vizinhas competem entre si; a presença de pequenas bolhas pode limitar outras em crescimento; uma imagem unimodal (com apenas um máximo e nenhum outro ponto crítico) terá apenas uma bolha que tomará toda a imagem.
- A classificação de uma imagem em bolhas é incompleta; vários pontos da imagem não correspondem a bolha alguma e serão chamados pontos de fundo;

É possível que pontos da imagem pertençam tanto a uma bolha branca quanto a uma bolha preta; para evitar este acontecimento, deveríamos mudar a definição e parar de baixar a maré não somente quando a bolha em crescimento colide com outra bolha, mas também quando ela muda de topologia. Veja exemplo abaixo.



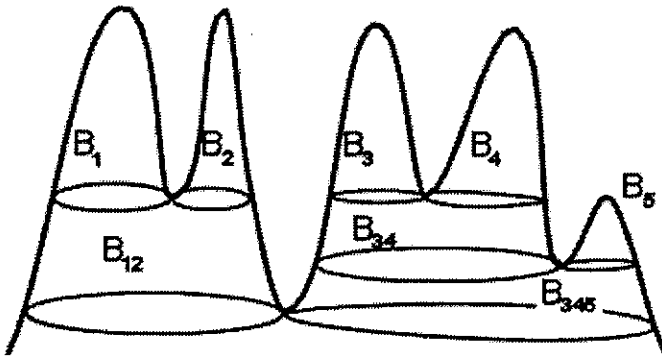
O mínimo local pertence a uma bolha branca e uma preta. Uma ampliação revela o momento da transição topológica.

Exemplo 14.3 Para a função F da figura acima, o máximo local inicia uma bolha "branca"; à medida que a maré baixa, a bolha cresce. No entanto, o primeiro ponto de sela atingido (um ponto no "anel" superior) não deterá a bolha – ambos os lados do ponto de sela pertencem à mesma bolha e, portanto, $\overline{G^P} = H^P$ também no nível z do ponto de sela. O que ocorre é que a bolha muda

de topologia – a partir deste momento, há um buraco em seu suporte. A maré continua baixando e acaba incluindo o mínimo local (quando o buraco se fecha) – e ainda assim nossa definição por caminhos indica que a bolha continuará crescendo (possivelmente tomando toda a imagem). Por outro lado, é claro que o mínimo local central é a semente de uma bolha “preta”.

14.1.1 Árvores de bolhas

Uma alternativa à definição acima corresponde a “não parar de baixar a maré mesmo que bolhas tenham colidido”. Ao invés, cada vez que duas bolhas colidem, iniciamos uma nova bolha. Tal estrutura produz uma árvore de bolhas, como na figura abaixo:

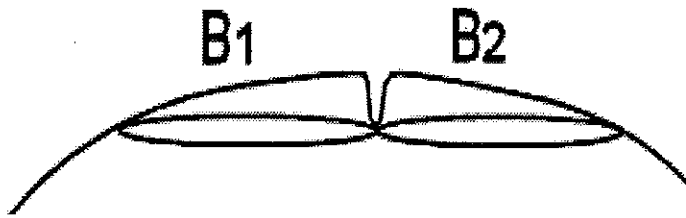


Uma árvore de bolhas

Note que cada arco da árvore de bolhas tem um volume, uma região suporte e uma área. Um ponto sela não só sinaliza o fim do crescimento de bolhas como definido anteriormente – ele também inicia um novo arco (uma nova bolha). As bolhas definidas na subseção anterior tornam-se as folhas desta árvore.

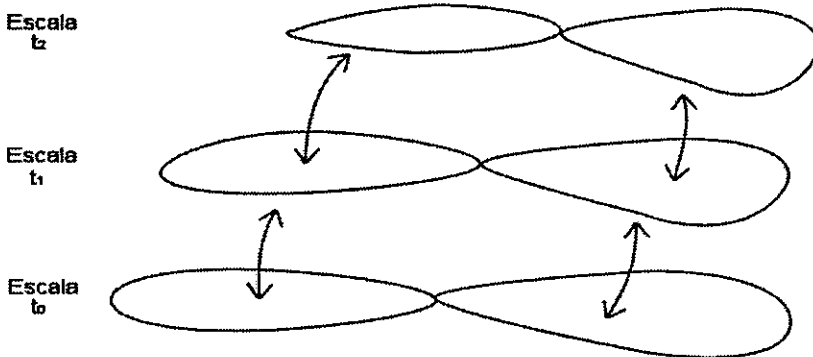
14.2 Hierarquia multi-escala

Infelizmente, a definição de bolhas da seção anterior tem um problema: ela é muito sensível a pequenas variações na imagem F . Nesta seção, indicamos como resolver tal problema utilizando uma hierarquia multi-escala. As demonstrações das propriedades que citaremos a seguir é parte da belíssima Teoria das Singularidades – não as apresentaremos neste livro, mas o leitor interessado em um esboço pode consultar [28]; esperamos apresentar os aspectos relevantes de tal teoria aos espaços de escala num outro trabalho.



Um ruído mata uma bolha potencial criando duas pequenas

Tentaremos criar bolhas multi-escala já que o processo de difusão que cria os espaços de escala é razoavelmente imune à presença de ruídos, pois estes tendem a corresponder a bolhas menos importantes que devem desaparecer rapidamente com a escala. Intuitivamente, uma bolha existente numa determinada escala sobreviverá em geral a uma pequena mudança de escala (talvez movendo ou deformando um pouco no processo). A idéia é ligar bolhas correspondentes em diferentes escalas para criar um objeto tetrádico – a bolha multi-escala².



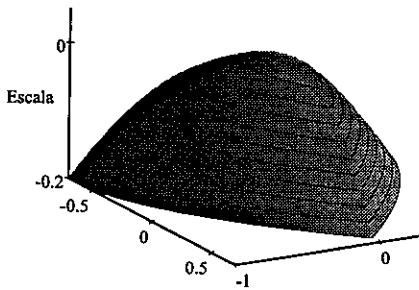
Ligando bolhas para criar bolhas multi-escala

No entanto, há exceções, isto é, situações onde uma bolha muda catastroficamente quando passa por uma determinada escala. Chamamos esses eventos de **eventos de bolha**. Genericamente, apenas quatro situações podem ocorrer (a evolução dos *suportes* das bolhas é ilustrado nas figuras a seguir para dois deles, os outros dois correspondem a inverter o eixo de escala):

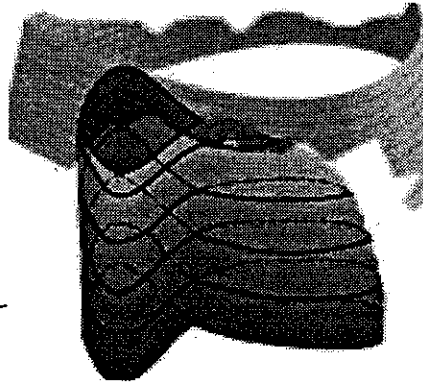
- **Destruição de uma bolha:** é comum que uma bolha diminua de tamanho à medida que a escala cresce e desapareça; acontece quando um máximo e um ponto de sela (não-compartilhado) de F colidem e desaparecem.

²Note que, nas figuras onde mostramos ligações entre bolhas para formar a bolha multi-escala, representamos apenas os suportes das bolhas em cada escala; assim, é possível ver uma espécie de "corte tridimensional" da bolha multi-escala.

- **Fusão de bolhas:** às vezes, duas bolhas com um ponto de sela comum desaparecem para dar lugar a uma bolha muito maior (na verdade, uma das bolhas do par desaparece e a outra cresce de maneira descontínua); acontece quando um máximo e um ponto de sela compartilhado colidem e desaparecem;
- **Criação de bolhas:** pode ser que uma pequena bolha apareça “do nada”; esta situação corresponde à geração de um máximo e um ponto sela que se separam a partir de um ponto comum.
- **Divisão de bolhas:** uma bolha pode desaparecer para dar lugar à duas bolhas menores que compartilham um ponto de sela; esta situação corresponde à criação de um par máximo-sela onde o ponto sela é compartilhado e já foi esboçada no exemplo 5.13.



Evento destruição



Evento fusão

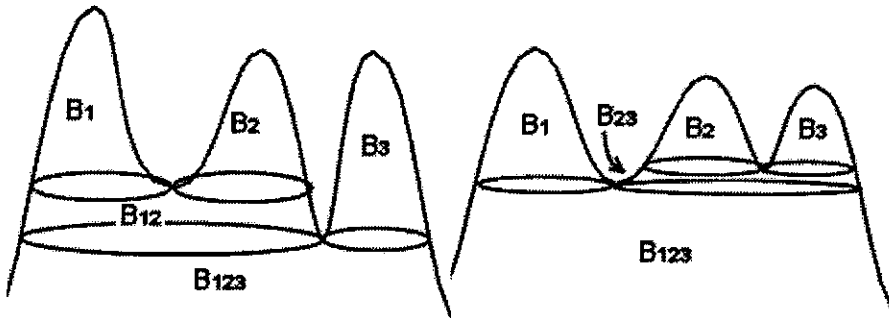
Pode-se mostrar que, no momento da transição, o ponto de máximo e o ponto sela correspondem a um único ponto da imagem onde a Hessiana é degenerada (veja a seção 15.2 a seguir); isto é, na escala t do evento de bolha e no ponto P onde o máximo local e o ponto sela colidem (ou surgem):

$$\frac{\partial F_t}{\partial x}(P) = \frac{\partial F_t}{\partial y}(P) = 0$$

$$\left[\frac{\partial^2 F_t}{\partial x^2} \frac{\partial^2 F_t}{\partial y^2} - \left(\frac{\partial^2 F_t}{\partial x \partial y} \right)^2 \right]_{\text{em } P} = 0$$

Em outras palavras, bolhas em escalas “sucessivas” podem ser unidas no processo de criação da bolha multi-escala enquanto as condições acima não forem satisfeitas. Por outro lado, escolhemos não ligar bolhas através dos eventos acima na formação de bolhas multi-escala; assim, uma bolha multi-escala será limitada em escala por dois tais eventos. O tempo de vida de uma bolha multi-escala será definido como a diferença entre a escala onde ela surge e a escala onde ela desaparece.

Os mesmos eventos se aplicam a árvores de bolhas. É possível se mostrar que somente as bolhas das folhas produzem eventos (que são os mesmos indicados acima). A única outra mudança significativa na estrutura das árvores de bolhas através das escalas é uma ocasional reordenação, que ocorre quando os valores de F_t em dois pontos de sela próximos mudam de posição relativa:



Pontos sela trocam de posição...

...reordenando a árvore de bolhas.

14.3 Importância de bolhas multi-escala

A estrutura descrita na seção anterior gera uma miríade de bolhas multi-escala, especialmente em escalas pequenas. Para que tal estrutura seja útil, devemos determinar uma medida de importância nas nossas bolhas multi-escala (para que possamos eliminar estruturas indesejáveis antes de passá-las para módulos mais “inteligentes” do nosso sistema de visão). Tal medida será certamente dependente da aplicação em questão, mas nesta seção apresentaremos algumas opções básicas.

Basicamente, queremos uma medida de importância baseada no tamanho espacial de uma bolha multi-escala, em seu contraste e em seu tempo de vida. A idéia mais simples é tomar o **volume tetradimensional da bolha multi-escala B** baseado nestas quantidades

$$\int_{t_0}^{t_1} \int_{\text{Suporte}(B(t))} \int_{z_{\text{base}}(B(t))}^{F_t(x,y)} dz dx dy dt = \int_{t_0}^{t_1} V_B(t) dt$$

onde t_0 e t_1 são as escalas onde B surge e desaparece, e $B(t)$ é a seção da bolha multi-escala B na escala t . No entanto, isto assume que as unidades de escala e de espaço são equivalentes e as mesmas para qualquer escala. Será que uma bolha que tem volume tridimensional 1 durante 2 unidades de tempo é tão importante quanto uma que tenha volume 2 durante 1 unidade de tempo? Será que o tempo de vida de 2 unidades de tempo é igualmente importante em escalas grandes e pequenas? Enfim, qual a métrica a ser utilizada no cálculo do volume da bolha multi-escala?

14.3.1 Tempo de vida efetivo de uma bolha multi-escala

Em primeiro lugar, considere o uso do tempo de vida de uma bolha como parte da medida de sua importância. Infelizmente, pequenos ruídos que se encontrem numa região da imagem onde haja uma gradual variação de cinza (digamos, praticamente linear) tendem a sobreviver um grande tempo sem que pareçam importantes per se. Mais ainda, de certa maneira, eventos de bolha são raros *demais* em escalas maiores, favorecendo muito a estabilidade das bolhas lá existentes – mais do que o desejado. Aliás, esta é exatamente a razão pela qual um espaço de escala é usualmente calculado em escalas diádicas (isto é, $t = 2^k$) inicialmente.

Se parece natural utilizar escalas diádicas ou exponenciais, por que não usar então uma escala efetiva dada pelo **logaritmo do tempo de vida**? De fato, pode-se mostrar que uma bolha encontrada em escala t tem comprimento característico aproximadamente proporcional a $\ln t$. Assim, se vamos utilizar o volume de uma bolha mantendo as unidades das coordenadas espaciais, parece natural mudar a variável escala para $\tau = \ln t$ (vide [12]). Este método é análogo a considerar uma equação do calor acelerada pela mudança de variáveis $t = e^\tau$:

$$\frac{\partial F_\tau}{\partial \tau} = \frac{\partial F_t}{\partial t} e^\tau = e^\tau \nabla^2 F_\tau \Rightarrow F_\tau = G_{e^\tau} * f_0$$

O tempo de vida efetivo e a importância de uma bolha multi-escala criada em escala t_0 e destruída em escala t_1 seriam então

$$\begin{aligned} \tau_1 - \tau_0 &= \ln t_1 - \ln t_0 = \ln \frac{t_1}{t_0} \\ I_{\text{Ef}}(B) &= \int_{\tau_0}^{\tau_1} V_B(\tau) d\tau = \int_{t_0}^{t_1} \frac{V_B(t)}{t} dt \end{aligned}$$

Esta idéia funciona no caso contínuo mas gera um desastre quando da discretização do espaço de escala. Afinal, considere bolhas que existam em escalas bem pequenas, isto é, imagine que $t_0 \rightarrow 0$. No caso contínuo, um sinal genérico terá cada vez mais bolhas que se recombinaem cada vez mais rápido (à medida que $t_0 \rightarrow 0$) e, essencialmente, $t_1 \rightarrow 0$ e $V_B(t) \rightarrow 0$ também; assim, há esperança de que (mesmo que o tempo de vida efetivo seja infinito) pelo menos a integral de $I_{\text{Ef}}(B)$ convirja, e bolhas de pequena escala não necessariamente assumem importâncias massivas.

No entanto, no caso discreto, há bolhas que existem em $t_0 = 0$ e que não somem imediatamente; o tamanho de tais bolhas é de pelo menos um pixel, e seu contraste não é necessariamente pequeno; a importância associada a cada uma dessas bolhas será infinita – definitivamente não o comportamento que desejamos.

Enfim, o problema de determinar a importância relativa de diferentes bolhas é complicado, ainda mais no caso discreto. A abordagem de [28] é assumir que a taxa relativa de desaparecimento de bolhas deva ser constante à medida que a escala varia. Por exemplo, é possível mostrar que se o sinal original

é um ruído branco unidimensional não discreto proveniente de um processo estacionário normal, então tomar a escala efetiva como $\tau = \ln t$ satisfaz esse critério. Se o sinal original é um ruído branco discreto, [29] mostra que

$$\tau = -\ln \left(1 + \frac{2}{\pi} \arcsin \left(1 - \frac{1}{2t} \frac{P_i[0]}{P_i[1]} - 1 \right) \right) \quad (14.1)$$

seria a escala efetiva a ser tomada com este critério. Não temos conhecimento de cálculos semelhantes no caso multidimensional, mas resultados empíricos apontam para $\tau = \ln t$ como uma boa medida de escala efetiva exceto para escalas pequenas.

14.3.2 Volume efetivo de uma bolha

Do mesmo modo que mudamos o parâmetro escala para encontrar uma escala efetiva, teremos de mudar a métrica espacial de acordo com a escala (se todas as escalas devem ser tratadas igualmente). Bolhas em escalas maiores tendem a ter volume maior. De fato, resultados experimentais mostram que as médias da área do suporte e do contraste de uma bolha B_P satisfazem

$$\begin{aligned} A_{\text{Suporte}}(t) &\approx t \\ F(P) - z_{\text{base}}(P) &\approx \frac{1}{\sqrt{t}} \end{aligned}$$

(por exemplo, analise o comportamento de uma bolha Gaussiana simples). A distribuição de probabilidade do volume tridimensional de uma bolha tem média e desvio padrão dados por

$$\begin{aligned} \mu_{V_B}(t) &\approx \sqrt{t} \\ \sigma_{V_B}(t) &\approx \sqrt{t} \end{aligned}$$

Isto sugere que o *volume efetivo* de uma bolha (em uma escala t) seja renormalizado como

$$\tilde{V}_{B_P}(t) = \frac{V_{B_P}(t) - \sqrt{t}}{\sqrt{t}}$$

para que os volumes de bolhas em diferentes escalas sejam tratados igualmente. Infelizmente, a expressão acima pode apresentar valores negativos; para evitar isto, [28] sugere o uso de

$$V_{\text{Ef}}(t) = \begin{cases} 1 + \tilde{V}_{B_P}(t), & \text{se } \tilde{V}_{B_P} \geq 0 \\ e^{\tilde{V}_{B_P}}, & \text{caso contrário} \end{cases}$$

Neste caso, a bolha “média” tem volume efetivo 1 enquanto bolhas menores tem volume efetivo decaindo para zero.

14.3.3 Conclusão

Resumindo a discussão desta seção, uma possibilidade para o cálculo de *importância* ou *volume efetivo de uma bolha multi-escala* seria

$$V_B = \int_{t_0}^{t_1} V_{\text{Ef}}(t) \frac{dt}{t}$$

onde

$$V_{\text{Ef}}(t) = \begin{cases} \frac{V_{B_P}(t)}{\sqrt{t}}, & \text{se } V_{B_P}(t) \geq \sqrt{t} \\ \exp\left(\frac{V_{B_P}(t)}{\sqrt{t}} - 1\right), & \text{caso contrário} \end{cases}$$

e $V_{B_P}(t)$ é o volume tridimensional da bolha B_P que é a seção da bolha multi-escala B na escala t . Esta expressão seria usada ao invés do volume tetradimensional normal, tanto em sua versão contínua como na discreta, e dá resultados satisfatórios para t_0 não muito pequeno. No caso discreto unidimensional, o ajuste de tempo efetivo dado pela equação 14.1 é teoricamente uma expressão que também funciona em escalas pequenas. No caso multidimensional discreto, não temos uma boa resposta que funcione também para pequenas escalas – este é ainda um problema em aberto.

14.4 Exercícios

1) Encontre as trajetórias dos pontos críticos (em função da escala t) dos seguintes espaços de escala unidimensionais, classificando e localizando no espaço e em escala os eventos (de bolha) que ocorrem:

- $F_t(x) = x^3 + 6xt$;
- $F_t(x) = x^4 + 20x^3t + 60t^2$;
- $F_t(x) = Ae^{-t} \cos x + Be^{-4t} \cos 2x$;

2) Se $F_t(x)$ é um espaço de escala unidimensional, $H_t(x, y) = F_t(x) \pm y^2 \pm 2t$ é um espaço de escala bidimensional. Qual a correspondência entre os pontos críticos de H_t e F_t ? Dê posição relativa e classificação (que dependerá, entre outras coisas, do sinal de y^2).

Capítulo 15

Singularidades*

O objetivo deste capítulo é introduzir as ferramentas básicas da Teoria das Singularidades e utilizá-las para mostrar a existência de bolhas e classificar os eventos de bolhas mencionados no capítulo anterior. Este capítulo é bastante matemático e pode ser omitido se o leitor decidir simplesmente “ter fé” naqueles resultados. Por outro lado, acreditamos que esta é uma ótima oportunidade para aprender a usar estas ferramentas teóricas. De fato, vários resultados interessantes podem ser obtidos apenas do Teorema da Função Implícita – é a partir dele que começamos nossa discussão.

O leitor que gostar desta introdução e desejar ver a “estória completa” deve consultar [37].

15.1 O Teorema da Função Implícita

O teorema básico a ser usado no resto deste capítulo é

Teorema 15.1 (Teorema da Função Implícita) *Seja $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^p$ uma função suave. A superfície de nível $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}_0, \mathbf{y}_0)$ é localmente um gráfico $\mathbf{y} = \tilde{\mathbf{y}}(\mathbf{x})$ (com $\mathbf{y}_0 = \tilde{\mathbf{y}}(\mathbf{x}_0)$ e $\tilde{\mathbf{y}}$ suave) sempre que a aproximação linear correspondente $Df|_{(\mathbf{x}_0, \mathbf{y}_0)}(\mathbf{x}, \mathbf{y}) = 0$ for um gráfico $\mathbf{y} = \mathbf{y}(\mathbf{x})$.*

Comentário 15.2 *Uma função é dita suave quando ela for infinitamente diferenciável. Note que $\mathbf{x} \in \mathbb{R}^m$ e $\mathbf{y} \in \mathbb{R}^n$ podem ter várias componentes. O Teorema da Função Implícita pode ser escrito sob condições mais fracas, mas isto será suficiente para nossas aplicações. Se escrevermos as p componentes de f explicitamente como*

$$(f_1(x_1, \dots, x_m, y_1, \dots, y_n), f_2(x_1, \dots, x_m, y_1, \dots, y_n), \dots, f_p(x_1, \dots, x_m, y_1, \dots, y_n))$$

então Df no ponto $(\mathbf{x}_0, \mathbf{y}_0)$ é a transformação linear dada pela matriz $p \times$

$(m+n)$ abaixo

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0, y_0) & \frac{\partial f_1}{\partial x_2}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial y_n}(x_0, y_0) \\ \frac{\partial f_2}{\partial x_1}(x_0, y_0) & \frac{\partial f_2}{\partial x_2}(x_0, y_0) & \cdots & \frac{\partial f_2}{\partial y_n}(x_0, y_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x_0, y_0) & \frac{\partial f_p}{\partial x_2}(x_0, y_0) & \cdots & \frac{\partial f_p}{\partial y_n}(x_0, y_0) \end{pmatrix}$$

Exemplo 15.3 Seja $f(x, y) = x^2 + y^2 - 1$. Note que $Df|_{(x_0, y_0)} = (2x_0, 2y_0)$ e portanto $Df(x, y) = 2xx_0 + 2yy_0 = 0$ pode ser resolvido para y ($y = -\frac{xx_0}{y_0}$) se e somente se $y_0 \neq 0$. Assim, podemos resolver $f(x, y) = f(x_0, y_0)$ para y em função de x sempre que $y_0 \neq 0$. Este fato poderia, é claro, ser verificado sem o uso do teorema; afinal

$$\begin{aligned} x^2 + y^2 - 1 &= x_0^2 + y_0^2 - 1 \Rightarrow \\ &\Rightarrow y(x) = \pm \sqrt{x_0^2 + y_0^2 - x^2} \Rightarrow \\ &\Rightarrow \frac{dy}{dx} = \pm \frac{-x}{\sqrt{x_0^2 + y_0^2 - x^2}} \end{aligned}$$

mostra que $y(x)$ é uma função diferenciável em $x = x_0$ a menos que $y_0 = 0$. Note que a escolha do sinal depende do sinal de y_0 , já que queremos $y(x_0) = \pm \sqrt{y_0^2} = y_0$ - no entanto, uma vez escolhido este sinal, a definição de $y(x)$ é única.

Exemplo 15.4 O TFI não é do tipo "se e somente se". Por exemplo, seja $f(x, y) = x^3 - y^9$. Note que $Df|_{(x_0, y_0)}(x, y) = 3x_0^2x - 9y_0^8y = 0$ não pode ser escrito na forma $y = y(x)$ quando $x_0 = y_0 = 0$. Assim, não podemos aplicar o TFI e não sabemos a priori se é possível ou não escrever $f(x, y) = 0 \Rightarrow y = \tilde{y}(x)$. No entanto, vê-se diretamente que a função suave $\tilde{y}(x) = x^3$ é a única solução de $f(x, y) = 0$ perto de $(0, 0)$.

Exemplo 15.5 Trace uma figura para convencer-se de que uma curva de nível de uma função de duas variáveis $f(x, y)$ pode ser escrita localmente como $y = \tilde{y}(x)$ com $\tilde{y}(x_0) = y_0$ sempre que $f_y(x_0, y_0) \neq 0$. Para provar este fato, calcule a aproximação linear de f no ponto em questão:

$$\begin{aligned} Df|_{(x_0, y_0)}(x, y) &= f_x(x_0, y_0)x + f_y(x_0, y_0)y = 0 \Rightarrow \\ &\Rightarrow y = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}x \end{aligned}$$

sempre que $f_y \neq 0$; então, aplique o Teorema da Função Implícita.

15.2 Trajetórias de Pontos Críticos

Nesta seção, analisaremos o comportamento de um ponto crítico e uma imagem F à medida que a escala muda. Em primeiro lugar, exibimos uma definição um pouco mais formal de "ponto crítico" e "ponto crítico não-degenerado".

Definição 15.6 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$; dizemos que $\mathbf{x}_0 \in \mathbb{R}^n$ é um ponto crítico de f sempre que $\nabla f = \vec{0}$ em \mathbf{x}_0 . Dizemos que \mathbf{x}_0 é não-degenerado se a Hessiana de f*

$$Hf = \begin{pmatrix} f_{x_1x_1} & f_{x_2x_1} & \cdots \\ f_{x_1x_2} & f_{x_2x_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

for não singular em \mathbf{x}_0 , isto é, se $\det Hf \neq 0$.

Agora estamos prontos para acompanhar a trajetória de um ponto crítico de $F_t(x, y)$ enquanto ele não é degenerado. A demonstração a seguir chamará de $P(x_1(t), y_1(t))$ a trajetória do ponto crítico de uma função $F_t(x, y)$ que satisfaz a equação do calor. Também gostaríamos de obter uma expressão para a velocidade de P em função de F_t .

Teorema 15.7 *Se \mathbf{x}_0 é um ponto crítico não-degenerado de um espaço de escala $F_{t_0}(\mathbf{x})$ (na escala $t = t_0 > 0$) então há um caminho suave $(\mathbf{x}_1(t), t)$ ($t_1 < t < t_2$) tal que $\mathbf{x}_1(t)$ é ponto crítico não-degenerado de $F_t(\mathbf{x})$ na escala t e $\mathbf{x}_1(t_0) = \mathbf{x}_0$. Mais ainda, tem-se*

$$\dot{\mathbf{x}}_1(t) = - (H)^{-1} (\nabla (\nabla^2 F_t))$$

onde H é a matriz Hessiana de F_t no ponto $\mathbf{x}_1(t)$ e o vetor gradiente (calculado apenas nas coordenadas espaciais de \mathbf{x}) do Laplaciano $\nabla (\nabla^2 F_t)$ também é calculado em $\mathbf{x}_1(t)$.

Demonstração. É uma aplicação direta do TFI. Vamos demonstrá-lo no caso bidimensional, a demonstração geral sendo completamente análoga (apenas com notação mais feia, como se vê acima). Escreveremos (x, y) ao invés de \mathbf{x} . Defina

$$f(x, y, t) = \nabla F_t(x, y) = (\partial_x F_t(x, y), \partial_y F_t(x, y))$$

Queremos escrever a curva de nível $f(x, y, t) = f(x_0, y_0, t_0) = 0$ como $x = x_1(t)$ e $y = y_1(t)$. Para tanto, analisamos

$$Df = \begin{pmatrix} \partial_{xx} F_t & \partial_{xy} F_t & \partial_{xt} F_t \\ \partial_{xy} F_t & \partial_{yy} F_t & \partial_{yt} F_t \end{pmatrix}$$

Como (x_0, y_0) é ponto crítico não-singular de F_t , o determinante menor 2×2 à esquerda de Df é não-nulo. Portanto $Df|_{(x_0, y_0)}(x, y) = 0$ é um sistema linear que pode ser facilmente resolvido como $x = x(t)$ e $y = y(t)$. O TFI garante então que o caminho $(x_1(t), y_1(t))$ é suave e satisfaz as condições desejadas.

Para a estimativa da velocidade, note que

$$\begin{aligned} \partial_x F_t(x_1(t), y_1(t), t) &\equiv 0 \\ \partial_y F_t(x_1(t), y_1(t), t) &\equiv 0 \end{aligned}$$

(esta é, de fato, a definição do ponto P). Diferenciando com relação a t (e omitindo os argumentos)

$$\begin{aligned}\partial_{xx}F_t \dot{x} + \partial_{xy}F_t \dot{y} + \partial_{xt}F_t &= 0 \\ \partial_{xy}F_t \dot{x} + \partial_{yy}F_t \dot{y} + \partial_{yt}F_t &= 0\end{aligned}$$

isto é

$$(HF_t) \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = -\nabla(\partial_t F_t) \Rightarrow \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = -(HF_t)^{-1} \cdot \nabla(\partial_t F_t)$$

Note que até aqui todo o raciocínio acima poderia ser feito não só para a equação do calor mas para qualquer evolução de F_t ! Somente agora é que usaríamos a definição do espaço de escala $\partial_t F_t = \nabla^2 F_t$ para finalizar o resultado. ■

Comentário 15.8 (Existência de Bolhas Multi-Escala) *Como cada bolha é gerada por um ponto de máximo local, concluímos que bolhas continuam existindo em escalas sucessivas enquanto tal máximo local existir. Note que um ponto crítico não pode mudar de tipo sem “degenerar” primeiro! De fato, um máximo local tem $\det HF_t > 0$ e $\partial_{xx}F_t < 0$; para que ele se torne um ponto de sela, devemos ter $\det HF_t < 0$ e portanto ele deve passar por um momento degenerado onde $\det HF_t = 0$; para que ele se torne um mínimo, devemos ter $\partial_{xx}F_t > 0$ e ele deve passar antes por um ponto onde $\partial_{xx}F_t = 0 \Rightarrow \det HF_t < 0$ (que seria um ponto de sela). Assim, o teorema acima garante que bolhas em escalas sucessivas podem ser ligadas para formar bolhas multi-escala enquanto o ponto crítico que gera a bolha for não-degenerado. Os únicos eventos de bolha ocorrerão quando $\det HF_t = 0$ no tal ponto.*

Note que a velocidade do ponto crítico pode aumentar violentamente à medida que nos aproximamos de um evento de bolha (onde HF_t é singular, isto é, $\det H$ se aproxima de 0)! Isto sugere um **cuidado extra ao seguir bolhas perto de eventos** – as bolhas podem se mover, crescer ou diminuir muito rapidamente!

No caso da equação do calor unidimensional o raciocínio acima leva a

$$v_P = -\frac{F'''(P)}{F''(P)}$$

para pontos críticos P de $F_t(x)$.

Também vale a pena citar que o método acima pode ser estendido para outros detalhes da imagem. Por exemplo, arestas de sinais unidimensionais são freqüentemente definidas como pontos críticos de $F'(x)$. Usando o espaço de escala de $F'(x)$ na expressão acima, concluímos que um ponto P de aresta de um sinal F tem velocidade

$$v_P = -\frac{F''''(P)}{F'''(P)}$$

isto é, esta é a velocidade de deslocamento de arestas à medida que a variância da Gaussiana utilizada na suavização cresce. Esta expressão reforça ainda mais que não é possível simplesmente borrar a imagem e encontrar detalhes numa “escala” maior; como estes se movem com a suavização (possivelmente com velocidades imensas), é necessário seguir os detalhes de volta a escalas menores para obtermos sua localização correta na imagem original.

15.3 Classificação de singularidades

Definição 15.9 *Sejam U, V dois abertos em \mathbb{R}^n . Uma função $f : U \rightarrow \mathbb{R}^n$ com $f(U) = V$ é um difeomorfismo quando*

i) f é suave;

ii) f tem uma inversa $g : V \rightarrow \mathbb{R}^n$, isto é, $f \cdot g = \mathbb{I}_V$ e $g \cdot f = \mathbb{I}_U$;

iii) g é suave.

Definição 15.10 *Um difeomorfismo local em um ponto \mathbf{x} é uma função cuja restrição a uma vizinhança de \mathbf{x} é um difeomorfismo. Um difeomorfismo local é freqüentemente chamado de troca de coordenadas.*

Comentário 15.11 *O Teorema da Função Implícita tem como consequência o Teorema da Função Inversa: se $f : U \rightarrow \mathbb{R}^m$ é suave, $\mathbf{x} \in U$ e $Df|_{\mathbf{x}}$ é não-singular, então f é um difeomorfismo local em \mathbf{x} . A demonstração deste fato é deixada como exercício.*

Definição 15.12 *Duas funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (em $\mathbf{x}_0 \in \mathbb{R}^n$) e $g : \mathbb{R}^n \rightarrow \mathbb{R}$ (em $\mathbf{y}_0 \in \mathbb{R}^n$) são ditas (localmente) equivalentes se elas forem idênticas a menos de uma constante e de uma troca de coordenadas (local) em \mathbb{R}^n , isto é,*

$$f \sim g \Leftrightarrow f(\mathbf{x}) = g(H(\mathbf{x})) + C$$

onde C é uma constante e $H(\mathbf{x})$ é um difeomorfismo com $H(\mathbf{x}_0) = \mathbf{y}_0$.

Não é difícil ver que a relação acima é de fato uma relação de equivalência. De fato:

- Tome $H(\mathbf{x}) = \mathbf{x}$ e $C = 0$. Então $g(H(\mathbf{x})) + C = g(\mathbf{x})$ e concluímos que $g \sim g$ para qualquer g (em qualquer ponto \mathbf{x}_0);
- Se $f(\mathbf{x}) = g(H(\mathbf{x})) + C$, então

$$f(H^{-1}(\mathbf{x})) - C = (g(H(H^{-1}(\mathbf{x}))) + C) - C = g(\mathbf{x})$$

note que, H^{-1} é um difeomorfismo sempre que H o for e $H^{-1}(\mathbf{y}_0) = \mathbf{x}_0$; assim $f \sim g \Leftrightarrow g \sim f$.

- Se $f(\mathbf{x}) = g_1(H_1(\mathbf{x})) + C_1$ e $g_1(\mathbf{x}) = g_2(H_2(\mathbf{x})) + C_2$ então claramente $f(\mathbf{x}) = g_2(H_2(H_1(\mathbf{x}))) + C_2 + C_1$; é fácil ver que $H = H_2 \cdot H_1$ é um difeomorfismo. Assim $f \sim g_1$ e $g_1 \sim g_2 \Rightarrow f \sim g_2$.

Note que os gráficos de f e g são difeomorfos; de fato,

$$F : (\mathbf{x}, \alpha) \mapsto (H(\mathbf{x}), \alpha + C)$$

leva o gráfico $(\mathbf{x}, f(\mathbf{x}))$ no gráfico $(\mathbf{y}, g(\mathbf{y}))$.

A relação de equivalência acima nos permite classificar as funções (em cada um de seus pontos). Vamos fazê-lo a seguir:

15.3.1 Caso 1D

No caso 1D, podemos mostrar que quase toda função $f(x)$ no ponto $x_0 \in \mathbb{R}$ é equivalente a um monômio $g(x)$ no ponto $x = 0$, onde:

- $g(x) = 0$ se f é constante numa vizinhança de x_0 ;
- $g(x) = x$ se x_0 não é um ponto crítico (isto é, $f'(x_0) \neq 0$);
- $g(x) = \pm x^2$ se x_0 é um ponto crítico não-degenerado (isto é, $f'(x_0) = 0 \neq f''(x_0)$); o sinal corresponde ao sinal de $f''(x_0)$;
- $g(x) = x^3$ se x_0 é um ponto crítico degenerado com $f'(x_0) = f''(x_0) = 0 \neq f'''(x_0)$;
- $g(x) = \pm x^4$ se $f'(x_0) = f''(x_0) = f'''(x_0) = 0 \neq f^{iv}(x_0)$; o sinal é o sinal de $f^{iv}(x_0)$;
- etc etc etc

Esta classificação não diz nada a respeito de pontos onde *todas* as derivadas de f são 0 mas f não é constante (como é o caso de $f(x) = e^{-1/x^2}$ em $x = 0$). Estas funções pertencerão a outras classes.

Demonstração. Translações de coordenadas fazem com que $x_0 = 0$ e $f(x_0) = 0$. O caso em que f é constante é trivial; assim, seja $f^{(k)}(0)$ a primeira derivada não-nula de f em 0. Assuma $f^{(k)}(0) > 0$. Aplicando a regra de L'Hôpital repetidas vezes

$$\lim_{x \rightarrow 0} \frac{f(x)}{x^k} = \lim_{x \rightarrow 0} \frac{f'(x)}{kx^{k-1}} = \dots = \lim_{x \rightarrow 0} \frac{f^{(k)}(x)}{k!} = \frac{f^{(k)}(0)}{k!} > 0$$

já que todos os limites intermediários são da forma $\frac{0}{0}$. Assim, $f(x) \geq 0$ numa vizinhança de $x = 0$. Tome

$$H(x) = (f(x))^{\frac{1}{k}}$$

Note que

$$H'(0) = \lim_{x \rightarrow 0} \frac{(f(x))^{\frac{1}{k}}}{x} = \left(\lim_{x \rightarrow 0} \frac{f(x)}{x^k} \right)^{\frac{1}{k}} = \left(\frac{f^{(k)}(0)}{k!} \right)^{\frac{1}{k}} \neq 0$$

e, pelo Teorema da Função Inversa, H é um difeomorfismo local. Mas

$$f(x) = (H(x))^k$$

e, portanto, $f \sim x^k$. Caso $f^{(k)} < 0$, a demonstração acima mostra que $-f \sim x^k$, isto é, $f \sim -x^k$. Mas para k ímpar é fácil ver que $x^k \sim -x^k$ através do difeomorfismo $H(x) = -x$. ■

15.3.2 Caso ND

A classificação acima é surpreendentemente mais complicada para n variáveis. No entanto, para uma função $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ainda podemos dizer que:

- Se x_0 é um ponto não-crítico de $f(x)$, então $f \sim g$ onde $g(x_1, x_2, \dots, x_n) = x_1$;
- Se x_0 é um ponto crítico não-degenerado de $f(x)$, então $f \sim g$ onde g é da forma $g(x_1, x_2, \dots, x_n) = x_1^2 + x_2^2 + \dots + x_i^2 - x_{i+1}^2 - x_{i+2}^2 - \dots - x_n^2$ (i é um número entre 0 e n);

Note que a situação é bem mais complicada para pontos críticos degenerados. Por exemplo, $f(x, y) = x^2y + y^9$ não é equivalente a x^2y em $(0, 0)$, como podemos ver comparando $f(x, y) = 0 \Leftrightarrow y = 0$ com $x^2y = 0 \Leftrightarrow x = 0$ ou $y = 0$.

15.4 Genericidade de Singularidades

Na seção anterior, classificamos todas as singularidades que vêm de pontos não-degenerados, e mencionamos que o caso degenerado é bem mais complicado. No entanto, casos degenerados são “raros” ou “instáveis” e não deveriam ter muita importância. Nesta seção, procuraremos definir o que queremos dizer por “estabilidade” de uma singularidade.

Em geral, dizemos que uma situação ou objeto é estável (em se tratando de uma propriedade) se pequenas alterações da situação ainda possuem a tal propriedade. No nosso caso, os objetos são funções suaves e a propriedade em questão é o tipo de singularidade que ela possui.

Definição 15.13 (Estabilidade Estrutural) *Uma função suave $f: \mathbb{R}^n \rightarrow \mathbb{R}$ é dita (estruturalmente) estável num ponto crítico x_0 se, para qualquer função suave g pequena (no sentido de que todas suas derivadas parciais são pequenas), tem-se $f \sim f + g$.*

Em outras palavras, $f(x)$ será estável se, dada qualquer família a um parâmetro $F: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ suave, denotada $F_s(x) = F(x, s)$ com $F_0 = f$, tem-se

$$F_s \sim F_0 = f \text{ para } |s| < \varepsilon$$

Na prática, é mais fácil provar que f é instável encontrando uma família F_s tal que $F_s \not\sim f$ para s arbitrariamente pequeno.

Exemplo 15.14 A função $f(x) = x^2$ tem um ponto crítico não-degenerado em $x = 0$. Seja $F_s(x) = x^2 + tx$. Note que todas as funções da família F_s tem um tal ponto crítico não-degenerado:

$$F'_s(x) = 0 \rightarrow 2x + t = 0 \rightarrow x = -\frac{t}{2}$$

$$F''_s(x) = 2$$

Será que a singularidade em $x = 0$ de x^2 é estável? Para mostrar este fato, considere uma família qualquer $F_s(x)$ com $F_0(x) = x^2$. Esperamos que $F'_s(x) = 0$ possa ser resolvido como $x = x(s)$ para s pequeno (o que é garantido pelo TFI já que $DF'_s|_{(0,0)}(x, y) = 2x + Ks$; apenas K depende da família); por outro lado, como $F''_0(0) = 2 > 0$, espera-se que $F''_s(x(s)) > 0$ pelo menos para s pequeno. Assim, o mínimo local de x^2 em $x = 0$ é estável.

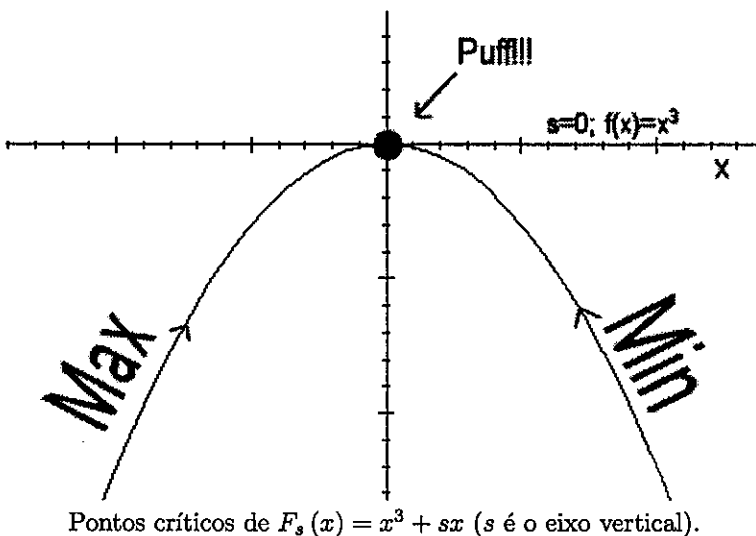
Exemplo 15.15 Por outro lado, o ponto crítico $x = 0$ de $f(x) = x^3$ é instável. De fato, tome $F_s(x) = x^3 + sx$. Note que os pontos críticos satisfazem

$$3x^2 + s = 0 \Leftrightarrow x = \pm\sqrt{\frac{-s}{3}}$$

e só existem para $s < 0$. Note que

$$F''_s = 6x$$

e portanto $x = -\sqrt{-s/3}$ é um máximo local e $x = \sqrt{-s/3}$ é um mínimo local para $s < 0$. Assim, à medida que s cresce, o mínimo e o máximo ficam próximos, colidem (em $s = 0$) e desaparecem.



Para funções, um ponto crítico não-degenerado é estável, enquanto um ponto crítico degenerado é instável.

De fato, se $F_s(x)$ é uma família de funções de n variáveis, então encontrar um ponto crítico de $F_s(x)$ é resolver as n equações dadas por $\nabla F_s(x) = 0$. Como x tem n componentes, o teorema da função implícita nos garante que podemos resolver $x = x(s)$ (isto é, resolver para as n variáveis x_1, x_2, \dots, x_n em função de s); a condição que o TFI necessita equivale exatamente a pedir que a Hessiana de F_s seja não-degenerada! Assim, as funções “vizinhas” $F_s(x)$ terão $x = x(s)$ também como ponto crítico não-degenerado.

Por outro lado, um ponto crítico degenerado é um que satisfaz, além de $\nabla F_s(x) = 0$, $\det HF_s(x) = 0$; tipicamente, estas $n + 1$ equações a $n + 1$ variáveis (incluindo s aqui como variável) levam a soluções isolados do tipo $(x, s) = (x_0, s_0)$. Se $(x_0, 0)$ é um destes pontos, não esperamos encontrar pontos próximos a ele que também sejam degenerados.

Exemplo 15.16 $F_s(x, y) = x^2 + y^2 + sx$. Note que $F_0(x, y) = x^2 + y^2$ tem um mínimo em $(0, 0)$ que é um ponto crítico não-degenerado ($\det HF_s = 4$ é constante e diferente de 0). É fácil ver que o ponto crítico de F_s é $(x(s), y(s)) = (-\frac{s}{2}, 0)$, comprovando que todos os F_s têm um ponto de mínimo não degenerado, que se move ao longo do eixo x .

15.5 Famílias de Funções

Para famílias de funções, a propriedade de que apenas pontos críticos não-degenerados são estáveis não vale mais. Por exemplo, dada uma família de funções $F_t(x)$, a condição $F'(x) = F''(x) = 0$ é agora estável (note que isto é diferente de um ponto crítico degenerado de $F(x, t)$ já que há uma direção privilegiada, x , e uma direção ignorada, t). De fato, crie uma família a um parâmetro de tais famílias, digamos $F_{s,t}(x)$. Um ponto crítico degenerado é obtido resolvendo

$$F'_{s,t}(x) = 0$$

$$F''_{s,t}(x) = 0$$

o que nos permite encontrar $x = x(s)$ e $t = t(s)$. Assim, famílias próximas à família original $F_t(x)$ também terão pontos críticos degenerados! O ponto crítico degenerado move-se na família à medida que a modificamos (mudamos s), mas sempre existirá!

Por outro lado, como uma direção é privilegiada, devemos também criar uma definição de equivalência de famílias (não queremos usar $F_t(x) \sim G_t(x) \Leftrightarrow F(x, t) \sim G(x, t)$ já que isto permite a “mistura” das variáveis x e t). Temos:

Definição 15.17 Duas famílias de funções $f_t : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ (em $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$) e $g_u : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ (em $(y_0, u_0) \in \mathbb{R}^n \times \mathbb{R}$) são ditas (localmente) equivalentes se elas forem equivalentes para valores correspondentes de seus parâmetros (correspondência essa dada por um difeomorfismo entre t e u) de maneira que

o difeomorfismo e a constante de cada equivalência dependam suavemente do parâmetro; em outras palavras

$$f_t \sim g_u \Leftrightarrow f_t(\mathbf{x}) = g_{e(t)}(H_t(\mathbf{x})) + C(t)$$

onde $e(t)$ é um difeomorfismo local com $e(t_0) = u_0$; $C(t)$ é suave; e $H_t(\mathbf{x})$ é um difeomorfismo local para cada t , depende suavemente de \mathbf{x} e t e satisfaz $H_{t_0}(x_0) = y_0$.

15.5.1 Teorema de Classificação de Thom

Sob esta definição, uma família de funções a um parâmetro $F_t(\mathbf{x})$ tipicamente tem

- Pontos não-singulares (onde $\nabla F_t \neq \mathbf{0}$): em tais pontos, tem-se $F_t \sim G_t$ onde $G_t(x_1, x_2, \dots, x_n) = x_1$;
- Pontos singulares não-degenerados ou pontos de Morse (onde $\nabla F_t = \mathbf{0}$ mas $\det HF_t \neq 0$): ali, $F_t \sim G_t$ onde $G_t(x_1, x_2, \dots, x_n) = \sum \pm x_i^2$;
- Pontos tipo “dobra” (onde $\nabla F_t = \det HF_t = 0$): ali, $F_t \sim G_t$ onde $G_t(x_1, x_2, \dots, x_n) = x_1^3 + 3x_1t + \sum_{i=2}^n \pm x_i^2$.

A título de curiosidade, o teorema mais geral neste sentido é o teorema de Thom para famílias de funções a r parâmetros. Ele diz que genericamente, uma família de funções suaves a r parâmetros ($r \leq 5$) $F_t(\mathbf{x}) : \mathbb{R}^m \times \mathbb{R}^r \rightarrow \mathbb{R}$ é estável e em cada ponto equivalente a uma das seguintes formas:

- Pontos não-singulares: $F_t \sim x_1$;
- Pontos de Morse: $F_t \sim \sum \pm x_i^2$;
- “Fold” ou A_2 (Dobras): $F_t \sim x_1^3 + 3x_1t_1 + (M)$;
- “Cusp” ou A_3 (Cúspide): $F_t \sim \pm (x_1^4 + t_2x_2^2 + t_1x_1) + (M)$;
- “Swallowtail” ou A_4 (Rabo de andorinha):
 $F_t \sim x_1^5 + t_3x_1^3 + t_2x_2^2 + t_1x_1 + (M)$;
- “Butterfly” ou A_5 (Borboleta):
 $F_t \sim \pm (x_1^6 + t_4x_1^4 + t_3x_1^3 + t_2x_1^2 + t_1x_1) + (M)$;
- “Wigwam” ou A_6 : $F_t \sim x_1^7 + t_5x_1^5 + t_4x_1^4 + t_3x_1^3 + t_2x_1^2 + t_1x_1 + (M)$;
- Elíptico umbílico ou D_4^- : $F_t \sim x_1^2x_2 - x_2^3 + t_3x_1^2 + t_2x_2 + t_1x_1 + (N)$;
- Hiperbólico umbílico ou D_4^+ : $F_t \sim x_1^2x_2 + x_2^3 + t_3x_1^2 + t_2x_2 + t_1x_1 + (N)$;
- Parabólico umbílico ou D_5 :
 $F_t \sim \pm (x_1^2x_2 + x_2^4 + t_4x_2^2 + t_3x_1^2 + t_2x_2 + t_1x_1) + (N)$;

- Segundo Elíptico Umbílico ou D_6^- :
 $F_t \sim x_1^2 x_2 - x_2^5 + t_5 x_2^3 + t_4 u_2^2 + t_3 x_1^2 + t_2 x_2 + t_1 x_1 + (N)$;
- Segundo Hiperbólico Umbílico ou D_6^+ :
 $F_t \sim x_1^2 x_2 + x_2^5 + t_5 x_2^3 + t_4 u_2^2 + t_3 x_1^2 + t_2 x_2 + t_1 x_1 + (N)$;
- Simbólico Umbílico ou E_6 :
 $F_t \sim \pm (x_1^3 + x_2^4 + t_5 x_1 x_2^2 + t_4 x_2^2 + t_3 x_1 x_2 + t_2 x_2 + t_1 x_1) + (N)$;

onde $(M) = \sum_{i=2}^n \pm x_i^2$ e $(N) = \sum_{i=3}^n \pm x_i^2$

Note que todas elas têm mais de um parâmetro exceto as que já havíamos mencionado. Para 6 ou mais parâmetros, não há estabilidade no sentido da equivalência que definimos através de difeomorfismos (apesar de existir estabilidade no sentido de tipos de pontos críticos).

15.6 Aplicação aos Espaços de Escala

Como espaços de escala são famílias de funções a um parâmetro, esperamos encontrar tipicamente apenas pontos não-singulares, pontos de Morse (que geram bolhas multi-escala) e pontos tipo dobra (que geram eventos de bolha). Os dois primeiros tipos já foram bem analisados; para os pontos dobra, analisemos

$$G_t(x) = x^3 + 3xt$$

$$G_t(x, y) = x^3 + 3xt \pm y^2$$

nos casos 1D e 2D respectivamente. As singularidades vêm de

$$\partial_x G_t = 3x^2 + 3t = 0 \rightarrow x = \pm\sqrt{-t}$$

$$\partial_y G_t = 2y = 0 \rightarrow y = 0$$

sendo a segunda equação necessária apenas no caso 2D. Para verificar os tipos de singularidades que colidem, calculamos

$$\partial_{xx} G_t = 6x$$

$$\partial_{xy} G_t = 0; \partial_{yy} G_t = \pm 2$$

assim, no caso 1D temos um máximo $x = -\sqrt{-t}$ e um mínimo $x = \sqrt{-t}$ que colidem e desaparecem após $t = 0$. No caso 2D, $\det HG_t = \pm 12x$ e podemos ter máximo+sela ou sela+mínimo colidindo e desaparecendo.

Como esta é a única forma genérica em que um evento pode ocorrer, elas definem a única maneira na qual uma bolha pode sumir ou desaparecer quando a escala aumenta. A única questão agora é saber se a sela que participa do evento pertence a alguma outra bolha além daquela determinada pelo ponto extremo que desaparece. Caso haja uma segunda bolha, temos o evento fusão; se não houver outra bolha envolvida, temos uma destruição. Finalmente, a princípio não sabemos se na equivalência $F_t \sim G_t$ os parâmetros de F e G crescem na

mesma direção (isto depende do sinal de $e'(t)$ na definição de equivalência de famílias); se $e'(0) < 0$, a direção dos eventos no espaço de escala F_t é a inversa da direção analisada com $G_t = x^3 + 3xt$ ($\pm y^2$); assim, a princípio, podemos ter também os eventos **divisão** ou **criação**.

Teorema 15.18 *No caso 1D, a criação de pontos críticos não pode ocorrer.*

Demonstração. De fato, se um espaço de escala num ponto (x_0, t_0) é equivalente a $G_t(x) = x^3 + 3xt$ em $(0, 0)$, então

$$x^3 + 3xt = F(a(x, t), e(t)) + c(t)$$

onde colocamos a dependência do parâmetro dentro de F para evitar erros nas derivações. Escreveremos F_a e F_t para as derivadas parciais de F com relação à primeira e segunda variáveis. Queremos calcular $e'(0)$, que aparece tomando a derivada mista¹

$$\partial_{xt} : 3 = F_{aa}a_xa_t + F_{at}a_xe_t$$

Precisamos portanto calcular a_x para poder obter e_t em função de F . Tome derivadas de ordem até 3 em relação a x :

$$\partial_x : 3x^2 + 3t = F_a a_x$$

$$\partial_{xx} : 6x = F_{aa}a_x^2 + F_a a_{xx}$$

$$\partial_{xxx} : 6 = F_{aaa}a_x^3 + 3F_{aa}a_{xx}a_x + F_a a_{xxx}$$

No ponto $(x, t) = (0, 0)$ sabemos que $a(0, 0) = x_0$ e $e(0) = t_0$, então ali

$$0 = F_a a_x$$

$$0 = F_{aa}a_x^2 + F_a a_{xx}$$

$$6 = F_{aaa}a_x^3 + 3F_{aa}a_{xx}a_x + F_a a_{xxx}$$

(derivadas de F em (x_0, t_0) ; de a em $(0, 0)$ e de e em $t = 0$). Então, não é difícil ver que $F_a = F_{aa} = 0$ (em (x_0, t_0)); isto já era esperado, pois queríamos que x_0 fosse um ponto crítico degenerado de F_{t_0} onde o evento de bolha acontece. Assim, podemos calcular

$$a_x^3 = \frac{6}{F_{aaa}(x_0, t_0)}$$

Substituindo na expressão que contém e_t obtemos

$$e_t = \frac{3}{F_{at}} \sqrt[3]{\frac{F_{aaa}}{6}}$$

¹ $e'(t)$ também aparece se tomarmos ∂_t , mas a expressão encontrada também envolve $c'(t)$ que não pode ser encontrado em outra derivada parcial.

Esta expressão vale (no ponto $(x, t) = (x_0, t_0)$) para qualquer família F_t que seja equivalente a $G_t = x^3 + 3xt$ em $(0, 0)$. No caso específico de espaços de escala, temos adicionalmente que $F_t = F_{aa}$, isto é

$$e'(0) = \frac{\sqrt[3]{36}}{2} (F_{aaa})^{-\frac{2}{3}}$$

e concluímos que $e'(0) > 0$, isto é, as direções da escala t em F_t e do parâmetro t de G_t sempre coincidem! Portanto, o único evento possível é a destruição de bolhas! ■

Teorema 15.19 *No caso 2D, a criação de bolhas pode ocorrer num ponto singular degenerado; usando o sistema pq de coordenadas, isto ocorrerá quando $F_p = F_q = F_{pp} = F_{pq} = 0$, e $F_{ppp}(F_{ppp} + F_{pqq}) < 0$.*

Demonstração. Começando de

$$x^3 + 3xt \pm y^2 \sim F(a(x, y, t), b(x, y, t), e(t)) + c(t)$$

calculamos primeiramente o gradiente

$$\partial_x : 3x^2 + 3t = F_a a_x + F_b b_x$$

$$\partial_y : \pm 2y = F_a a_y + F_b b_y$$

como ambas as expressões são 0 em $(x, y, t) = (0, 0, 0)$, concluímos que $F_a = F_b = 0$ no ponto (x_0, y_0, t_0) em questão (já que $a_x b_y - a_y b_x \neq 0$). Escrevendo as derivadas de segunda ordem

$$\partial_{xx} : 6x = F_{aa} a_x^2 + 2F_{ab} a_x b_x + F_{bb} b_x^2 + F_a a_{xx} + F_b b_{xx}$$

$$\partial_{xy} : 0 = F_{aa} a_x a_y + F_{ab} (a_x b_y + b_x a_y) + F_{bb} b_x b_y + F_a a_{xy} + F_b b_{xy}$$

$$\partial_{yy} : \pm 2 = F_{aa} a_y^2 + 2F_{ab} a_y b_y + F_{bb} b_y^2 + F_a a_{yy} + F_b b_{yy}$$

Para simplificar, vamos usar o sistema de coordenadas pq para que $F_{aa} = F_{ab} = 0$ (isto é sempre possível num ponto crítico degenerado, veja lista de exercícios). Assim, no ponto em questão

$$0 = F_{bb} b_x^2; \quad 0 = F_{bb} b_x b_y; \quad \pm 2 = F_{bb} b_y^2$$

e concluímos que $b_x = 0$ e $F_{bb} = \pm \frac{2}{b_y^2}$. Finalmente, escrevemos a terceira derivada com relação a x

$$\begin{aligned} \partial_{xxx} : 6 = F_{aaa} a_x^3 + 3F_{aab} a_x^2 b_x + 3F_{abb} a_x b_x^2 + F_{bbb} b_x^3 + \\ + 3F_{aa} a_x a_{xx} + 3F_{ab} (a_{xx} b_x + a_x b_{xx}) + 3F_{bb} b_x b_{xx} + \dots \end{aligned}$$

Substituindo os valores acima, vemos que todos os termos que envolvem derivadas de F em primeira e segunda ordem mais os termos com b_x desaparecem; ficamos apenas com

$$6 = F_{aaa} a_x^3 \Rightarrow a_x = \left(\frac{6}{F_{aaa}} \right)^{\frac{1}{3}}$$

exatamente como no caso 1D. A derivada mista torna-se

$$\begin{aligned} \partial_{xt} : 3 = F_{aa}a_x a_t + F_{ab}b_x a_t + F_{ab}a_x b_t + F_{bb}b_x b_t + \\ + F_{at}a_x e_t + F_{bt}b_x e_t + F_a a_{xt} + F_b b_{xt} \end{aligned}$$

e todos os termos contendo b_x e derivadas de F de ordem menor que 2 desaparecem, exceto

$$3 = F_{at}a_x e_t \Rightarrow e_t = \frac{3}{F_{at}} \left(\frac{F_{aaa}}{6} \right)^{\frac{1}{3}}$$

Na prática, o uso do sistema de coordenadas pq transformou o problema 2D no problema 1D. Colocando agora a equação do calor

$$e_t = \sqrt[3]{\frac{9F_{aaa}}{2(F_{aaa} + F_{abb})^3}}$$

Note que é possível ter $e_t < 0$ (quando o ponto crítico degenerado tiver $f_{aa} = f_{ab} = 0$ e f_{aaa} com sinal diferente de $f_{aaa} + f_{abb}$). ■

Exemplo 15.20 *O exemplo mais simples que cria extremos é*

$$\begin{aligned} f(a, b) &= -a^3 + 4ab^2 - 4b^2 \\ f_t(a, b) &= -a^3 + 4ab^2 - 4b^2 - 8t + 2at \end{aligned}$$

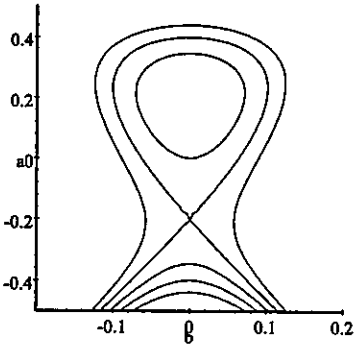
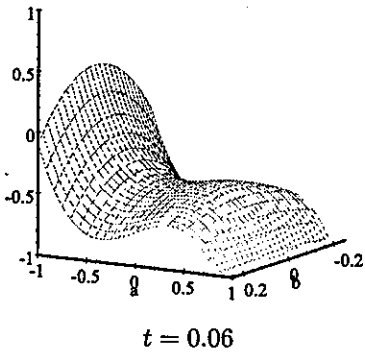
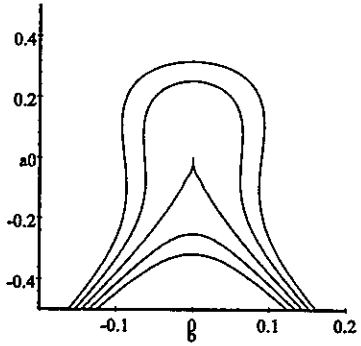
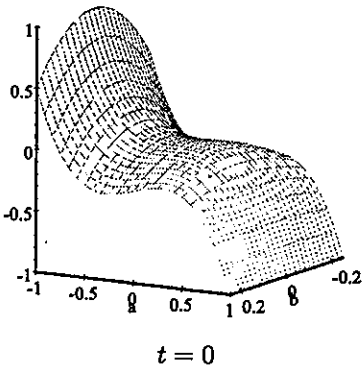
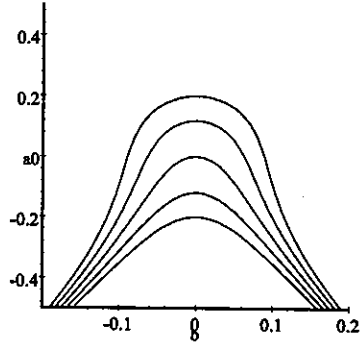
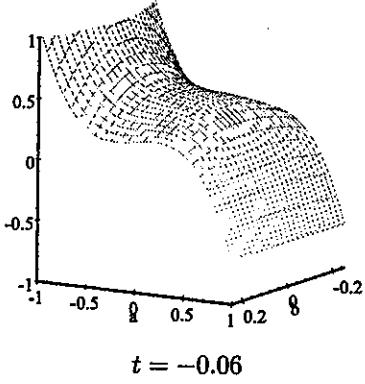
Note que f_t tem pontos críticos para

$$\begin{aligned} -3a^2 + 4b^2 + 2t &= 0 \\ 8ab - 8b &= 0 \end{aligned}$$

isto é, em $(a, b) = \left(\pm\sqrt{\frac{2t}{3}}, 0 \right)$ e $(a, b) = \left(1, \pm\sqrt{\frac{3-2t}{2}} \right)$. O determinante da Hessiana é

$$\det H f_t = (-6a)(8a - 8) - (8b)^2 = -48a^2 - 64b^2 + 48a$$

Assim, o primeiro par de pontos críticos é um par sela-máximo para t pequeno ($t < 1.5$ para ser exato); o segundo par são dois pontos sela longe da origem. Abaixo estão os gráficos de f_t perto da origem para $t = -0.06$, $t = 0$ e $t = 0.06$



15.7 Exercícios

1) Suponha que uma função $f(x)$ em $x = x_0$ é equivalente à função $g(x) = x^3$ em $x = 0$ (no sentido deste capítulo). Mostre que $f'(x_0) = f''(x_0) = 0$ mas $f'''(x_0) \neq 0$. É verdade que $f(x_0) = 0$?

2) Suponha que uma família de funções $F_t(x)$ em $(x, t) = (x_0, t_0)$ é equivalente a $G_t(x) = x^3 + 3xt$ em $(x, t) = (0, 0)$ (no sentido deste capítulo). Mostre que $\partial_x F_{t_0}(x_0) = \partial_{xx} F_{t_0}(x_0) = 0$ mas $\partial_{xxx} F_{t_0}(x_0) \neq 0$. É verdade que $\partial_t F_{t_0}(x_0) = 0$?

3) Considere uma mudança de coordenadas por rotação

$$R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ : (x, y) \rightarrow (x \cos \theta + y \sin \theta, y \cos \theta - x \sin \theta)$$

Seja $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$; defina $g = f \circ R_\theta$. Mostre que $\det(Hf) = 0 \iff \det(Hg) = 0$ (isto é, o sinal do determinante da Hessiana

$$Hf = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$$

é invariante por rotações) onde os determinantes das Hessianas são tomadas em pontos correspondentes. Em particular, mostre que se $\det(Hf) = 0$ num ponto então pode-se tomar um sistema de coordenadas p, q tal que $f_{pp} = f_{pq} = 0$ neste ponto. [Nota: pontos que determinam eventos de bolhas são em geral pontos críticos degenerados, e então $\det(Hf) = 0$ num deles; o exercício mostra que podemos supor $f_p = f_q = f_{pp} = f_{pq} = 0$ num desses pontos usando uma simples rotação de coordenadas.].

4) Mostramos que se a família $F_t(x, y)$ em (x_0, y_0, t_0) é equivalente à família $G_t(x, y) = x^3 + 3xt \pm y^2$, então o difeomorfismo de parâmetros $e(t)$ satisfaz

$$e_t(t_0) = \frac{3}{F_{pt}} \left(\frac{F_{ppp}}{6} \right)^{\frac{1}{3}}$$

Para montar um espaço de escala com não-criação de extremos, teríamos que garantir que $e_t(t_0) > 0$, isto é, gostaríamos que $F_{pt} = F_{ppp}$ (algo)². No espaço de escala Gaussiano

$$\partial_t F = F_{pp} + F_{qq} \Rightarrow e_t(t_0) = \left(\frac{9F_{ppp}}{2(F_{ppp} + F_{qqq})^3} \right)^{\frac{1}{3}}$$

e o evento "criação de bolhas" pode ocorrer quando $(F_{ppp} + F_{qqq})$ e F_{ppp} têm sinais contrários. Agora, defina um espaço de escala não-linear via

$$\partial_t F = \tilde{\kappa} = F_{xx} F_y^2 - 2F_{xy} F_x F_y + F_{yy} F_x^2$$

e calcule $F_{pt} = \partial_p \tilde{\kappa}$ num ponto P onde $F_p = F_{pq} = F_{pp} = 0$ (use o exercício 4 do capítulo sobre Invariantes Geométricos). Note que um ponto crítico degenerado P satisfaz estas condições... o que podemos afirmar sobre a destruição de extremos em pontos críticos degenerados? [Muito cuidado!]

Capítulo 16

Algoritmos

Neste capítulo, apresentamos mais explicitamente os algoritmos e estruturas de dados necessários para o cômputo das bolhas e das bolhas multi-escala. Descreveremos como obter uma bolha branca em fundo preto, correspondendo a um par máximo-sela. De fato, enquanto um ponto de máximo será a semente de uma bolha, o ponto de sela correspondente será classificado como “imediatamente” fora de uma bolha para forçar bolhas distintas a serem disjuntas (lembre-se, uma sela pode ser compartilhada por duas bolhas distintas). A partir daí, mostraremos como ligar bolhas em diferentes escalas para formam bolhas multi-escala.

No que segue abaixo, assumimos que $F_t[m, n]$ é o espaço de escala de um sinal $f_0[m, n]$. Os algoritmos aqui apresentados e a estrutura de dados abaixo são modificações de [28].

16.1 Detecção de bolhas

16.1.1 Bolha a partir de máximo local

O algoritmo abaixo funciona para imagens genéricas, onde não há dois pixels com o mesmo valor de cinza:

1. Dado um máximo local P numa imagem F_t , este ponto corresponderá a uma bolha. Assim, inicialize a sua bolha B (aproveite para armazenar seu ponto máximo $e(B) = P$ e seu valor máximo $F(P)$); inicie uma lista de vizinhos L_1 com os pixels vizinhos a P , ordenando-a de acordo com os valores de F_t . Vale a pena levantar um flag¹ associado a esta bolha marcando-a como “ativa” ou “em formação”.

2. Considere o pixel Q , o mais alto na lista de vizinhos L_1 (maior valor de F_t) e analise os seus vizinhos; alguns deles já estarão marcados como pertencentes à bolha em questão; se os outros estiverem todos abaixo de Q , ponha-os na

¹Este flag somente será útil no caso de processamento paralelo de bolhas; veja nota na subseção a seguir.

lista de vizinhos L_1 (mantenha a ordem), remova Q da lista, coloque-o na bolha “ativa” e volte ao passo 2.

3. Caso contrário, o ponto Q é o ponto de sela correspondente à bolha em questão; marque todos os pontos que restaram na lista de vizinhos L_1 como pontos de fundo, assim como os pontos vizinhos de Q que estão abaixo de Q (os outros vizinhos de Q podem fazer parte de uma outra bolha ou também serem fundo); o flag da bolha muda para “inativa” ou “completa”; aproveite para calcular o contraste da bolha como $F(P) - F(Q)$.

Caso consideremos imagens onde possam haver pixels com exatamente o mesmo valor de cinza, devemos agrupar os pixels em platôs conexas primeiro e então aplicar o algoritmo acima aos platôs (note que até mesmo a definição de máximo local depende deste agrupamento – caso P pertença a um platô constante, somente analisando os vizinhos deste platô é que podemos verificar se P é um máximo local ou não). De qualquer forma, tais situações são não-gênicas quando da aplicação da Equação do Calor.

16.1.2 Bolha a partir de um ponto qualquer

Se um ponto R qualquer é dado e queremos descobrir a que bolha ele pertence (se é que ele pertence a alguma bolha), podemos fazer o seguinte.

1. Inicialize uma lista de pontos L_2 com apenas R e seus vizinhos; ordene-os por valores de F_i ;

2. Seja S o ponto mais alto da lista. Enquanto S não for um máximo local nem um ponto previamente classificado, simplesmente adicione os seus vizinhos à lista L_2 e volte ao passo 2.

A aplicação sucessiva do passo 2 criará um caminho de pontos que rapidamente sobe na direção de um máximo local da imagem (gerando uma lista de pontos que inclui o tal caminho e seus vizinhos), possivelmente alcançando uma bolha antes de chegar ao máximo local (neste caso o processo do passo 2 é interrompido)².

3. Se S não foi classificado, S é um máximo local mas sua bolha ainda não foi calculada. Aplique o processo da subseção anterior para encontrar a bolha de S .

4. Agora, de uma maneira ou de outra, S já está classificado como pertencente a alguma bolha completa; os pontos de L_2 que ainda não foram classificados não pertencem à bolha de S e são portanto pontos de fundo; isto pode ou não incluir R .

16.1.3 Todas as bolhas de uma imagem

Enquanto podemos aplicar os algoritmos acima para encontrar todas as bolhas de uma imagem, há um algoritmo mais eficiente que exige apenas uma

²Novamente, o flag que decide se a bolha está ativa ou não só será útil no caso de processamento paralelo. Assim, o passo 2 pode ser interrompido quando S fizer parte de uma bolha “ativa”; neste caso, simplesmente espere que o processamento da bolha de S se complete antes de iniciar o passo 4.

“passagem” pela imagem para calcular todas as suas bolhas. Tal classificação será feita de cima para baixo (o leitor deve imaginar a maré baixando como na definição de bolhas no capítulo anterior).

1. Ordene todos os pixels da imagem de acordo com os valores de F , criando uma lista L .

2. Seja S o ponto mais alto da lista L ;

2a. Se S é um máximo local, inicialize uma bolha que contenha S ; levante um flag³ para marcar esta bolha como “ativa” ou “em formação”.

Caso contrário, olhe para os vizinhos de S que estão acima de S ; todos já foram classificados.

2b1. Se todos os vizinhos acima de S pertencem à mesma bolha em formação, S também pertence a esta bolha em formação; esta é a única maneira de S pertencer a uma bolha;

2b2. Caso contrário, S não pode pertencer a nenhuma de suas bolhas vizinhas (se é que há alguma); se alguma das regiões vizinhas é uma bolha “em formação”, mude-a para “bolha completa” e marque S como seu ponto de sela; outras regiões vizinhas podem incluir bolhas “completas” ou fundo, mas essas não interessam – de qualquer maneira, S será um ponto de fundo.

3. Remova S da lista L ; se a lista não está vazia, volte ao passo 2.

Novamente, se há possibilidade de que pixels distintos tenham o mesmo valor de cinza, devemos agrupá-los primeiro em platôs conexos e usá-los no algoritmo acima ao invés de pixels.

16.1.4 Árvores de bolhas

O algoritmo da última seção pode ser facilmente modificado para calcular árvores de bolhas; basta fazer com que um ponto de sela S não somente feche uma ou mais bolhas delimitadas por S mas também inicie uma nova semente para uma nova bolha em formação; aproveite para criar os arcos da árvore que ligam esta nova bolha em formação àquelas que acabam de se completar.

16.2 Detecção de bolhas multi-escala

Uma vez que bolhas tenham sido detectadas em cada escala, temos o problema de ligá-las dentro de uma bolha multi-escala. Devemos lembrar que (pelo menos no caso contínuo) há genericamente apenas 4 tipos de eventos que podem impedir uma tal ligação: destruição, criação, divisão ou fusão de bolhas. A idéia básica do algoritmo é a seguinte:

1. Analise o espaço de escala em escalas sucessivas t_1 e t_2 e encontre as bolhas em cada uma delas;

2a. Para cada bolha B_1 em escala t_1 , verifique todas as bolhas em t_2 que são candidatas a serem ligadas a B_1 .

³Cuidado! Neste caso, este flag é necessário mesmo em processamento simples, não-paralelo. O leitor não deve se deixar levar pelas duas últimas notas de pé de página.

2b. Para cada bolha B_2 em escala t_2 , verifique todas as bolhas em t_1 que são candidatas a serem ligadas a B_2 .

3a. Se B_1 é a única par candidata para B_2 e vice-versa, faça uma **ligação** dentro de uma mesma bolha multi-escala.

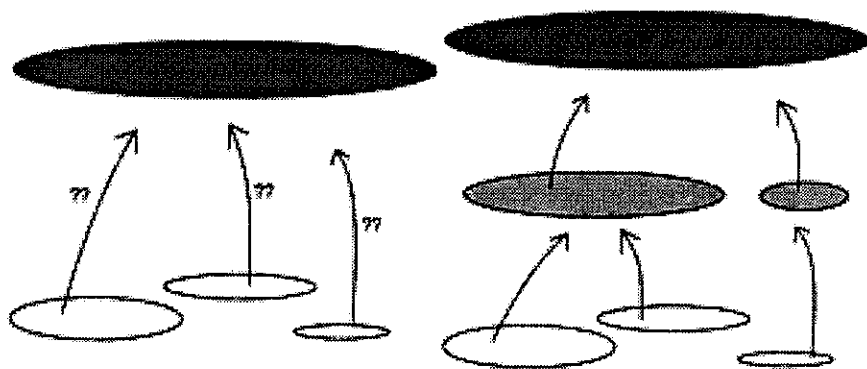
3b. Se B_1 não tem nenhuma par em escala t_2 , o evento **destruição** aconteceu. Feche a bolha multi-escala que contém B_1 por cima.

3c. Se B_2 não tem nenhuma par em escala t_1 , o evento **criação** aconteceu. Feche a bolha multi-escala que contém B_2 por baixo.

3d. Se B_1 tem duas pares candidatas em escala t_2 e cada uma delas tem apenas B_1 como par, o evento **divisão** aconteceu. Feche a bolha multi-escala que contém B_1 por cima e feche as duas outras bolhas multi-escala por baixo.

3e. Se B_2 tem duas pares candidatas em escala t_1 e cada uma delas tem apenas B_2 como par, o evento **fusão** aconteceu. Feche a bolha multi-escala que contém B_2 por baixo e feche as duas outras bolhas multi-escala por cima.

4. Caso contrário, pode ser que tenhamos um evento de bolha não genérico em nossas mãos. Mais provavelmente, o que acontece é simplesmente que há dois (ou mais) eventos genéricos acontecendo em rápida sucessão e não calculamos o espaço de escala entre os dois eventos. Assim, se isto acontecer, calcule o espaço de escala numa escala intermediária entre as duas consideradas e veja se isto separa os dois eventos; faça isto recursivamente até que os eventos se resolvam ou o passo de escala seja bem pequeno. Se mesmo com um passo de escala muito pequeno você foi incapaz de separar os eventos genéricos, desista – simplesmente ligue todas as bolhas envolvidas num **evento complexo**, separando as bolhas multi-escala através deste evento.



Para resolver este evento...

...precisamos calcular uma escala intermediária.

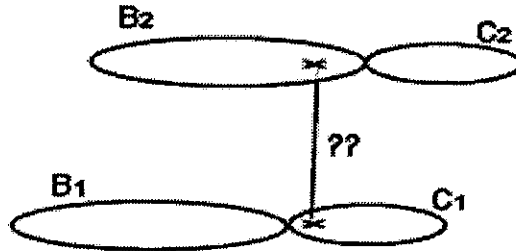
16.2.1 Como determinar pares candidatos?

Há várias maneiras de determinar se duas bolhas B_1 e B_2 calculadas em escalas sucessivas t_1 e t_2 são candidatas a pareamento ou não. Por exemplo,

poderíamos simplesmente verificar se os suportes de B_1 e B_2 se intersectam

$$B_1 \cap B_2 \neq \emptyset$$

No entanto, considere duas bolhas B_1 e C_1 em escala t_1 que tenham uma sela compartilhada e que se movem através das escalas; B_1 deve corresponder a uma bolha B_2 em escala ligeiramente maior t_2 , enquanto C_1 deve corresponder a uma bolha C_2 . No entanto, não importa quão fino seja o espaço de escala, existe um potencial para que C_1 sempre intersecte B_2 . Isto cria uma ambigüidade que provavelmente levará nosso algoritmo a classificar esta situação como um evento complexo ao invés de duas ligações simples! Nosso algoritmo falharia terrivelmente neste caso.



Sela compartilhada em movimento

Assim, consideraremos duas outras condições que ajudarão a determinar se B_1 e B_2 devem ser ligadas

$$e(B_1) \in B_2$$

$$e(B_2) \in B_1$$

Esta condição, mais o fato de que a situação acima ocorre exatamente quando as bolhas em questão têm uma sela compartilhada, devem ser levadas em consideração no algoritmo. Assim, modificamos o passo 3 do algoritmo acima para:

3a. Se B_1 intersecta somente B_2 e vice-versa (e o ponto extremo de cada uma delas pertence à outra, e ambas têm pontos de sela não compartilhados), faça uma **ligação** dentro de uma mesma bolha multi-escala.

3b. Se B_1 não intersecta ninguém em escala t_2 , o evento **destruição** aconteceu. Feche a bolha multi-escala que contém B_1 por cima.

3c. Se B_2 não intersecta ninguém em escala t_1 , o evento **criação** aconteceu. Feche a bolha multi-escala que contém B_2 por baixo.

3d. Se B_1 intersecta apenas B_2 e C_2 em escala t_2 , e cada uma delas intersecta apenas B_1 (e B_2 e C_2 compartilham um ponto de sela, e os extremos delas estão em B_1), o evento **divisão** aconteceu. Feche a bolha multi-escala que contém B_1 por cima e feche as duas outras bolhas multi-escala por baixo.

3e. Se B_2 intersecta apenas B_1 e C_1 em escala t_1 , e cada uma delas intersecta apenas B_2 (e B_1 e C_1 compartilham um ponto de sela, e os extremos delas estão em B_2), o evento **fusão** aconteceu. Feche a bolha multi-escala que contém B_2 por cima e feche as duas outras bolhas multi-escala por baixo.

3f. Se B_1 tem um ponto de sela compartilhado com C_1 e B_2 tem um ponto de sela compartilhado com C_2 ; $B_1 \cap B_2 \neq \emptyset$; $C_1 \cap C_2 \neq \emptyset$; $e(B_1) \in B_2$; $e(B_2) \in B_1$; $e(C_1) \in C_2$; $e(C_2) \in C_1$; e, finalmente, se as bolhas em questão não intersectam nenhuma outra bolha que não seja uma destas quatro; então temos uma **ligação dupla** : B_1 corresponde a B_2 enquanto C_1 corresponde a C_2 .

16.2.2 Problemas e refinamentos

Um dos problemas com o algoritmo acima é que bolhas muito pequenas, tipicamente do tamanho de um pixel, não serão conectadas à medida que se movem. De fato, assim que a nossa discretização registrar o movimento desta bolha, ela pulará de um pixel e o algoritmo acima registrará uma destruição e uma criação ao invés da ligação normal. É possível que uma bolha maior também se mova rapidamente o suficiente para que pareça ser bolhas distintas em escalas sucessivas.

Para evitar este fato, sempre que houver uma criação de uma bolha pequena (tipicamente, de um pixel de tamanho) B_2 (que é um fato possível mas que deveria ser um tanto raro), verifique se há alguma bolha B_1 na escala imediatamente anterior num raio de um pixel em torno de B_2 . Se houver, e se B_1 é uma candidata a destruição, troque estes dois eventos (destruição + criação) por uma simples ligação multi-escala.

Note também que não discutimos como exatamente começar o processo de ligação. Tipicamente, calcula-se o espaço de escala em passos constantes de escala *efetiva*, ou seja, tipicamente, em escalas diádicas ou exponencialmente espaçadas. A partir daí, escolha se você quer percorrer as escalas de cima para baixo ou de baixo para cima, ou mesmo iniciar em algum ponto no meio caso já haja uma idéia de que bolha multi-escala se procura.

Tipicamente, um número suficiente de refinamentos de passos de escala tende a resolver eventos complexos em eventos genéricos. No entanto, mantenha em mente que isto funcionará para imagens genéricas – imagens sintéticas sem ruído não são genéricas e tem um grande potencial de gerar eventos genuinamente complexos!

Finalmente, o refinamento em escala do passo 4 do algoritmo acima não precisa necessariamente calcular todas as bolhas na nova escala – basta classificar as bolhas que estejam na região do evento complexo que está sendo decifrado. Por este motivo é que apresentamos três algoritmos para o cálculo de bolhas, dois deles sendo locais.

16.3 Estrutura de Dados

A estrutura de dados abaixo é uma adaptação da encontrada em [28], mas não deve ser considerada como a última palavra no assunto. Em particular, uma “lista” de pixels como indicado abaixo é uma estrutura ineficiente que não se utiliza da correlação espacial dos pixels na definição de platôs, regiões etc. (os alunos do Visgraf fazem uma boa sugestão: usar “bounding boxes” com matrizes binárias para indicar relações de pertinência devem ser mais eficientes tanto no armazenamento quanto na procura!).

A única estrutura não definida abaixo é a estrutura *Imagem*, que pode ser qualquer estrutura de dados destinada a conter uma única imagem (possivelmente já suavizada).

- **Pixel**
 Propriedades:
Coordenadas (x e y)
Nível de cinza (valor de F ali)
 Ponteiros para:
Platô (que contém este pixel)
- **Platô** (pixels contíguos com exatamente o mesmo valor)
 Propriedades:
Área (possivelmente o número de pixels no platô)
Nível de cinza (valor comum de F no platô)
Criticalidade (se este platô é um mínimo local, máximo local, sela ou nada)
 Ponteiros para:
Região (que contém este platô)
Lista de pixels (no platô)
Lista de platôs vizinhos (vizinhos espaciais deste platô)
- **Região** (grupos de platôs que formarão o suporte de uma bolha ou o fundo)
 Propriedades:
Área (possivelmente o número de pixels)
Centro de gravidade
Momentos de segunda ordem (dando uma elipse que aproxima a região)
Caixa limitante (coordenadas de um retângulo que contenha a região)
 Ponteiros para:
Lista de platôs (na região)
Bolha (que contém esta região, se alguma)
Fundo (que contém esta região)
- **Fundo**
 Propriedades:
Polaridade (branca, preta);
Área (raramente o volume é importante)
 Ponteiros para:

Fatia de escala (imagem $F_t(\cdot)$ que contém este fundo);

Lista de regiões.(do fundo);

- **Bolha**

Propriedades:

Polaridade (branca, preta);

Atividade (em formação ou completa);

Volume

Ponteiros para:

Fatia de escala (imagem $F_t(\cdot)$ que contém esta bolha);

Platô crítico (a semente da bolha, seu máximo ou mínimo local);

Lista de selas (que limitam esta bolha);

Região suporte.(da bolha);

- **Sela**

Ponteiros para:

Platô (que corresponde a esta sela)

Lista de bolhas (das quais esta sela será limitante)

- **Fatia de escala** (a imagem $F_t(\cdot)$ em uma escala t e todas as suas estruturas)

Propriedades:

Valor da escala

Ponteiros para:

Imagem (o sinal bidimensional $F_t(\cdot)$, suavização de $f_0(\cdot)$)

Lista de bolhas brancas

Lista de bolhas pretas

Fundo branco

Fundo preto

Próxima fatia de escala (para a escala mais alta mais próxima)

Prévia fatia de escala (para a escala mais baixa mais próxima)

- **Bolha multi-escala**

Propriedades:

Polaridade (branca ou preta)

Importância (volume ou volume efetivo)

Ponteiros para:

Lista de bolhas (que compõem esta bolha m.e.)

Evento início

Evento fim

Bolha representante (uma de suas bolhas, escolhida numa escala representante; pode ser a bolha de maior volume efetivo)

- **Evento**

Propriedades:

Tipo (criação, destruição, fusão, divisão, complexo, escala mínima, escala máxima)

Posição (x , y e escala aproximados, ou intervalo de escalas, onde o evento

acontece)

Ponteiros para:

Lista de bolhas de cima (que se formam com este evento)

Lista de bolhas de baixo (que somem com este evento)

- **Esboço primal** (simplesmente junta a estrutura toda)

Ponteiros para:

Lista de fatias de escala

Lista de bolhas multi-escala brancas

Lista de bolhas multi-escala pretas

Lista de eventos para bolhas brancas

Listas de eventos para bolhas pretas

Parte V

Referências

Capítulo 17

Implementação de convoluções

Um dos detalhes mais irritantes dos espaços de escala (e em geral de Processamento de Sinais) é que a teoria está feita para sinais contínuos ou discretos infinitos enquanto os sinais do processamento computacional são em geral finitos e discretos. Neste apêndice, gostaríamos de citar as diferentes maneiras de lidar com este problema e esboçar os algoritmos que implementam as convoluções discretas correspondentes.

No que se segue, discutiremos o caso unidimensional (mais uma vez, o caso multidimensional é análogo). Assim, dado um sinal $f : I_N = \{0, 1, \dots, N - 1\} \rightarrow \mathbb{R}$ e um núcleo de convolução $T : \mathbb{R} \rightarrow \mathbb{R}$, como dar sentido à expressão “ $f * T$ ”? Como calculá-la, pelo menos nas posições originais de f , isto é, em I_N ? Como implementá-las?

17.0.1 Extensões de f

Podemos interpretar $f * T$ como $g * T$ onde $g : \mathbb{Z} \rightarrow \mathbb{R}$ é uma boa representação de f . A maneira mais comum de fazer isto é tomar $g \equiv f$ em I_N e então estender f de uma das seguintes formas:

- **Periodicamente:** simplesmente tome g periódica de período N

$$g[n] = f_{n \bmod N}$$

- **Por espelhamento:** para evitar descontinuidades nas bordas do sinal, concatene uma reflexão de f a uma de suas bordas e só então use a extensão

periódica

$$g[n] = \begin{cases} f_n & \text{para } 0 \leq n \leq N-1 \\ f_{(2N-1)-n} & \text{para } N \leq n \leq 2N-1 \\ g[n \bmod 2N] & \text{, caso contrário} \end{cases}$$

(bordas repetidas – período $2N$)

$$g[n] = \begin{cases} f_n & \text{para } 0 \leq n \leq N-1 \\ f_{(2N-2)-n} & \text{para } N \leq n \leq 2N-3 \\ g[n \bmod (2N-2)] & \text{, caso contrário} \end{cases}$$

(bordas não repetidas – período $2N-2$)

- **Por constante:** simplesmente preencha g com um valor constante independente de f (freqüentemente tomado como zero):

$$g[n] = \begin{cases} f_n & \text{para } 0 \leq n \leq N-1 \\ C (= 0) & \text{, caso contrário} \end{cases}$$

- **Por média:** preencha g com o valor da média de f

$$g[n] = \begin{cases} f_n & \text{para } 0 \leq n \leq N-1 \\ \mu_f = \frac{1}{N} \sum_{j=0}^{N-1} f_j & \text{, caso contrário} \end{cases}$$

Outras possibilidades incluem fazer com que haja um decaimento de f além de seus bordos em direção a algum valor constante (0 ou a média de f , em geral).

17.0.2 Algoritmos para convoluções

É freqüente que T seja um núcleo de convolução de suporte pequeno, digamos, com m elementos (como por exemplo a maioria dos núcleos utilizados para derivadas discretas). Neste caso, a melhor maneira é simplesmente usar a definição explícita da convolução¹

$$(g * T)[j] = \sum_{k=0}^{m-1} T_k g_{j-k}$$

que envolverá $o(m)$ operações para cada j ; em geral, só interessam os índices $0 \leq j \leq N-1$, e a convolução como um todo é calculada em $o(mN)$ operações. Note que apenas $N+m-1$ elementos de g são necessários para este cálculo.

Caso T tenha suporte grande (ou, como acontece freqüentemente, infinito), um dos seguintes métodos deve ser utilizado.

¹A expressão assume que T tem 0 como primeiro índice; caso contrário, basta transladar a resposta.

Extensão periódica

Se, por “sorte”, o núcleo de convolução T também for periódico de mesmo período que f (por exemplo, quando T for outro sinal ou imagem), podemos interpretar $f * T$ simplesmente como a convolução circular $f *_c T$ onde T é restrito a um de seus períodos. A implementação pode ser feita diretamente do somatório que define a convolução circular ou mais eficientemente com o auxílio da Transformada de Fourier Discreta (comentário 3.9)

$$(f *_c T)_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{n-1} f_k T_{k \bmod n} \quad o(N^2) \text{ operações}$$

$$f *_c T = \sqrt{N} \text{IFFT}(\text{FFT}(f) \cdot \text{FFT}(T)) \quad o(N \log N) \text{ operações}$$

Aliás, mesmo que T seja infinito e não periódico, $g * T$ será periódica e só precisa ser calculada em I_n . De fato, de acordo com a proposição 3.11, um período de $g * T$ é dado por

$$“f * T” = \text{IFFT} \left(\text{FFT}(f) \cdot \hat{T} \left(\frac{w}{N} \right) \right)$$

onde a multiplicação é tomada componente a componente, indexadas por w . Por exemplo, para o núcleo de Poisson simétrico

$$“f * P_t” = \text{IFFT} \left(e^{-4t \sin^2(\pi w/N)} \cdot \text{FFT}(f) \right) \quad o(N \log N) \text{ operações}$$

$$(f * P_t)_n = \frac{1}{N} \sum_{w,k=0}^{N-1} f_k \exp \left(\frac{2\pi i w (n-k)}{N} - 4t \sin^2 \left(\frac{\pi w}{N} \right) \right)$$

onde a última expressão permite os cálculos dos coeficientes em f_k a priori e é boa para processamento paralelo.

Extensão por espelhamento

Simplesmente use um dos períodos de g (de tamanho $2N$ ou $2N - 2$) no processo descrito na última subseção. Usando o algoritmo FFT para as Transformadas de Fourier Discretas, são realizadas $o(N \log N)$ operações.

Extensão por zeros

A convolução pode ser calculada diretamente através da definição da convolução de maneira semelhante à que usamos quando o núcleo era de suporte finito e pequeno

$$(g * T)[j] = \sum_{k=0}^{N-1} T_{j-k} f_k$$

Note que $g * T$ não tem suporte finito. Se desejarmos apenas os seus valores em I_N , usaremos $o(N^2)$ operações.

O cálculo de $g * T$ em I_N necessita dos valores de T_j apenas para $2N - 1$ valores de j , a saber, $-N + 1 \leq j \leq N - 1$. Assim, pode-se encarar $g * T$ como uma convolução entre dois núcleos de suporte finito e usar o comentário 3.10: seja \tilde{g} o sinal finito de tamanho $3N - 2$ obtido estendendo os N valores de f com zeros; seja \tilde{T} o sinal finito obtido estendendo os $2N - 1$ valores de T acima com zeros:

$$\tilde{g} = \left(f_0 \ f_1 \ f_2 \ \dots \ f_{N-1} \ \underbrace{0 \ 0 \ 0 \ \dots \ 0}_{2N-2 \text{ zeros}} \right); \quad \tilde{g}_0 = f_0$$

$$\tilde{T} = \left(T_{-N+1} \ T_{-N+2} \ \dots \ T_{N-1} \ \underbrace{0 \ 0 \ 0 \ \dots \ 0}_{N-1 \text{ zeros}} \right); \quad \tilde{T}_0 = T_{-N+1}$$

Então, $g * T$ e $\tilde{g} *_c \tilde{T}$ coincidem em N de seus valores, isto é

$$(g * T)[n] = \left(\tilde{g} *_c \tilde{T} \right)_{N-1+n} \quad \text{para } 0 \leq n \leq N - 1$$

$$g * T \approx \hat{g} *_c \tilde{T} = \sqrt{3N - 2} \text{IFFT} \left(\text{FFT}(\tilde{g}) \cdot \text{FFT}(\tilde{T}) \right)$$

que também pode ser feito em $o(N \log N)$ operações.

Extensão por média ou outro valor constante

Se g é obtido a partir de f usando uma extensão constante (digamos C , onde possivelmente $C = \mu_f$), então

$$g = h + C\mathbb{I}$$

onde h é obtida a partir de $f - C$ estendendo-o por zeros, e \mathbb{I} é a seqüência constante igual a 1:

$$h = (\dots 0 \ 0 \ 0 \ 0 \ f_0 - C \ f_1 - C \ f_2 - C \ \dots \ f_{N-1} - C \ 0 \ 0 \ 0 \ \dots)$$

$$\mathbb{I} = (\dots 1 \ 1 \ 1 \ 1 \ \dots)$$

Assim,

$$g * T = h * T + C\mathbb{I} * T = h * T + \left(C \sum_{k=-\infty}^{\infty} T_k \right) \mathbb{I}$$

onde $h * T$ pode ser calculado em $o(N \log N)$ operações aplicando o método da subseção anterior a $f - C$. Em particular, para núcleos normalizados (um caso comum), tem-se

$$"f * T" = g * T = (f - C) * T + C$$

Núcleos Recursivos

Existem núcleos de suporte infinito cuja convolução pode ser implementada de maneira recursiva. Por exemplo, o núcleo geométrico (normalizado)

$$T = (1 \ \beta \ \beta^2 \ \dots \ \beta^n \ \dots) \cdot (1 - \beta)$$

tem suporte infinito. Para obter

$$f_{out} = f_{in} * T$$

note que

$$\begin{aligned} \varphi_{out} = \varphi_{T * f_{in}} &= \varphi_T \varphi_{in} = \frac{1 - \beta}{1 - \beta z} \varphi_{in} \Rightarrow \\ \Rightarrow (1 - \beta z) \sum f_{out}[n] z^n &= (1 - \beta) \sum f_{in}[n] z^n \Rightarrow \\ \Rightarrow f_{out}[n] - \beta f_{out}[n - 1] &= (1 - \beta) f_{in}[n] \Rightarrow \\ \Rightarrow f_{out}[n] &= \beta f_{out}[n - 1] + (1 - \beta) f_{in}[n] \end{aligned}$$

nos dá uma **relação recursiva simples** para o cálculo de f_{out} !

Que outros núcleos possuem propriedades semelhantes? Em geral, podemos calcular $f_{out} = f_{in} * T$ a partir de uma relação recursiva da esquerda para a direita do tipo

$$f_{out}[n] = \sum_{j=1}^B b[j] f_{out}[n - j] + \sum_{j=-A}^A a[j] f_{in}[n - j] \quad (17.1)$$

se e somente se a função geradora do núcleo T em questão pode ser escrita como um quociente de polinômios²

$$\varphi_T = \frac{\sum_{-A}^A a_j z^j}{1 - \sum_1^B b_j z^j}$$

Definição 17.1 Um núcleo T é RECURSIVO quando $f_{out} = T * f_{in}$ corresponde à relação recursiva 17.1.

Que núcleos são recursivos? Se quebrarmos o quociente de polinômios de φ_T em frações parciais, esperamos encontrar

$$\varphi_T = \sum_j \frac{a_j z^{k(j)}}{1 - \beta_j z}$$

²Para ser exato, o numerador é uma série de potências finita. Note que uma relação recursiva da direita para a esquerda também pode ser escrita desta forma! Por exemplo, a relação recursiva $g_n = g_{n+1} + 3g_{n+2} + 6f_n + 3f_{n+1}$ pode ser reescrita como $g_{n+2} = -\frac{1}{3}g_{n+1} + \frac{1}{3}g_n - 2f_n - f_{n+1}$. O leitor deve se convencer de que a direção da recursão é irrelevante na análise que se segue.

ou seja, T é uma soma de núcleos semelhantes ao já apresentado núcleo geométrico $T_\beta = (1 \ \beta \ \beta^2 \ \dots \ \beta^n \ \dots)$, cada um possivelmente amplificado por $a/(1-\beta)$ e transladado k posições.

Ignore a amplificação e a translação por enquanto; note que podemos ter β complexo, digamos $\beta = re^{i\theta}$; neste caso, reescrevemos o núcleo acima como

$$T_\beta = (1 \ r \cos \theta + ir \sin \theta \ r^2 \cos 2\theta + ir^2 \sin 2\theta \dots)$$

As únicas combinações lineares de tais núcleos que produzem coeficientes reais correspondem às combinações lineares reais dos seguintes núcleos

$$T_c = \frac{T_\beta + T_{\bar{\beta}}}{2} = (1 \ r \cos \theta \ r^2 \cos 2\theta \ r^3 \cos 3\theta \ \dots) \Rightarrow \\ \Rightarrow T_c(n) = r^n \cos n\theta, \ n \geq 0$$

$$T_s = \frac{T_\beta - T_{\bar{\beta}}}{2i} = (0 \ r \sin \theta \ r^2 \sin 2\theta \ r^3 \sin 3\theta \ \dots) \Rightarrow \\ \Rightarrow T_s(n) = r^n \sin n\theta, \ n \geq 0$$

As Transformadas Z correspondentes são

$$\varphi_c = \frac{\varphi_\beta + \varphi_{\bar{\beta}}}{2} = \frac{1}{2} \left(\frac{1}{1-\beta z} + \frac{1}{1-\bar{\beta} z} \right) = \frac{1 - (r \cos \theta) z}{1 - (2r \cos \theta) z + r^2 z^2}$$

$$\varphi_s = \frac{\varphi_\beta - \varphi_{\bar{\beta}}}{2i} = \frac{1}{2i} \left(\frac{1}{1-\beta z} - \frac{1}{1-\bar{\beta} z} \right) = \frac{(r \sin \theta) z}{1 - (2r \cos \theta) z + r^2 z^2}$$

Resumindo, temos 3 núcleos recursivos básicos reais (aqui, normalizados) cujas funções geradoras, médias, variâncias e correspondentes relações recursivas para $f_{out} = T * f_{in}$ estão listadas abaixo:

- Geométrico

$$T = (\dots 0 \ 1 \ \beta \ \beta^2 \ \beta^3 \ \dots) \cdot (1 - \beta)$$

$$\varphi = \frac{1 - \beta}{1 - \beta z}; \ \mu = \frac{\beta}{1 - \beta}; \ \sigma^2 = \frac{\beta}{1 - \beta}$$

$$f_{out}[n] = \beta f_{out}[n-1] + (1 - \beta) f_{in}[n]$$

- Cosseno recursivo

$$T_c = (1 \ r \cos \theta \ r^2 \cos 2\theta \ r^3 \cos 3\theta \ \dots) \cdot \frac{1 - 2r \cos \theta + r^2}{1 - r \cos \theta}$$

$$\varphi_c = \frac{1 - (r \cos \theta) z}{1 - (2r \cos \theta) z + r^2 z^2} \cdot \frac{1 - 2r \cos \theta + r^2}{1 - r \cos \theta}$$

$$\mu_c = r \cdot \frac{(1 + r^2) \cos \theta - 2r}{(1 - 2r \cos \theta + r^2)^2}$$

$$f_{out}[n] = ((2r \cos \theta) f_{out}[n-1] - r^2 f_{out}[n-2]) + \frac{1 - 2r \cos \theta + r^2}{1 - r \cos \theta} (f_{in}[n] - (r \cos \theta) f_{in}[n-1])$$

- Seno recursivo

$$T_s = (r \sin \theta \quad r^2 \sin 2\theta \quad r^3 \sin 3\theta \quad \dots) \cdot \frac{1 - 2r \cos \theta + r^2}{r \sin \theta}$$

$$\varphi_s = \frac{1 - 2r \cos \theta + r^2}{1 - (2r \cos \theta)z + r^2 z^2}$$

$$\mu_s = \frac{1 - r^2}{1 - 2r(\cos \theta) + r^2}$$

$$f_{out}[n] = ((2r \cos \theta) f_{out}[n-1] - r^2 f_{out}[n-2]) + (1 - 2r \cos \theta + r^2) f_{in}[n]$$

Qualquer núcleo recursivo será uma soma de núcleos dos tipos descritos acima (possivelmente transladados e amplificados individualmente). Note que o método exige que se estabeleçam valores de f_{in} para iniciar a recursão, o que só pode ser feito diretamente se f_{in} é uma extensão por zeros ou pela constante C . Se T é normalizado, basta tomar os valores de f_{out} para iniciar a recursão como C ; caso contrário

$$f_{out}[n] = C \sum_{j=-\infty}^{\infty} T[j] = C \frac{\sum_j a_j}{1 + \sum_j b_j} \text{ para } n < 0$$

onde $b[j]$ e $a[j]$ são como em 17.1. Caso f_{in} seja um outro tipo de extensão, algoritmos mais sofisticados são necessários.

Na prática, o método recursivo providenciará uma maneira eficientíssima de calcular a convolução $f_{in} * T$, com número de operações por índice aproximadamente igual ao número de coeficientes a_j e b_j não-nulos.

Mesmo que o núcleo de convolução a ser usado não seja recursivo, podemos aproximá-lo por um núcleo recursivo e então usar este método. Por exemplo, de acordo com [7], uma amostragem da Gaussiana G_t pode ser bem aproximada utilizando uma das seguintes expressões

$$e^{-\frac{1}{4t}x^2} \approx e^{-0.891 \frac{|x|}{\sqrt{t}}} \left(0.9629 \cos \left(0.5974 \frac{|x|}{\sqrt{t}} \right) + 1.942 \sin \left(0.5974 \frac{|x|}{\sqrt{t}} \right) \right)$$

$$e^{-\frac{1}{4t}x^2} \approx 1.898 e^{-1.1 \frac{|x|}{\sqrt{t}}} - e^{-1.069 \frac{|x|}{\sqrt{t}}} \left(0.8929 \cos \left(1.043 \frac{|x|}{\sqrt{t}} \right) - 1.021 \sin \left(1.043 \frac{|x|}{\sqrt{t}} \right) \right)$$

$$e^{-\frac{1}{4t}x^2} \approx e^{-1.261\frac{|x|}{\sqrt{t}}} \left(1.68 \cos \left(0.4468 \frac{|x|}{\sqrt{t}} \right) + 3.735 \sin \left(0.4468 \frac{|x|}{\sqrt{t}} \right) \right) \\ - e^{-1.218\frac{|x|}{\sqrt{t}}} \left(0.6803 \cos \left(1.412 \frac{|x|}{\sqrt{t}} \right) + 0.2598 \sin \left(1.412 \frac{|x|}{\sqrt{t}} \right) \right)$$

A primeira permite que se escreva o núcleo Gaussiano G_t como a soma de um par de núcleos cosseno recursivos (possivelmente amplificado; $r = e^{-0.891/\sqrt{t}}$ e $\theta = 0.5974/\sqrt{t}$; o primeiro será para $x \geq 0$ e o outro será seu simétrico) com um par de seno recursivos (amplificado; $r = e^{-0.891\sqrt{t}}$ e $\theta = 0.5974/\sqrt{t}$), menos uma constante (para compensar pela dupla incidência no ponto $x = 0$); a segunda dá um par de cada um dos três tipos; a terceira dá dois pares de cossenos e dois pares de senos.

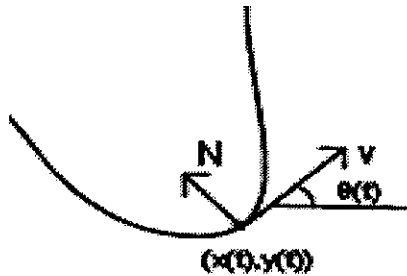
Capítulo 18

Curvatura de Curvas Planas

18.1 Expressão em coordenadas cartesianas

Dada uma curva plana parametrizada por $(x(t), y(t))$, o seu vetor velocidade $\vec{v} = \dot{x} \vec{i} + \dot{y} \vec{j}$ forma um certo ângulo $\theta(t)$ com a horizontal (eixo \overrightarrow{Ox}). Definimos a curvatura desta curva plana num determinado ponto como a taxa de variação deste ângulo por unidade de comprimento medida na curva, isto é,

$$v = |\vec{v}| = \frac{ds}{dt} = \sqrt{\dot{x}^2 + \dot{y}^2}$$
$$\tan \theta = \frac{\dot{y}}{\dot{x}} \text{ ou } \theta = \arctan \left(\frac{\dot{y}}{\dot{x}} \right)$$
$$\kappa = \frac{d\theta}{ds}$$



Curva, velocidade e o ângulo θ

Para calcular κ em função da parametrização, simplesmente use

$$\begin{aligned}\kappa &= \frac{d\theta}{dt} \cdot \frac{dt}{ds} = \frac{d\left(\arctan\left(\frac{\dot{y}}{\dot{x}}\right)\right)}{dt} \frac{1}{v} = \\ &= \frac{1}{1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2} \frac{\dot{y}\ddot{x} - \ddot{y}\dot{x}}{(\dot{x})^2} \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \Rightarrow \\ &\Rightarrow \kappa = \frac{\dot{y}\ddot{x} - \ddot{y}\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}\end{aligned}$$

Em particular, note que a definição de κ depende apenas do formato da curva e não da velocidade em que a percorremos, isto é, dada uma curva, κ não depende da parametrização escolhida¹.

Exemplo 18.1 *Uma possível parametrização de um círculo de raio R é*

$$x(t) = R \cos wt; \quad y(t) = R \sin wt$$

Neste caso, uma utilização direta da expressão para κ nos dá

$$\begin{aligned}\kappa &= \frac{(-Rw^2 \sin wt)(-Rw \sin wt) - (-Rw^2 \cos wt)(Rw \cos wt)}{\left((-Rw \sin wt)^2 + (Rw \cos wt)^2\right)^{3/2}} = \\ &= \frac{R^2 w^3}{R^3 w^3} = \frac{1}{R}\end{aligned}$$

confirmando a idéia intuitiva de que κ é maior quando a curva é mais fechada. Note que nem a curva nem κ dependem de w (que apenas muda a parametrização). Vale a pena notar que em geral κ é o inverso do “raio de curvatura” de uma curva.

18.2 Parametrização por comprimento de arco

Se parametrizamos a curva por comprimento de arco (isto é, $v = \sqrt{\dot{x}^2 + \dot{y}^2} = 1$), então o vetor velocidade é também o vetor tangente unitário \vec{T} ; se denotarmos o vetor unitário normal por \vec{N} , temos

$$\begin{aligned}\dot{x}^2 + \dot{y}^2 = 1 &\stackrel{\partial t}{\Rightarrow} \dot{x}\ddot{x} + \dot{y}\ddot{y} = 0 \\ \vec{T} = \vec{v} = \dot{x} \vec{i} + \dot{y} \vec{j} &\Rightarrow \frac{d\vec{T}}{dt} = \vec{a} = \ddot{x} \vec{i} + \ddot{y} \vec{j}\end{aligned}$$

¹Para ser exato, a curvatura depende somente da direção em que a curva é percorrida; se usarmos uma parametrização que reverte essa direção, a expressão da curvatura muda de sinal. É comum se exigir que tal parametrização siga o sentido anti-horário para curvas fechadas.

e portanto vemos que $\vec{v} = \vec{T}$ e $\vec{a} = \frac{d\vec{T}}{dt}$ são perpendiculares, isto é, $\frac{d\vec{T}}{dt}$ é um múltiplo de \vec{N} . Para encontrar o fator multiplicativo, note que

$$\left. \begin{aligned} \vec{N} &= -\dot{y} \vec{i} + \dot{x} \vec{j} \\ \kappa &= \dot{y}\dot{x} - \dot{x}\dot{y} \end{aligned} \right\} \Rightarrow \kappa \vec{N} = (-\dot{x}\dot{y}\dot{y} + \dot{x}\dot{y}^2) \vec{i} + (\dot{x}^2\dot{y} - \dot{x}\dot{x}\dot{y}) \vec{j} = \\ = (\dot{x}^2 + \dot{y}^2) \dot{x} \vec{i} + (\dot{x}^2 + \dot{y}^2) \dot{y} \vec{j} = \dot{x} \vec{i} + \dot{y} \vec{j} = \frac{d\vec{T}}{dt}$$

De maneira análoga, podemos encontrar também $\frac{d\vec{N}}{dt}$, obtendo as equações de Frenet para curvas planas parametrizadas por comprimento de arco

$$\begin{aligned} \frac{d\vec{T}}{dt} &= \kappa \vec{N} \\ \frac{d\vec{N}}{dt} &= -\kappa \vec{T} \end{aligned}$$

É comum definir-se curvatura a partir das expressões acima, ou como

$$\kappa = \left| \frac{d\vec{T}}{dt} \right|$$

18.3 Curvatura de curvas de nível

Dada uma função $F(x, y)$ que assuma valores reais, como calcular a curvatura da curva de nível $F(x, y) = C$ num determinado ponto regular² de seu domínio? Notemos que (utilizando índices para derivadas parciais) $\nabla F = F_x \vec{i} + F_y \vec{j}$ está na direção normal à curva de nível, portanto o vetor $\vec{T} = F_y \vec{i} - F_x \vec{j}$ é tangente à mesma. Assim, podemos encontrar uma parametrização $(x(t), y(t))$ com tal velocidade, isto é

$$\dot{x} = F_y(x(t), y(t)); \quad \dot{y} = -F_x(x(t), y(t))$$

Podemos então calcular

$$\begin{aligned} \ddot{x} &= F_{xy}\dot{x} + F_{yy}\dot{y} = F_{xy}F_y - F_{yy}F_x \\ \ddot{y} &= -F_{xx}\dot{x} - F_{xy}\dot{y} = -F_{xx}F_y + F_{xy}F_x \end{aligned}$$

onde omitimos da notação o ponto $(x(t), y(t))$ onde todas as derivadas parciais tem de ser calculadas.

Estamos prontos para calcular a curvatura de uma curva de nível de F num ponto $P = (x(t), y(t))$. De fato, basta utilizar a expressão da subseção anterior

$$\begin{aligned} \kappa &= \frac{\dot{y}\dot{x} - \dot{x}\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} = \frac{(-F_{xx}F_y + F_{xy}F_x)(-F_y) - (F_{xy}F_y - F_{yy}F_x)F_x}{(F_x^2 + F_y^2)^{3/2}} \Rightarrow \\ \Rightarrow \kappa &= \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)^{3/2}} \end{aligned}$$

²Isto é, um ponto não crítico ou um ponto P tal que $|\nabla F(P)| \neq 0$.

e temos uma expressão para κ em função de derivadas (cartesianas) de F .

Note que uma rotação ou uma translação dos eixos cartesianos não muda o formato das curvas de nível de F e, portanto, não muda a expressão de κ . Já que temos liberdade em escolher tais coordenadas, por que não escolher um sistema de coordenadas locais que simplifique as expressões encontradas? Para tanto, dada uma função $F(x, y)$ e um ponto regular $P(x_0, y_0)$, definiremos o **sistema de coordenadas gradiente** (denotado por u, v) colocando a origem no ponto P , o eixo $\vec{O}u$ tangente à curva de nível de F passando por P e a direção $\vec{O}v$ na direção do gradiente de F no ponto P .

Segue imediatamente que, no ponto P , temos $F_u = 0$ e $F_v = |\nabla F|$. Como a expressão de κ é a mesma neste sistema de coordenadas, podemos substituir $F_u = 0$ na expressão de κ para obter simplesmente³

$$\kappa = \frac{F_{uu}}{F_v}$$

18.4 Curvatura das curvas integrais do gradiente

Dada uma função $F(x, y)$, podemos considerar as curvas integrais do campo vetorial $|\nabla F|^4$. Dado um ponto $P(x_0, y_0)$, como calcular a curvatura da curva integral que passa por tal ponto?

De maneira análoga ao cálculo de κ , podemos imaginar uma parametrização de tais curvas cuja velocidade seja exatamente o vetor gradiente, isto é

$$\dot{x}(t) = F_x(x(t), y(t)); \quad \dot{y}(t) = F_y(x(t), y(t))$$

Daqui tiramos

$$\begin{aligned} \ddot{x} &= F_{xx}\dot{x} + F_{xy}\dot{y} = F_{xx}F_x + F_{xy}F_y \\ \ddot{y} &= F_{xy}\dot{x} + F_{yy}\dot{y} = F_{xy}F_x + F_{yy}F_y \end{aligned}$$

³É importante notar que as coordenadas u e v dependem do ponto P escolhido; podemos imaginar que a expressão de κ é válida para qualquer ponto do domínio de F se imaginarmos que as coordenadas u e v variam de acordo com o ponto escolhido. Note que $F_u = 0$ somente no ponto P ; portanto, não podemos escrever que $F_{uu} = 0$. Em outras palavras, ou pense que $F_u = 0$ é válido somente em P para u fixo (e fique à vontade para diferenciar qualquer coisa com relação a u), ou pense que $F_u = 0$ em todos os pontos P do domínio (mas então u varia com o ponto P escolhido, e não faz sentido diferenciar com respeito a u). O truque é manter ambas as interpretações em mente ao mesmo tempo...

⁴Intuitivamente, obtida "conectando" os vetores $|\nabla F|$ no plano. Esta curva também pode ser imaginada como tangente a $|\nabla F|$ (ou normal às curvas de nível de F) em todos os seus pontos.

e substituindo na expressão da curvatura a partir de uma parametrização

$$\begin{aligned}\mu &= \frac{(F_{xy}F_x + F_{yy}F_y)F_x - (F_{xx}F_x + F_{xy}F_y)F_y}{(F_x^2 + F_y^2)^{3/2}} \Rightarrow \\ \Rightarrow \mu &= \frac{(F_x^2 - F_y^2)F_{xy} + F_xF_y(F_{yy} - F_{xx})}{(F_x^2 + F_y^2)^{3/2}}\end{aligned}$$

que é a expressão de μ em função das derivadas parciais de F .

Novamente, uma rotação ou translação de eixos não pode alterar a forma das curvas integrais do gradiente de F nem, portanto, a expressão de μ . Então podemos usar o sistema gradiente de coordenadas (u, v) com $F_u = 0$ para simplificar a expressão de μ

$$\mu = -\frac{F_{uv}}{F_v}$$

Capítulo 19

Funções Modificadas de Bessel

Neste apêndice, mostramos alguma propriedades das funções $I_n(x)$ (e portanto do núcleo de Poisson $P_t[n] = e^{-2t} I_n(2t)$).

Definição 19.1 *Definimos a função modificada de Bessel de ordem n no ponto α como o coeficiente em z^n da expansão em série de potências de $e^{\frac{\alpha}{2}(z+\frac{1}{z})}$, isto é*

$$e^{\frac{\alpha}{2}(z+\frac{1}{z})} = \sum_{n=-\infty}^{\infty} I_n(\alpha) z^n$$

Proposição 19.2 $I_n(0) = \delta(n) = \begin{cases} 1, & \text{se } n = 0 \\ 0, & \text{se } n \neq 0 \end{cases}$

Demonstração. Basta tomar $\alpha = 0$ na expansão acima. Então

$$1 = \sum_{n=-\infty}^{\infty} I_n(0) z^n$$

e comparando as expressões temos o resultado. ■

Proposição 19.3

$$I_n(\alpha) = \sum_{k=0}^{\infty} \frac{(\alpha/2)^{2k+|n|}}{k! (k+|n|)!}$$

Demonstração. Basta usar as expansões de $e^{\frac{\alpha}{2}z}$ e de $e^{\frac{\alpha}{2}z^{-1}}$ (válidas em volta de $|z| = 1$)

$$e^{\frac{\alpha}{2}z} = \sum_{k=0}^{\infty} \left(\frac{\alpha}{2}\right)^k \frac{z^k}{k!}$$

$$e^{\frac{\alpha}{2}z^{-1}} = \sum_{k=0}^{\infty} \left(\frac{\alpha}{2}\right)^k \frac{z^{-k}}{k!}$$

e então multiplicá-las. Para $n > 0$, o termo em z^n será

$$I_n(\alpha) z^n = \sum_{k=0}^{\infty} \left(\frac{\alpha}{2}\right)^{k+n} \frac{z^{k+n}}{(k+n)!} \left(\frac{\alpha}{2}\right)^k \frac{z^{-k}}{(k)!} =$$

$$= \sum_{k=0}^{\infty} \left(\frac{\alpha}{2}\right)^{2k+n} \frac{z^n}{k!(k+n)!}$$

Como claramente tem-se $I_n(\alpha) = I_{-n}(\alpha)$, o resultado está provado. ■

Proposição 19.4 $I_n(\alpha)$ satisfaz a relação recursiva

$$I_{n-1} - I_{n+1} = \frac{2n}{\alpha} I_n$$

Demonstração. Se diferenciarmos a expansão acima com relação a z , obtemos

$$\frac{\alpha}{2} \left(1 - \frac{1}{z^2}\right) \sum I_n(\alpha) z^n = \sum n I_n(\alpha) z^{n-1}$$

e o termo em z^{n-1} indica que

$$I_{n-1} - I_{n+1} = \frac{2n}{\alpha} I_n$$

■

Afirmção 19.5 A derivada de $I_n(\alpha)$ satisfaz

$$I'_n = \frac{I_{n-1} + I_{n+1}}{2}$$

Demonstração. Diferenciando a definição de $I_n(\alpha)$ com relação a α , obtemos

$$\frac{1}{2} \left(z + \frac{1}{z}\right) \sum I_n(\alpha) z^n = \sum I'_n(\alpha) z^n$$

e o termo em z^n indica que

$$I'_n = \frac{I_{n-1} + I_{n+1}}{2}$$

■

Corolário 19.6 $I'_n(0) = \begin{cases} \frac{1}{2}, & \text{se } n = \pm 1 \\ 0, & \text{se } |n| \neq 1 \end{cases}$

Proposição 19.7 $I_n(x)$ é solução da equação diferencial

$$\alpha^2 I''_n + \alpha I'_n - (n^2 + \alpha^2) I_n = 0$$

Demonstração. De fato, como

$$I'_n = \frac{I_{n-1} + I_{n+1}}{2}$$

temos

$$\begin{aligned} I''_n &= \frac{I'_{n-1} + I'_{n+1}}{2} = \frac{I_{n-2} + 2I_n + I_{n+2}}{4} = \\ &= \frac{1}{4} \left(\left(I_n + \frac{2(n-1)}{\alpha} I_{n-1} \right) + 2I_n + \left(I_n - \frac{2(n+1)}{\alpha} I_{n+1} \right) \right) = \\ &= I_n - \frac{I_{n-1} + I_{n+1}}{2\alpha} + \frac{n}{2\alpha} (I_{n-1} - I_{n+1}) = \\ &= I_n - \frac{I'_n}{\alpha} + \frac{n^2}{\alpha^2} I_n \end{aligned}$$

onde utilizamos também a relação recursiva 3 vezes. Isto é, I_n pode ser definida como uma solução da equação diferencial

$$\alpha^2 I''_n + \alpha I'_n - (n^2 + \alpha^2) I_n = 0$$

com condições iniciais

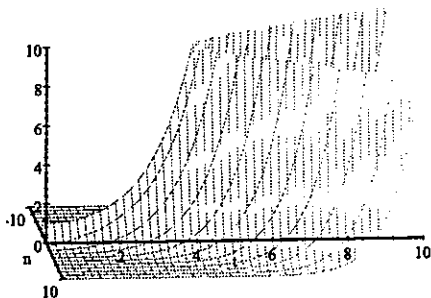
$$I_n(0) = 0; \quad I'_n(0) = 0 \text{ se } n \neq \pm 1, 0$$

$$I_0(0) = 1; \quad I'_0(0) = 0 \text{ se } n = 0$$

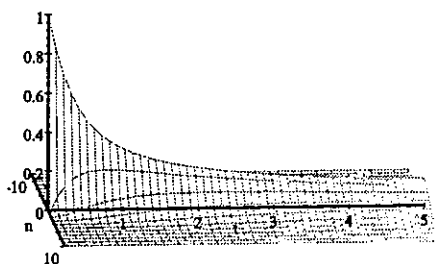
$$I_n(0) = 0; \quad I'_n(0) = \frac{1}{2} \text{ se } n = \pm 1$$

■

Seguem abaixo os gráficos das funções $I_n(t)$ e $P_t[n] = e^{-2t} I_n(2t)$; várias das propriedades mencionadas acima podem ser vistas nos gráficos. Em particular, note que os cortes em $t = 0$ indicam que $I_n(0) = P_0[n] = \delta[n]$.



$I_n(t)$



$P_t[n] = e^{-2t} I_n(2t)$

Bibliografia

- [1] Luis Alvarez, Pierre-Louis Lions, and Jean-Michel Morel, *Image selective smoothing and edge detection by nonlinear diffusion, ii*, SIAM Journal of Numerical Analysis (1992), no. 29, 845–866.
- [2] Francine Catté, Françoise Dibos, and Georges Koepfler, *A morphological scheme for mean curvature motion and applications to anisotropic diffusion and motion of level sets*, SIAM Journal of Numerical Analysis **32** (1995), no. 6, 1895–1909.
- [3] Yun-Gang Chen, Yoshikazu Giga, and Shun'ichi Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, Journal of Differential Geometry (1991), no. 33, 749–786.
- [4] A. J. Chorin, *Curvature and solidification*, Journal of Computational Physics (1985), no. 57, 472+.
- [5] Charles K. Chui, *An introduction to wavelets*, Academic Press, San Diego, CA, 1992.
- [6] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bulletin of the American Mathematical Society **27** (1992), no. 1, 1–67.
- [7] Rachid Deriche, *Recursively implementing the gaussian and its derivatives*, Tech. Report 1893, INRIA – Institut National de Recherche en Informatique et en Automatique, 2004 route des Lucioles, BP93, 06902 Sophia-Antipolis, France, April 1993, Programme 4, Robotique, Image et Vision.
- [8] Charles Eames and Ray Eames, *Powers of ten*, Video, 1968.
- [9] Bart M. Ter Haar Romeny (Ed.), *Geometry-driven diffusion in computer vision*, Kluwer Academic Publishers, Dordrecht, 1994.
- [10] L. C. Evans and J. Spruck, *Motion of level sets by mean curvature. i*, Journal of Differential Geometry (1991), no. 33, 635–681.
- [11] Adam Finkelstein and Marisa Range, *Image mosaics*, 4th International Conference on Raster Imaging and Digital Typography (St. Malo, France)

- (Roger D. Hersch, Jacques André, and Heather Brown, eds.), Springer, April 1998, Lecture Notes in Computer Science, Vol. 1375, pp. 11–22.
- [12] L. M. J. Florack, B. M. Ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, *Scale and the differential structure of images*, Image and Vision Computing **10** (1992), 376–388.
- [13] M. Gage and R. S. Hamilton, *The heat equation shrinking convex plane curves*, Journal of Differential Geometry (1986), no. 23, 69–96.
- [14] Michael E. Gage, *An isoperimetric equation with application in curve shortening*, Duke Mathematical Journal **50** (1983), no. 4, 1225–1228.
- [15] ———, *Curve shortening makes convex curves circular*, Inventiones Mathematicae (1984), no. 76, 357–364.
- [16] M. A. Grayson, *The heat equation shrinks embedded plane curves to round points*, Journal of Differential Geometry (1987), no. 26, 285–314.
- [17] Frédéric Guichard and Jean Michel-Morel, *Image iterative smoothing and p.d.e.'s*, School on Mathematical Problems in Image Processing (2000).
- [18] I. I. Hirschamnn and D. V. Widder, *The convolution transform*, Princeton University Press, Princeton, New Jersey, 1955.
- [19] Rafael Iório Júnior and Valéria de Magalhães Iório, *Equações diferenciais parciais: Uma introdução*, IMPA, Rio de Janeiro, 1988.
- [20] S. Karlin, *Total positivity*, Stanford University Press, 1968.
- [21] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker, *On the evolution of curves via a function of curvature, i: the classical case*, Journal of Mathematical Analysis and Applications (1992), no. 163, 438–458.
- [22] Benjamin B. Kimia, *Conservation laws and theory of shape*, PhD dissertation, McGill University, 1989, McGill Center for Intelligent Machines.
- [23] Benjamin B. Kimia, Allen R. Tannenbaum, and Steven W. Zucker, *Towards a computational theory of shape: an overview*, Technical Report CIM-89-13, McGill University, Department of Electrical Engineering, Montreal, Canada, June 1989.
- [24] ———, *Shapes, shocks and deformations i: the components of shape and the reaction-diffusion space*, Technical Report LEMS-105, Brown University, Division of Engineering, June 1992.
- [25] A. Klinger, *Pattern and search statistics*, Optimizing Methods in Statistics (J. S. Rustagi, ed.), Academic Press, New York, 1971.
- [26] L. Landau, *On the theory of slow combustion*, ACTA Physiocochemica (1944), no. 19, 77+, URSS.

- [27] J. S. Langer, *Instabilities and pattern formation in crystal growth*, Rev. Mod. Physics (1980), no. 52, 1+.
- [28] T. Lindeberg, *Scale-space theory in computer vision*, The Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1994.
- [29] Tony Lindeberg, *Discrete scale space theory and the scale space primal sketch*, Ph.D. thesis, Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, May 1991.
- [30] David Marr, *Vision*, Freeman and Co., New York, 1982.
- [31] O. M. Miller and R. J. Voskuil, *Thematic-map generalization*, Geographical Review 54 (1964), 13–19.
- [32] Predrag Neskovic and Benjamin B. Kimia, *Three-dimensional shape representation from curvature dependent surface evolution*, Technical Report LEMS-128, Brown University, Division of Engineering, November 1993.
- [33] M. Nitzberg and T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, PAMI 14 (1992), no. 8, 826–833.
- [34] Powers of Ten Interactive, *Guide to powers of ten interactive*, Website, <http://www.powersof10.com/powers/space/space.html>.
- [35] Stanley Osher and James A. Sethian, *Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations*, Journal of Computational Physics (1988), no. 79, 12–49.
- [36] P. Perona, T. Shiota, and J. Malik, *Anisotropic diffusion*, Geometry-Driven Diffusion in Computer Vision (B. M. Ter Haar Romeny, ed.), Kluwer, Dordrecht, Netherlands, 1994, pp. 72–92.
- [37] Tim Poston and Ian Stewart, *Catstrophe theory and its applications*, Pitman Publishing, London, 1978.
- [38] Murray H. Protter and Hans F. Weinberger, *Maximum principles in differential equations*, Partial Differential Equation Series, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1967.
- [39] A. Rosenfeld and M. Thurston, *Edge and curve detection for visual scene analysis*, IEEE Transactions on Computers C-20 (1971), no. 5, 562–569.
- [40] Guillermo Sapiro and Allen Tannenbaum, *Affine invariant scale-space*, International Journal of Computer Vision 11 (1992), no. 1, 25–44.
- [41] I. J. Schoenberg, *Some analytical aspects of the problem of smoothing*, Courant Anniversary Volume, Studies and Essays, New York, 1948, pp. 351–370.

- [42] ———, *On pòlya frequency functions II: Variation-diminishing integral operators of the convolution type*, Acta. Sci. Math. **12** (1950), 97–106.
- [43] ———, *On smoothing operators and their generating functions*, Bulletin of the American Mathematical Society **59** (1953), 199–230.
- [44] J. A. Sethian, *An analysis of flame propagation*, PhD dissertation, University of California, Berkeley, June 1982, CPAM Rep. 79.
- [45] ———, *Numerical algorithms for propagating interfaces: Hamilton-jacobi equations and conservation laws*, Journal of Differential Geometry (1990), no. 31, 131–161.
- [46] ———, *Level set methods and fast marching methods*, 2nd ed., Cambridge University Press, Cambridge, UK, 1999.
- [47] Kaleem Siddiqi and Benjamin B. Kimia, *Parts of visual form: Computational aspects*, Technical Report LEMS-114, Brown University, Division of Engineering, November 1992.
- [48] Projeto Visgraf, *Image processing course - final project*, Website, 1998, <http://www.visgrafimpa.br/Courses/ipcg-mosaic-1998/>.

Impresso na Gráfica do



pelo Sistema Xerox / 5390