Luiz San Martin
e Mauro S. de F. Marques

# Cálculo Estocástico

LUIZ SAN MARTIN e MAURO S. DE F. MARQUES
Inst. de Mat. e Ciência da Computação
Univ. Estadual de Campinas
Cidade Universitária Zeferino Vaz
13081 – Campinas-SP



ISBN
85-244-0059-5

**Conselho Nacional de Desenvolvimento Científico e Tecnológico**

**INSTITUTO DE MATEMÁTICA PURA E APLICADA**
Estrada Dona Castorina, 110
22.460 – Rio de Janeiro-RJ

# Prefácio

Ao longo dos anos em decorrência dos trabalhos de Itô, Kunita, Watanabe, Mallivan, da escola francesa de Meyer e muitos outros a teoria de cálculo estocástico consolidou-se sendo hoje um ramos de pesquisa com características próprias. Estas notas surgiram do interesse dos autores em estudar essa teoria com vistas às suas diferentes aplicações. Nossa intenção é que este trabalho seja um ponto de partida para a produção de texto mais elaborado e definitivo.

# Índice

# 1. Elementos de Teoria de Probabilidade e Processos Estocásticos.

Neste capítulo apresentaremos alguns elementos básicos de teoria de probabilidade e processos estocásticos necessários ao desenvolvimento dos capítulos posteriores. O material apresentado é padrão e um estudo mais completo pode ser obtido, por exemplo através dos livros de Billingsley [B], Shiryayev [S], Doob [D] e Gihman e Skorohod [G-S].

Nosso ponto de partida será a definição axiomática de probabilidade de Kolmogorov. Recordaremos alguns resultados de teoria de medida, mas assumimos que o leitor possui, como pré-requisito noção desta teoria ao nível dos livros de Halmos [H] ou Royden R|.

## 1. Conceitos Básicos de Teoria de Probabilidade.

### 1.1 Espaços de probabilidade.

Um espaço de probabilidade é uma tripla ordenada $(\Omega, \mathcal{F}, P)$ onde

i) $\Omega$ é um conjunto arbitrário não vazio;

ii) $\mathcal{F}$ é uma $\sigma$-algebra de subconjuntos de $\Omega$;

iii) $P$ é uma medida de probabilidade.

Na linguagem probabilística os pontos $\omega \in \Omega$ representam os resultados possíveis de um experimento aleatório, os subconjuntos $A \in \mathcal{F}$ são chamados eventos e a probabilidade $P$ é uma aplicação que atribui graus de incerteza aos eventos de $\mathcal{F}$.

Um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ é o objeto primário da teoria de probabilidade. Assim sendo, faremos precisos os conceitos envolvidos em sua definição.

Seja $\Omega$ um conjunto não vazio. Uma coleção de subconjuntos de $\Omega$ é chamada uma álgebra se satisfaz as seguintes propriedades:

i) $\Omega \in \mathcal{A}$;

ii) Se $A \in \mathcal{A}$, então $A^c \in \mathcal{A}$;

iii) Se $A_1, \ldots, A_n \in \mathcal{A}$, então $\bigcup_{k=1}^{n} A_k \in \mathcal{F}$.

Uma $\sigma$-álgebra $\mathcal{F}$ de $\Omega$ é uma álgebra de subconjuntos de $\omega$ que satisfaz:

iv) Se $A_1, A_2, \ldots$ é uma sequência em $\mathcal{F}$, então $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

O par $(\Omega, \mathcal{F})$ é chamado em espaço mensurável.

**Proposição 1.1:** Seja $\mathcal{E}$ uma coleção arbitrária de subconjuntos de $\Omega$. Então, sempre existe uma mínima álgebra ($\sigma$-álgebra) que contém $\mathcal{E}$, chamada álgebra ($\sigma$-álgebra) gerada por $\mathcal{E}$ e denotada por $\mathcal{A}(\mathcal{E})(\sigma(\mathcal{E}))$.
Dem: exercício

Quando $\Omega$ é um espaço topológico e $\mathcal{E}$ a classe de todos os conjuntos abertos de $\Omega, \sigma(\mathcal{E})$ é chamada $\sigma$-álgebra de Borel e seus elementos, de conjuntos Borelianos denotados por $\mathcal{B}(\Omega)$.

Uma função de conjuntos $\mu$, definida numa álgebra $\mathcal{A}$ de subconjuntos de $\Omega$ com valores em $[0, \infty]$ ($\mu : \mathcal{A} \to [0, \infty]$) é chamada uma pré-medida finitamente aditiva na álgebra $\mathcal{A}$ se

$$\mu\Big(A \cup B\Big) = \mu(A) + \mu(B)$$

para todo par de conjuntos disjuntos $A, B, \in \mathcal{A}$. $\mu$ é dita ser $\sigma$-aditiva em $\mathcal{A}$ se para toda sequência $A_1, A_2 \ldots$ de elementos mutuamente disjuntos em $\mathcal{A}$, tal que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

$$\mu\Big(\cup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

$\mu$ é chamada <u>finita</u> se $\mu(\Omega) < \infty$ e $\sigma$-finita se existe uma sequência $A_1, A_2, \ldots$, em $\mathcal{A}, \bigcup_n^{\infty} A_n = \Omega$ e $\mu(A_n) < \infty$ para todo $n$. É fácil ver que $\mu(\phi) = 0$, $P(A) \leq P(B)$ se $A \subset B$ e $P(A \bigcup B) \leq P(A) + P(B)$ para $A, B \in \mathcal{A}$

A verificação direta da $\sigma$-aditividade de uma pré-medida finitamente aditiva em uma álgebra $\mathcal{A}$ em geral não é simples. A próxima proposição nos fornece um critério útil para este fim.

**Proposição 1.2:** Seja $\mu$ uma pré-medida finitamente aditiva numa álgebra $\mathcal{A}$, com $\mu(\Omega) < \infty$. Então $\mu$ é $\sigma$-aditiva se e somente se ela é contínua superiormente no vazio, isto é, para toda sequência $A_1 A_2, \ldots$ em $\mathcal{A}$, $A_n \supset A_{n+1}(A_n \searrow)$, $\bigcap_{n=1}^{\infty} A_n = \phi$ e $\lim_{n \to \infty} \mu(A_n) = 0$.

<u>Dem:</u> Seja $\mu$ $\sigma$-aditiva e $A_n \searrow \bigcap_n A_n = \phi$. Defina $A_0 = \Omega$. Claro $A_n^c \subset A_{n+1}^c(A_n^c \nearrow)$ e $A_n^c \nearrow \bigcup_k A_k^c$.

$$
\begin{aligned}
\mu(\Omega) &= \mu\Big(\bigcup_k A_k^c\Big) = \mu\Big(\bigcup_k (A_k^c \backslash A_{k-1}^c)\Big) \\
&= \sum_{k=1}^{\infty} \mu\Big(A_k^c \backslash A_{k-1}^c\Big) \\
&= \sum_{k=1}^{\infty} \Big(\mu(A_k^c) - \mu(A_{k-1}^c)\Big) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \Big(\mu(A_k^c) - \mu(A_{k-1}^c)\Big) \\
&= \lim_{n\to\infty} \mu(A_n^c)
\end{aligned}
$$

Portanto:

$$
\begin{aligned}
\mu(\phi) &= 0 = \mu(\Omega) - \lim_{n\to\infty} \mu(A_n^c) = \\
&= \lim_{n\to\infty} \Big(\mu(\Omega) - \mu(A_n^c)\Big) \\
&= \lim_{n\to\infty} \mu(A_n),
\end{aligned}
$$

mostrando que a condição é necessária.

Para mostrar suficiência, seja $A_1, A_2, \dots$ uma sequência de elementos mutuamente disjuntos de $\mathcal{A}$

$$
\begin{aligned}
\mu\Big(\cup_{k=1}^{\infty} A_k\Big) &= \mu\Big(\big(\cup_{k=1}^{n} A_k\big) \cup \big(\cup_{k=n+1}^{\infty} A_k\big)\Big) = \\
&= \sum_{k=1}^{n} \mu\Big(A_k\Big) + \mu\Big(\cup_{k=n+1}^{\infty} A_k\Big)
\end{aligned}
$$

Como $\bigcup_{k=n+1}^{\infty} A_k \searrow \phi$, temos que.

$$\begin{aligned}\sum_{k=1}^{\infty}\mu(A_k) &= \lim_{n\to\infty}\sum_{k=1}^{n}\mu(A_k) = \\ &= \lim_{n\to\infty}\Big(\mu\big(\bigcup_{k=1}^{\infty}A_k\big) - \mu\Big(\bigcup_{k=n+1}^{\infty}A_k\big)\Big) = \\ &= \mu\Big(\bigcup_{k=1}^{\infty}(A_k) - \lim_{n\to\infty}\mu\Big(\bigcup_{k=n+1}^{\infty}A_k\Big) = \\ &= \mu\Big(\bigcup_{k=1}^{\infty}A_k\Big),\end{aligned}$$

ou seja $\mu$ é $\sigma$-aditiva. $\square$

Uma medida é uma função de conjuntos $\mu$, definida numa $\sigma$-álgebra $\mathcal{F}$ com valores em $[0,\infty]$ ($\mu : \mathcal{F} \to [0,\infty]$) que é $\sigma$-aditiva.

Uma medida de probabilidade $P$, ou simplesmente uma probabilidade, é uma medida com $P(\Omega) = 1$. Equivalentemente, uma medida de probabilidade $P$ é uma função real de conjuntos, definida em uma $\sigma$-álgebra $\mathcal{F}$ de subconjuntos de um conjunto não vazio $\Omega$ que satisfaz:

i) $P(A) \geq 0$, para todo $A \in \mathcal{F}$ (positividade);

ii) $P(\Omega) = 1$ (normalidade);

iii) $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$, se $A_n \in \mathcal{F}, n = 1, 2\ldots$ e $A_n \bigcap A_m = \phi$ para $n \neq m$ ($\sigma$-aditividade).

As condições i), ii) e iii) acima são conhecidas como axiomas de Kolmogorov. É fácil ver que uma medida de probabilidade cumpre as seguintes propriedades:

**a)** $P(\phi) = 0$;

**b)** $P(A \bigcap B) = P(A) + P(B) - P(A \bigcap B)$ (aditividade forte);

c) $P(A) \leq P(B)$, se $A \subset B$ (monotonicidade);

d) $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$ (sub-aditividade);

e) $\lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n)$, se $A_n \nearrow$ (continuidade inferior);

f) $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$, se $A_n \downarrow$ (continuidade superior).

Um evento $A$ é dito ocorrer quase certamente (q.c.) ou quase seguramente (q.s.) se $P(A) = 1$.

Assumiremos com frequência nos próximos capítulos que o espaço de probabilidade $(\Omega, \mathcal{F}, P)$ é completo significando que: se $A \in \mathcal{F}$, $P(A) = 0$ e $B \subset A$, então $B \in \mathcal{F}$. Isto não chega a ser uma restrição pois todo espaço de probabilidade pode ser estendido a um espaço completo $(\Omega, \overline{\mathcal{F}}^P, \overline{P})$, onde $B \in \overline{\mathcal{F}}^P$ se existem $A, C \in \mathcal{F}$ com $A \subset B \subset C$ e $P(C \backslash A) = 0$ e $\overline{P}(B) = P(A) = P(C)$ (Ver Royden [R]).

## 1.2 Exemplos de espaços de probabilidade.

As construções de exemplos de espaços de probabilidade que faremos a seguir são baseadas no seguinte teorema de teoria da medida.

**Teorema de Caratheodory.** Seja $\Omega$ um conjunto não vazio, $\mathcal{A}$ uma álgebra de subconjuntos de $\Omega$ e $\mathcal{F} = \sigma(\mathcal{A})$ a $\sigma$-álgebra gerada por $\mathcal{A}$. Seja $\mu_0$ uma pré-medida $\sigma$-aditiva e $\sigma$-finita em $\mathcal{A}$. Então existe uma única medida $\mu$ definida em $\mathcal{F}$ que é uma extensão de $\mu_0$, isto é

$$\mu(A) = \mu_0(A)$$

para todo $A \in \mathcal{A}$.

O Teorema de Carathedory requer uma pré-medida na álgebra que seja $\sigma$-finita. Removida esta condição existem exemplos "patológicos" da não unicidade de $\mu$. No caso de medidas de probabilidade, a condição é automaticamente satisfeita pois $P(\Omega) = 1$.

**Exemplo 1:** Probabilidades em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Seja $\mathbb{R} = (-\infty, \infty)$ o conjunto dos reais e $\mathcal{B}(\mathbb{R})$ a correspondente $\sigma$-álgebra de Borel. Consideremos a classe $\mathcal{A}$ formada pelo conjunto vazio e por uniões finitas de intervalos disjuntos, abertos a direita e fechados a esquerda, isto é, intervalos de forma $(a, b], -\infty \leq a < b \leq \infty$, com $(a, \infty]$ tomado como $(a, \infty)$. É fácil ver que $\mathcal{A}$ é uma álgebra e que $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{A})$. Note que $\mathcal{A}$ não é uma $\sigma$-álgebra.

**Definição 1.3:** Uma função $F : \mathbb{R} \to \mathbb{R}$ é chamada <u>função distribuição de probabilidade</u> se ela satisfaz as seguintes propriedades:

i) $F$ é não decrescente;

ii) $F$ é contínua a direita;

iii) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ e $F(+\infty) = \lim_{x \to \infty} F(x) = 1$.

**Teorema 1.4:** Seja $F$ uma função distribuição de probabilidade. Então existe uma única medida de probabilidade $P$ definida em $\mathcal{B}(\mathbb{R})$ que satisfaz

$$P\Big((a, b]\Big) = F(b) - F(a) \tag{1.1}$$

para $a, b, -\infty \leq a < b \leq \infty$. Reciprocamente se $P$ é uma medida de probabilidade em $\mathcal{B}(\mathbb{R})$

$$F(x) = P\Big((-\infty, x]\Big)$$

é uma função distribuição de probabilidade em $\mathbb{R}$ que satisfaz (1.1).

Dem: Seja $A \in \mathcal{A}$. Então,

$$A = \bigcup_{k=1}^{n} (a_k, b_k]$$

$$-\infty \leq a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_n < b_n \leq \infty .$$

Defina $P_0$ em $\mathcal{A}$ por

$$P_0(A) = \sum_{k=1}^{n} \Big( F(b_k) - F(a_k) \Big)$$

É fácil ver que $P_0$ esta bem definida (no sentido que independe da representação de $A$) e que é uma pré-medida finitamente aditiva em $\mathcal{A}$ com $P_0(\mathbb{R}) = 1$.

Para mostrarmos que $P$ é $\sigma$-aditiva em $\mathcal{A}$ usaremos a Proposição 1.2. Para tanto, seja $A_n, n = 1, 2 \ldots \in \mathcal{A}, A_n \searrow \phi$.

Suponhamos primeiramente que $A_n \subset K, K$ compacto $n = 1, \ldots,$. Como $A_n$ é uma união finita de intervalos da forma $(a, b]$, $F$ é contínua a direita e $F(\infty) = 1$, para cada $A_n$, existe $B_n \in \mathcal{A}$, $\overline{B}_n \subset A_n$ e $P_0(A_n \backslash B_n) \leq \varepsilon 2^{-n}$, onde $\varepsilon > 0$. Dado que $\bigcup_n A_n = \phi$, temos que $\bigcup_n \overline{B}_n = \phi$. Portanto $\{\overline{B}_n^c ; n = 1, 2, \cdots\}$ é uma cobertura aberta de $K$ e existe $n_0$ tal que $K \subset \bigcap_{k=1}^{n_0} \overline{B}_k^c \subset \bigcup_{k=1}^{n_0} B_k^c$. Usando este resultado e a inclusão $A_{n+1} \subset A_n$, obtemos para $n \geq n_0$,

$$
\begin{aligned}
P_0(A_n) &= P_0\Big(A_n \bigcap K\Big) \leq P_0\Big(A_n \bigcap (\textstyle\bigcup_{k=1}^{n_0} B_k^c)\Big) \\
&= P_0\Big(\textstyle\bigcup_{k=1}^{n_0}\Big(A_n \bigcap B_k^c\Big)\Big) \\
&\leq \textstyle\sum_{k=1}^{n_0} P\Big(A_n \bigcap B_k^c\Big) \\
&\leq \textstyle\sum_{k=1}^{n_0} P\Big(A_k \bigcap B_k^c\Big) \\
&= \textstyle\sum_{k=1}^{n_0} P\Big(A_k \backslash B_k\Big) \\
&= \textstyle\sum_{k=1}^{n_0} \varepsilon 2^{-k} < \varepsilon
\end{aligned}
$$

O que mostra que $P_0(A_n) \to 0$ quando $n \to \infty$.

Se os $A_n's$ não estão necessariamente contidos em um compacto $K$, para $\varepsilon > 0$ seja $N > 0$ tal que

$$P_0\Big((-N, N]\Big) > 1 - \varepsilon/2.$$

(Este $N$ existe pois $P_0((-N, N]) = F(N) - F(-N) \to 1$, quando $N \to \infty$).
Temos então que

$$
\begin{aligned}
P_0(A_n) &= P_0\Big(A_n \bigcap (-N, N] + P_0\Big(A_n \bigcap [-N, N]^c\Big) \\
&\leq P_0\Big(A_n \bigcap (-N, N]\Big) + \varepsilon/2
\end{aligned}
$$

Usando o caso anterior com $A_n \bigcap (-N, N]$ no lugar de $A_n$ obtemos que $P_0(A_n) \to 0$ que $n \to \infty$. Pela proposição 2, $P_0$ é $\sigma$-aditiva. Como $P_0(\mathbb{R}) = 1$, a primeira parte do teorema segue pelo Teorema de Caratheodory.

A recíproca é uma consequência da monotonicidade, continuidade superior e inferior da medida de probabilidade $P$:

i) Se $x_1 < x_2$, então $(-\infty, x_1] \subset (-\infty, x_2]$, logo

$$F(x_1) = P\Big((-\infty, x_1]\Big) \leq P\Big((-\infty, x_2]\Big) = F(x_2)$$

o que mostra que $F$ é não decrescente;

ii) Se $x_n \searrow x$, então $(-\infty, x_n] \searrow (-\infty, x]$ e

$$F(x+) = \lim_{x_n \searrow x} F(x_n) = \lim_{x_n \searrow x} P\Big((-\infty, x_n]\Big) = P\Big((-\infty, x]\Big) = F(x);$$

iii) Analogamente, se $x_n \searrow -\infty$ e $y_n \nearrow +\infty$ então,

$$(-\infty, x_n) \downarrow \phi \text{ e } (-\infty, y_n) \nearrow (-\infty, \infty).$$

Portanto,

$$F(-\infty) = \lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P((-\infty, x_n]) = P(\phi) = 0$$

e

$$F(-\infty) = \lim_{n \to \infty} F(y_n) = \lim_{n \to \infty} P((-\infty, y_n]) = P(I\!R) = 1. \qquad \square$$

O Teorema 1.4 mostra uma correspondência biunívoca entre funções de distribuição de probabilidade e medidas de probabilidade em $\mathcal{B}(I\!R)$. Na verdade, temos que medidas de probabilidade em $\mathcal{B}(I\!R)$ são casos particulares de medidas de Lebesgue-Stieltjes, isto é, medidas que atribuem valores finitos a intervalos limitados (ver Royden [R]). Em particular, podemos construir de maneira análoga a medida de Lebesgue em $\mathcal{B}(I\!R)$ como a única medida que cumpre, para todo $a, b \in I\!R$ e $-\infty < a < b < \infty, \lambda\big((a, b]\big) = b - a$.

Alguns exemplos clássicos de funções distribuição de probabilidade são:

1. distribuição uniforme em $[0, 1)$

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

2. distribuição de Bernoulli com parâmetro $p \in [0, 1]$

$$F(x) = \begin{cases} 0 & x < 1 \\ (1-p) & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

3. distribuição exponencial com parâmetro $\theta > 0$

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\theta x} & x \ge 0 \end{cases}$$

4. distribuição de Poisson com parâmetro $\lambda > 0$

$$F(x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{[x]} & \frac{e^{-\lambda}\lambda^k}{k!} \end{cases}$$

onde $[x]$ denota a parte interna de $x$,

5. distribuição normal (ou Gaussiana) com parâmetro $\mu \in I\!R$ e $\sigma > 0$

$$F(x) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{x} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} dt \; .$$

Uma função distribuição de probabilidade é chamada <u>discreta</u> se existem pontos $x_1, x_2, \ldots$ tais que $\Delta F(x_k) = F(x_k) - F(x_n-) > 0$ e $\sum_k \Delta F(x_k) = 1$. É chamada <u>absolutamente contínua</u> se existe $f : I\!R \to I\!R$, não negativa tal que $F(x) = \int_{-\infty}^{x} f(t)dt$. Uma tal $f$ é chamada função densidade de probabilidade.

**Exemplo 2:** Probabilidades em $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Seja $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$ o produto cartesiano de $\mathbb{R}$. Considere os retângulos de $\mathbb{R}^n$ da forma $R = I_1 \times I_2 \times \ldots \times I_n$, onde $I_k = (a_k, b_k]$ são intervalos como no Exemplo 1. Pode ser mostrado que a classe $\mathcal{A}$ formada pelo conjunto vazio e por uniões finitas de retângulos disjuntos da forma acima é uma álgebra e que $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{A})$.

Uma função $F : \mathbb{R}^n \to \mathbb{R}$ é chamada uma função distribuição de probabilidade conjunta (ou multivariada) se ela cumpre as seguintes propriedades:

i) $F$ é não decrescente em cada argumento;

ii) $\Delta^{(1)}_{a1,b1}\ \Delta^{(2)}_{a2,b2} \ldots \Delta^{(n)}_{an,bn}\ F(x_1, \ldots, x_n) \geq 0$, para todo $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$ com $a_1^i < b_i$, onde $\Delta^{(i)}_{a_i b_i}\ F(x_1, \ldots, x_n) - F(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots x_n) - F(x_i, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n)$ e

iii) $F(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots x_n) = \lim\limits_{x_i \to \infty} F(x_1, \ldots, x_n) = 0,\ i = 1, 2, \ldots n$ e

$$F(+\infty, +\infty, \ldots, +\infty) = \lim_{\substack{x_1 \to \infty \\ x_n \to \infty}} F(x_1, \ldots, x_n) = 1.$$

De maneira inteiramente análoga ao Teorema 1.4 podemos provar a seguinte versão multivariada para construção de medidas de probabilidade em $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

**Teorema 1.5:** Seja $F$ uma função distribuição de probabilidade conjunta em $\mathbb{R}^n$. Então existe uma única medida de probabilidade $P$ em $B(\mathbb{R})$ que satisfaz:

$$P\Big((a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n]\Big) =$$

$$= \Delta^{(1)}_{a1,b1}\ \Delta^{(2)}_{a2,b2} \ldots \Delta^{(n)}_{an,bn}\ F(x_1, \ldots, x_n). \tag{1.2}$$

Reciprocamente, se $P$ é uma medida de probabilidade em $\mathcal{B}(\mathbb{R}^n)$,

$F(x_1,\ldots,x_n) = P((-\infty,x_1]\times\ldots\times(-\infty,x_n])$ define uma distribuição de probabilidade conjunta em $I\!R^n$ que satisfaz (1.2).

Um exemplo importante de distribuição multivariada é dado por

$$F\Big(x_1,\ldots,x_n\Big) = F_1(x_1\Big)\ldots F_n(x_n),$$

onde $F_1,\ldots,F_n$ são distribuições em $(I\!R)$. Neste caso, se $P_1,\ldots,P_n$ são as respectivas probabilidades em $\mathcal{B}(I\!R)$ a probabilidade de $P$ definida por $F$ em $\mathcal{B}(I\!R^n)$ é a única que satisfaz:

$$P\Big(A_1\times A_2\times\cdots\times A_n\Big) = P_1(A_1)P_2(A_2)\ldots P_n(A_n), \tag{1.3}$$

onde $A_i \in B(I\!R), i = 1,2,\ldots,n$.

Uma maneira de mostrarmos (1.3) é aplicar o Teorema de Caratheodory com $\mathcal{A}$ sendo a álgebra formada por uniões finitas e disjuntas de retângulos da forma:

$$\bigcup_{i=1}^{k} A_{i,1},\times\ldots\times A_{i,n}$$

com $A_{i,j}\in\mathcal{B}(I\!R)$ e $P_0$ definida por

$$P_0\Big(\textstyle\bigcup_{i=1}^{k} A_{i,1}\times\ldots\times A_{i,n}\Big)$$

$$= \textstyle\sum_{i=1}^{k} P_0\Big(A_{i,1},\times\ldots\times A_{i,n}\Big)$$

$$= \textstyle\sum_{1=i}^{k} P_1\Big(A_{i,1})P_2(A_{i,2})\ldots P_{\tilde{n}}(A_{i,n})$$

A unicidade das respectivas extenções segue uma vez que $\mathcal{B}(I\!R^n) = \sigma(\mathcal{A})$.

Uma medida de probabilidade definida por (1.3) é chamada de medida de probabilidade produto e é denotada por $P = P_1 \times P_2 \times \cdots \times P_n$. Probabilidades produto estão relacionadas com o importante conceito de independência em teoria de probabilidade que definiremos posteriormente.

Outros exemplos de distribuições multivariadas podem ser criados definindo

$$F\left(x_1, \ldots, x_n\right) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f\left(t_1, \ldots, t_n\right) dt_n \ldots dt_1,$$

onde $f$, chamada de função densidade de probabilidade conjunta, é não negativa e

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1.$$

Entre exemplos de densidades conjuntas, destaca-se o caso normal ou Gaussiano (não degenerado). Neste caso $f$ é dada por

$$f(x_1, \ldots, x_n) = \left[\frac{\det(\sum)}{2\pi}\right]^{1/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=i}^{n} (x_i - m_i)\sigma^{i,j}(x_j - m_j)\right),$$

onde $m_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$ e $\sum$ é uma matriz $n \times n$, positiva definida com $\sum^{-1} = (\sigma^{i,j})_{i,j=1,\ldots,n}$. O vetor $m = (m_1 \cdots, m_n)$ é chamado vetor média e $\sum$ matriz de covariância.

**Exemplo 3:** Probabilidades em $(\mathbb{R}^{\mathbb{T}}, C(\mathbb{R}^{\mathbb{T}}))$. Seja $\mathbb{T}$ um conjunto não vazio e $\mathbb{R}^{\mathbb{T}}$ o conjunto de todas as funções $x = (x(t); t \in \mathbb{T})$ definidos em $\mathbb{T}$ com valores em $\mathbb{R}$.

Um subconjunto de $\mathbb{R}^{\mathbb{T}}$ da forma

$$C\left(t_1, \ldots, t_n; B\right) = \left\{x \in \mathbb{R}^{\mathbb{T}} : \left(x(t_1), \ldots, x(t_n)\right) \in B\right\},$$

onde $n \in \mathbb{N}, t_1, \ldots, t_n \in \mathbb{T}$ e $B \in \mathcal{B}(\mathbb{R}^n)$, é chamado cilindro (ou conjunto de dimensão finita) de $\mathbb{R}^{\mathbb{T}}$ em $t_1, \ldots, t_n$ com base $B$. Seja $\mathcal{A}$ a classe de todos os cilindros de $\mathbb{R}^{\mathbb{T}}$ e $\mathcal{C}(\mathbb{R}^{\mathbb{T}}) = \sigma(\mathcal{A})$.

Para cada subconjunto $u$ finito e não vazio de $\mathbb{T}(u = \{t_1, \cdots, t_n\} \subset \mathbb{T}, |u| =$ card $(u) = n < \infty)$. Seja $\pi_u$ a projeção de $\mathbb{R}^{\mathbb{T}}$ em $\mathbb{R}^{|u|}(\pi_u : \mathbb{R}^{\mathbb{T}} \to \mathbb{R}^{|u|}, \pi_u(x) = (x(t_1), \ldots x(t_n)))$. Para $v = \{t_{i1}, \ldots, t_{i_k}\} \subset u : \{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$ e $|v| = k \leq n$, seja $\pi_{u,v}$ a projeção de $\mathbb{R}^{|u|}$ em $\mathbb{R}^{|n|}(\pi_{u,v} : \mathbb{R}^{(u)} \to \mathbb{R}^{(n)}, \pi_{u,v}(x_1, \ldots, x_n) = (x_{i_1}, \ldots, x_{i_n}))$. As aplicações $\pi_u$ e $\pi_{u,v}$ são respectivamente $\mathcal{C}(\mathbb{R}^{\mathbb{T}}) - \mathcal{B}(\mathbb{R}^{|u|})$ e $\mathcal{B}(\mathbb{R}^{|u|}) - \mathcal{B}(\mathbb{R}^{|v|})$ mensuráveis (uma aplicação $T$ do espaço mensurável $(\Omega, \mathcal{F})$ no espaço mensurável $(\Omega', \mathcal{F}')$ é dita ser $\mathcal{F} - \mathcal{F}'$ mensurável, ou simplesmente $\mathcal{F}'$ mensurável se $\mathcal{F}$ é fixo, se $T^{-1}(A) \in \mathcal{F}$ para todo $A \in \mathcal{F}'$), e

$$\pi_{u,v} \circ \pi_u = \pi_n \tag{1.4}$$

para todo $v \subset u$.

Note que usando a notação acima temos que:

$$C(u; B) = C(t_1, \ldots, t_n; B) = \pi_u^{-1}(B).$$

**Lema 1.6:** Se $C(u_1; B_2)$ e $C(u_2; B_2)$ são cilindros de $\mathbb{R}^{\mathbb{T}}$ então existe $u \subset \mathbb{T}$ com $|u| < \infty$ e $D_1, D_2 \in \mathcal{B}(\mathbb{R}^{|u|})$ tal que

$$C(u_1; B_1) = C(u; D_1) \quad \text{e} \quad C(u_2; B_2) = C(u; D_2).$$

Dem: $C(u_1; B_1) = \pi_{u_i}^{-1}(B_1) \;\; i = 1, 2$. Se $u = u_1 \bigcap u_2$, então $\pi_{u_i} = \pi_{u,u_i} \circ \pi_u$

$$
\begin{aligned}
C(u_i;B_i) &= \pi_{u_i}^{-1}(B_i) = \left(\pi_{u,u_i}\circ\pi_u\right)^{-1}(B_i) \\
&= \pi_u^{-1}\left(\pi_{u,u_i}^{-1}(B_i)\right) = \pi_u^{-1}(D_i) \\
&= C(u;D_i)
\end{aligned}
$$

para $i = 1, 2$. □

**Lema 1.7:** Seja $\mathcal{A}$ a classe de todos os cilindros de $\mathbb{R}^{\mathbb{T}}$ : $\mathcal{A} = \{C; C = C(u;B), u \subset \mathbb{T}, |u| < \infty;\ B \in \mathcal{B}(\mathbb{R}^{|u|})$. Então $\mathcal{A}$ é uma álgebra.

Dem: i) $\mathbb{R}^{\mathbb{T}} = C(\{t\};\mathbb{R}) \in \mathcal{A}$, qualquer $t \in \mathbb{T}$. ii) $C(u;B)^c = C(u;B^c) \in \mathcal{A}$. iii) $C_i = C(u_i;B_i)$, $i = 1, 2, \ldots, n$. Pelo Lema 1.6 existe $u \subset \mathbb{T}$, com $|u| < \infty$ e $D_1, \ldots, D_n \in \mathcal{B}(\mathbb{R}^{|u|})$ tal que $C_i = C(u;D_i)$. Segue que $\bigcup_{i=1}^n C_i = C(u;\bigcup_{i=1}^n D_i) \in \mathcal{A}$. □

**Definição 1.8:** Uma família de medidas de probabilidade $\{P_u : u \subset \mathbb{T}, |u| < \infty, u \neq \phi, P_u$ uma probabilidade em $\mathcal{B}(\mathbb{R}^{|u|})\}$ é dita ser consistente no sentido de Kolmogorov se:

$$
P_u\left(\pi_{u,n}^{-1}(B)\right) = P_v(B) \tag{1.5}
$$

para todo $v \subset u$ e $B \in B(\mathbb{R}^{|n|})$.

Nosso próximo teorema nos mostra como construir medidas de probabilidade na $\sigma$-álgebra de $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$ gerada pelos cilindros de $\mathbb{R}^{\mathbb{T}}$. Antes disto porém, necessitamos de um resultado geral sobre medidas de probabilidade definidas em subconjuntos Borelianos de um espaço métrico. Ésta propriedade já apareceu de forma implícita no Teorema 1.4 através da continuidade a direita da função distribuição.

**Proposição 1.9:** Seja $\Omega$ um espaço métrico e $P$ uma medida de probabilidade em $\mathcal{B}(\Omega)$. Então, para $A \in \mathcal{B}(\Omega)$ e $\varepsilon > 0$ existe $O$ aberto e $F$ fechado tal que $F \subset A \subset O$ tal que $P\,(O\backslash F) < \varepsilon$. Se $\Omega$ é completo e separável, $P(A) = \sup\{P(K) : K$ compacto, $K \subset A\}$ para todo $A \in \mathcal{B}(\Omega)$ (1.5).

Dem: Seja $\mathcal{R} = \{A \in B(\Omega) : A$ satisfaz $(1.5)\}$ Claro, $\phi \in \mathcal{R}$ e $\Omega \in \mathcal{R}$, pois ambos são fechados e abertos. Seja $A \in \mathcal{B}$ e $\varepsilon > 0$. Então existe $O_\varepsilon$ aberto e $F_\varepsilon$ fechado tal que $F_\varepsilon \subset A \subset O_\varepsilon$ e $\mu(O_\varepsilon \backslash F_\varepsilon) < \varepsilon$. Tomando complementos $O_\varepsilon^c \subset A^c \subset F_\varepsilon^c$ e $F_\varepsilon^c \backslash O_\varepsilon^c = F_\varepsilon^c \bigcap O_\varepsilon = O_\varepsilon \bigcap F_\varepsilon^c = O_\varepsilon \backslash F_\varepsilon$. Portanto $A^c \in \mathcal{R}$, pois $O_\varepsilon^c$ é fechado, $F_\varepsilon^c$ é aberto e $P(F_\varepsilon^c \backslash O_\varepsilon^c) < \varepsilon$. Agora seja $A_n, n = 1, 2, \ldots$ uma sequência em $\mathcal{R}$. Então existe $O_{n,\varepsilon}$ aberto, $F_{n,\varepsilon}$ fechado, tal que $F_{n,\varepsilon} \subset A_n \subset O_{n,\varepsilon}$ e $P(O_{n,\varepsilon} \backslash F_{n,\varepsilon}) < \frac{\varepsilon}{2^{n+1}}$ para todo $n \geq 1$. É fácil ver que existe $n_0$ tal que

$$\mu\Big(\bigcup_{n=1}^{\infty} F_{n,\varepsilon} \backslash \bigcup_{n=1}^{n_0} F_{n\varepsilon}\Big) < \varepsilon/2.$$

Seja $F_\varepsilon = \bigcup_{n=1}^{n_0} F_{n,\varepsilon}$ e $O_\varepsilon = \bigcup_{n=1}^{\infty} O_{n,\varepsilon}$. Claro, $F_\varepsilon$ é fechado, $O_\varepsilon$ é aberto e $F_\varepsilon \subset \bigcup_{n=1}^{\infty} A_n \subset O_\varepsilon$.

Segue que

$$\begin{aligned} P(O_\varepsilon \backslash F_\varepsilon) &\leq \textstyle\sum_{n=1}^{\infty} P(O_{n,\varepsilon} \backslash F_{n,\varepsilon}) + P(\bigcup_{n=1}^{\infty} F_{n,\varepsilon} \backslash F_\varepsilon) \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Portanto $\bigcup_n A_n \in \mathcal{R}$, o que prova que $R$ é uma $\sigma$-álgebra.

Se $A$ aberto, então

$$F_n = \left\{x : \inf_{y \in A_n^c} d(x, y) \geq 1/n\right\}$$

é fechado, $F_n \subset F_{n+1}$ e $\bigcup_n F_n = A$. Portanto, se $\varepsilon > 0$ existe $n_0$ tal que $P(A \backslash \bigcup_{n=1}^{n_0} A_n) < \varepsilon$.

$F_\varepsilon = \bigcup_{n=1}^{n_0} F_n$ é fechado e $F_\varepsilon \subset A \subset O_\varepsilon = A$. Segue que $A \in \mathcal{R}$, o que mostra que $\mathcal{R} = \mathcal{B}(\Omega)$, pois $R$ é uma $\sigma$-álgebra que contém todos os abertos e $\mathcal{R} \subset \mathcal{B}(\Omega)$ que é mínima. Isto mostra a primeira parte da proposição.

Para mostrarmos a segunda afirmação, seja $\varepsilon > 0$. Então, existe $K_0$ compacto tal que $P(K_0) > 1 - \varepsilon$. De fato, como $\Omega$ é separável ele pode ser coberto por uma família contável de bolas fechadas de raio arbitrário $\mathcal{R}$. Para cada $n$, seja $B_{n,k}, k = 1, 2, \ldots$ uma sequência de bolas abertas de raio $1/n$ que cobre $\Omega$.

$$P(\Omega) = \lim_{N\to\infty} P\Big(\bigcup_{k=1}^{N} B_{n,k}\Big) = 1.$$

Portanto, existe $N_n$ tal que

$$P\Big(\bigcap_{k=1}^{N_n} B_{n,k}^c\Big) < \varepsilon/2^{n+1}. \tag{1.6}$$

Seja $K_0 = \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{N_n} B_{n,k}$. Para cada $k > 0, K$ pode ser coberto por um número finito de bolas com raio $\mathcal{R}$, portanto $K_0$ é compacto pois $\Omega$ é separável e $K_0$ totalmente limitado. Por (1.6) temos que

$$\begin{aligned} P(K_0^c) &= P\Big(\bigcup_{n=1}^{\infty} \bigcap_{k=1}^{N_n} B_{n,k}^c\Big) \le \\ &\le \sum_{n=1}^{\infty} P\Big(\bigcap_{K=1}^{N_n} B_{n,1}^c\Big) \le \textstyle\sum_{n=1}^{\infty} \varepsilon/2^{n+1} = \varepsilon/2, \end{aligned}$$

ou seja

$$P(K_0) > 1 - \varepsilon/2.$$

Agora seja $B \in \mathcal{B}(\Omega)$, pela primeira parte, existe $F$ fechado $F \subset B$ tal que $P(B\backslash F) = P(B) - P(F) < \varepsilon/2$. Seja $K \coloneqq F \bigcap K_0$, claro, $K$ é compacto, $K \subset B$ e

$$P(F) - P(K) = P(F/K) = P\left(F \bigcap K_0^c\right) \leq P(K_0^c) < \varepsilon/2.$$

Segue que

$$P(B) < P(F) + \varepsilon/2 < P(K) + \varepsilon$$

e como $\varepsilon$ é arbitrário, $P(B) = \sup\left\{P(K) : K \text{ compacto}, K \subset A\right\}$. □

**Teorema 1.10:** (Teorema de Extensão de Kolmogorov). Seja $\{P_u; u \subset \mathbb{T}, |u| < \infty$ e $P_u$ probabilidade em $\mathcal{B}(\mathbb{R}^u)\}$ uma família consistente de medidas de probabilidade. Então existe uma única medida de probabilidade $P$ em $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$ que satisfaz

$$P\Big(C(u;B)\Big) = P_u(B) \tag{1.7}$$

para todo $u = \{t_1, \ldots, t_n\} \subset \mathbb{T}$, $|u| = n < \infty$ e $B \in \mathcal{B}(\mathbb{R}^{|u|})$.

Dem: Como no Teorema 1.4, faremos uso do Teorema de Caratheodory. A relação (1.7) sugere a definição de uma pré-medida $P_0$ na álgebra de cilindros $\mathcal{A}$, da forma $P_0(C(u;B)) = P_u(B)$. Temos que mostrar primeiramente que $P_0$ independe da representação do cilindro. Para $C \in \mathcal{A}$ sejam $u$ e $v$ subconjuntos finitos de $\mathbb{T}, B \in \mathcal{B}(\mathbb{R}^{|u|})$ e $D \in \mathcal{B}(\mathbb{R}^{|u|})$ tais que

$$C = C(u;B) = C(v;D)$$

ou seja

$$C = \pi_u^{-1}(B) = \pi_v^{-1}(D).$$

Se $w = u \bigcup v$, então existe $E \in B(\mathbb{R}^{|w|})$ tal que

$$C = C(w;E) = \pi_w^{-1}(E).$$

Por outro lado,

$$E = \pi_w\left(\pi_w^{-1}(E)\right) = \pi_w(C) = \pi_w\left(\pi_w^{-1}(B)\right)$$

e como $u \subset w$, segue por (1.3) que

$$E = \pi_w\left(\pi_w^{-1}\left(\pi_{w,u}^{-1}(B)\right)\right) = \pi_{w,u}^{-1}(B)$$

Pela condição de consistência (1.4)

$$P_w(E) = P_w\left(\pi_{w,u}^{-1}(B)\right) = P_u(B) \ .$$

Analogamente $P_w(E) = P_v(D)$. Portanto $P_0$ é bem definida em $\mathcal{A}$.

Para mostrarmos que $P_0$ é finitamente aditiva, sejam $C_1, \ldots, C_n$ cilindros disjuntos. Pelo Lema 1.6 existe $u \subset \mathbb{T}$, $|u| < \infty$ e conjuntos disjuntos $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}^{|u|})$ tal que

$$C_i = C(u; B_i) = \pi_u^{-1}(B_i), i = 1, 2, \ldots, n \ .$$

Da definição de $P_0$ segue que

$$\begin{aligned} P_0\left(\cup_{i=1}^n C_i\right) &= P_0\left(\cup_{i=1}^n \pi_u^{-1}(B)\right) = P_0\left(\pi_u^{-1}(\cup_{i=1}^n B_i)\right) \\ &= P_u\left(\cup_{i=1}^n B_i\right) = \sum_{i=1}^n P_u(B_i) \\ &= \sum_{i=1}^n P_0\left(\pi_u^{-1}(B_i)\right) = \sum_{i=1}^n P_0(C_i). \end{aligned}$$

Portanto $P_0$ é finitamente aditiva.

Como $P_0(\mathbb{R}^{\mathbb{T}}) = 1$, para aplicarmos o Teorema de Caratheodory temos somente que mostrar que $P_0$ é $\sigma$-aditiva, ou, equivalentemente pela Proposição 2, que $P_0$ é contínua superiormente no vazio. Para tanto, consideramos $C_1, C_2, \ldots \in \mathcal{A}$ e $C_n \searrow \phi$. Suponhamos que $P_0(C_n)$ não converge a zero, isto é existe $\varepsilon > 0$ tal que $P(C_n) \geq \varepsilon$.

Como os $C_n$ são cilindros, os correspondentes conjuntos de índices podem ser tomados de forma que

$$C_n = C\Big(u_n; B_n\Big)$$

$u_n = \{t_1, t_2, \ldots, t_{N_n}\} \subset \{t_1, t_2, \ldots, t_{N_{n+1}}\} = u_{n+1}, B_n \in B(\mathbb{R}^{|u|})$.

$P_0(C_n) = P_{u_n}(B_n)$. Pela Proposição 1.9, existe $K_n$ compacto, $K_n \subset B_n$ tal $P_u(B_n \backslash K_n) < \varepsilon 2^{-(n+1)}$. Se $L_n = C(u_n; K_n)$, então $C_n \backslash L_n = C(u_n; B_n \backslash K_n)$ e $P_0(C_n \backslash L_n) < \varepsilon 2^{-(n+1)}$. Seja $D_n = \bigcap_{k=1}^{n} L_k$. Então $D_n \subset L_n \subset C_n$ e $P_0(C_n \backslash D_n) < \varepsilon/2$. Portanto, $P(D_n) > \varepsilon/2 > 0$ e $D_n \neq \phi$.

Seja $x^{(n)} \in D_n \subset \mathbb{R}^{\mathbb{T}}$. Para $k \leq n, x^{(n)} \in D_n \subset D_k \subset L_k = C(u_k; K_k)$ o que implica que $(x^{(n)}(t_1), \ldots, x^{(n)}(t_{N_k})) \in K_k$. Como $K_k$ é limitado, a sequência $x^{(n)}(t_{N_k}), x^{(2)}(t_{N_k}), \ldots, x^{(n)}(t_{N_k}), \ldots$ é limitada para cada $k$. Pelo método da diagonal (Billingsley (1575), Theorem 25.13) existe uma subseqüência $n_1 < n_2 < \ldots$ tal que $(x^{(ni)}(t_{N_k})$ converge quando $i \to \infty$. Seja $x \in \mathbb{R}^{\mathbb{T}}$ tal que $x(t_{N_k}) = \lim_{i \to \infty} x^{(n)}(t_{N_k})$. Como por construção

$$\Big(x^{(n_i)}(t_{N_1}), \ldots, x^{(n_i)}(t_{N_k})\Big) \xrightarrow[i \to \infty]{} \Big(x(t_{N_k}), \ldots, x(t_{N_k}\Big),$$

temos que $(x(t_{N_1}), \ldots, x(t_{N_k})) \in K_k$. Segue que $x \in L_k \subset C_k$, para todo $k$ e portanto $\bigcap_k C_k \neq \phi$, o que contradiz a hipótese que $C_n \downarrow \phi$. Portanto $P_0$ é $\sigma$-aditiva.

Segue pelo Teorema de Caratheodory que $P_0$ se estende univocamente a uma probabilidade $P$ em $C(\mathbb{R}^{\mathbb{T}})$. □

Um caso importante é aquele em que $\mathbb{T} = \mathbb{N}$, o conjunto dos naturais. $\mathbb{R}^{\mathbb{T}} = \mathbb{R}^{\mathbb{N}}$ é o espaço das sequências reais $x = (x_n : n \in \mathbb{N})$ e é usualmente denotado por $\mathbb{R}^\infty$. $\mathbb{R}^\infty$ é um espaço métrico, com métrica dada por $d(x,y) = \sum_{n=1}^{\infty} 2^{-n}(x_n - y_n)$ e $\mathcal{C}(\mathbb{R}^\infty) = \mathcal{B}(\mathbb{R}^\infty)$.

**Corolário 1.11:** Seja $P_1, P_2, \ldots$ uma sequência de medidas de probabilidade em $\mathcal{B}(\mathbb{R}^1)$, $\mathcal{B}(\mathbb{R}^2), \cdots$ que satisfaz a condição

$$P_{n+1}\Big(B \times \mathbb{R}\Big) = P_n(B),$$

$B \in \mathcal{B}(\mathbb{R}^n)$. Então existe uma única medida de probabilidade $P$ em $\mathcal{B}(\mathbb{R}^\infty)$ que cumpre

$$P\Big(C(1, 2, \ldots, n; B)\Big) = P_n(B),$$

$B \in \mathcal{B}(\mathbb{R}^n)$, $n = 1, 2, \ldots$.
<u>Dem:</u> exercício □

Outro exemplo de probabilidades em $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$ que merece destaque é o seguinte:

**Corolário 1.12:** Se para cada $t \in \mathbb{T}$, $P_t$ é uma medida de probabilidade em $\mathcal{B}(\mathbb{R})$ existe uma única medida de probabilidade de $P$ em $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$ tal que para cada $u = \{t_1, \ldots, t_n\} \subset \mathbb{T}$,

$$P\Big(\pi_u^{-1}(B)\Big) = \Big(P_{t_1} \times P_{t_2} \times \ldots \times P_{t_n}\Big)(B),$$

$B \in \mathcal{B}(\mathbb{R}^{|u|})$.
<u>Dem:</u> exercício

A probabilidade $P$ no Corolário 1.12 é usualmente denotada por

$$P = \times_{t \in \mathbb{T}} P_t$$

Uma discussão mais detalhada do espaço mensuravel $(\mathbb{R}^{\mathbb{T}}, \mathcal{C}(\mathbb{R}^{\mathbb{T}}))$ é apresentada nos livros de Shiryayev [S] e Billingsley [B]. Gostariamos no entanto de enfatizar alguns pontos.

Se $\mathbb{T}$ é não contável e $A \in \mathcal{C}(\mathbb{R}^{\mathbb{T}})$, pode ser mostrado (Shiryayev [S]. Teorema II.2.3) que $A$ é determinado por um número contável de restrições, isto é, existem $t_1, t_2, \ldots$ em $\mathbb{T}$ e $B \in \mathcal{B}(\mathbb{R}^{\infty})$ tal que $A = \{x \in \mathbb{R}^{\mathbb{T}} : (x(t_n) : n \in \mathbb{N}) \in B\}$. Assim sendo, conjuntos importantes como $\{x \in \mathbb{R}^{\mathbb{T}} : x \text{ é continua}\}$, $\{x \in \mathbb{R}^{\mathbb{T}} : x \text{ é limitado }\}$, etc.., não são mensuráveis com respeito a $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$ (não pertencem a $\mathcal{C}(\mathbb{R}^{\mathbb{T}})$). Isto nos leva a considerar espaços de função menores que $\mathbb{R}^{\mathbb{T}}$ como no próximo exemplo.

Outro ponto a ser destacado é que o Teorema de Extensão de Kolmogorov continua válido se em lugar de $(\mathbb{R}^{\mathbb{T}}, \mathcal{C}(\mathbb{R}^{\mathbb{T}}))$ tomamos $(\times_{t \in \mathbb{T}} \Omega_t, \times_{t \in \mathbb{T}} \mathcal{F}_t)$, onde $\times_{t \in \mathbb{T}} \Omega_t = \{w = (w(t); \in \mathbb{T}) : w(t) \in \Omega_t\}, X_{t \in \mathbb{T}} \ \mathcal{F}_t$ é a $\sigma$-álgebra gerada pelos conjuntos $\{w \in \ X_{t \in \mathbb{T}} \Omega_t : w(t_1) \in B_1, \ldots, w(t_n) \in B_n\}$, $B_i \in \mathcal{F}_i, \Omega_t$ é um espaço métrico separável e $\mathcal{F}_t = B(\Omega_t)$. Em particular, Corolários 1.11 e 1.12 são válidos para qualquer $(\Omega_t, \mathcal{F}_t)$.

**Exemplo 4:** Probabilidades em $C[0,1]$. Seja $C[0,1]$ o espaço das funções reais contínuas definidas no intervalo $[0,1]$. $C[0,1]$ é um espaço métrico com métrica dada por

$$d(f,g) = \sup_{0 \le t \le 1} \Big| f(t) - g(t) \Big| \,.$$

A classe $\mathcal{A}$ formada pelo vazio e por uniões finitas de conjuntos disjuntos da forma

$$C\Big(t_1, \ldots, t_n; I_1 \times \ldots \times I_n\Big) = \Big\{ f \in C[0,1] : \Big(f(t_1), \ldots, f(t_n)\Big) \in I_1 \times \ldots \times I_n \Big\}$$

onde $n \in \mathbb{N}, t_1, \ldots, t_n \in [0,1], I_i = (a_i, b_i], -\infty \leq a_i < b_i \leq \infty, i = 1, 2, \ldots, n$, é uma álgebra $\sigma(\mathcal{A}) = \mathcal{B}(C[0,1])$.

A demonstração do resultado seguinte é similar aos teoremas anteriores (ver por exemplo Kuo [K]).

**Teorema 1.13:** (Medida de Wiener) Existe uma única medida de probabilidade $P_w$ em $\mathcal{B}(C[0,1])$ tal que para todo $n \in \mathbb{N}, t_i \in [0,1]$ e $I_i = (a_1, b_i] - \infty \leq a_i < b_i \leq \infty, i = 1, 2, \ldots, n$ (sem perda de generalidade seja $t_1 < t_2 < \ldots < t_n$),

$$P_w(C(t_1, \ldots, t_n; I_1 \times \ldots \times I_n)) =$$

$$= \left[(2\pi)^n t_1(t_2 - t_1) \ldots (t_n - t_{n-1})\right]^{-1/2} \int_{a_1}^{b_1} \ldots \int_{a_i}^{b_i}$$

$$\exp\left(-1/2\left[\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \ldots + \frac{(x_n - x_{n-1})^2}{(t_n - t_{n-1})}\right]dx_n \ldots dx_1.\right.$$

**Definição 1.14:** A medida de probabilidade $P_w$ é chamada medida de Wiener.

## 1.3 Independência.

O conceito de independência a ser definido a seguir particulariza a teoria de probabilidade como um ramo distinto na teoria geral de medida.

**Definição 1.15:** Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade, dizemos que os eventos $A$ e $B$ em $\mathcal{F}$ são <u>independentes</u> se:

$$P\left(A \bigcap B\right) = P(A)P(B).$$

Uma classe de eventos $\mathcal{E} \subset \mathcal{F}$ é chamada uma classe de eventos independentes se para toda coleção finita de eventos $A_1, A_1, \ldots A_n$ em $\mathcal{E}$,

$$P\Big(\bigcap_{k=1}^{n} A_k\Big) = \Pi_{k=1}^{n} P(A_k).$$

**Definição 1.16:** Se $\mathcal{E}_\lambda \subset \mathcal{F}$ é uma classe de eventos para $\lambda$ pertencente a um conjunto de índices $\wedge$, dizemos que $\{\mathcal{E}_\lambda : \lambda \in \wedge\}$ é uma família de classes independentes se para cada seleção de $A_\lambda \in \mathcal{E}_\lambda$, a classe $\{A_{\lambda:\lambda\in\wedge}\}$ é de eventos independentes.

**Lema 1.17:** Se $A_1, A_2 \ldots, A_n$ são independentes, então $A_1^c, A_2^c, \ldots, A_n^c$ são independentes.

**Lema de Borel-Cantelli:** Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade, $(E_n : n \in I\!N)$ uma sequência de eventos e $E = \overline{\lim_{n\to\infty}} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$.
a) $P(E) = 0$, se $\sum_{n=1}^{\infty} P(E_n) = 0$;
b) Para $E_n's$ independentes. $P(E) = 1$, se $\sum_{n=1}^{\infty} P(E_n) = \infty$.
Dem: a) $E = \bigcap_{n=1}^{\infty} F_n$, onde $F_n = \bigcup_{k=1}^{\infty} E_k$. Para $n \geq 1$, $0 \leq P(F_n) \leq \sum_{m=n}^{\infty} P(E_n) \to 0$ quando $n \to \infty$. Como $F_u \downarrow E$, temos que $P(E) = \lim_{n\to\infty} P(F_n) = 0$.
b) $E^c = \bigcup_{k=n}^{\infty} E_n^c$ e $E_1^c, E_2^c, \ldots$ são independentes. Seja $1 \leq n < N$.

$$\begin{aligned} P\Big(\bigcup_{k=n}^{\infty} E_k^c\Big) &\leq P\Big(\bigcup_{k=n}^{N} E_k^c\Big) \\ &= \pi_{k=n}^{N} P\Big(E_k^c\Big) \\ &= \pi_{k=1}^{N}\Big(1 - P(E_k)\Big) \\ &\leq \exp\Big(-\textstyle\sum_{k=n}^{N} P(E_k)\Big). \end{aligned}$$

Fazendo $N \to \infty$

$$P\Big(\bigcap_{k=n}^{\infty} E_k^c\Big) \leq \exp\Big(-\sum_{k=n}^{\infty} P(E_k)\Big).$$

Como $\sum_{k=n}^{\infty} P(E_k) = \infty$, segue que $P(\bigcap_{m=n}^{\infty} E_m^c) = 0$. Portanto $P(E^c) = 0 = 1 - P(E)$. $\square$

## 1.4 Elementos aleatórios e variáveis aleatórias.

Sejam $(\Omega, \mathcal{F}, P)$ e $(\Omega', \mathcal{F}')$ um espaço de probabilidade e um espaço mensurável respectivamente. Uma aplicação $X : \Omega \to \Omega'$ que é $\mathcal{F} - \mathcal{F}'$ mensurável, $(X^{-1}(B) = \{w \in \Omega; X(w) \in B\} \in \mathcal{F}$ para todo $B \in \mathcal{F}')$, é chamada um elemento aleatório com valores em $\Omega'$ (notação: $(X : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{F}'))$. Quando $\Omega' = \mathbb{R}(\mathbb{R}^n)$ e $\mathcal{F}' = \mathcal{B}(\mathbb{R})$ $(\mathcal{B}(\mathbb{R}^n)$ o elemento aleatório $X$ é chamado variável (vetor) aleatória.

Um elemento aleatório $X$ com valores em $\Omega'$ induz uma medida de probabilidade $P_X$ em $(\Omega' \mathcal{F}')$ através da relação $P_X(B) = P(X^{-1}(B))\cdot P_X$ é chamada distribuição de $X$. Reciprocamente, se $P$ é uma medida de probabilidade em $(\Omega', \mathcal{F}')$ existe um elemento aleatório com valores em $\Omega'$ tal que $P_X = P$. Para tanto seja $(\Omega, \mathcal{F}) = (\Omega', \mathcal{F}')$ e $X(\omega) = \omega$. Segue do Exemplo 1 (2) que uma variável (vetor) aleatória $X$ esta associada de maneira única a uma função distribuição em $\mathbb{R}(\mathbb{R}^n)$ que é chamada função distribuição de $X$. É fácil ver que se $F$ é a função distribuição de uma variável aleatória $X$,

$$F(x) = P\Big(X^{-1}(-\infty, x)\Big) = P\Big(X \leq x\Big).$$

A seguinte proposição é util para mostrar que uma aplicação é um elemento aleatório.

**Proposição 1.18:** Seja $X : \Omega \to \Omega'$ e $\mathcal{E}$ uma classe de subconjuntos de $\Omega'$ tal que $\mathcal{F}' = \sigma(\mathcal{E})$. Uma condição necessária e suficiente para que $X$ seja um elemento aleatório é que $X^{-1}(E) \in \mathcal{F}$ para todo $E \in \mathcal{E}$.

Dem: A necessidade da condição é clara. Para provar que a condição é suficiente definimos

$$\mathcal{D} \doteq \left\{B \in \mathcal{F}' = \sigma(\mathcal{E}); X^{-1}(B) \in \mathcal{F}\right\}.$$

E fácil mostra que $\mathcal{D}$ é uma $\sigma$-algebra. Como $\mathcal{E} \subset \mathcal{D}$, temos que $\sigma(\mathcal{E}) \subset \mathcal{D}$. Portanto $\mathcal{D} = \sigma(\mathcal{E}) = \mathcal{F}'$. □

**Corolario 1.19:** Uma condição necessária e suficiente para $X : \Omega \to \mathbb{R}$ ser uma variável aleatória é que $[X \leq x] = X^{-1}((-\infty, x]) \in \mathcal{F}$ para todo $x \in \mathbb{R}$.

Dem: exercício. □

**Proposição 1.20:** Seja $\varphi : \Omega' \to \mathbb{R}, \mathcal{F} - \mathcal{B}(\mathbb{R})$ mensurável e $X : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$. Então $Y = \varphi \circ X$ é uma variável aleatória.

Dem: exercício. □

**Corolario 1.21:** Seja $X$ uma variável aleatória. Então $X^p, X^+ = \max(X,0)$, $X^- = -\min(X,0)$ e $|X|$ são variáveis aleatórias.

Dem: exercício. □

**Corolario 1.22:** Se $(X_1, X_2)$ é um vetor aleatório então são variáveis aleatórias:

i) $X_i i = 1,2$

ii) $X_1 \pm X_2$;

iii) $\alpha X_i, \alpha \in \mathbb{R}$;

iv) $X_1X_2$;

v) $X_1/X_2$, se $P(X_2>0)=1$.

Dem: exercício

Quando trabalhamos com operações que envolvem limites de sequências de variáveis aleatórias podemos nos deparar com a ocorrencia de valores $+\infty$ e $-\infty$. Por esta razão é conveniente consideramos, $\overline{I\!R}=[-\infty,\infty]$. Um elemento aleatório $X:(\Omega,\mathcal{F})\to(\overline{I\!R},\mathcal{B}(\overline{I\!R}))$ é chamado variável aleatória estendida.

**Proposição 1.23:** Seja $(X_n; n\in I\!N)$ uma sequência de variáveis aleatórias estendidas. Então; também são variáveis aleatórias estendidas $\sup_n X_n$, $\inf\limits_n X_n$, $\overline{\lim}_{n\to\infty} X_n$ e $\underline{\lim}_{n\to\infty} X_n$.

Dem: exercício

Uma variável aleatória $X$ em $(\Omega,\mathcal{F},P)$ é chamada simples se

$$X(\omega)=\sum_{k=1}^{n} x_k \mathbb{1}_{A_k}(\omega),$$

onde $A_1,A_2,\ldots,A_n\in\mathcal{F}$ e $\mathbb{1}_A$ é a função indicadora do conjunto $A$ ($\mathbb{1}_A(\omega)=1$ se $\omega\in A$ e igual a zero se $\omega\notin A$). O próximo lema, apesar de simples, é importante para o desenvolvimento da próxima seção.

**Lema 1.29:** Toda variável aleatória $X$ (incluindo as estendidas) é limite pontual de uma sequência de variáveis aleatórias simples $X_n$ tal que $|X_n|\leq|X|$.

Dem: Seja $X\geq 0$. Defina para $n=1,2,\ldots$

$$X_n(\omega) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{A_{n,k}}(\omega) + n\mathbb{1}_{A_{n,n2^n+1}}(\omega),$$

onde

$$A_{n,k} = \left\{\omega : \frac{k-1}{2^n} \leq X(\omega) < \frac{2}{2^n}\right\} \text{ e } A_{n,n2^n+1} = \left\{\omega : X(\omega) \geq n\right\}.$$

É fácil ver que $X_n$ é não decrescente e $X_n(\omega) \nearrow X(\omega)$.

Para $X$ arbitrária, o resultado segue pois $X = X^+ - X^-, |X| = X^+ + X^-$ e $X^+$ e $X^-$ são não negativas. □

Para $X : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{F}')$ um elemento aleatório, definimos

$$\sigma(X) = \left\{X^{-1}(B) : B \in \mathcal{F}'\right\}$$

É fácil ver que $\sigma(X)$ é uma $\sigma$-álgebra contida em $\mathcal{F}$ e é a menor sub-$\sigma$-álgebra de $\mathcal{F}$ que faz $X$ um elemento aleatório. $\sigma(X)$ é chamada <u>$\sigma$-álgebra gerada por $X$</u>.

**Proposição 1.25:** Se $Y$ é uma variável aleatória que é mensurável com respeito a $\sigma(X)$, existe uma função $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ Borel mensurável tal que $Y = \varphi_0 X$.

<u>Dem:</u> Seja $Y = \mathbb{1}_A, A \in \sigma(X)$. Então,

$$Y(\omega) = \mathbb{1}_A(\omega) = \mathbb{1}_{X^{-1}(B)} = \mathbb{1}_B(X(\omega)) = (\mathbb{1}_B \circ X)(\omega) \quad \text{i.e} \quad \varphi = \mathbb{1}_B.$$

Para $Y$ arbitrário o resultado e uma consequência do Lema 1.24. □

Dizemos que $X_n, n = 1, 2, \dots$ são <u>elementos aleatórios independentes</u> se $\sigma(X_n), n = 1, 2, \dots$ são $\sigma$-álgebras independentes, ou seja, para $B_1, B_2, \dots, B_k$ em $\mathcal{F}', k < \infty$, os eventos $X_1^{-1}(B_1), \ldots, X_k^{-1}(B_k)$ são independentes.

**Teorema 1.26:** Uma condição necessária e suficiente para que as variáveis aleatórias $X_1, X_2, \ldots X_n$ sejam independentes é

$$P\Big(X_1 \leq x_1, \ldots, X_n \leq x_n\Big) = P\Big(X_1 \leq x_1\Big) \ldots P\Big(X_n \leq x_n\Big)$$

para todo $(x_1, \ldots, x_n) \in I\!R^n$.
Dem: exercício

Finalizando a discussão do conceito de independência de elementos aleatórios observamos que o Corolário 1.11 garante a existência de uma sequência $X_1, X_2, \ldots$ de elementos aleatórios independentes. Para tanto basta tomar

$$P_n = P_{X_1} \times P_{X_2} \times \ldots \times P_{X_n}.$$

Este comentário mostra a conexão entre independência de elementos aleatórios e medidas produto.

A convergência de elementos aleatórios é um importante capítulo da teoria de probabilidade, no entanto nesta monografia usaremos apenas conceitos amplamente estudados em teoria da medida e integração. Assim sendo, somente definiremos aqui o material necessário usando a linguagem probabilística.

Uma sequência de variáveis aleatórias $\{X_n : n \in I\!N\}$ é dita ser convergente com probabilidade 1 (ou q.s.) para uma variável aleatória $X$ (e escrevemos $X_n \xrightarrow{q.c} X$), se $P(\underline{\lim}_{n\to\infty} X_n(\omega) < \overline{\lim} X_1(\omega))$. A sequência é dita ser convergente em probabilidade a $X$ ($X_n \xrightarrow{P} X$), se $P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$, para todo $\varepsilon > 0$. Finalmente um outro conceito importante é o de convergência em média $p$ : $X_n \xrightarrow[n\to\infty]{L_p} X$, se $E|X_n - X|^p \xrightarrow[n\to\infty]{} 0$.

# 5. Esperança e Esperança Condicional.

## 5.1 Esperança matemática. Integral abstrata de Lebesgue.

Nesta secção $(\Omega, \mathcal{F}, P)$ é um espaço de probabilidade fixo e, exceto quando mencionado, as variáveis aleatórias assumem valores em $\overline{I\!R}$.

$X$ é uma variável aleatória simples (ou discreta com um número de valores finito) se:

$$X(\omega) = \sum_{k=1}^{n} x_n \mathbb{1}_{A_n}(\omega),$$

onde $A_n = \{\omega : X(\omega) = x_n\} \in \mathcal{F}$. A esperança matemática (ou simplesmente esperança) de $X$ denotada por $E(X)$, é o número real estendido

$$E(X) = \sum_{k=1}^{n} x_n P(A_n).$$

A esperança de $X$ é portanto a integral abstrata de Lebesgue de função mensurável simples $X$ com relação a medida $P$. Como na teoria geral de integração de Lebesgue, podemos definir a esperança matemática de uma variável aleatória não negativa $X$ como:

$$EX = \lim_{n\to\infty} EX_n\ ,$$

onde $(X_n, n = 1, 2, \ldots)$ é uma sequência não decrescente de variáveis aleatórias simples $X_n \nearrow X$ (ver Lema 1.24). O valor do limite é independente da sequência $(X_n; n = 1, 2, \ldots)$.

No caso geral, dizemos que a esperança $EX$ de uma variável aleatória arbitrária existe se

$$\min\Big(EX^+, EX^-\Big) < \infty.$$

Neste caso definímos

$$EX = EX^+ - EX^-.$$

Assim sendo, todas as propriedades da integral de Lebesgue com respeito a uma medida finita, são satisfeitas pela esperança matemática.

Seguindo a teoria geral de integração, usaremos em adição a $EX$ e $E(X)$ as notações $\int XdP$, $\int X(\omega)\ P(d\omega)$, etc.

Recordemos que uma família de variáveis aleatórias $\{X_\alpha : \alpha \in A\}$, é dita ser uniformemente integral se

$$\sup_\alpha \int_{\{|X_\alpha|>\lambda\}} |X_\alpha| dP \xrightarrow[\lambda\to\infty]{} 0.$$

Além das usuais desigualdades estudadas em teoria de integração: Hölder, Cauchy-Buniakowski, Minkowski, etc.., daremos aqui duas desigualdades, estudadas usualmente em teoria de probabilidade.

**Proposição 1.27:** (Desigualdade de Chebyshev). Se $E(|X|) < \infty$, então para todo $\varepsilon > 0$

$$\varepsilon P(|X|) > \varepsilon) \leq E(X)$$

<u>Dem:</u> $E(|X|) = \int |X|dP = \int_{|X|>\varepsilon} |X|dP + \int_{|X|<\varepsilon} |X|dP \geq$

$$\geq \int_{|X|>\varepsilon} |X|dP > \varepsilon \int_{|X|>\varepsilon} dP = \varepsilon P(|X| > \varepsilon).$$

**Proposição 1.28:** (Desigualdade de Jensen). Se $E(|X|) < \infty$, $\varphi$ é uma função convexa em um intervalo contendo a imagem $X$ e $E(|\varphi(X)|) < \infty$, então

$$\varphi\big(E(X)\big) \leq E\big(\varphi(X)\big).$$

Dem: Seja $L(x) = ax+b$ uma reta tangente a curva $(x, \varphi(x))$ no ponto $\big(E(X), \varphi E(X)\big)$. Claro, $L(x) \leq \varphi(x)$. Então $aX(\omega) + b \leq \varphi\big(X(\omega)\big)$, e portanto $aE(X) + b \leq E\big(\varphi(X)\big)$. O resultado segue uma vez que o lado esquerdo da última desigualdade é exatamente $\varphi\big(E(X)\big)$. □

## 5.2 Esperança condicional.

Fixamos um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e seja $A \in \mathcal{F}$ um evento com $P(A) > 0$. A probabilidade condicional de $B$ com relação a $A$, denotada por $P(B/A)$, é definida por

$$P(B/A) = \frac{P(A \cap B)}{P(A)}.$$

Na modelagem probabilística a probabilidade de $P$ mede o grau de incerteza associado com um dado evento $B$. Informações adicionais a respeito do experimento, como por exemplo a ocorrência de um evento $A$, podem modificar o grau de incerteza associado ao evento $B$. Quando tomamos a probabilidade condicional com respeito a $A$ é como se tivessemos tomando $A$ como o novo espaço básico no lugar de $\Omega$ e normalizando, para fazermos a probabilidade de $A$ igual a 1.

No caso dos eventos $A$ e $B$ serem independentes temos que $P(B/A) = P(B)$.

Se $X$ é uma variável aleatória com $E|X| < \infty$ definimos a esperança de $X$ com respeito a $A$, por $E(X\mathbb{1}_A)/P(A)$, que é simplesmente a esperança usual com respeito a medida de probabilidade $P(\cdot/A)$.

A definição de probabilidade condicional acima pode ser facilmente estendida para o condicionamento com respeito a uma partição mensurável, finita ou contável, $\mathcal{D} = \{A_1, A_2, \dots\}$ de $\Omega$ $(A_i \cap A_j = \phi, \bigcup A_i = \Omega, A_i \in \mathcal{F})$ com $P(A_i) > 0$. Para tanto, definimos a probabilidade condicional de $B$ com respeito a $\mathcal{D}$, denotada por $P(B/\mathcal{D})$, como sendo a variável aleatória,

$$P\Big(B/\mathcal{D}\Big)(\omega) = \sum_{c\geq 1} P\Big(B/\mathcal{D}\Big)\mathbb{1}_{D_i}(\omega)$$

que é claramente mensurável com respeito a $\sigma(\mathcal{D})$. Neste caso definimos $E(X/\mathcal{D}) = \sum_{i\geq 1} E(X/D_i)\mathbb{1}_{D_i}$, se $E|X| < \infty$.

A extensão do conceito de probabilidade condicional envolvendo eventos de probabilidade zero é mais delicada e a definição que daremos a seguir depende do Teorema de Radon-Nikodym. Recordemos que se $P_1$ e $P_2$ são duas medidas de probabilidade em $(\Omega, \mathcal{F})$, $P_1$ é absolutamente contínua com respeito a $P_2$ ($P_1 \ll P_2$) se $P_1(A) = 0$ para todo $A \in \mathcal{F}$ para o qual $P_2(A) = 0$. Neste caso o Teorema de Radon-Nikodym estabelece que existe uma função $\mathcal{F}$ mensurável $dP_1/dP_2$, não negativa e única q.c. $P_2$ que cumpre

$$P_1(A) = \int_A \frac{dP_1}{dP_2}\, dP_2$$

para todo $A \in \mathcal{F}$.

Seja $(\Omega, \mathcal{F}, P)$ um, espaço de probabilidade fixo, $\mathcal{G}$ uma sub-$\sigma$-álgebra de $\mathcal{F}$ ($\mathcal{G} \subseteq \mathcal{F}$) e $X$ uma variável aleatória não negativa. Definimos esperança condicional de $X$ com respeito a $\sigma$-álgebra $\mathcal{G}$, como sendo qualquer função $Y : \Omega \to \mathbb{R}$ que satisfaz:

i) $Y$ é $\mathcal{G}$ mensurável e

ii) $\int_A X\, dP = \int_A Y\, dP$, para todo $A \in \mathcal{G}$.

Tal $Y$ será denotado por $E(X/\mathcal{G})$ (ou $E(X/\mathcal{G})(\omega)$)

A existência da esperança condicional $E(X/\mathcal{G})$ é garantida pelo Teorema de Radon Nikodym. Para tanto observemos que,

$$Q(A) = \int_A X\, dP$$

é uma medida em $(\Omega, \mathcal{F})$ que é absolutamente contínua com respeito a $P$. Se restringirmos estas duas medidas a $\mathcal{G}$, segue por Radon-Nikodym que existe $dQ/dP$, $\mathcal{G}$ mensurável tal que $\int_A X dP = \int_A \frac{dQ}{dP}\, dP$ para todo $A \in \mathcal{G}$. Tomemos então $E(X/\mathcal{G}) = \frac{dQ}{dP}$.

Quando $X = \mathbb{1}_B$, $B \in \mathcal{F}$, $E(\mathbb{1}_B/\mathcal{G})$ é denotada por $P(A/\mathcal{G})$ e é chamada probabilidade condicional do evento $B$ com respeito a $\sigma$-álgebra $\mathcal{G}$. Note que $P(\cdot/\mathcal{G})(\omega)$ não é em geral uma medida de probabilidade em $\mathcal{F}$ para cada $\omega$, pois $P(B/\mathcal{G})$ é definido somente em um conjunto de probabilidade 1 que depende de $B$. Para que $P(\cdot/\mathcal{G})(\omega)$ seja uma medida de probabilidade em $\mathcal{F}$, ela tem que satisfazer, por exemplo

$$P\Big(\bigcup_{n=1}^{\infty} B_n/\mathcal{G}\Big)(\omega) = \sum_{n=1}^{\infty} P\big(B_n/\mathcal{G}\big)(\omega)$$

para toda sequência de eventos disjuntos em $\mathcal{F}$. Como o número de sequências disjuntas é em geral não contável e intersecção não contável de conjuntos de probabilidade 1 não obrigatoriamente tem probabilidade 1 ou é mensurável, $P(\cdot/\mathcal{G})(\omega)$ não é necessariamente uma medida de probabilidade.

No caso em que $\mathcal{G} = \{\phi, A, A^c, \Omega\}$, $0 < P(A) < 1$, é fácil ver que:

$$P(B/\mathcal{G}) = P(B/A)1_A + P(B/A^c)1_{A^c}.$$

Outro caso particular que merece ser comentado é quando $\mathcal{G} = \{\phi, \Omega\}$, neste caso

$$E\big(X/\mathcal{G}\big) = E\big(X\big).$$

A definição da esperança condicional de uma variável aleatória não necessariamente não negativa com respeito a uma $\sigma$-álgebra $\mathcal{G}$ é definida por

$$E\big(X/\mathcal{G}\big) = E\big(X^+/\mathcal{G}\big) - E\big(X^-/\mathcal{G}\big)$$

se $\min\big(E(X^+/\mathcal{G}), E(X^-/\mathcal{G})\big) < \infty$.

Esta definição de esperança condicional não é geralmente adotada. Ela aparece no livro de Shiryayev (1984), e tem a vantagem de não assumir a existência da $E(X)$. Quando $EX$ existe e $\mathcal{G} = \{\phi, \mathcal{F}\}$ temos que $E(X/\mathcal{G}) = E(X)$. No entanto $E(X)$ pode não existir ($EX^+ = EX^- = \infty$) e se $\mathcal{G} = \mathcal{F}$, $E(X/\mathcal{F}) = X^+ - X^-$.

A definição mais usual de esperança condicional com respeito a $\sigma$-algebras assume que $E|X| \leq \infty$ e usa o Teorema de Radon-Nikodym diretamente na medida sinal $\int_A XdP$.

Se consideramos a $\sigma$-álgebra $\mathcal{G}$ como sendo gerada por um elemento aleatório $Y$, $\mathcal{G} = \sigma(Y)$, então $E(X/\mathcal{G})$ será denotada por $E(X/Y)$ e é chamada esperança condicional de $X$ com respeito a $Y$. Neste caso como $E(X/Y)$ é $\sigma(Y)$ mensurável $E(X/Y) = \varphi \circ Y$, onde $\varphi$ é Borel mensurável. A função $\varphi(y)$ é denotada por $E(X/Y = y)$ é chamada esperança condicional de $X$ dado $Y = y$. Note que $E(X/Y)$ é uma função definida em $\Omega$ enquanto $E(X/Y = \cdot)$ é uma função definida em $\Omega'$, o domínio do elemento aleatório $Y$.

O seguinte lema é útil no cálculo de esperanças condicionais.

**Lema 1.29:** Seja $X$ uma variável aleatória, $E|X| < \infty$ e $\mathcal{G} = \sigma(\mathcal{A})$, onde $\mathcal{A}$ é uma álgebra. Então $\mathcal{G}$ é uma versão e $E(X/\mathcal{G})$, se

i) $g$ é $\mathcal{G}$ mensurável e

ii) $\int_A g\, dP = \int_A XdP$, para todo $A \in \mathcal{A}$.

Dem: Teorema de Caratheodory

A definição de esperança condicional com respeito a um $\sigma$-álgebra estende a definição dada anteriormente de esperança condicional com respeito a uma partição. Consideremos $\mathcal{D} = \{D_1, D_2 \ldots\}$ uma partição mensurável, finita ou contável, com $P(D_i > 0)$.

**Teorema 1.30:** Se $\mathcal{G} = \sigma(\mathcal{D})$ e $X$ é uma variável aleatória com $E|X| < \infty$, então

$$E\big(X/\mathcal{G}\big) = \sum_{i \geq q} E\big(X/D_i\big)\mathbb{1}_{Di} = \sum_{i>1} \frac{E(X\mathbb{1}_{Di})}{P(D_i)}\mathbb{1}_{Di}$$

Dem: Como $E(X/\mathcal{G})$ é $\sigma(\mathcal{D})$ mensurável, temos que $E(X/\mathcal{G}) = \sum_{1\geq 1} d_i \mathbb{1}_{D_i}$. Para $A \in \sigma(\mathcal{D})$,

$$\int_A X dP = \int_{\cup_{i\geq 1}(A\cap D_i)} X dP$$

$$\stackrel{T.C.D.}{=} \sum_{i\geq 1} \int_{A\cap D_i} \int_X dP =$$

$$= \sum_{i\geq 1} \int_{A\cap D_i} E(X/\sigma(\mathcal{D}))dP =$$

$$= \sum_{i\geq 1} d_i \int_{A\cap D_i} dP =$$

$$= \sum_{i\geq 1} \int_A d_i \mathbb{1}_{Di} dP = \int_A \sum_{i\geq 1} d_i \mathbb{1}_{D_i} dP \qquad \square$$

## Propriedades de Esperança Condicional.

**Teorema 1.31:** Seja $X$ uma variável aleatória com $E|X| < \infty$ e seja $\mathcal{G} \subset \mathcal{F}$ uma $\sigma$-álgebra.

(a) Se $X$ é $\mathcal{G}$-mensurável, então $E(X/\mathcal{G}) = X$ q.c..

(b) Se $X = C$ q.c., onde $C$ é uma constante então $E(X/\mathcal{G}) = C \quad$ q.c..

(c) Se $X \geq 0$ q.c., então $E(X/\mathcal{G}) \geq 0$ q.c..

(d) $|E(X/\mathcal{G})| \leq E(|X|/\mathcal{G})$ q.c..

(e) Seja $Y$ uma outra variável aleatória com $E|Y| < \infty$, $a$ e $b$ constantes. Então, $E(aX + bY/\mathcal{G}) = aE(X/\mathcal{G}) + bE(Y/\mathcal{G})$ q.c..

(f) $E(E(X/\mathcal{G})) = E(X)$

(g) Se $\mathcal{G} = \{\phi, \Omega\}$, então $E(X/\mathcal{G}) = E(X)$.

Dem: Segue imediatamente de definição. $\square$

Teoremas limites como convergência monotona, fator e convergência dominada tem versões "condicionais" que resumiremos no seguinte teorema.

**Teorema 1.32:** Seja $(X_n; n \geq 1)$ uma sequência de variáveis aleatórias estendidas.

(a) Se $|X_n| \leq Y$, $EY < \infty$ e $X_n \xrightarrow[n\to\infty]{} X$ q.c.; então $E\big(|X_n - X|/\mathcal{G}\big) \xrightarrow[n\to\infty]{} 0$ q.c..

(b) Se $X_n \geq 0$ e $X_n \nearrow X$ q.c.; então $E\big(X_n/\mathcal{G}\big) \nearrow E\big(X/\mathcal{G}\big)$ q.c..

(c) Se $X_n \geq 0$, então $E\big(\underline{\lim}_{n\to\infty} X_n/\mathcal{G}\big) = \underline{\lim}_{n\to\infty} E\big(X_n/\mathcal{G}\big)$ q.c..

Dem:

(a) $E(|X_n - X|/\mathcal{G}) \leq E(\sup\limits_{m\geq n}|X_m - X|/\mathcal{G})$. Como $\sup\limits_{m\geq n}|X_m - X|$ é não crescente, $\lim\limits_{n\to\infty} E(\sup\limits_{m\geq n}|X_n - X|/\mathcal{G})$ existe q.c.. Segue que

$$
\begin{aligned}
0 \leq\ & \int \lim_{n\to\infty} E\Big(\sup_{m\geq n}|X_m - X|/\mathcal{G}\Big)dP \leq \\
\leq\ & \int E\Big(\sup_{m\geq n}|X_m - X|/\mathcal{G}\Big)dP = \\
=\ & \int \sup_{m\geq n}|X_m - X|dP \xrightarrow[n\to\infty]{} 0,
\end{aligned}
$$

onde a convergência a zero segue do teorema da convergência dominada uma vez que

$$\sup_{m\to n} |X_n - X| \leq 2Y.$$

(b) Como $E(X_n/\mathcal{G})$ é uma sequência não decrescente q.c., $\lim_{n\to\infty} E(X_n/\mathcal{G})$ existe q.c.. Para todo $A \in \mathcal{G}$

$$\int_A X_n dP = \int_A E(X_n/\mathcal{G})\ dP$$

e pelo teorema da convergência monótona

$$\int_A X dP = \int_A \lim X_n dP = \int_A \lim E(X_n/\mathcal{G})dP.$$

Consequentemente, para todo $A \in \mathcal{G}$

$$\int_A E(X/\mathcal{G})dP = \int_A \lim_{n\to\infty} E(X_n/\mathcal{G})dP$$

ou seja,

$$E(X/\mathcal{G}) = \lim_{n\to\infty} E(X_n/\mathcal{G})$$

(c) $\inf_{m\geq n} X_m \nearrow \underline{\lim}_{n\to\infty} X_n$. Aplicando (b), temos

$$E(\underline{\lim}_{n\to\infty} X_n/\mathcal{G}) = \lim_{n\to\infty} E(\inf_{m\geq n} X_m/g) = \underline{\lim}_{n\to\infty} E(\inf_{m\geq n} X_m/\mathcal{G}) \leq \underline{\lim}_{n\to\infty} E(X_n/\mathcal{G}).\square$$

Outras propriedades importantes da esperança condicional são apresentadas nas seguintes proprosições.

**Teorema 1.33:** Seja $X$ uma variável aleatória com $E|X| < \infty$. Se $\mathcal{G}_1$ e $\mathcal{G}_2$ são sub-$\sigma$-álgebras de $\mathcal{F}$ com $\mathcal{G}_1 \subset \mathcal{G}_2$, então

$$E\Big(E\big(X/\mathcal{G}_2\big)/\mathcal{G}_1\Big) = E\Big(E\big(X/\mathcal{G}_1\big)/\mathcal{G}_2\Big) = E\big(X/\mathcal{G}_1\big).$$

<u>Dem:</u> Como $\mathcal{G}_1 \subset \mathcal{G}_2$ e $E\big(X/\mathcal{G}_1\big)$ é $\mathcal{G}_2$ mensurável, $E(E(X/\mathcal{G}_1)/\mathcal{G}_2) = E(X/\mathcal{G}_1)$ q.c. Agora para todo $A \in \mathcal{G}_1$,

$$\int_A \Big(E\big(E(X)/\mathcal{G}2\big)/\mathcal{G}1\Big)dP = \int_A E\big(X/\mathcal{G}_2\big)dP$$

pela definição de esperança condicional. Como $\mathcal{G}_1 \subset \mathcal{G}_2$ temos que para

$$\int_A E\big(X/\mathcal{G}_2\big)dP = \int_A XdP = \int_A E\big(X/\mathcal{G}_1\big)dP$$

e portanto, $E\Big(E\big(X/\mathcal{G}_2\big)/\mathcal{G}_1\Big) = E\big(X/\mathcal{G}_1\big)$ $\square$

**Teorema 1.34:** Seja $X$ uma variável aleatória com $E|X| < \infty$ e suponhamos que $\sigma(X)$ é independente de $\mathcal{G}$. Então $E(X/\mathcal{G})$ é constante e igual a $E(X)$.

<u>Dem:</u> Como $\mathcal{G}$ e $\sigma(X)$ são independentes; $\mathbb{1}_E$ e $X$ são variáveis aleatórias independentes para todo $E \in \mathcal{G}$. Assim sendo

$$\int_E E(X/\mathcal{G})dP = \int_E XdP = \int \mathbb{1}_E\, XdP = \int 1_E dP \int XdP = \int_E dP\; EX = \int_E EX\; dP\,.$$

Portanto $E(X/\mathcal{G}) = EX$ q.c. □

**Corolário 1.35:** Para $X$ e $Y$ variáveis aleatórias independentes, $E(X/Y) = EX$.

**Teorema 1.36:** Seja $Y$ uma variável aleatória $\mathcal{G}$ mensurável, $E|Y| < \infty$ e $E|XY| < \infty$. Então

$$E\big(XY/\mathcal{G}\big) = YE\big(X/\mathcal{G}\big).$$

<u>Dem:</u> Primeiramente, seja $Y = \mathbb{1}_B$, para $B \in \mathcal{G}$. Então, para todo $A \in \mathcal{G}$

$$\int_A E(XY/\mathcal{G})dP = \int_A XYdP = \int_{A\cap B} XdP = \int_{A\cap B} E\big(X/\mathcal{G}\big)dP = \int_A 1_B\, E(X/\mathcal{G})dP\,.$$

Temos então que

$$E\big(XY/\mathcal{G}\big) = YE\big(X/\mathcal{G}\big) \text{ q.c.}$$

para $Y = \mathbb{1}_B$, $B \in \mathcal{G}$. O resultado permanece válido para variáveis aleatórias simples $Y = \sum_{k=1}^{n} y_k \mathbb{1}_{B_k}, B_k \in \mathcal{G}$. Para $Y$ uma *va.* $\mathcal{G}$ mensurável qualquer, seja $(Y_n; n \geq I\!N)$ uma sequência de $v'.s\ a's$ simples tal que $|Y_n| \leq Y|$ e $Y_n \to Y$. Segue que

$$E\big(XY_n/\mathcal{G}\big) = Y_n E\big(X/\mathcal{G}\big). \text{ q.c..}$$

Como $|XY_n| \leq |XY|$, pelo teorema da convergência monotona condicional

$$E\big(XY_n/\mathcal{G}\big) \xrightarrow[n\to\infty]{} E\big(XY/\mathcal{G}\big) \text{ q.c.}$$

Como $E|X| < \infty$ implica $E(X/\mathcal{G})$ finita q.c., temos então que:

$$Y_n E\big(X/\mathcal{G}\big) \xrightarrow[n\to\infty]{} Y E\big(X/\mathcal{G}\big) \quad \text{q.c.}$$

Portanto

$$E\big(XY/\mathcal{G}\big) = Y E\big(X/\mathcal{G}\big) \text{ q.c.}$$ □

## Probabilidade Condicional Regular.

Como comentamos anteriormente, se definimos $P(A/\mathcal{G})$ como sendo uma versão de $E(\mathbb{I}_A/\mathcal{G})$, a função $P(\cdot/\mathcal{G})(\omega) : \mathcal{F} \to \mathbb{R}$, $\omega \in \Omega$, não define uma probabilidade. Este fato leva a seguinte definição

**Definição 1.37:** Uma função $P : \Omega \times \mathcal{F} \to \mathbb{R}$ é chamada uma probabilidade condicional regular com respeito a uma $\sigma$-álgebra $\mathcal{G} \subset \mathcal{F}$ se:

i) $P(\omega, \cdot) : \mathcal{F} \to \mathbb{R}$, é uma medida de probabilidade em $\mathcal{F}$ para cada $\omega \in \Omega$.

ii) $P(\cdot, A) : \Omega \to \mathbb{R}$, é uma versão de $P(A/\mathcal{G})$, isto é $P(\cdot, A) = P(A/\mathcal{G})$ q.c..

**Teorema 1.38:** Seja $P(\cdot, \cdot)$ uma probabilidade condicional regular com respeito a $\mathcal{G}$, e $X$ uma variável aleatória com $E|X| < \infty$. Então

$$E\big(X/\mathcal{G}\big)(\omega) = \int_\Omega X(\tilde{\omega}) P(\omega, d\tilde{\omega}) \text{ q.c.} \tag{1.9}$$

Dem: Seja $X = \mathbb{1}_B$, $B \in \mathcal{F}$, então (1.9) é verdade por definição. Consequentemente (1.9) vale para variáveis aleatórias simples. Seja agora $Y \geq 0$ e $X_n, n = 1, 2, \ldots$ uma sequência de variáveis aleatórias simples $X_n \nearrow X$. Temos então

$$\lim E\big(X_n/\mathcal{G}\big) = E\big(X/\mathcal{G}\big) \text{ q.c.}\,.$$

Agora

$$\begin{aligned}\lim E(X_n/\mathcal{G})(\omega) &= \lim \int_\Omega X_n(\tilde{\omega}), P(\omega, d\tilde{\omega}) \\ &= \int_\Omega X(\tilde{\omega})\ P(\omega, d\tilde{\omega}),\end{aligned}$$

onde a última igualdade segue pelo Teorema de convergência monótona.

O caso geral segue da representação $X = X^+ - X^-$ $\square$

Em geral não podemos garantir a existência de uma probabilidade condicional regular e o seguinte conceito, se faz então necessário.

**Definição 1.38:** Seja $(\Omega, \mathcal{F})$ um espaço mensurável, $X$ um elemento aleatório com valores em $\Omega'$ e $\mathcal{G}$ uma sub-$\sigma$-álgebra de $\mathcal{F}$. Uma função $Q_X : \Omega \times \mathcal{F}' \to \mathbb{R}$ é chamada uma distribuição condicional regular de $X$ com respeito a $\mathcal{G}$ se:

a) $Q_X(\omega, \cdot)$ é uma medida probabilidade em $\mathcal{F}'$ para cada $\omega \in \Omega$.

b) $Q_X(\cdot, B)$ é uma versão de $P(X^{-1}(B)/\mathcal{G})$ para todo $B \in \mathcal{F}'$, isto é: $Q_X(\omega, B) = P(X^{-1}(B)/\mathcal{G})(\omega)$ q.c..

**Definição 1.39:** Seja $X$ uma variável aleatória. Uma função $F_X : \Omega \times \mathbb{R} \to \mathbb{R}$ é chamada função distribuição condicional regular de $X$ com respeito a $\mathcal{G}$ se

a) $F(\omega,\cdot)$ uma função distribuição de probabilidade em $I\!R$ para cada $\omega \in \Omega$.

b) $F_X(\cdot,x)$ é uma versão de $P(X^{-1}((-\infty,x])/\mathcal{G})$ para cada $x \in I\!R$, isto é $F(\omega,x) = P(X^{-1}((-\infty,x]/\mathcal{G})(\omega)$ q.c.

Note que como $F_X(\omega,\cdot)$ é a função distribuição em $I\!R$, existe uma medida de probabilidade correspondente em $B(I\!R)$

**Teorema 1.40:** Uma função distribuição condicional regular de uma variável aleatória $X$ sempre existe.

Dem: Seja $\{r_n : n \in I\!R\}$ o conjunto dos racionais e $P(X \leq r_n/\mathcal{G})$ a probabilidade condicional de $X^{-1}((-\infty,r_n])$ com respeito a uma $\sigma$-álgebra $\mathcal{G} \subset \mathcal{F}$. Para $m,n \geq 1$, seja

$$A_{m,n} = \Big\{\omega : P\big(X \leq r_n/\mathcal{G}\big) < P\big(X \leq r_m/\mathcal{G}\big)\Big\}$$

e

$$A = \bigcup_{r_n > r_m} A_{m,n}.$$

Como $P\big(X \leq r_m/\mathcal{G}\big) \leq P\big(X \leq r_n/\mathcal{G}\big)$ q.c. se $r_m < r_n$, temos que $P(A) = 0$. Seja

$$B_n = \left\{\omega : \lim_{\substack{r_m \downarrow r_n \\ r_m > r_n}} P\big(X \leq r_m/\mathcal{G}\big) \neq P\big(X \leq r_n/\mathcal{G}\big)\right\}$$

e

$$B = \bigcup_n B_n\ .$$

Pelo teorema da convergência monotona $P(B_n) = 0$ e portanto $P(B) = 0$. Similarmente, se

$$C_1 = \Big\{\omega : \lim_{r_n \uparrow \infty} P\big(X \leq r_n/\mathcal{G}\big) \neq 1\Big\}$$

e

$$D = \left\{\omega : \lim_{r_n \downarrow -\infty} P\left(X \leq r_n / \mathcal{G}\right) \neq 0\right\}$$

$P(C) = P(D) = 0.$

Seja $\Omega_0 = A \cup B \cup C \cup D$. Então para $\omega \in \Omega_0^c$, a função $x - P(X \leq x/\mathcal{G})(\omega) = \lim_{r_n \downarrow x} P(X \leq R_n/\mathcal{G})$ é uma função distribuição de probabilidades. Seja $F_X : \Omega \times I\!R \to I\!R$ definida por

$$F_X(\omega, x) = \lim_{r_n \downarrow x} P\left(X \leq r_n/\mathcal{G}\right)(\omega) 1\!\!1_{\Omega_0^c}(\omega) + F(x) 1\!\!1_{\Omega_0}(\omega),$$

onde $F$ é uma função distribuição de probabilidade qualquer. É fácil ver que $F_X(\cdot,\cdot)$ é uma função distribuição condicional regular de $X$ dado $\mathcal{G}$. Se $Q_X(\omega,\cdot)$ é a medida de probabilidade em $\mathcal{B}(I\!R)$ correspondente a $F_X(\omega,\cdot)$, então para mostrar que $Q_X(\cdot,\cdot)$ é uma distribuição condicional regular de $X$ dado $\mathcal{G}$, note primeiramente que

$$Q_X(\omega, B) = \int_B F(\omega, dx) \ ,$$

onde a integral é de Lebesgue-Stieltjs. Logo $Q_X(\omega,\cdot)$ é $\mathcal{G}$-mensurável.

Seja $\mathcal{E}$ a classe de todos os conjuntos $B \in \mathcal{B}(I\!R)$ tal que

$$\int_A Q_X\left(\omega, X^{-1}(B)\right) P(d\omega) \ = \int_A 1\!\!1_{X^{-1}(B)} dP = P\left(A \bigcap X^{-1}(B)\right). \tag{1.10}$$

É fácil mostrar que $\mathcal{E}$ é uma $\sigma$-álgebra e que $(-\infty, x] \in \mathcal{E}$ para todo $x \in I\!R$. Portanto $\mathcal{E} = \mathcal{B}(I\!R)$ é a relação (1.10) vale para todo $B \in \mathcal{B}(I\!R)$. □

**Teorema 1.41:** Seja $X$ uma variável aleatória e $\varphi$ uma função Borel mensurável com $E(|\varphi(X)|) < \infty$. Então

$$E\Big(\varphi(X)/\mathcal{G}\Big)(\omega) = \int_{\mathbb{R}} \varphi(x) Q_X\Big(\omega, dx\Big).$$

Dem: exercício

A existência da distribuição condicional regular é possível para elementos aleatórios assumindo valores em espaço mensuráveis que são "semelhantes" a $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

**Definição 1.42:** Um espaço mensurável $(\Omega', \mathcal{F}')$ é chamado um espaço de Borel se existe uma aplicação $T : \Omega' \to \mathbb{R}$, um a um, tal que:

i) $T(\Omega') = \{T(\omega') : \omega' \in \Omega'\} \in \mathcal{B}(\mathbb{R})$.

ii) $T$ é $\mathcal{F}' - \mathcal{B}(\mathbb{R})$ mensurável.

iii) $T^{-1}$ definida em $T(\Omega')$ é $\mathcal{B}(T(\Omega')) - \mathcal{F}'$ mensurável.

**Teorema 1.43:** Seja $X$ um elemento aleatório com valores em um espaço de Borel $(\Omega', \mathcal{F}')$. Então existe uma distribuição condicional regular de $X$ com respeito a uma sub-$\sigma$-álgebra $\mathcal{F}$.

Dem: Seja $T$ como na Definição 1.42. Então $T \circ X$ é uma variável aleatória e $Q_{T\circ X}$ existe com respeito a $\mathcal{G}$. Seja $Q_X(\omega, B) = Q_{T\circ X}(\omega, T(B))$, $B \in \mathcal{F}'$. Pela definição 1.43 iii), $T(B) \in \mathcal{B}(T(\Omega'))$ e portanto $Q_X$ é bem definida. Claro $Q_X(\omega, \cdot)$ é uma medida em $\mathcal{F}'$ para todo $\omega \in \Omega$. Para $B \in \mathcal{F}'$, como $T$ é um a um

$$\begin{aligned} Q_X(\omega, B) = Q_{T\circ X}(\omega, T(B)) &= P\Big(T(X) \in T(B)/\mathcal{G}\Big) \\ &= P\Big(X \in B/\mathcal{G}\Big) \text{ q.c.,} \end{aligned}$$

o que complete a prova do teorema. $\square$

Pode se mostrar que todo espaço métrico completo e separável é um espaço de Borel. Em particular $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ são espaços de Borel.

# 2. Conceitos Básicos de Processos Estocásticos.

## 2.1 Introdução.

Um processo estocástico é uma estrutura constituida de um espaço de probabilidade $(\Omega, \mathcal{F}, P)$, um conjunto não vazio $\mathbb{T}$ e uma aplicação $X : \mathbb{T} \times \Omega \to \overline{\mathbb{R}}$, tal que para cada $t \in \mathbb{T}$, a função $X(t, \cdot) : \Omega \to \overline{\mathbb{R}}$ é uma variável aleatória. Em outras palavras, um processo estocástico é uma coleção de variáveis aleatórias definidas num espaço de probabilidade $(\Omega\ \mathcal{F}, P)$, indexadas por um conjunto $\mathbb{T}$.

Para $t \in \mathbb{T}, X_t$ ou $X(t)$ denotarão a variável aleatória $X(t, \cdot)$, isto é $X(t, \cdot) = X(t) = X_t$. A coleção de variáveis aleatórias $\{X(t) : t \in \mathbb{T}\}$ também será de notada por $X$. $\mathbb{T}$ é chamado espaço de índices ou parâmetros. Para cada $\omega \in \Omega$, a função $X(\cdot, \omega) : \mathbb{T} \to \overline{\mathbb{R}}$ é chamada trajetória, ou realização, ou função amostral correspondente a $\omega$.

**Definição 1.44:** Dois processos estocásticos $X$ e $X'$ com o mesmo conjunto de índices $\mathbb{T}$, definidos em $(\Omega, \mathcal{F}, P)$ e $(\Omega', \mathcal{F}', P')$ respectivamente, são chamados indistinguiveis ou equivalentes se para todo $n \in \mathbb{N}$, $t_1, \ldots, t_n \in \mathbb{T}$ e $B \in \mathcal{B}(\overline{\mathbb{R}}^n)$,

$$P\Big((X(t_1), \ldots, X(t_n)) \in B\Big) = P'\Big((X'(t_1), \ldots, X'(t_n)\Big) \in B\Big).$$

Um conceito mais forte que equivalência é possível quando os processos estão definidos no mesmo espaço de probabilidade

**Definição 1.45:** Dois processos estocásticos, $X$ e $X'$ com o mesmo espaço de índices

$\mathbb{T}$, definidos em $(\Omega, \mathcal{F}, P)$ são chamados versões ou modificações um do outro se

$$P\Big(X(t) = X'(t)\Big) = 1$$

para todo $t \in \mathbb{T}$, isto é $X(t) = X'(t)$ q.c. para todo $t \in \mathbb{T}$.

Se $E\Big(|X(t)|\Big) < \infty$ para todo $t \in \mathbb{T}$, a função $m_X : \mathbb{T} \to \mathbb{R}$, definida por $m_X(t) = EX(t)$ é chamada função média de $X$. Se $E|X(t)|^2 < \infty$ para todo $t \in \mathbb{T}$ a função $C_X : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, definida por $C_X(t,s) = E((X(t) - m_X(t))(X(s) - m_X(s)))$ é chamada função covariância de $X$.

Seja $\mathbb{T}$ um espaço metrico. $X = \{X(t) : t \in \mathbb{T}\}$ tem trajetórias contínuas q.c. se

$$P\Big(\Big\{\omega \in \Omega : X(\cdot, \omega) \quad \text{é contínua em } \mathbb{T}\Big\}\Big) = 1.$$

Para todo $t \in \mathbb{T}$, $X$ é dito ser contínuo em probabilidade em $t$ se para toda sequência $t_n, n = 1, 2, \ldots$ em $\mathbb{T}, t_n \to t$ quando $t \to \infty$

$$P\Big(\Big\{\omega : |X(t_n, \omega) - X(t, \omega)| > \varepsilon\Big\}\Big) \xrightarrow[n \to \infty]{} 0$$

para todo $\varepsilon > 0$.

Se $X$ é de segunda ordem, isto é, $E|X(t)|^2 < \infty$ para todo $t \in \mathbb{T}$, dizemos que $X$ é contínuo em média quadrática em $t \in \mathbb{T}$ se para toda sequência $t_n, n = 1, 2 \ldots,$ $t_n \to t$ quando $n \to \infty$

$$E\Big(|X(t_n) - X(t)|^2\Big) \xrightarrow[n \to \infty]{} 0$$

isto é, a aplicação $t \to X(t)$ com valores no espaço de Hilbert $L_2(\Omega, \mathcal{F}, P)$ é contínua.

## 2.2 Representação canônica de um processo estocástico.

Um processo estocástico $X$ pode ser visto como sendo uma medida de probabilidade $\mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}})$. Para tanto considere a aplicação $\Phi_X : \Omega \to \overline{\mathbb{R}}^{\mathbb{T}}$, definida por

$$\Phi_X(\omega) = X(\cdot, \omega) .$$

$\Phi_X$ é $\mathcal{F} - \mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}})$ mensurável pois $\Phi_X^{-1}(C) \in \mathcal{F}$ para todo cilindro $C$. Portanto $\Phi_X$ é um elemento aleatório com valores em $\overline{\mathbb{R}}^{\mathbb{T}}$.

Para $A \in \mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}})$, seja

$$\mu_X(A) = P\Big(\Phi_X^{-1}(A)\Big) = P\Big(\big\{\omega; X(\cdot,\omega) \in A\big\}\Big) .$$

Se definimos a aplicação $X' : \mathbb{T} \times \overline{\mathbb{R}}^{\mathbb{T}} \to \overline{\mathbb{R}}$ por $X'(t,x) = x(t)$, é fácil ver que $X'$ define um processo estocástico em $(\overline{\mathbb{R}}^{\mathbb{T}}, \mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}}), \mu_X)$ e que

$$P\Big(\big(X(t_1), \ldots, X(t_n)\big) \in B\Big) = \mu_X\Big(\big(X'(t_1), \ldots X'(t_n)\big) \in B\Big)$$

para todo $n \in \mathbb{N}$, $t_1, \ldots, t_n \in \mathbb{T}$ e $B \in \mathcal{B}(\overline{\mathbb{R}}^n)$. Em outras palavras $X$ e $X'$ são equivalentes.

**Definição 1.46:** A estrutura $\{\mathbb{T}, (\overline{\mathbb{R}}^{\mathbb{T}}, \mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}}), \mu_X), X'(t,x) = x(t)\}$ é chamada representação canônica do processo $X$.

As representações canônicas de dois processos estocásticos com o mesmo espaço de índices diferem somente na medida de probabilidade. Processos equivalentes tem a mesma representação canônica.

## 2.3 Especificação de um processo estocástico pelas suas distribuições de dimensão finita.

Seja $X = \{X(t); t \in \mathbb{T}\}$ um processo estocástico definido em $(\Omega, \mathcal{F}, P)$. Para todo subconjunto finito $u = \{t_1\ t_2, \ldots t_n\}$ de $\mathbb{T}$, a aplicação mensurável $\omega \to (Z(t_1,\omega), \ldots, X(t_n,\omega))$, induz uma medida de probabilidade $P_u$ em $(\overline{\mathbb{R}}^n, \mathcal{B}(\overline{\mathbb{R}}^n))$ : $P_u(A) = P[\omega \in \Omega : (X(t_1,\omega), \ldots X(t_n,\omega)) \in A]$, para todo $A \in \mathcal{B}(\overline{\mathbb{R}}^n)$.

**Definição 1.47:** O conjunto de medidas de probabilidade $\{P_u; u$ subconjunto finito de $\mathbb{T}\}$ é chamado <u>conjunto das distribuições de dimensão finita do processo $X$</u>.

Temos portanto que dois processos estocásticos equivalentes tem o mesmo conjunto de distribuições finitamente dimensionais.

É natural perguntarmos quando um conjunto de medidas de probabilidade é um conjunto de distribuições de dimensão finita de um processo estocástico. Esta pergunta é respondida pelo Teorema de Kolmogorov enunciado a seguir na linguagem de processos estocásticos.

**Teorema de Kolmogorov.** Seja $\mathbb{T}$ um conjunto não vazio e $\mathcal{P}\{P_u; u \subset \mathbb{T}, |u| < \infty, P_u$ probabilidade em $(\overline{\mathbb{R}}^{|u|}, \mathcal{B}(\mathbb{R}^{|u|}))\}$ uma família de medidas de probabilidade. Uma condição necessária e suficiente para que $\mathcal{P}$ seja o conjunto de distribuições de dimensão finita de um processo estocástico é que

$$P_u = P_n \pi_{u,v}^{-1}$$

para tanto $u \subset v \subset \mathbb{T}, v$ finito.

<u>Dem:</u> Seja $u = \{t_{i_1}, \ldots, t_{i_n}\} \subset \{t_1, \ldots t_n\} = v$. Seja $\mathcal{P}$ o conjunto de distribuições de

dimensão finita de um processo $X$. Se $A \in B(\mathbb{R}^k)$

$$P_u(A) = P\Big((X(t_{i_1}), \ldots, X(t_{1_n})) \in A\Big) = P\Big(\pi_{u,v}\big(X(t_1), \ldots, X(t_1)\big) \in A\Big)$$

$$= P\Big((X(t_1), \ldots, X(t_n)) \in \pi_{u,v}^{-1}(A)\Big) = P_v\Big(\pi_u^{-1}(A)\Big) .$$

Portanto a condição é necessária.

Para mostrar suficiencia temos que o Teorema de Kolmogorov para medidas nos garante a existência de uma probabilidade $\mu$ em $(\overline{\mathbb{R}}^{\mathbb{T}}, C\overline{\mathbb{R}}^{\mathbb{T}})$ tal que $P_u = \mu\pi_u^{-1}$. Em $(\overline{\mathbb{R}}^{\mathbb{T}}, C(\overline{R}^{\mathbb{T}}), \mu)$ consideremos o processo estocástico (canônico) $X(t,x) = x(t)$. Claro, as distribuições finitamente dimensionais de $X$ são dadas por $P_u = \mu\pi_u^{-1}$. De fato,

$$P_u(A) = \mu\Big(\pi_u^{-1}(A)\Big) = \mu\Big(\{x \in \mathbb{R}^{\mathbb{T}} : \pi_u(x) \in A\}\Big) =$$

$$= \mu\Big((x(t_1), \ldots, x(t_n)) \in A\Big) \qquad \square$$

## Exemplos de Processos Estocásticos.

a) Processos Gaussianos. Um processo estocástico $X = \{X(t); t \in \mathbb{T}\}$ com função média uma função arbitrária $m_X : \mathbb{T} \to \mathbb{R}$ e função covariância uma função não negativa definida arbitrária $R_X : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, que tem como distribuições finitamente dimensionais

$$P_u(A) = \int_A p_u(v_1, \ldots, v_n) dv_1 \ldots dv_n, \; A \in \mathcal{B}(\mathbb{R}) \text{ e } u = \{t_1, \ldots, t_2\}$$

onde

$$p_u(v_1, \ldots, v_n) = \Big[(2\pi)^n \det(C_u)\Big]^{-1} \exp\Big(-\frac{1}{2}(v - m_u)' C_u^{-1}(v - m_u)\Big)$$

onde $v' = (v_1, \ldots, v_n), m'_u = (m_X(t_1), \ldots, m_X(t_n))$ e $C_u = \left(C_X(t_i, t_j)\right)_{i,j=1,2,\ldots,n}$ é chamado em processo Gaussiano.

Pode ser mostrado que um processo $X = \{X(t); t \in \mathbb{T}\}$ é Gaussiano se e somente se todas as combinações lineares $a_1X(t_1) + \ldots + a_nX(t_n)$ são variáveis aleatórias normais.

b) Processos com valores independentes. Um processo estocástico com distribuições finitamente dimensionais dadas por

$$P_{t_1,\ldots,t_n}\left(A_1 \times \ldots \times A_n\right) = P_{t_1}(A_1) \ldots P_{t_n}(A_n)$$

onde $\{P_t : t \in \mathbb{T}\}$ é uma coleção arbitrária de medidas de probabilidade em $(\overline{\mathbb{R}}, B(\overline{\mathbb{R}}))$ é chamado processo com valores independentes.

c) Processos com incremento independentes. Um processo estocástico $\{X(t); t \in [a, b) \subset \overline{\mathbb{R}}\}$ é dito ter incrementos independentes se para todo $a \leq t_0 < t_1 < \ldots < t_n < b$,

$$X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})$$

são variáveis aleatórias independentes.

As distribuições finitamente dimensionais são especificadas por:

$$\begin{aligned} Q_a &= P_{X(a)} \\ Q_{s,t} &= P_{X(t)-X(s)} \quad a \leq s < t < b \end{aligned}$$

como

$$P_{t_1,\ldots,t_n}\left((-\infty, x_1] \times \ldots \times (-\infty, x_n]\right) =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{r=1}^{n} \mathbb{1}_{[0,\infty)}\Big(x_k - \sum_{i=0}^{k} v_i\Big) Q_a(dv_0) Q_{a,t_1}(dv_1) \ldots Q_{t_{n-1},t_n}(dv_n) \ .$$

Note que, para $a \leq t_1 < t_2 < t_1 < b$, $X(t_3) - X(t_1) = \big(X(t_3) - X(t_2)\big) + \big(X(t_2) - X(t_1)\big)$ e portanto $Q_{t_1,t_3} = Q_{t_1,t_2} * Q_{t_2,t_3}$

d) Processos Estacionários. Um processo $X = \{X(t); t \in \mathbb{I} \subset \mathbb{R}\}$, $\mathbb{I}$ um intervalo, é chamado estacionário, se para todo $t_1, t_2, \ldots, t_n \in \mathbb{I}$ e $h$ tal que $t_1+h, \ldots, t_n+h \in \mathbb{I}$ temos que

$$P_{t_1 \ldots, t_n} = P_{t_1+h, \ldots, t_n+h}$$

e) Processos Markovianos. Um processo estocástico $\{X = X(t) : t \in \mathbb{T}\}, \mathbb{T} \subset \mathbb{R}$ é chamado Markoviano, se para todo $n \geq 1, t_1 < t_2 < \ldots < t_n < s < t$ em $\mathbb{T}$ e $A \in \mathcal{B}(\overline{\mathbb{R}})$.

$$P\big(X(t) \in A / X(t_n), \ldots, X(t_n), X(s)\big) = P\big(X(t) \in A / X(s)\big) \ \text{q.c..}$$

## 2.5 Processos estocásticos como elementos aleatórios.

A aplicação $\Phi_X(\omega) = X(\cdot, \omega) = (X(t,\omega); t \in \mathbb{T})$ usada na definição da representação canônica de um processo estocástico $X$ pode ser vista como um elemento aleatório com valores em $\overline{\mathbb{R}}^{\mathbb{T}}$. Como comentamos anteriormente, a $\sigma$-álgebra $\mathcal{C}(\overline{\mathbb{R}}^{\mathbb{T}})$ é muito pequena e quando $\mathbb{T}$ não é contável muitos conjuntos importantes de funções não pertencem a esta família.

Se as trajetórias de um processo estocástico $X$ pertencem a um certo espaço de funções $\Omega' \subset \overline{\mathbb{R}}^{T}$, é natural perguntar quando o processo define um elemento aleatório com valores em $\Omega'$ e se a probabilidade induzida por este elemento aleatório determina as distribuições finitamente dimensionais do processo $X$.

**Proposição 1.48:** Seja $\Omega' \subset \mathbb{R}^{\mathbb{T}}$ e $\mathcal{F}'$ uma $\sigma$-algebra de subconjuntos de $\Omega'$. Seja $X = \{X(t) : t \in \mathbb{T}\}$ um processo estocástico tal que $X(\cdot, \omega) \in \Omega'$. Defina $\Psi_X : \Omega \to \Omega'$, por $\Psi_X(\omega) = X(\cdot, \omega)$. Então:

i) $\Psi_X$ é um elemento aleatório com valores em $\Omega'$ se e somente se $\Psi_X^{-1}(\mathcal{B}) \in \mathcal{F}$, para todo $B \in \mathcal{F}'$.

ii) Neste caso, $\nu_X = P\Psi_X^{-1}$ determina de maneira única as distribuições finitamente dimensionais de $X$ se, e somente se $\overline{\sigma\Phi_X)}^P \in \overline{\sigma(\Psi_X)}^P$.

Dem: i) óbvio.

ii) Suponhamos $\overline{\sigma(\Phi_X)}^P \subset \overline{\sigma(\Psi_X)}^P$. Então, $A = \{\omega : (X(t_1), \ldots, X(t_n)) \in B\} \in \overline{\sigma(\Psi_X)}^P$. Ou seja $A = \Psi_X^{-1}(B) \cup N$, onde $B \in \mathcal{F}'$ e $P(N) = 0$. Portanto $P(A) = P(\Psi_X^{-1}(B)) = \nu_x(B)$, isto é $\nu_X$ determina as distribuições finitamente dimensionais.

Reciprocamente, como $\nu_X$ determina a medida $P$ somente para conjuntos em $\overline{\sigma\left(\Psi_X\right)}^P$, ela determina as distribuições finitamente dimensionais se conjuntos da forma $\{\omega : X(t_1), \ldots, X(t_n)\} \cup N$, onde $P(N) = 0$, pertencem a $\overline{\sigma(\Psi_X)}^P$. Portanto $\overline{\sigma(\Psi_X)}^P \subset \overline{\sigma(\Psi_X)}^P$. □

**Exemplo:** Seja $\Omega' = C[0,1]$ e $\mathcal{F}' = B(C[0,1])$. Para $X = \{X(t) : t \in [0,1]\}$ um processo estocástico tal que $P(\{\omega \in \Omega : X(\cdot, \omega) \in C[0,1]\}) = 1$, defina

$$\Psi_X(\omega) = \begin{cases} X(\cdot, \omega) & \text{se } X(\cdot, \omega) \in C[0,1] \\ \\ 0 & \text{caso contrário.} \end{cases}$$

Então,

$$\sigma\left(\Psi_X\right) = \left\{\Phi_X^{-1}\left(C\right) \cap \left\{\omega : X(\cdot,\omega) \in C[0,1]\right\}; C \in \mathcal{C}\left(\mathbb{R}^{\mathbb{T}}\right)\right\}.$$

Como $\mathcal{B}(C[0,1])$ também é gerada pelos cilindros de $C[0,1]$, isto é, conjuntos da forma $A = \{x \in C[0,1] : (x(t_1),\ldots,x(t_n)) \in B\}, B \in \mathcal{B}(\mathbb{R})$, e

$$\Psi_X^{-1}(A) = \left\{\omega \in \Omega : \left(X(t_1,\omega),\ldots,X(t_n,\omega)\right) \in B\right\} \bigcap \left\{\omega : X(\cdot,\omega) \in C[0,1]\right\},$$

segue que $\Psi_X^{-1}(A) \in \mathcal{F}$. Portanto $\Psi_X$ é um elemento aleatório e dado que $P(\{\omega : X(\cdot,\omega) \in C[0,1]\}) = 1$, $\Psi_X$ determina as distribuições finitamente dimensionais.

Temos então que todo processo estocástico com trajetórias em $C[0,1]$ q.c. pode ser considerado como um elemento aleatório com valores em $C[0,1]$ ou equivalentemente como uma medida em $\mathcal{B}(C[0,1])$.

## 2.6 Separabilidade de um processo estocástico.

Seja $X = \{X(t) : t \in \mathbb{T}\}$ um processo estocástico com $\mathbb{T}$ um intervalo arbitrário ou um espaço métrico separável. Conjuntos do tipo

$$\left\{\omega \in \Omega; X(t,\omega) \in B, \forall\, t \in \mathbb{I}\right\} = \bigcap_{t \in \mathbb{I}} \left\{X(t) \in B\right\},$$

onde $B \in \mathcal{B}(\mathbb{R})$ e $\mathbb{I} \subset \mathbb{T}$ é não contável, não necessariamente pertencem a $\mathcal{F}$. Funções do tipo

$$\inf\left\{X(t) : t \in \mathbb{I}\right\}$$

não são necessariamente variáveis aleatórias. O mesmo problema ocorre com conjuntos como

$$\Big\{\omega \in \Omega : X(\cdot,\omega) \text{ é contínua}\Big\}$$

Além disto, mesmo que os conjuntos e funções acima sejam eventos e variáveis aleatórias, as distribuições finitamente dimensionais não permitem (em geral) que suas probabilidades e distribuições sejam calculadas. Isto pode ser visto através do seguinte exemplo: Seja $\Omega = [0,1], \mathcal{F} = \mathcal{B}([0,1])$, $P =$ Lebesgue, $\mathbb{T} = [0,1]$. Defina

$$X(t,\omega) = 0 \quad \text{para todo} \quad t \in [0,1] \text{ e } \omega \in \Omega$$

e

$$Y(t,\omega) = \begin{cases} 1 & \text{se } t = \omega \\ \\ 0 & \text{caso contrário.} \end{cases}$$

Para todo $t \in [0,1]$

$$P\Big(X(t) = Y(t)\Big) = Leb\Big([0,t)\cup(t,1]\Big) = 1,$$

portanto $X$ e $Y$ tem as mesmas distribuições finitamente dimensionais. Por outro lado $\sup\limits_{t\in[0,1]} X(t) = 0$, $\sup\limits_{t\in[0,1]} Y(t) = 1$ e $X$ tem trajetórias contínuas e $Y$ discontínuas.

Entre processos estocásticos equivalentes alguns tem trajetórias mais suaves, onde os problemas discutidos acima podem ser melhor estudados.

**Definição 1.49:** Seja $X\{X(t) : t \in \mathbb{T}\}$ um processo estocástico onde $\mathbb{T}$ é um espaço métrico. $X$ é chamado <u>separável</u> se existe $S$ um subconjunto contável e denso de $\mathbb{T}$ e um conjunto $N$ com $P(N) = 0$, tal que se $\omega \notin N$ e $t \in \mathbb{T}$, existe uma sequência $(s_n; n \in \mathbb{N})$ em $S$ (dependendo de $\omega$ e $t$) tal que

$$s_n \to t \quad \text{e} \quad X(s_n,\omega) \to X(t,\omega)\ .$$

$S$ é chamado conjunto separador e $N$ conjunto negligível.

**Proposição 1.50:** Seja $X$ um processo estocástico separável com conjunto separador $S$ e conjunto negligivel $N$. Se $F \subset \overline{I\!R}$ é fechado e $O \subset I\!T$ é aberto, então

$$\Big\{X(t) \in F : t \in O \cap I\!T\Big\} \Big\backslash \Big\{X(t) \in F : t \in O \cap S\Big\} \subset N.$$

Portanto,

$$\Big\{X(t) \in F : t \in O \cap I\!T\Big\} \in \overline{\mathcal{F}} \quad \text{e}$$

$$P\big(X(t) \in F : t \in O \cap I\!T\big) = P\big(X(t) \in F : t \in O \cap S\big)$$

Dem: Seja

$$A(S) = \Big\{X(t) \in F : t \in O \cup S\Big\}$$

$$A(I\!T) = \Big\{X(t) \in F : t \in O \cup I\!T\Big\}$$

Claro $A(I\!T) \subset A(S)$. Mostraremos que $A(S) \cap N^c \subset A(I\!T)$, o que implica que $A(I\!T) \backslash A(S) \subset N$.

Seja $\omega \in A(S) \cap N^c$. Fixemos $t \in O \cap I\!T$. Por hipótese existem $s_1, s_2 \ldots \in S$ com $s_n \to t$ e $X(s_n,\omega) \to X(t,\omega)$ quando, $n \to \infty$. Como $t \in O$ existe $n_0$ tal que $s_n \in O$ para $n \geq n_0$, ou seja $s_n \in O \cap S$ e $X(s_n,\omega) \in F$. Como $F$ é fechado e $X(s_n \omega) \to X(t,\omega), X(t,\omega) \in F$. Portanto $\omega \in A(I\!T)$. □

**Corolário 1.51:** Seja $X$ um processo estocástico e $S$ como na Proposição 1.50. Então,

i) $\sup_{t\in O\cap \mathbb{T}} X(t) = \sup_{t\in O\cap S} X(t)$ q.c.;

ii) $\inf_{t\in O\cap \mathbb{T}} X(t) = \inf_{t\in O\cap S} X(t)$ q.c.;

iii) $\overline{\lim}_{\substack{t\to u\\ t\in \mathbb{T}}} X(t) = \overline{\lim}_{\substack{t\to u\\ t\in S}} X(t)$ q.c.;

iv) $\underline{\lim}_{\substack{t\to u\\ t\in \mathbb{T}}} X(t) = \underline{\lim}_{\substack{t\to u\\ t\in S}} X(t)$ q.c.;

v) $\{\omega : X(\cdot,\omega)$ é contínua $\} \subset \overline{\mathcal{F}}$.

<u>Dem:</u> exercício. □

Segue da definição de separabilidade que todo processo estocástico com trajetórias contínuas q.c. é separável com conjunto separador tomado como qualquer subconjunto contável e denso $\mathbb{T}$.

Mostraremos agora que todo processo estocástico tem uma modificação separável baseado no trabalho de Cohn [C].

**Lema 1.52:** Seja $f_n, n = 1, 2\ldots$ uma sequência de funções reais estendidas definidas em um conjunto $\Omega$. Então existe um função real estendida $f$ em $\Omega$ tal que:

i) $f(\omega) = \lim_{n\to\infty} f_n(\omega)$,

ii) $f$ é $\sigma(f_n : n = 1, 2\ldots)$ mensurável (portanto se as $f_n$ são mensuráveis com respeito a $\mathcal{F}$, $f$ é mensurável com respeito $\mathcal{F}$).

<u>Dem:</u> Considere em $\overline{\mathbb{R}}$ a métrica $d(x,y) = \sum_{n=1}^{\infty} \text{Leb}((x,y)\bigcap(-n,n))\backslash 2n2^n$. Para cada $n = 1, 2, \ldots$, seja $\{I(n,i), i = 1, 2, \ldots, N_n\}$ uma partição de $\overline{\mathbb{R}}$ onde $I_{n,i}$ são intervalos com $\text{diam}(I_{n,i}) < 1/n$. Assumimos também $\sup I(n,i) < \inf I(n, i+1)$ e que a partição para $n+1$ é um refinamento da partição para $n$.

Para cada, $n$ defimos uma função $g$ com valores inteiros por

$$g_n(\omega) = \min\{i : I(n,i) \text{ contém infinitos pontos da sequência } (f_k(\omega)\}.$$

Então para cada $x \in X$, $\{I(n, g_n(\omega))\}$ é decrescente e $\text{diam} I(n, g_n(\omega)) \searrow 0$. Portanto $\bigcap_{n=1}^{\infty} I\Big(n, g_n(\omega)\Big) = \{y\}$. Seja $f(\omega) = y$. Então $f : \Omega \to \overline{\mathbb{R}}$. Como todo intervalo aberto em torno de $f(\omega)$ contém um dos intervalos $I(n, g_n(\omega))$ e portanto infinitos pontos de $(f_k(\omega))$, segue que $f(\omega)$ é um ponto limite da sequência $(f_k(\omega))$.

Cada $g_n$ é $\sigma(f_n : n = 1, 2 \ldots)$ mensurável pois $g_n(\omega) \geq i$ se e somente se existe um número finito de pontos de $(f_k(\omega))$ em $I(n,j) j = 1, \ldots, i-1$. Portanto, $\{\omega \in \Omega : g_n(\omega) \geq i\} = \bigcup_{m=1}^{\infty} \bigcup_{k_1, \ldots k_m \in \mathbb{N}} \{\omega \in \Omega : f_{k_1}(\omega), \ldots, f_{k_m}(\omega) \in \bigcup_{j=1}^{i-1} I(n,j), f_k \notin \bigcup_{j=1}^{j-1} I(n,j), k \neq k_1, \ldots, k_m\}$ pertencem a $\sigma(f_n : n = 1, 2, \ldots)$. A mensurabilidade de $f$ segue da relação

$$\Big\{\omega : f(\omega) < \lambda\Big\} = \bigcup_{\substack{n,i \\ \overline{I(n,i)} \subset [-\infty, \lambda)}} \Big\{\omega \in \Omega : g_n(\omega) = i\Big\} \in \sigma\Big(f_n : n = 1, 2, \ldots\Big) \qquad \square$$

**Teorema 1.53:** (Doob [D]. Todo processo estocástico tem uma modificação separável.

<u>Dem:</u> Primeiramente mostraremos que existe um subconjunto contável e denso $S$ de $\mathbb{T}$ e para cada $t \in \mathbb{T}$ um conjunto $N_t$ com $P(N_t) = 0$ tal que, se $\omega \in N_t$, então $X(t, \omega)$ é um ponto limite de $X(s, \omega)$ quando $s \to t$, $s \in S$.

Seja $\mathcal{V}$ a classe de todas as uniões finitas de intervalos abertos e fechados de $\overline{\mathbb{R}}$ com fronteiras racionais ou $\pm\infty$. $\mathcal{V}$ é contável e todo fechado de $\overline{\mathbb{R}}$ é uma intersecção contável de conjuntos de $\mathcal{V}$. Seja $\mathcal{S}$ a classe de todas as bolas com centro em um conjunto contável e denso de $\mathbb{T}$. Para cada $D \in \mathcal{V}$ e $O \in \mathcal{S}$

$$\lambda(F, O) = \inf_{\substack{A \subset O \\ A \text{ contável}}} P\Big(X(t) \in F : t \in A\Big).$$

Considerando uma sequência aproximando infinito e tomando uniões dos correspondentes conjuntos contáveis de $t's$, obtemos um conjunto contável $S(F,O)$ de $\mathbb{T}$ tal que

$$\lambda(F,O) = P\Big(X(t) \in F : t \in S(F,O)\Big).$$

Seja $S = \bigcup_{\substack{F\in\mathcal{V} \\ O\in\mathcal{S}}} S(F,O)$. Então $S$ é um subconjunto contável de $\mathbb{T}$ e se necessário acrescentamos pontos para faze-lo também denso.

Para cada $t \in \mathbb{T}$ e $O \in \mathcal{S}$ com $t \in O$, seja

$$N_t(F,O) = \Big\{X(s) \in F : s \in S \cap O, X(t) \notin F\Big\}.$$

Então $P(N_t(F,O)) = 0$. De fato, isto é claro para $t \in S \cap O$. Se $t \notin S$ e $P(N_t(F,O)) > 0$ temos, como $S(F,O) \subset O \cap S$, que

$$P\Big(X(s) \in F : s \in S(F,O)\Big) \geq P\Big(X(s) \in F : s \in S \cap O\Big)$$

$$= P\Big(X(s) \in F : s \in S \cup O,\ X(t) \in F\Big) + P\Big(N_t(F,0)\Big)$$

$$> P\Big(X(s) \in F, s \in S \cup O, X(t) \in F\Big)$$

o que contradiz a definição de $\lambda(F,O)$.

Definindo $N_t = \bigcup_{\substack{F\in\mathcal{V} \\ O\in\mathcal{S}}} N_t(F,O)$, claro $N_t \in \mathcal{F}$ e $P(N_t) = 0$. Temos também que se $\forall\ t \in \mathbb{T}, \forall\ F \in \mathcal{V}$ e $\forall\ O \in \mathcal{S}$ com $t \in O$, $\omega \notin N_t$ e $X(s,\omega) \in F\ \forall s \in S \cap O$, então $X(t,\omega) \in F$. Isto permanece válido se colocamos no lugar de $F$, qualquer fechado de $\overline{\mathbb{R}}$, pois cada conjunto fechado $F$ de $\overline{\mathbb{R}}$ é de forma $F = \bigcup_{n=1}^{\infty} F_n$, onde $F_n \in \mathcal{V}$. Então para $t \in \mathbb{T}, \omega \in N_t$, $O \in \mathcal{S}$ com $t \in O$ e $F_{\omega,t} = \overline{\{X(s,\omega) : s \in S \cap O\}}$, temos que

$$X(t,\omega) \in F_{\omega,t}$$

para todo $t \in \mathbb{T}$ e $\omega \notin N_t$. Segue que $X(t,\omega)$ é um ponto limite de $\{X(s,\omega); S \in S \cap O\}$. Definindo $Y = \{Y(t) : t \in \mathbb{T}\}$ por:

i) $Y(t,\omega) = X(t,\omega)$, para $t \in S$ e $\omega \in \Omega$;

ii) $Y(t,\omega) = X(t,\omega)$, para $t \notin S$ e $\omega \notin N_t$ e

iii) para cada $t \in S$ e $\omega \in N_t$, selecione uma sequência $\{s_n\}$ em $S$ tal que $s_n \to t$. Pelo Lema 1.52 existe um limite de $\{X(s_n\omega)\}$ mensurável, seja $Y(t,\omega)$ este limite.

$Y(t)$ é uma variável aleatória para cada $t$, na verdade $Y(t)$ é $\sigma(X(t) : t \in \mathbb{T})$ mensurável, e $P(X(t) = Y(t)) = 1$ para todo $t \in \mathbb{T}$.

Portanto, $Y$ é uma modificação de $X$. $Y$ é separável com conjunto separador $S$ e conjunto negligivel $\phi$, pois para cada $t \in \mathbb{T} \backslash S$,

i) se $\omega \in N_t$, $Y(t,\omega)$ é limite de $X(s_n,\omega) = Y(s_n,\omega)$ e

ii) se $\omega \notin N_t, Y(t,\omega) = X(t,\omega)$ e um limite de $X(s,\omega) = X(s,\omega)$ quando $s \to t$ através $S$. □

## 2.7 Processos estocásticos mensuráveis.

Seja $X = \{X(t)t \in \mathbb{T}\}$ um processo estocástico com $\mathbb{T} \subset \mathbb{R}$ um intervalo. As trajetórias $X(\cdot,\omega)$ não são necessariamente Borel ou Lebesgue mensuráveis e assim sendo, integrais com

$$\int_{\mathbb{T}} X(t,\omega)dt \quad \text{ou} \quad \int_{\mathbb{T}} X(t,\omega)f(t)dt$$

($f$ mensurável) não estarão em geral definidas e não são necessariamente variáveis aleatórias.

**Definição 1.54:** Seja $X = \{X(t); t \in \mathbb{T}\}$ um processo estocástico. Se $(\mathbb{T}, \Im)$ é um espaço mensurável, $X$ é dito ser mensurável se a aplicação $X: \mathbb{T} \times \Omega \to \overline{\mathbb{R}}$ é mensurável com respeito a $\sigma$-álgebra $\Im \times \mathcal{F}$.

Se $\mathbb{T}$ é um intervalo, $\Im = \mathcal{B}(\mathbb{T})$, pelo teorema de Fubini, todas as trajetórias $X(\cdot, \omega)$ são Borel mensuráveis e se $\int_{\mathbb{T}} |X(t,\omega)|dt < \infty$ q.c. ou $X(t,\omega) \geq 0$ q.c., a integral $\int_{\mathbb{T}} X(t,\omega)dt$ define uma variável aleatória. Note também que se $E|X(t)|^p < \infty$ então:

$$E \int_{\mathbb{T}} |X(t,\omega)|^p dt = \int_{\mathbb{T}} E|X(t,\omega)|^p dt < \infty$$

o que implica $\int |X(t,\omega)|^p dt < \infty$ q.c. e $E \int_{\mathbb{T}} X^p(t,\omega)dt = \int_T EX^p(t)dt$.

**Definição 1.55:** Seja $\mathbb{T} \subset \mathbb{R}$ um intervalo, e $\{\mathcal{F}_t : t \in \mathbb{T}\}$ uma família não decrescente de sub $\sigma$-álgebras de $\mathcal{F}/(s, t \in \mathbb{T}, s < t, \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F})$. Um processo $X = \{X(t) : t \in \mathbb{T}\}$ é dito adaptado a $\{\mathcal{F}_t : t \in \mathbb{T}\}$ se $X(t)$ é $\mathcal{F}_t$ mensurável para cada $t \in \mathbb{T}$. $X$ é dito ser progressivamente mensurável se para cada $t \in \mathbb{T}$ a aplicação $(s,\omega) \to X(s,\omega)$ de $\mathbb{T} \cap (-\infty, t] \times \Omega$ em $\overline{\mathbb{R}}$ é $\mathcal{B}(\mathbb{T} \bigcap (-\infty, t]) \times \mathcal{F}$-mensurável.

Note que $X = \{X(t) : t \in \mathbb{T}\}$ é sempre adaptado a família $\mathcal{F}_t = \sigma(X(t) : s \leq t)$ e se $X$ é progressivamente mensurável então ele é mensurável.

**Lema 1.56:** Seja $V$ o conjunto de todas as variáveis aleatórias em $(\Omega, \mathcal{F}, P)$. Então

$$d(X,Y) = E\left(\frac{|X-Y|}{1+|X-Y|}\right) \quad X, Y \in V$$

define uma pseudo-distância em $V$ que metriza convergência em probabilidade.

Dem: exercício

**Teorema 1.57:** Se $X = \{X(t) : t \in \mathbb{T}\}$ com $\mathbb{T}$ um intervalo, é adaptado a $\{\mathcal{F}_t; t \in \mathbb{T}\}$ e contínuo em probabilidade, então ele tem uma modificação separável e progressivamente mensurável com respeito a $\{\mathcal{F}_t : t \in \mathbb{T}\}$.

Dem: Seja $\mathbb{T} = [0, \infty)$ (a prova é análoga para qualquer outro intervalo). Para cada $n = 1, 2 \ldots$ considere a partição de $[0, n]$ : $0 = t_{n,0} < t_{n,1} < \ldots < t_{n,N_n} = n$ tal que $|t_{n,k} - t_{n,k-1}| < 1/n$ e $P\Big(|X(t) - X(s)| > 1/n\Big) < \frac{1}{2^n}$, para todo $t, s \in [t_{n,k-1}, t_{n,k}), k = 1, 2, \ldots, N_n$. Isto é possível pois $X$ é contínuo em probabilidade em $[0, \infty)$ e portanto uniformemente contínuo em probabilidade em $[0, n]$ para cada $n$. Escolhendo a partição tal que

$$E\left(\frac{|X(t) - X(s)|}{1 + |X(t) - X(s)|}\right) < \frac{1}{(n+1)2^n}$$

para $t, s \in [t_{n,k-1} t_{n,k}), k = 1, 2, \ldots, N_n$, temos

$$P\left(|X(t) - X(s)| > \frac{1}{n}\right) = P\left(\frac{|X(t) - X(s)|}{1 + |X(t) - X(s)|} < \frac{1}{n+1}\right)$$

$$\leq (n+1)E\left(\frac{|X(t) - X(s)|}{1 + |X(t) - X(s)|}\right) < \frac{1}{2^n}\,.$$

Sem perda de generalidade podemos assumir

$$\Big\{t_{n,k}\ k = 1, \ldots, N_n\Big\} \subset \Big\{t_{n+1,k}\ x = 1, \ldots, N_{n+1}\Big\}\,.$$

Defina:

$$X_n(t,\omega) = \begin{cases} X(t_{k-1},\omega) & t_{k,k-1} < t < t_{n,k},\ k = 1,\ldots,N_n \\ X(n,\omega) & t \geq n\ . \end{cases}$$

$X_n$ é progressivamente mensurável com respeito a $\{\mathcal{F}_t : t \geq 0\}$, pois

$$X_n(t,\omega) = \sum_{k=1}^{N_n} X(t_{k-1},\omega)\mathbb{1}_{[t_{n,k-1},t_{n,k})}(t) + X(n,\omega)\mathbb{1}_{[n,\infty)}$$

e suas restrições a $[0,T]$ tem como último termo

$$X(t_n,n)(\omega)\mathbb{1}_{[t_{n,k},T]}(t) \qquad \text{se} \quad t_{n,k} \leq T < t_{n,n+1}$$

ou

$$X(n,\omega)1_{[n,T]}(t) \qquad \text{se} \quad n \leq T$$

que são claramente $\mathcal{B}([0,T]) \times \mathcal{F}_t$-mensuráveis.

Usando Lema 1.52 com $\mathbb{T} \times \Omega$ no lugar de $\Omega$ e $f_n = X_n$, escolhemos um processo mensurável $Y$ que é de fato progressivamente mensurável. Como $Y(t,\omega)$ é um ponto limite da sequência $\{X_n(t,\omega); n = 1,2,\ldots\}$, $Y$ é separável com $S = \bigcup_n\{t_{n,k} : k = 1,\ldots N_n\}$ e $N = \phi$.

Finalmente para cada $t \in [0,\infty)$, se $t \in [t_{n,k-1},t_{n,k})$, então $X_n(t) = X(t_{n,k-1})$ e

$$\sum_{n=1}^{\infty} P(|X_n(t) - X(t)| > 1/n) < \sum_{n=1}^{\infty} \frac{1}{2^n} < \infty.$$

Pelo Lema de Borel-Cantelli, $P\left(\overline{\lim}_{n\to\infty}\{|X_n(t) - X(t)| > 1/n\}\right) = 0$, e portanto $X_n(t) \to X(t)$ q.c.. Como para todo $t \in [0,\infty)$ e $\omega \in \Omega$, $Y(t,\omega)$ é um limite da sequência $\{X_n(t,\omega); n = 1,2\ldots\}$, temos que para todo $t$, $Y(t) = X(t)$ q.c., mostrando que $Y$ é uma modificação de $X$. □

O próximo teorema apresenta um resultado geral sobre mensurabilidade de um processo estocástico. Para a sua demonstração indicamos os trabalhos de Chung e Doob [C-D] e Cohn [Co]. Outros resultados importantes são as publicações de Cambanis [Ca] e Hoffman-Jorgensen [H-J].

**Teorema 1.57:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ com $\mathbb{T}$ um intervalo, adaptado a $\{\mathcal{F}_t : t \in \mathbb{T}\}$. As seguintes afirmações são equivalentes:

i) $X$ tem uma modificação mensurável.

ii) $X$ tem uma modificação separável e mensurável.

iii) A aplicação $t \to X(t)$ de $\mathbb{T}$ em $V$($V$ como no Lema 1.56) é Borel mensurável.

iv) $X$ tem uma modificação separável e progressivamente mensurável.

Se $\mathbb{T}$ é um espaço métrico completo e separável, i) ii) e iii) acima são equivalentes.

## 2.8 Continuidade de trajetórias.

Como definimos anteriormente, um processo $X = \{X(t) : t \in [a,b]\}$ é contínuo q.c. se existe $\Omega_c \in \mathcal{F}$, com $P(\Omega_c) = 1$ tal que para $\omega \in \Omega_c$ $X(\cdot,\omega)$ é contínua em $[a,b]$.

O "modulos" de continuidade de um trajetória é definido por

$$Z(h,\omega) = \sup_{\substack{t,s \in [a,b] \\ |t-s|<h}} |X(t,\omega) - X(s,\omega)|.$$

Claro $Z(\cdot,\omega)$ é não crescente e $X(\cdot,\omega)$ é contínuo (uniformemente contínuo) se e somente se $Z(h,\omega) \to 0$ quando $h \to 0$. Como $X$ tem uma versão separável, $Y(h)$ é um processo estocástico. Portanto $X$ tem trajetórias contínuas se e somente se $Z(h) \to 0$ q.c. (ou $Z(h) \to 0$ em probabilidade) quando $h \to 0$.

**Teorema 1.58:** Se $X\{X(t) : t \in [a,b]\}$ é separável e para algum $\delta > 0$ e todo $h \in (0,\delta)$ com $t, t+h \in [c,b]$,

$$P\Big((\dot{X}(t+h)-X(t))>g(h)\Big)\leq q(h),$$

onde $g$ e $q$ são funções positivas, não decrescentes, definidas em $(0,\delta)$ e satisfazendo

$$\int_0^\delta g(h)\frac{dh}{h}<\infty \quad \text{e} \quad \int_0^\delta q(h)\frac{dh}{h^2}<\infty,$$

então $X$ tem trajetórias contínuas com probabilidade 1 e existe uma variável aleatória $Y, 0< Y<\infty$ q.c. tal que

$$\sup_{|t-s|<h}|X(t,\omega)-X(s,\omega)|\leq\frac{4}{\ln 2}\int_0^h g(h)\frac{du}{u}, \quad \text{se } h\leq y(\omega) \text{ q.c.}$$

**Observações:** **a)** É conveniente que as funções $g$ e $q$ sejam definidas e satisfaçam as propriedades para $h\in(0,b-c]$. Isto pode ser obtido estendendo $g$ para qualquer função não decrescente e $q$ fazendo $q(h)=\max\{q(\delta-),1\}$ para $\delta\leq h\leq b-a$. **b)** Como $g$ e $q$ são não decrescentes, as condições de integrabilidade são equivalentes as seguintes condições de somabilidade.

$$\sum_{n=1}^{\infty}g\Big(\frac{1}{2^n}\Big)<\infty \quad \text{e} \quad \sum_{n=1}^{\infty}2^n q\Big(\frac{1}{2^n}\Big)<\infty\,.$$

<u>Dem:</u> Sem perda de generalidade tomemos $[a,b]=[0,1]$. Como $g(h)\downarrow 0$ e $q(h)\downarrow 0$ quando $h\downarrow 0$,

$$P\Big((X(t+h)-X(t)\Big)\geq g(h))\leq q(h)$$

implica que $X$ é contínuo em probabilidade. Portanto, qualquer conjunto contável e denso de $[0,1]$ é um conjunto separador para $X$. Tomemos

$$S=\left\{\frac{k}{2^n}:k=0,1,\ldots,2^n,n=1,2,\ldots\right\}.$$

Segue que

$$\sup_{\substack{t,s\in[0,1]\\|t-s|<h}} |X(t)-X(s)| = \sup_{\substack{t,s\in S\\|t-s|<h}} |X(t)-X(s)| \quad \text{q.c.}$$

Seja

$$X_m = \sup_{0\leq k\leq 2^m} \left|X\left(\frac{k+1}{2^m}\right) - X\left(\frac{k}{2^m}\right)\right| \geq 0$$

Temos então que

$$\begin{aligned} P\Big(X_m \geq g\Big(\frac{1}{2^m}\Big)\Big) &\leq \sum_{k=0}^{2^m-1} P\Big(\Big|X\Big(\frac{k+1}{2^m}\Big) - \Big|X\Big(\frac{k}{2^n}\Big)\Big) \geq g\Big(\frac{1}{2^n}\Big) \\ &\leq \sum_{k=0}^{2^N-1} q\Big(\frac{1}{2^m}\Big) = 2^m q\Big(\frac{1}{2^m}\Big) \end{aligned}$$

e portanto

$$\sum_n P\Big(X_n \geq g\Big(\frac{1}{2^n}\Big)\Big) \leq \sum_n 2^n q\Big(\frac{1}{2^n}\Big) < \infty.$$

Por Borel-Cantelli

$$P\Big(\overline{\lim}_{n\to\infty}\Big\{X_n \geq \Big(\frac{1}{2^n}\Big)\Big\}\Big) = 0$$

Definindo $\Omega_c = \Big(\overline{\lim}_{n\to\infty}\{X_n \geq g(\frac{1}{2^n}\}\Big)^c, P(\Omega_c) = 1$. Segue que para todo $\omega \in \Omega_c$, existe $N(\omega)$ tal que

$$X_n(\omega) \leq g\Big(\frac{1}{2^n}\Big) .$$

Como $\sum_n g\left(\frac{1}{2^n}\right) < \infty$, segue que $sum_n X_n < \infty$ q.c..

Agora, se $t$ é um racional diádico, $t \in S$ e para algum par $(k,n)$, $|t - \frac{k}{2^n}| < \frac{1}{2^n}$, então $t = \frac{k}{2^n} \pm \sum_{m=n+1}^{n+n'} \frac{a_m}{2^m}$, onde $a_m = 0$ ou 1. Se definimos

$$s_i = \frac{k}{2^n} \pm \sum_{m=n+1}^{n+i} \frac{a_m}{2^m}, i = 1, \dots, n'$$

e

$$s_0 = \frac{k}{2^n}$$

teremos

$$\left|X(t) - X\Big(\frac{k}{2^n}\Big)\right| \le \sum_{i=1}^{n'-1} \left|X(s_{i+1}) - X(s_i)\right| .$$

Se $a_{i+1} = 0$, então $s_{i+1} = s_i$ e $|X(s_{i+1}) - X(s_i)| = 0$. Se $a_{i+1} = 1$, então $s_{i+1} = s_i \pm \frac{1}{2^{n+i+1}}, s_i = \frac{l}{2^{n+i+1}}$ e $|X(s_{i+1}) - X(s_i)| \le X_{n+i+1}$.
Segue que

$$\left|X(t) - X\Big(\frac{k}{2^n}\Big)\right| \le \sum_{i=1}^{n-1} X_{n+i+1} \le \sum_{m \ge n} X_m .$$

Se $t, s \in S$ e $|t - s| < \frac{1}{2^n}$, então existe $k$ tal que $|t - \frac{k}{2^n}| < \frac{1}{2^n}$ e $|s - \frac{k}{2^n}| < \frac{1}{2^n}$. Portanto,

$$|X(t) - X(s)| \le |X(t) - X(\frac{k}{2^n})| + |X(s) - X(\frac{k}{2^n})| \le 2 \sum_{m>n} X_m .$$

Temos então

$$\sup_{\substack{s,t \in [0,1] \\ |t-s| < 1/2^n}} |X(t) - X(s)| = \sup_{\substack{s,t \in S \\ |t-s| < 1/2^n}} |X(t)| - X(s)| \le 2 \sum_{m>n} X_m \searrow 0$$

q.c. quando $n \to \infty$. Como $Z(h)$ é não decrescente

$$\sup_{\substack{t,s \in [0,1] \\ |t-s| < h}} |X(t) - X(s)| \downarrow 0 \text{ q.c. quando } h \downarrow 0.$$

Portanto $X$ tem trajetória contínua com probabilidade 1.

Para mostrar a segunda parte do teorema, observe que $\Omega_c = (\overline{\lim}_{n\to\infty}\{X_n \geq g(\frac{1}{2^n})\})^c = \left(\bigcup_{n\leq 1}\{X_n \geq g(\frac{1}{2^n})\}\right)^c \bigcup_{k\geq 1}[\bigcup_{m\geq k}\{X_m \geq g(\frac{1}{2^m})\}/\bigcup_{m>k}\{X_m \geq g(\frac{1}{2^m})\}]$

Defina

$$N(\omega) = \begin{cases} k & \text{em } \bigcup_{m\geq k} g\left(\frac{1}{2^n}\right)\setminus\bigcup_{m>k} X_m \geq g\left(\frac{1}{2^m}\right)\Big) \\ 0 & \text{em } \left(\bigcup_n\left\{X_n \geq g\left(\frac{1}{2^n}\right)\right\}\right)^c \bigcap \Omega_c^c \end{cases}$$

Claro

$$X_n(\omega) \leq g\Big(\frac{1}{2^n}\Big)$$

e

$$\sum_{m>n} X_n(\omega) \leq g\Big(\frac{1}{2^m}\Big).$$

Se $\omega \in \Omega_c \bigcap N_0^c$, onde $N_0$ é o conjunto de probabilidade zero associado ao conjunto separador $S$ do processo $X$ e $n > N(\omega)$, então

$$\begin{aligned} \sup_{\substack{s,t\in[0,1] \\ |t-s|<\frac{1}{2^n}}} |X(t,\omega) - X(s,\omega)| &= \sup_{\substack{t,s\in S \\ |t-s|<\frac{1}{2^n}}} |X(t,\omega) - X(s,\omega)| \leq 2\sum_{m\geq n} X_m(\omega) \\ &\leq 2\sum_{m>n} g\left(\tfrac{1}{2^n}\right). \end{aligned}$$

Para $\omega \in \Omega_c \bigcap N_0^c$ fixo e $0 < h < n2^{-N(\omega)}$

$$\sup_{\substack{t,s\in[0,1] \\ |t-s<h}} \Big|X(t,\omega) - X(s,\omega)\Big| \leq 2 \sup_{\substack{t,s\in S \\ |t-s|<\frac{h}{2}}} \Big|X(t,\omega) - X(s,\omega)\Big|$$

$$\leq 2 \sup_{\substack{t,s \in S \\ |t-s| < \frac{1}{2^n}}} |X(t,\omega) - X(s,\omega)|, \qquad \frac{n}{2} < \frac{1}{2^n} < h$$

$$\leq 4 \sum_{m>n} g\left(\frac{1}{2^m}\right) \leq \frac{4}{\ln 2} \int_0^{\frac{1}{2^n}} g(u) \frac{du}{u} \ , \ n > N(\omega)$$

$$\leq \frac{4}{\ln 2} \int_0^h g(u) \frac{du}{u} \ .$$

O resultado segue, definindo $Y(\omega) = 2^{-N(\omega)}$ que é positiva e finita q.c. $\square$

**Teorema 1.59:** (Kolmogorov) Se $X\{X(t) : t \in [a,b]\}$ é separável e para todo $h \in (0,\delta), \delta > 0$ com $t, t+h \in [a,b]$ temos

$$E\Big(|X(t+h) - X(t)|^\alpha\Big) \leq Ch^{1+\beta}$$

para constantes positivas $C, \alpha$ e $\beta$. Então para todo $0 < \gamma < \beta/\alpha$,

$$\frac{1}{h^\gamma} \sup_{\substack{t,s \in [a,b] \\ |t-s| < h}} |X(t) - X(s)| \to 0 \text{ q.c.}$$

quando $h \downarrow 0$. Segue que as trajetórias de $X$ são q.c. Lipschitz de ordem $\gamma < \beta/\alpha$ e em particular $X$ tem trajetórias contínuas com probabilidade 1.

Dem: Fixemos constantes $\varepsilon > 0$ e $0 < \gamma < \beta\delta/\alpha$. Então as funções

$g(h) = \varepsilon h^\gamma$ e $q(h) = C\varepsilon^{-\alpha} h^{1+\beta-\alpha\gamma}$ satisfazem as condições do Teorema 1.58. Temos também que, para todo $h \in (0,\delta)$ com $t$ e $t+h \in [a,b]$,

$$\begin{aligned} P\Big(|X(t+h) - X(t)| > \varepsilon h^\gamma\Big) &\leq \frac{E(|X(t+h) - X(t)|^\alpha}{(\varepsilon h^\gamma)^\alpha} \\ &\leq \frac{Ch^{1+\beta}}{\varepsilon^\alpha h^{1+\beta-\alpha\gamma}} \\ &= C\varepsilon^{-\alpha} n^{1+\beta-\alpha\gamma}. \end{aligned}$$

Com $\int_0^h g(u)\frac{du}{u} = \varepsilon \int_0^h u^{\gamma-1}du = \varepsilon h^\gamma$, segue pelo Teorema 1.58 que $\sup\limits_{\substack{t,s\in[a,b]\\|t-s|<h}} |X(t,\omega) - X(s,\omega)) \leq \frac{4}{\ln 2}\varepsilon h^\gamma$ se $h < Y(\omega)$ q.c.. Portanto como $\varepsilon > 0$ é arbitrário,

$$\lim_{h\to 0} \frac{1}{h^\alpha} \sup_{\substack{t,s\in[a,s]\\|t-s|<h}} |X(t) - X(s)| = 0 \quad \text{q.c.} \qquad \square$$

## 2.9 Processo de Wiener.

Um tipo de processo que será intensivamente usado nos próximos capítulos e talvez o mais estudado na literatura é chamado processo de Wiener ou movimento Browniano.

**Definição 1.60:** Um processo de Wiener ou movimento Browniano é um processo estocástico $\{W(t) : t \in \mathbb{T}\}, \mathbb{T} = [0,\infty)$ ou $[0,T]$, definido em $(\Omega, \mathcal{F}, P)$ que satisfaz
i) $P(W(0) = 0) = 1$
ii) $W$ tem incrementos independentes: Se $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n$, então $W(t_0), W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})$ são variàveis aleatórias independentes.
iii) Para $0 \leq s < t$, o incremento $W(t) - W(s)$ tem distribuição normal (Gaussiana) com média 0 e variância 1:

$$P\big(W(t) - W(s) \in A\big) = \int_A \frac{1}{\sqrt{2\pi(t-s)}} \exp\Big(-\frac{x^2}{2(t-s)}\Big)dx.$$

A existência de um processo de Wiener é garantida pelo teorema de extensão de Kolmogorov. Para tanto, observemos que um tal processo, se existe, deve satisfazer para $0 = t_0 < t_1 < \ldots < t_n$, $x_0 = 0$ e $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$

$$P\big(W(t_1) \in A_1, W(t_2) \in A_2, \ldots, W(t_n) \in A_n\big) =$$

$$= \int_{A_1}\int_{A_2}\cdots\int_{A_n} \Pi_{i=1}^n \left\{ \Big[2\pi(t_i - t_{i-1})\Big]^{-1/2} \exp\left(-\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right)\right\} dx_n \ldots dx_1.$$

Como esta família de medidas de probabilidade é consistente no sentido de Kolmogorov, a existência de um processo de Wiener esta garantida Teorema de extensão de Kolmogorov.

É fácil mostrar que:

$$E\Big(W(t)\Big)=0,\ \ E\Big(W^2(t)\Big)=t \ \text{ e } \ E\Big(W(t)W(s)\Big)=\min(t,s)$$

**Lema 1.61:** Se $Y$ é uma variável aleatória com distribuição normal com média 0, então para todo $\alpha>0$

$$E\Big(|Y|^\alpha\Big)=K(\alpha)\Big[|Y|^2\Big]^{\alpha/2}.$$

<u>Dem:</u> exercício

**Proposição 1.62:** Um processo de Wiener tem uma modificação com trajetórias contínuas com probabilidade 1.

<u>Dem:</u> Considere uma versão separável de $W$. Pelo Lema 1.61 para $\alpha>0$, $h>0$

$$\begin{aligned} E\Big(|W(t+h)-W(t)|^\alpha\Big) &= K(\alpha)\Big[E\Big(|W(t+h)-W(t)|^2\Big)\Big]^{\alpha/2}= \\ &= K(\alpha)h^{\alpha/2}=k(\alpha)h^{1+(\alpha/2-1)}\,. \end{aligned}$$

Tomando $\alpha>2$ e $\beta=\alpha/2-1$

$$E\Big(|W(t+h)-W(t)|^\alpha\Big)=C\ \ h^{1+\beta}.$$

Segue pelo Teorema 1.59 que $W$ tem trajetórias contínuas com probabilidade 1. $\square$

# 2. Teoria de Martingales.

## 2.1 Introdução.

Daremos aqui os resultados básicos sobre uma classe de processos estocásticos chamados martingales. O conjunto de índices $\mathbb{T}$ sera $[0,\infty)$ ou $[0,T]$ com $T < \infty$. Ocasionalmente outros conjuntos de índices poderão ser considerados e neste caso a escolha de $\mathbb{T}$ sera explicitada. O espaço de probabilidade $(\Omega, \mathcal{F}, P)$ sera em geral assumido completo.

Além do espaço de probabilidade e do conjunto de índices, consideremos também na definição de martingales uma família de sub-$\sigma$-algebras de $\mathcal{F}$ indexados por $\mathbb{T}$, $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$, que satisfaz:

i) $\mathcal{F}_s \subset \mathcal{F}_t$, se $s < t$ ($\mathbb{F}$ é não decrescente)

ii) $\mathcal{F}_0$ contem todos os subconjuntos de $\mathcal{F}$ com probabilidade 0 ($\mathbb{F}$ é completa)

Uma família sub-$\sigma$-algebra de $\mathcal{F}$ satisfazendo i) e ii) acima é chamada um filtração ou filtragem. Se para $t \in \mathbb{T}, \mathcal{F}_t = \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$, então a filtração $\mathbb{F}$ é dita ser contínua a direita. Um processo $\{X(t) : t \in \mathbb{T}\}$ é dito ser adaptado a $\mathbb{F}$ se $X(t)$ é $\mathcal{F}_t$ mensurável.

Variáveis aleatórias positivas satisfazendo certas condições de mensurabilidade em relação a $\mathbb{F}$ são importantes no estudo de martingales.

**Definição 2.1:** Uma variável aleatória $\tau$ assumindo valores em $\mathbb{T}$ é chamada um tempo parada com respeito a $\mathbb{F}$, se $[\tau \leq t] = \{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ para todo $t \in \mathbb{T}$

Tempos de parada cumprem propriedade que listaremos a seguir.

**Proposição 2.2:** (a) $[\tau < t] \in \mathcal{F}_t$ e $[\tau = t] \in \mathcal{F}_t$, se $\tau$ é um tempo de parada. (b) Se $\mathbb{F}$ é contínua a direita, então $\tau$ é um tempo de parada se e somente se $[\tau < t] \in \mathcal{F}_t$ para todo $t \in \mathbb{T}$.

Dem: a) Como $\tau$ é um tempo de parada $[\tau \leq s] \in \mathcal{F}_s \subset \mathcal{F}_t$, para $t > s$. Portanto $[\tau < t] = \bigcup_n [\tau \leq t - 1/n] \in \mathcal{F}_t$ e $[\tau = t] = [\tau \leq t] \setminus [\tau < t] \in \mathcal{F}_t$.

b) Se $[\tau < t] \in \mathcal{F}_t$ para todo $t$, então $[\tau < t + \frac{1}{n}] \in \mathcal{F}_{t+\frac{1}{n}} \subset \mathcal{F}_{t+\frac{1}{m}}$ para $m \leq n$. Segue que $[\tau \leq t] = \bigcap_{n \geq m} [\tau < t + 1/n] \in \mathcal{F}_{t+1/m}$ para todo $m$ e conseguentemente

$$\left[\tau \leq t\right] \in \bigcap_m \mathcal{F}_{t+1/m} = \mathcal{F}_t \ ,$$

pois $\mathbb{F}$ é contínua a direita. A necessidade da condição segue de (a). □

**Proposição 2.3:** Seja $\tau_1, \tau_2, \ldots$ uma sequência de tempos de parada com respeito a $\mathbb{F}$. Então são tempos de parada com respeito a $\mathbb{F}$

a) $\tau_1 + \tau_2$,

b) $\min_{k \leq n} \tau_k$,

c) $\max_{k \leq n} \tau_k$ e

d) $\sup_n \tau_n$.

Se $\mathbb{F}$ é contínua a direita, também são tempos de parada

e) $\inf_n \tau_n$

f) $\underline{\lim}_{n \to \infty} \tau_n$

g) $\overline{\lim}_{n \to \infty} \tau_n$

Dem: Provaremos (b), (d) e (e) deixando os restantes como exercício.

(b) Segue da relação $[\min_{k\leq n}\tau_k \leq t] = \bigcup_{k\leq n}\{\tau_k \leq t\} \in \mathcal{F}_t$.

(d) Segue da relação $[\sup_n \tau_n \leq t] = \bigcap_n\{\tau_n \leq t\} \in \mathcal{F}_t$.

Para mostrarmos (e) usamos a Proposição 2.2.b e a relação

$$[\inf_n \tau_n < t] = \bigcup_n [\tau_n < t] \in \mathcal{F}_t. \qquad \square$$

Como toda constante é um tempo de parada, segue da Proposição 2.3 que todo tempo de parada $\tau$ pode ser aproximado por uma sequência de tempos de parada limitados como $\tau = \lim_{n\to\infty} \min(\tau, n)$. Um segundo tipo de aproximação é apresentado a seguir com $\mathbb{T} = [0,\infty)$. Para $\mathbb{T} = [0,T]$ resultado análogo é válido.

**Proposição 2.4:** Seja $\tau$ um tempo de parada com respeito a $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$. Então existe uma sequência $(\tau_n : n \in \mathbb{N})$ de tempos de parada com respeito a $\mathbb{F}$, assumido valores em um conjunto contável, que a próxima $\tau$ superiormente, isto é:

i) $\tau_n \geq \tau$ q.c

ii) $\tau_n \downarrow \tau$ quando $n \to \infty$

iii) $\tau_n(\omega) \in \{x_1, x_2, \ldots\}, x_i \in [0,\infty]$.

Dem: Seja

$$\tau_n(\omega) = \begin{cases} \frac{k}{2^n} & \text{se } \frac{(k-1)}{2^n} \leq \tau(\omega) < \frac{k}{2^n}, k = 1,2\ldots \\ \infty & \text{se } \tau(\omega) = \infty\,. \end{cases}$$

Então para $t \geq 0$,

$$
\begin{aligned}
\{\tau_n \leq t\} &= \bigcup_{k/2^n \leq t}\left[\tau_n = k/2^n\right] = \\
&= \bigcup_{k/2^n \leq t}\left[\frac{k-1}{2^n} \leq \tau < \frac{k}{2^n}\right] \in \mathcal{F}_t \,.
\end{aligned}
$$

Portanto $\tau_n$ é um tempo de parada e $\tau_n \geq \tau$. Para cada $\omega$ tal que $\tau(\omega) < \infty, 0 \leq \tau_n(\omega) - \tau(\omega) < 1/2^n$. Então $\lim\limits_{n\to\infty} \tau_n(\omega) = \tau(\omega)$ para todo $\omega$. Se $\tau_n(\omega) = 1/2^n$, então $\tau_{n+1}(\omega) = (2k-1)/2^{n+1}$ ou $2k/2^{n+1}$. Assim $\tau_n(\omega) \geq \tau_{n+1}(\omega)$ em todo $\Omega$. Segue da definição que $\tau_n$ assume um número contável de valores. □

Um tempo de parada não pode em geral ser aproximado inferiormente por uma sequência de de tempos de parada. Se existe de tempos de parada $\tau_n$, tal que $\tau_n \nearrow \tau$ q.c. e $\tau_n < \tau$ q.c., então $\tau$ é chamado <u>predizível</u>.

Interpretando $t$ como tempo, podemos interpretar $\mathcal{F}_t$ como a informação conhecida até o tempo $t$, isto é claro no caso $\mathcal{F}_t = (\{X(s) : s \leq t\}) = \mathcal{F}_t^X$, onde $X = \{X(t) : t \in \mathbb{T}\}$ é um processo estocástico. Para $\tau$ um tempo de parada com respeito a $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$, definimos <u>a $\sigma$-algebra de eventos anteriores a $\tau$</u>, como

$$
\mathcal{F}_\tau = \left\{A \in \mathcal{F}_\infty : A \bigcap [\tau \leq t] \in \mathcal{F}_t \ \text{ para todo } t \in \mathbb{T}\right\},
$$

onde $\mathcal{F}_\infty = \sigma\left(\bigcup_{t\in\mathbb{T}} \mathcal{F}_t\right)$ .

**Proposição 2.5:** Para $\tau$ tempo de parada com respeito a $\mathbb{F}$,

a) $\mathcal{F}_\tau$ é $\sigma$-algebra

b) $\tau$ é $\mathcal{F}_\tau$ mensurável

c) Se $\sigma$ é um outro tempo de parada com respeito a $\mathbb{F}$ e $\sigma \leq \tau$ q.c., então $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

<u>Dem:</u> a) Claro $\Omega \in \mathcal{F}_\tau$, pois $\Omega \bigcap \{\tau \leq t\} \in \mathcal{F}_t$ para todo $t$. Se $A \bigcap \{\tau \leq t\} \in \mathcal{F}_t$, então

$A^c \bigcap [\tau \leq t] = [\tau \leq t] \backslash (A \bigcap [\tau \leq t]) \in \mathcal{F}_t$. Portanto se $A \in \mathcal{F}_\tau$, então $A^c \in \mathcal{F}_\tau$. Se $A_n \bigcap [\tau \leq t] \in \mathcal{F}_t, n = 1, 2, \ldots$ $(\bigcup_n A_n) \bigcap [\tau \leq t] = \bigcup_n (A_n \bigcap [\tau \leq t]) \in \mathcal{F}_t$, o que mostra que $\mathcal{F}_\tau$ é fechado sob união finitas.

b) Se $x \geq 0, \{\tau \leq x\} \bigcap \{\tau \leq t\} = [\tau \leq \min(x,t)] \in \mathcal{F}_{\min(x,t)} \in \mathcal{F}_t$. e $[\tau \leq x] \in \mathcal{F}_x \subset \mathcal{F}_\infty$. Portanto $[\tau \leq x] \in \mathcal{F}_\tau$. Se $x < 0, [\tau \leq x] = \phi \in \mathcal{F}_\sigma$.

c) Se $A \in \mathcal{F}_\sigma$ e $\sigma \leq \tau$, então $A \bigcap [\tau \leq t] = A \bigcap [\sigma \leq t] \bigcap [\sigma < \tau] \in \mathcal{F}_t$ para todo $t$. Portanto $A \in \mathcal{F}_\tau$. □

Para um processo estocástico $X = \{X(t) : t \in \mathbb{T}\}$ e $\tau$ um tempo de parada com respeito a $\mathbb{F}$ com $P(\tau < \infty) = 1$, definimos $X^{(\tau)}$, o processo truncado em $\tau$, por $X^{(\tau)}(t) = X(t \wedge \tau)$.

**Proposição 2.6:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ um processo estocástico progressivamente mensurável com respeito a $\mathbb{F}$. Então $X^{(\tau)} = \{X^{(\tau)}(t) : t \in \mathbb{T}\}$ é um processo adaptado a $\mathbb{F}$ e $X(\tau)$ é uma variável aleatória $\mathcal{F}_\tau$-mensurável.

Dem: Para $t \in \mathbb{T}$ fixo, $\tau \wedge t$ é $\mathcal{F}_t$ mensurável. Segue que $\omega \longrightarrow (\tau(\omega) \wedge t, \omega)$ é $\mathcal{F}_t - \mathcal{B}([0,t]) \times \mathcal{F}_t$ mensurável e como $X$ é progressivamente mensurável, $(s,\omega) \longrightarrow X(s,\omega)$ também é $\mathcal{B}([0,t]) \times \mathcal{F}_t$ mensurável. Como $X^{(\tau)}(t) = X(\tau \wedge t)$ é a composição destas duas aplicações, ele é $\mathcal{F}_t$ mensurável.

Para $A \in \mathcal{B}(\mathbb{R}), [X(\tau) \in A] \bigcap [\tau \leq t] = [X(\tau \wedge t) \in A] \bigcap [\tau \leq t] \in \mathcal{F}_t$ e portanto $[X(\tau) \in A] \in \mathcal{F}_\tau$. □

## 2.2 Martingales, supermartingales e submartigales.

Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade completo, $\mathbb{T} = [0,\infty)$ ou $[0,T]$ ou $[0,\infty]$ e $\mathbb{F}$ uma filtração.

**Definição 2.7:** Um processo estocástico $X = \{X(t) : t \in \mathbb{T}\}$ é chamado um <u>martingale com respeito a $\mathbb{F}$</u>, ou um <u>$\mathbb{F}$-martingale</u> se:

i) $E|X(t)| < \infty$, para todo $t \in \mathbb{T}$,

ii) $X$ é adaptado a $\mathbb{F}$ e

iii) $E(X(t)/\mathcal{F}_s) = X(s)$ q.c., para $s \leq t,\ s,t \in \mathbb{T}$.

(Quando a filtração $\mathbb{F}$ é fixa, dizemos resumidamente $X$ é um martingale)

Se a igualdade em iii) é substituida por $\leq$ ou $\geq$ temos o que é chamado de <u>supermartingale e submartingale</u>, respectivamente. Note que se $X$ é um supermartingale, $-X$ é uma submartingale. Consequentemente muitos resultados necessitam ser provados somente para uma destas classes.

Um exemplo simples mas importante de martingale é obtido definindo $X(t) = E(Y/\mathcal{F}_t)$, onde $Y$ é uma variável aleatória com $E|Y| < \infty$. Outro exemplo interessante aparece quando consideramos em $(\Omega, \mathcal{F})$ duas probabilidades $P$ e $Q$ com $P << Q$ e $\mathbb{F}$ uma filtragem. Se restringirmos $P$ e $Q$ a $\mathcal{F}_t$ e denotarmos estas restrições por $P_t$ e $Q_t$ respectivamente, então $\left(\frac{dP_t}{dQ_t} : t \in \mathbb{T}\right)$ é um martingale (mostra estes resultados como exercício).

**Proposição 2.8:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ um supermartingale. Então

i) $EX(t)$ é não decrescente;

ii) $X$ é martingale se e somente se $EX(t)$ é independente de $t$;

iii) Se $f$ é uma função concava com $E|f(X(t))| < \infty$, então $\{f(X(t) : t \in \mathbb{T}\}$ é supermartingale;

iv) Se $Y = \{Y(t) : t \in \mathbb{T}\}$ é um outro supermartingale e $a \geq 0$, então $aX + Y$ e $X \times Y$ são supermartingales.

Dem: exercício $\square$

Seja $X = \{X(t) : t \in \mathbb{T}\}$ um submartingale e $\mathbb{T}_0$ um subconjunto finito de $\mathbb{T}$. Para $a < b$ defina:

$$\begin{aligned}
&\tau_0 = 0, \\
&\tau_1 = \min\{t \in \mathbb{T}_0 : X(t) \leq a\}, \\
&\tau_2 = \min\{t \in \mathbb{T}_0, t > \tau_1 : X(t) \geq b\}, \\
&\vdots \\
&\tau_{2k-1} = \min\{t \in \mathbb{T}_0, t > \tau_{2k-2} : X(t) \leq a\}, \\
&\tau_{2k} = \min\{t \in \mathbb{T}_0, t > \tau_{2k-1} : X(t) \geq b\}, \\
&\vdots
\end{aligned}$$

Se $\min\limits_{t \in \mathbb{T}_0} X(t) > a$, então $\tau_1 = \max \mathbb{T}_0$ e $\tau_2, \tau_3, \ldots$ não são definidos. Esta regra também se aplica aos $\tau's$ subsequentes.

O número de cruzamentos crescentes de um intervalo $(a, b)$ por um submartingale $X$ restrito a um conjunto finito $\mathbb{T}_0$ é definido por:

$$U(a, b; \mathbb{T}_0) = \max\Big\{k : \tau_{2k} \text{ é definido}\Big\}$$

**Proposição 2.9:** Seja $X$ um submartingale e $\mathbb{T}_0$ um subconjunto finito de $\mathbb{T}, \mathbb{T}_0 \subset [0, T], T < \infty$. Então

$$E\Big(U(a, b; \mathbb{T}_0)\Big) \leq \frac{E\big(X(T) - a\big)^+}{b - a}.$$

Dem: Como o número de cruzamentos crescentes do intervalo $(a,b)$ pelo submartingale $X$ restrito a $\mathbb{T}_0$ corresponde ao número de cruzamentos crescentes do intervalo $(0,b-a)$ pelo submartingale $\{(X(t)-a)^+ : t \in \mathbb{T}\}$, podemos assumir sem perder generalidade que $X$ é não negativo, $a=0$ e $b>0$.

Seja $\mathbb{T}_0 = \{t_1,\ldots,t_n\}, Y_k = X(t_k), Y_0 = 0$ e

$$\chi_i = \begin{cases} 1 & \text{se} \quad \tau_m < i < \tau_{m+1} \quad \text{para } m \text{ impar,} \\ \\ 0 & \text{se} \quad \tau_m < i < \tau_{m+1} \quad \text{para } m \text{ par.} \end{cases}$$

Então

$$b\, U\Big(0,b;\mathbb{T}_0\Big) = \sum_{i=1}^{n} \chi_i \Big(Y_i - Y_{i-1}\Big) \;\; \text{q.c.},$$

e

$$\Big\{\chi_i = 1\Big\} = \bigcup_{\substack{m \\ \text{impar}}} \Big([\tau_m < i] \backslash [\tau_{m+1} < 1]\Big) .$$

Segue que:

$$\begin{aligned}
b\, EU(0,b;\mathbb{T}_0) &= E\Big(\sum_{i=1}^{n}\chi_i(Y_i - Y_{i-1})\Big) = \\
&= \sum_{i=1}^{n}\int_{\chi_i=1}\Big(Y_i - Y_{i-1}\Big)dP = \\
&= \sum_{i=1}^{n}\int_{\chi_i=1}E\Big(Y_i - Y_{i-1}/\mathcal{F}_{t_{i-1}}\Big)dP = \\
&= \sum_{i=1}^{n}\int_{\chi_i=1}\Big(E\Big(Y_i/\mathcal{F}_{t_{i-1}}\Big) - Y_{i-1}\Big)dP \leq \\
&\leq \sum_{i=1}^{n}\int\Big(E\Big(Y_i/\mathcal{F}_{t_{i-1}}\Big) - Y_{i-1}\Big)dP = \\
&= \sum_{i=1}^{n}\Big(E(Y_i) - E(Y_{i-1})\Big) = \\
&= EY_n = EX(t_n) \leq EX(T) \qquad \square
\end{aligned}$$

Se $\mathbb{T}_0$ é um conjunto contável de $\mathbb{T}$, o número de cruzamentos crescentes de um intervalo $(a,b)$ por um submartingale $X$ restrito a $\mathbb{T}_0$ pode ser definido por

$$U\Big(a,b;\mathbb{T}_0\Big) = \lim_{n\to\infty} U\Big(a,b;\mathbb{T}_n\Big),$$

onde $(\mathbb{T}_n : n \in \mathbb{N})$ é uma sequência crescente de subconjuntos finitos de $\mathbb{T}_0, \bigcup_n \mathbb{T}_n = \mathbb{T}_0$.

**Corolário 2.10:** Seja $X$ um submartingale e $\mathbb{T}_0$ um subconjunto contável de $\mathbb{T}$, $\mathbb{T}_0 \subset [0,T], T < \infty$. Então

$$EU\Big(a,b;\mathbb{T}_0\Big) \leq \frac{E\Big((X(T)-a)^+\Big)}{b-a}$$

Dem: Pelo teorema da convergência monótona

$$EU\left(a,b;\mathbb{T}_0\right) = \lim_{n\to\infty} EU\left(a,b;\mathbb{T}_n\right) \leq \frac{E\left(X(T)-a\right)^+}{b-a}$$

onde $\mathbb{T}_n \nearrow \mathbb{T}, \mathbb{T}_n$ finito. □

Outros resultados envolvendo submartingales restritos a um subconjunto de índices finito ou contável são dados nas proposições seguintes.

**Proposição 2.11:** Seja $X$ um submartigale e $\mathbb{T}_0$ um subconjunto finito de $\mathbb{T}$. Se $\tau$ e $\sigma$ são dois tempos de parada com valores em $\mathbb{T}_0$ e $P(\sigma < \tau) = 1$, então

$$X(\sigma) \leq E\left(X(\tau)/\mathcal{F}_\sigma\right) \quad \text{q.c..}$$

Dem: $E|X(\tau)|$ é finita. De fato, para $\mathbb{T}_0 = \{t_1 < t_2 < \ldots < t_n\}$,

$$\begin{aligned} E|X(\tau)| &= \sum_{k=1}^{n} \int_{\{\tau=t_k\}} |X(\tau)|dP = \sum_{k=1}^{n} \int_{\{\tau=t_k\}} |X(t_k)|dP \\ &\leq \sum_{k=1}^{n} \int_{\Omega} |X(t_k)|dP = \sum_{k=1}^{n} E|X(t_k)| < \infty. \end{aligned}$$

Temos que mostrar que para todo $A \in \mathcal{F}_\sigma$

$$\int_A E\left(X(\tau)/\mathcal{F}_\sigma\right)dP = \int_A X(\tau)dP \geq \int_A X(\sigma)dP.$$

Como $A = \bigcup_{k=1}^{n} A \bigcap \{\sigma = t_k\}$ é suficiente mostrar que

$$\int_{A\cup\{\sigma=t_k\}} X(\tau)dP \geq \int_{A\cap\{\sigma=t_k\}} X(\sigma)dP.$$

Dado que $P(\sigma < \tau) = 1$, temos

$$\int_{A\cap\{\sigma=t_k\}} X(\tau)dP = \int_{A\cap\{\sigma=t_k\}\cap(\sigma<\tau)} X(\tau)dP = \sum_{i=k+1}^{\infty} \int_{A\cap\{\sigma=t_k\}\cap\{\tau=t_i\}} X(\tau)dP =$$

$$= \sum_{i=k+1}^{\infty} \int_{A\cap\{\sigma=t_k\}\cap\{\tau=t_i\}} X(t_i)dP \ .$$

Como $A\bigcap\{\sigma = t_k\}\bigcap\{\tau = t_i\} \in \mathcal{F}_{t_k}$,

$$\int_{A\cap\{\sigma=t_k\}} X(\tau)dP = \sum_{i=k+1}^{\infty} \int_{A\cap\{\sigma=t_k\}\cap\{\tau=t_i\}} E\Big(X(t_i)/\mathcal{F}_{tk}\Big)dP$$

$$\geq \sum_{i=k+1}^{\infty} \int_{A\cap\{\sigma=t_k\}\cap\{\tau=t_i\}} X(t_k)dP =$$

$$= \int_{A\cap\{\sigma=t_k\}} X(t_k)dP = \int_{A\cap\{\sigma=t_k\}} X(\sigma)sP$$

Agora

$$\int_A E(X(\tau)/\mathcal{F}_\sigma)dP = \int_A X(\tau)dP = \sum_{k=1}^{\infty} \int_{A\cap\{\sigma=t_k\}} X(\sigma)dP \geq \int_A X(\sigma)dP$$

Portanto

$$E(X(\tau)/\mathcal{F}_\sigma) \geq X(\sigma) \qquad \text{q.c..}$$ □

**Proposição 2.12:** Seja $X$ um submartingale, $\mathbb{T}_0 \subset \mathbb{T}$ contável, $\mathbb{I} = [a,b] \subset \mathbb{T}$ e $\lambda > 0$. Então

$$P\Big(\sup_{t\in \mathbb{I}\cap \mathbb{T}_0} X(t) \geq \lambda\Big) \leq \lambda^{-1} EX^+(b)$$

e

$$P\Big(\inf_{t\in \mathbb{I}\cap \mathbb{T}_0} X(t) \leq -\lambda\Big) \leq \lambda^{-1}\Big(EX^+(b) - EX(a)\Big).$$

<u>Dem:</u> Inicialmente consideremos $\mathbb{T}_0$ finito. Seja $\tau = \inf\{t \in \mathbb{T}_0 \bigcap \mathbb{I} : X(t) \geq \lambda\}$, $\tau = b$ se $\sup\{t \in \mathbb{T}_0 \bigcap \mathbb{I} : X(t)\} < \infty$. Claro

$$\begin{aligned} EX(b) &\leq EX(\tau) = \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)\geq\lambda} X(\tau)dP + \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)<\lambda} X(\tau)dP \\ &\geq \lambda P\Big(\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t) \geq \lambda\Big) + \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)<\lambda} X(b)dP. \end{aligned}$$

Portanto

$$\begin{aligned} \lambda P\Big(\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t) \geq \lambda\Big) &\leq EX(b) + \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)<\lambda} X(b)dP \\ &= \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)\geq\lambda} X(b)dP \leq \int_{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t)\geq\lambda} X^+(b)dP \\ &\leq EX^+(b). \end{aligned}$$

Se $\mathbb{T}_0$ é contável, consideremos $\mathbb{T}_1 \subset \mathbb{T}_2 \subset \ldots$ uma sequência crescente tal que $\bigcup_n \mathbb{T}_n = \mathbb{T}_0$. Então

$$\Big\{\sup_{t\in \mathbb{T}_n\cap \mathbb{I}} X(t) \geq \lambda\Big\} \nearrow \Big\{\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t) \geq \lambda\Big\}$$

Segue que

$$P\left(\sup_{t\in \mathbb{T}_0\cap \mathbb{I}} X(t) \;\geq \lambda\right) \leq \lim_{n\to\infty} \lambda^{-1} EX^+(b)$$

$$= \lambda^{-1} EX^+(b).$$

A demonstração da segunda desigualdade é análoga. □

A suposição que $\mathbb{T}_0$ é finito ou contável pode ser removida nas proposições acima e em outros resultados se condições são assumidas a respeito da filtração e trajetórias do processo. Antes de provarmos estes resultados enunciaremos por conveniência alguns resultados relacionados com estas novas condições. Para as demonstrações destes fatos nos referimos Meyer [M2].

**Fato 2.13:** Se $X = \{X(t) : t \in \mathbb{T}\}$ é um supermartingale. Então as trajetórias de $X$ tem q.c. limites a direita e a esquerda em todo ponto $t \in \mathbb{T}$ no seguinte sentido: Seja $S$ um conjunto contável e denso de $\mathbb{T}$. Então

$$\lim_{\substack{s\to t\\ s<t\\ s\in S}} X(s,\omega) \quad \text{e} \quad \lim_{\substack{s\to t\\ s>t\\ s\in S}} X(s,\omega)$$

existem q.c. para todo $t \in \mathbb{T}$

**Fato 2.14:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ um supermartingale com respeito a uma filtração $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$. Se $\mathbb{F}$ é contínua a direita, então $X$ tem uma versão contínua a direita se e somente se a função média $EX(t)$ é contínua a direita. Em particular todo martingale tem uma versão contínua a direita.

**Lema 2.15:** Se $X$ é uma variável aleatória com $E|X| < \infty$ e $\{\mathcal{F}_\lambda : \lambda \subset \Lambda\}$ uma família de sub-$\sigma$-algebra, então $\{E(X/\mathcal{F}_\lambda) : \lambda \in \Lambda\}$ é uniformemente integrável.

Dem: exercício. □

**Teorema 2.16:** Seja $X\{X(t) : t \in \mathbb{T}\}$ um submartingale contínuo a direita com respeito a $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ e sejam $\tau_1$ e $\tau_2$ tempos de parada com respeito a $\mathbb{F}$. Então para cada $T > 0$

$$E\Big(X(\tau_2 \wedge T)/\mathcal{F}_{\tau 1}\Big) \geq X\Big(\tau_1 \wedge \tau_2 \wedge T\Big) .$$

Se $P(\tau_2 < \infty) = 1, E|X(\tau_2)| < \infty$ e $\lim\limits_{T\to\infty} E(|X(T)|\mathbb{1}_{(\tau_2 > T)}) = 0$. Então

$$E\Big(X(\tau_2)/\mathcal{F}_{\tau_1}\Big) \geq X\Big(\tau_1 \wedge \tau_2\Big) .$$

Dem: Para $i = 1, 2$, seja

$$\tau_{i,n} = \begin{cases} \dfrac{k+1}{2^n} & \text{se } \dfrac{k}{2^n} \leq \tau_i < \dfrac{k+1}{2^n}, \\ \infty & \text{se } \tau_i = \infty . \end{cases}$$

Claro $\tau_{i,n}$ é um tempo de parada com respeito a $\mathbb{F}$. Se $a \in \mathbb{R}$ e $T > 0$, pela Proposição 2.11

$$E\Big(X(\tau_{2,n} \wedge T) \vee a/\mathcal{F}_{\tau_{1,n}}\Big) \geq X\Big(\tau_{1,n} \wedge \tau_{1,n} \wedge T\Big) \vee a$$

Como $\tau_1 < \tau_{1,n}, \mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_{1,n}}$, segue que

$$E\Big(X(X_{2,n} \wedge T) \vee a/\mathcal{F}_{\tau_1}\Big) \geq E\Big(X(\tau_{1,n} \wedge \tau_{1,2} \wedge T) \wedge a/\mathcal{F}_{\tau_1}\Big) .$$

Novamente, Proposição 2.11 implica

$$a \leq X\Big(\tau_{2,n} \wedge T\Big) \vee a \leq E\Big(X(t) \vee a/\mathcal{F}_{\tau_{2,n}}\Big) .$$

Segue do Lema 2.15 que as sequências $(X(\tau_{2,n} \wedge T) \vee a : n \in \mathbb{N})$ e $(X(\tau_{1,n} \wedge \tau_{2,n} \wedge T) \vee a)$ são uniformemente integráveis. Fazendo $n \to \infty$, a continuidade de a direita de $X$ e a integrabilidade uniforme das sequências implicam que

$$\begin{aligned} E\Big(X(\tau_2 \wedge T) \vee a/\mathcal{F}_{\tau_1}\Big) &\geq E\Big(X(\tau_1 \wedge \tau_2 \wedge T) \vee a/\mathcal{F}_{\tau_1}\Big) \\ &= X\Big(\tau_1 \wedge \tau_2 \wedge T\Big) \vee a \,. \end{aligned}$$

O primeiro resultado segue fazendo $a \to -\infty$ e o segundo, usando as hipóteses adicionais e fazendo $T \to \infty$. □

**Teorema 2.17:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ um submartingale contínuo a direita, $\mathbb{I} \subset [a, b] \subset \mathbb{T}$ e $\lambda > 0$. Então

$$P\left(\sup_{t \in \mathbb{I}} X(t) \geq \lambda\right) \leq \lambda^{-1} EX^+(b)$$

e

$$P\left(\inf_{t \in \mathbb{I}} X(t) \leq -\lambda\right) \leq \lambda^{-1}\Big(EX^+(b) - EX(a)\Big)$$

Dem: Como as trajetórias são contínuas a direita, os evento

$$\left\{\sup_{t \in \mathbb{I}} X(t) \geq \lambda\right\} = \left\{\sup_{t \in \mathbb{I} \cap \mathbb{Q}} X(t) \geq \lambda\right\}$$

e

$$\left\{\inf_{t \in \mathbb{I}} X(t) \leq -\lambda\right\} = \left\{\inf_{t \in \mathbb{I} \cap \mathbb{Q}} X(t) \leq \lambda\right\}$$

pertencem a $\mathcal{F}$($\mathbb{Q}$ = conjunto dos racionais) e as desigualdades seguem da Proposição 2.12. □

**Teorema 2.18:** Seja $X = \{X(t) : t \in \mathbb{T}\}$ um submartingale não negativo e contínuo a direita. Então para $\mathbb{I} = [a,b] \dot{\subset} \mathbb{T}$ e $p > 1$

$$E \sup_{t\in\mathbb{I}} \Big[X(t)\Big]^p \leq \left(\frac{p}{p-1}\right)^p E\Big[X(b)\Big]^p$$

<u>Dem:</u> Seja $\lambda > 0$ e $\tau = \inf\{t : X(t) > \lambda\}$. $\tau$ é um tempo de parada com respeito a $\mathcal{F}_{t+}$ (mostrar este fato como exercício). A continuidade a direita de $X$ implica que $X(\tau) \geq \lambda$ se $\lambda < \infty$. Portanto

$$\left\{\sup_{t\in\mathbb{I}} X(t) > \lambda\right\} \subset \left\{\tau \leq b\right\} \subset \left\{\sup_{t\in\mathbb{I}} X(t) \geq \lambda\right\}$$

e os três eventos tem mesma probabilidade exceto para um número contável de $\lambda$'s. Como $E\Big(X(\tau \wedge b)\Big) \leq E\Big(X(b)\Big)$, segue que

$$\lambda P\Big(\tau \leq b\Big) \leq E\Big(X(\tau)\mathbb{1}_{\{\tau\leq b\}}\Big) \leq E\Big(X(b)\mathbb{1}_{\{\tau\leq b\}}\Big) .$$

Seja $\varphi : \mathbb{T} \to \mathbb{R}$, diferenciável com $\varphi' \geq 0$ e $\varphi(0) = 0$. Defina $Z = \sup_{t\in[a,b]} X(t)$. Então para $c > 0$,

$$E\Big(\varphi(Z \wedge c)\Big) = \int_0^c \varphi'(x) P\Big(Z > x\Big) dx$$

$$\leq \int_0^c \varphi'(a) E\Big(X(b)\mathbb{1}_{(Z\geq x)}\Big) dx$$

$$= E\Big(X(b) \int_0^{Z\wedge c} \varphi'(x) x^{-1} dx\Big) .$$

Se $\varphi(x) = x^p$ para $p > t$, então

$$E\Big((Z \wedge c)^p\Big) \leq \frac{p}{p-1} E\Big(X(b)(Z \wedge c)^{p-1}\Big)$$

$$\leq \frac{p}{p-1} E\Big(X(b)^p\Big)^{1/p} E\Big[(Z \wedge c)^p\Big]^{\frac{p-1}{p}} .$$

Portanto,

$$E\Big[(Z \wedge c)^p\Big] \leq \frac{p}{p-1} E\Big(X(b)^p\Big)^{1/p} .$$

O resultado segue fazendo $c \to \infty$. □

Quando $\mathbb{T} = [0,\infty)$, sob condições não muito restritivas submartingales, martingales e também supermartingales convergem com probabilidade 1 quando $t \to \infty$. Este resultado, apesar de não ser utilizado no texto, é extremamente útil e tem ramificações importantes em probabilidade e estatística.

**Teorema 2.19:** Seja $X = \{X(t) : t \in [0,\infty)\}$ um submartingales com trajetórias contínuas a direita. Se $\sup_t EX^+(t) < \infty$, então existe uma variável aleatória $X(\infty)$ com $E|X(\infty)| < \infty$ tal que

$$X(t) \longrightarrow X(\infty)$$

quando $t \to \infty$, com probabilidade 1.

Dem: Para todo par de reais $a < b$ e $n \geq 0$ segue do Corolário 2.10

$$\begin{aligned} E\Big(U(a,b;\mathbb{Q}\cap[0,n])\Big) &\leq \frac{EX^+(n) + a^-}{b-a} \\ &\leq \frac{\sup EX^+(t) + a^-}{b-a} < \infty . \end{aligned}$$

Portanto $U(a,b,\mathbb{T}) = \bigcup(a,b,\mathbb{Q}) = \lim_{n\to\infty} U(a,b,\mathbb{Q}\cap[0,n]\Big) < \infty$ q.c.. Segue que, se

$$E_{a,b} = \left\{ \varliminf_{t\to\infty} X(t) < a < b < \varlimsup_{n\to\infty} X(t) \right\},$$

então

$$E_{a,b} \subset \left\{ U(a,b;\mathbb{T}) = \infty \right\}$$

e

$$P(E_{a,b}) = 0 \quad \text{para todo } \; a < b \; .$$

Assim, se

$$E = \bigcup_{\substack{a,b \\ a<b \\ a,b \text{ racionais}}} Ea,b = \left\{ \varliminf_{t\to\infty} X(t) < \varlimsup_{t\to\infty} X(t) \right\},$$

então $P(E = 0)$. Segue que $X(\infty) = \lim\limits_{t\to\infty} X(t)$ existe q.c.. Seja $t_n \nearrow \infty$. Então

$$E|X(\infty)| = E\left(\varliminf_{n\to\infty} |X(t_n)|\right) \leq \varliminf_{n\to\infty} E|X(t_n)| =$$

$$\varliminf_{n\to\infty} \left(2EX^+(t_n) - EX(t_n)\right) \leq 2 \varliminf_{n\to\infty} EX^+(t_n) - EX(t_1) < \infty$$

Portanto $E|X(\infty)| < \infty$. $\square$

## 2.3 Decomposição de Doob-Meyer.

Seja $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ uma filtragem. Um $\mathbb{F}$-submartingale $X = \{X(t) : t \in \mathbb{T}\}$ com trajetórias contínuas a direita pertence a classe DL se para cada $t \in \mathbb{T}, \{X(t \wedge \tau) : \tau$ tempo de parada com respeito a $\mathbb{T}\}$ é uniformemente integrável. Em particular, se $X$ é um martigale ou se $X$ é positivo, então $X$ pertence a classe $DL$.

Um processo $A = \{A(t) : t \in \mathbb{T}\}$ é chamado processo crescente com respeito a $\mathbb{F}$ se:

i) $A$ é adaptado a $\mathbb{F}$,

ii) $A$ tem trajetórias não decrescentes q.c.,

iii) $A(0) = 0$ q.c. $E|A_t| < \infty$ para cada $t \in \mathbb{T}$.

Seja $X = \{X(t) : t \in \mathbb{T}\}$ mensurável e não negativo. Para cada $\omega \in \Omega$ e $t \in \mathbb{T}$, considere a integral de Lebesgue-Stieltjes

$$Y(t,\omega) = \int_{[0,t]} X(t,\omega)dA(t,\omega) \equiv \int_0^t X(t,\omega)dA(t,\omega).$$

Se $X$ é progressivamente mensurável com respeito a $\mathbb{F}$, então o processo $Y = \{Y(t) : t \in \mathbb{T}\}$ é adaptado a $\mathbb{F}$, progressivamente mensurável, com trajetórias contínuas e para cada tempo de parada $\tau$ com respeito a $\mathbb{F}$, $Y(\tau) = \int_0^\tau X(t)dA(t)$ é $\mathcal{F}_\tau$ mensurável.

Um processo crescente é chamado natural se

$$E\int_0^t X(s)dA(s) = E\int_0^t X(s-)dA(s)$$

para todo $t \in \mathbb{T}$ e cada martingale $X$ não negativo, limitada e contínua direta.

**Teorema 2.20** (Doob-Meyer). Seja $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ uma filtragem contínua a direita. Um $\mathbb{F}$-submartingale $X = \{X(t) : t \in \mathbb{T}\}$ pertencente a classe $DL$ tem uma decomposição de Doob-Meyer

$$X(t) = M(t) + A(t) \ ,$$

onde $M$ é um $\mathbb{F}$-martingale e $A$ é um processo crescente único (exceto por equivalência) que satisfaz

$$E\int_0^{t\wedge\tau} Y(s)dA(s) = E\int_0^{t'} Y(s-)dA(s)$$

para todo $I\!\!F$-martingale $Y$ contínuo a direita e não negativo, $\tau$ tempo de parada com respeito a $I\!\!F$ e $t \in I\!\!T$.

Dem: Para $\varepsilon > 0$, seja

$$X_\varepsilon(t) = E\left(\frac{1}{\varepsilon}\int_0^\varepsilon X(t+s)ds/\mathcal{F}_t\right).$$

Então $X_\varepsilon$ é um submartingale e $\lim_{\varepsilon\to 0} E\Big(|X_\varepsilon(t) - X(t)|\Big) = 0, \quad t \geq 0$. Seja

$$Y_\varepsilon(t) = \frac{1}{\varepsilon}E\Big(X(t+\varepsilon) - X(t)/\mathcal{F}_t\Big)$$

e

$$A_\varepsilon(t) = \int_0^t Y_\varepsilon(s)ds\ .$$

Como $X$ é um submartingale, $Y_\varepsilon$ é crescente. Seja

$$M_\varepsilon(t) = X_\varepsilon(t) + A_\varepsilon(t).$$

Então $M_\varepsilon$ é uma martingale, pois para $s < t$ e $B \in \mathcal{F}_s$,

$$\int_B \Big(M_\varepsilon(s+t) - M_\varepsilon(s)\Big)dP =$$

$$\int_B \left[\frac{1}{\varepsilon}\int_0^\varepsilon \Big(X(s+t+u) - X(t+u)\Big)du - \frac{1}{\varepsilon}\int_s^{s+t}\Big(X(u+\varepsilon) - X(u)du\Big)\right]dP = 0.$$

Temos também que $\{A_\varepsilon(t) : 0 < \varepsilon \leq 1\}$ é uniformemente integrável para cada $t \geq 0$. De fato, seja $\tau_{\varepsilon,x} = \inf\Big\{t : A_\varepsilon(t) \geq x\Big\}$. Então

$$E(A_\varepsilon(t) - x \wedge A_\varepsilon(t)) = E(A_\varepsilon(t) - A_\varepsilon(\tau_{\varepsilon,x} \wedge t)) =$$

$$= E(X_\varepsilon(t) - X_\varepsilon(\tau_{\varepsilon,x} \wedge t)) = E\Big[X_\varepsilon(t) - X_\varepsilon(\tau_{\varepsilon,x} \wedge t)\mathbb{1}_{\{\tau_{\varepsilon,x}<t\}}\Big]$$

$$= \frac{1}{\varepsilon}\int_0^\varepsilon E\Big[\Big(X(t+s) - X(\tau_{\varepsilon,x} \wedge t + s)\Big)\Big]ds. \qquad (*)$$

Como $P(\tau_{\varepsilon,x} < t) \leq x^{-1}$, temos que $EA_\varepsilon(t) \leq x^{-1}E(X(t+\varepsilon) - X(0))$. A integrabilidade uniforme de $\{X(\tau \wedge (t+1)) : \tau\ \mathbb{F}\text{-tempo de parada}\ \}$ ($X$ é de classe $D.L.$) implica que $(*)$ converge a zero uniformemente em $0 < \varepsilon \leq 1$ quando $x \to \infty$. Portanto $\{A_\varepsilon(t) : 0 < \varepsilon \leq 1\}$ é uniformemente integrável para cada $t \geq 0$. Temos então que para cada $t$, existe $\varepsilon_n \to 0$ e $A(t)$ tal que:

$$E\Big(A_\varepsilon(t)\mathbb{1}_B\Big) \longrightarrow E\Big(A(t)\mathbb{1}_B\Big)$$

quando $\varepsilon_n \longrightarrow 0$, para cada $B \in \mathcal{F}$. Pelo método de diagonalização podemos tomar $\varepsilon_n$ tal que

$$E\Big(A_{\varepsilon_n}(t)\mathbb{1}_B\Big) \longrightarrow E\Big(A(t)\mathbb{1}_B\Big)$$

para todo $t \in \mathbb{Q} \bigcap \mathbb{T}$, quando $n \to \infty$.

Seja $s < t, s, t \in \mathbb{Q} \bigcap \mathbb{T}$ e $B = \{A(t) < A(s)\}$. Então

$$0 \geq E\Big(A(t) - A(s)\Big)\mathbb{1}_B = \lim_{n\to\infty} E\Big(A_{\varepsilon_n}(t) - A_{\varepsilon_n}(s)\Big)\mathbb{1}_B \geq 0,$$

o que mostra que $A(s) \leq A(t)$ q.c.. Para $s, t \in \mathbb{Q} \bigcap \mathbb{T}$ e $B \in \mathcal{F}_t$,

$$E\Big(X(t+s) - A(t+s) - X(t) + A(t)\Big)\mathbb{1}_B$$

$$= \lim_{n\to\infty} E\Big(M_{\varepsilon_n}(t+s) - M_{\varepsilon_n}(t)\Big)\mathbb{1}_B = 0\ .$$

Portanto se $M(t) = X(t) - A(t)$ para $t \in \mathbb{Q} \bigcap \mathbb{T}$, segue que $E(M(t)/\mathcal{F}_s) = M(s)$ para $s \leq t, s, t \in \mathbb{Q} \bigcap \mathbb{T}$. Pela continuidade a direita de $\mathbb{F}$ e $X$, podemos estender $M$ como um $\mathbb{F}$-martingale contínuo a direita e segue que $A$ tem uma modificação crescente e contínua a direita.

Temos também para um $\mathbb{F}$-martingale não negativo $Y$ que

$$E\Big(Y(t)A(t)\Big) = \lim_{n\to\infty} E\Big(Y(t)A_{\varepsilon_n}(t)\Big) =$$

$$= \lim_{n\to\infty} E\int_0^t Y(s-)dA_{\varepsilon_n}(s) =$$

$$= \lim_{n\to\infty} E\int_0^t Y(s-)\varepsilon_n^{-1}E\Big(X(s+\varepsilon_n) - X(s)/\mathcal{F}_s\Big)ds$$

$$= \lim_{n\to\infty} E\int_0^t Y(s-)\varepsilon_n^{-1}\Big(A(s+\varepsilon_n) - A(s)\Big)ds$$

$$= \lim_{n\to\infty} E\int_0^{t+\varepsilon_n}\int_0^t Y(s-)\varepsilon_n^{-1}\mathbb{1}_{(s,s+\varepsilon_n)}(u)dsdA(u)$$

$$= E\int_0^t Y(u-)dA(u)\ .$$

O mesmo argumento vale com $t \wedge \tau$ no lugar de $t$.

Para mostrarmos a unicidade, sejam $A_1$ e $A_2$ dois processos crescentes que satisfazem o resultado do teorema. Então $A_1 - A_2$ é um martingale e para $Y$ um martingale limitado e contínuo a direita temos que:

$$\int_0^t Y(s-)d\Big(A_1 - A_2\Big)(s) = \int_0^t Y(s-)dA_1(s) - \int_0^t Y(s-)dA_2(s)$$

é um $\mathbb{F}$-martingale contínuo a direita. Portanto

$$E\Big(Y(t)A_1(t)\Big) = E\left(\int_0^t Y(s-)dA_1(s)\right) = E\int_0^t Y(s-)dA_2(s) = E\Big(Y(t)A_2(t)\Big)$$

note que $\Big(E\Big(A_1(t) - A_2(t)\Big) = E\Big(A_1(0) - A_2(0)\Big) = 0\Big)$. Seja $B = \{A_1(t) > A_2(t)\}$ e $Y(t)$ uma versão contínua a direita de $E(\mathbb{1}_B/\mathcal{F}_t)$. Então

$$E\Big(A_1(t) - A_2(t)\Big)\mathbb{1}_B = 0\ .$$

Analogamente se $B = \{A_2(t) > A_1(t)$. Segue que $A_1(t) = A_2(t)$ q.c.. □

## 2.4 Variação quadrática.

**Definição 2.21:** Um $\mathbb{F}$-martingale $M = \{M(t) : t \in \mathbb{T}\}$ é chamado quadrado integrável se $\sup_{t\in\mathbb{T}} M^2(t) < \infty$.

Para um martingale $M$ quadrado integrável e contínuo a direita, temos que $\{M^2(t) : t \in \mathbb{T}\}$ é um submartingale contínuo a direita e uniformemente integrável. Se $\mathbb{F}$ é contínua a direita, temos pelo Teorema 2.20 que existe um processo crescente único – neste caso denotado por $\langle M\rangle$ – tal que

$$M^2 - \langle M\rangle$$

é um martingale. O proceso $\langle M\rangle$ é chamado variação quadrática de $M$. Se $0 \leq s \leq t$

$$\begin{aligned} E\Big(\big(M(t) - M(s)\big)^2/\mathcal{F}_s\Big) &= E\Big(M^2(t) - M^2(s)/\mathcal{F}_s\Big) \\ &= E\Big(\langle M\rangle(t) - \langle M\rangle(s)/\mathcal{F}_s\Big)\ . \end{aligned}$$

Por exemplo, consideremos $W = \{W(t) : t \in [0,T]\}$ um processo de Wiener. Então $W$ é um martingale quadrada integrável com respeito a filtração $\mathcal{F}_t^W = \sigma(W(s) : s \leq t)$. É fácil ver que

$$W^2(t) - t$$

é um martingale. Portanto $\langle W\rangle(t) = t$.

Para $M$ e $N$ martingales quadrado integráveis, seja

$$\begin{aligned}\langle M,N\rangle &\equiv \frac{1}{2}\Big(\langle M+N\rangle - \langle M\rangle - \langle N\rangle\Big)\\ &= \frac{1}{2}\Big(\langle M+N, M+N\rangle - \langle M,M\rangle - \langle N,N\rangle\Big)\ .\end{aligned}$$

$\langle M,N\rangle$ é chamada <u>variação cruzada</u> de $M$ e $N$.

**Proposição 2.22:** Se $M$ e $N$ são martingales quadrado integráveis, então

$$MN - \langle M,N\rangle$$

é um martingale

<u>Dem:</u> $MN - \langle M,N\rangle = \frac{1}{2}((M+N)^2) - \langle M+N, M+N\rangle(M^2 - \langle M,M\rangle) - (N^2 - \langle N,N\rangle$ □

Se $\langle M,N\rangle = 0$, então $M$ e $N$ são ditos ortogonais. Note que se $\langle M,N\rangle = 0$, então $MN$ é um martingale.

Pode ser mostrado que se $M$ é contínuo e quadrado integrável, então

$$\langle M\rangle(t) = L_2 - \lim_{n\to\infty}\sum_k \Big(M(s_{n,k+1}) - M(s_{n,k})\Big)\ ,$$

onde $\{0 = s_{n,0} < s_{n,1}, < \ldots\}, n = 1,2,\ldots$ são partições de $[0,t]$ com $\max_k(s_{n,k+1} - s_{n,k}) \longrightarrow 0$, quando $n \to \infty$.

## 2.5 Martingales locais.

Um processo estocástico $M = \{M(t); t \in \mathbb{T}\}$ é chamado um <u>martingale local</u> com respeito a uma filtração $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$, se existe uma sequência $\tau_1 < \tau_2 < \ldots$ de tempo de parada com respeito a $\mathbb{F}$ com $\tau_n \to \infty$ q.c. tal que $M^{(\tau_n)} = \{M^{(\tau_1)}(t) =$

$M(t\wedge\tau_n) : t \in \mathbb{T}\}$ é um $\mathbb{F}$-martingale para $n = 1, 2, \ldots$. Analogamente podemos definir submartingales locais. Claro todo martingale e um martingale local.

**Teorema 2.23:** Se $M$ é um $\mathbb{F}$-martingale local com trajetórias contínuas a direita e $\tau$ um tempo de parada com respeito a $\mathbb{F}$, então $M^{(\tau)}$ é um martingale local com respeito a $\mathbb{F}$.

<u>Dem:</u> Seja $\tau_1 < \tau_2 < \ldots$ uma sequência de tempos de parada com respeito a $\mathbb{F}$ com $\tau_n \to \infty$ q.c. tal que $M^{(\tau_n)}$ é um martingale. Então $M^{(\tau\wedge\tau_n)}$ é um $\mathbb{F}$-martingale e portanto um martingale local □

**Teorema 2.24:** Seja $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ uma filtração contínua a direita e $X = \{X(t) : t \in \mathbb{T}\}$ um $\mathbb{F}$-submartingale local contínuo a direita. Então existe um único (exceto por equivalência) processo crescente $A$, contínuo a direita e adaptado a $\mathbb{F}$ que satisfaz:

i) $X - A$ é um $\mathbb{F}$-martingale local e

ii) $E\int_0^{t\wedge\tau} Y(s-)dA(s) = E\int_0^{t\wedge\tau}(s)dA(s)$, para todo $Y$, $\mathbb{F}$-martingale contínuo a direita e todo $\tau$ tempo de parada com respeito a $\mathbb{F}$.

<u>Dem:</u> $\tau_1 < \tau_2 \ldots$ a sequência de tempo parada com $\tau_n \to \infty$ tal que $X^{(\tau_n)}$ é um submartingale. Seja $\sigma_n = \inf\{t : X(t) \leq -n\}$. Então $X(\tau_n \wedge \sigma_n)$ é uma submartingale de classe $DL$ pois para qualquer tempo de parada $\tau$ com respeito a $\mathbb{F}$

$$\begin{aligned} X^{(\tau_n\wedge\sigma_n)}(t) \wedge (-n) &\leq X^{(\tau_n\wedge\sigma_n)}(t\wedge\tau) \\ &\leq E\Big(X^{(\tau_n\wedge\sigma_n)}(t)/\mathcal{F}_t\Big)\ . \end{aligned}$$

Seja $A_n$ o processo crescente na decomposição de Doob-Meyer de $X^{(\tau_n\wedge\sigma_n)}$. Então $A = \lim_{n\to\infty} A_n$. □

# 3. Integração Estocástica.

## 1. Introdução.

Além dos conceitos de mensurabilidade e mensurabilidade progressiva estudados no Capítulo 1, o desenvolvimento de uma teoria de integração estocástica da maneira como faremos aqui, necessita também da noção de preditabilidade. Para tanto, consideremos $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade completo e $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}, \mathbb{T} = [0, \infty)$ ou $[0, t]$, uma filtração contínua a direita.

**Definição 3.1:** Seja $\mathcal{D}$ a classe de todas as funções $f : \mathbb{T} \times \Omega \to \mathbb{R}$ que satisfazem:

i) $f$ é $\mathcal{B}(\mathbb{T}) \times \mathcal{F}$ - mensurável;

ii) $f(t, \cdot)$ é $\mathcal{F}_t$-mensurável para todo $t \in \mathbb{T}$ e

iii) $f(t, \cdot)$ é contínua a esquerda para todo $\omega \in \Omega$.

Seja $\mathcal{P} = \mathcal{P}(\mathbb{F})$ a menor sub-$\sigma$-álgebra de $\mathcal{B}(\mathbb{T}) \times \mathcal{F}$ para qual todas as funções em $\mathcal{D}$ são mensuráveis. Um processo estocástico $X = \{X(t) : t \in \mathbb{T}\}$ é chamado predizivel com respeito a $\mathbb{F}$ (ou $\mathbb{F}$-predizivel) se a aplicação $(t, \omega) \to X(t, \omega)$ é $\mathcal{P}$-mensurável.

Por exemplo, processos adaptados a $\mathbb{F}$ que são contínuos a esquerda são predizíveis. Outros exemplos de processos predizíveis são:

i) $X(t, \omega) = X_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{i=0}^{n} X_i(\omega) \mathbb{1}_{(ti, ti+1]}(t)$,

onde $0 = t_0 < t_1 < \ldots < t_n$ e $X_0, X_1, \ldots, X_n$ são variáveis aleatórias com $X_i$ $\mathcal{F}_{t_i}$ mensurável. Um tal processo é chamado simples.

ii) Seja $X(t, \omega) = \sum_{i=1}^{n} X_i(\omega) \mathbb{1}_{[t_1, t_{i+1})}$. um processo adaptado a $\mathbb{F}$ e contínuo a direita, onde $0 < t_1 < t_2 < \ldots < t_{n+1}$. Então,

$$Y(t,\omega) = X(t-,\omega) = \lim_{s\downarrow t} X(s,\omega)$$

é predizível.

**Lema 3.2** Seja

$\mathcal{E} = \Big\{(a,b] \times B : 0 \le a < b,\ b \in \mathbb{T}, B \in \mathcal{F}_a\Big\} \bigcup \Big\{\{0\} \times B : B \in \mathcal{F}_0\Big\}$

Então $P = \sigma(\mathcal{E})$.

<u>Dem:</u> $\mathbb{1}_{(a,b]\times B}, B \in \mathcal{F}_a$ e $\mathbb{1}_{\{0\}\times B}$, $B \in \mathcal{F}_0$ são claramente $\mathcal{P}$-mensuráveis. Portanto $\mathcal{E} \subset \mathcal{P}$ e $\sigma(\mathcal{E}) \subset \mathcal{P}$. Seja $f \in \mathcal{D}$. Então $f = \lim_{n\to\infty} f_n$, onde

$$f_n(t,\omega) = f(0,\omega)\mathbb{1}_{\{0\}\times\Omega}(t,\omega) + \sum_{i=0}^{N_n-1} f(t_{n,i},\omega)\mathbb{1}_{(t_{n,i},t_{n,i+1}]\times\Omega}(t,\omega),$$

$0 = t_{n,0} < t_{n,1} < \ldots < t_{n,N_n} = n$ e $\max|t_{n,i+1} - t_{n,i}| < 1/n$. Como $f \in \mathcal{D}, f(t_{n,i};)$ é $\mathcal{F}_{t_{n,i}}$-mensurável e como tal é um limite pontual de funções simples $\sum_j x_j \mathbb{1}_{B_j}, B_j \in \mathcal{F}_{t_{n,i}}$. Temos então que $f_n$ é $\sigma(\mathcal{E})$)-mensurável para $n = 1, 2\ldots$ e portanto $f$ e $\sigma(\mathcal{E})$-mensurável. Isto mostra que $\mathcal{P} \subset \sigma(\mathcal{E})$.

**Lema 3.3:** Seja $\mathbb{T} = [0,T]$ e $\mathcal{E}_T = \{(a,b] \times B : 0 \le a < b \le T, B \in \mathcal{F}_a\} \bigcup \{\{0\} \times B : B \in \mathcal{F}_0\}\}$. Se $\mathcal{A}_T$ é a classe de todas as uniões finitas de elementos disjuntos de $\mathcal{E}_T$, então $\mathcal{A}_T$ é uma álgebra de subconjuntos de $[0,T] \times \Omega$.

<u>Dem:</u> exercício

Processos predizíveis são aqueles para os quais integrais estocásticas serão definidas. Em geral a filtração $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ não pode ser assumida contínua a direita. O conceito de processo predizível nos permite contornar este problema. Para mostrarmos este fato necessitamos do seguinte lema.

**Lema 3.4:** (Doob [D]), pág. 601). Seja $\mathcal{E}$ uma classe de subconjuntos de conjunto arbitrário $\Theta$ tal que a classe de uniões finitas e disjuntas de elementos de $\mathcal{E}$ é uma álgebra. Seja $\mathcal{H}$ uma classe de funções reais definidas em $\Theta$ que satisfazem:

i) $\mathbb{1}_A \in \mathcal{H}$ para todo $A \in \mathcal{E}$;

ii) $\mathcal{H}$ é um espaço vetorial e

iii) $\mathcal{H}$ é fechado com respeito a limites pontuais.

Então $\mathcal{H}$ contém todas as funções $\sigma(\mathcal{E})$ mensuráveis.

**Proposição 3.5:** Seja $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ uma filtração e $X = \{X(t) : t \in \mathbb{T}\}$ um processo predizível com respeito a $\mathbb{F}^+ = \{\mathcal{F}_{t+}; t \in \mathbb{T}\}$. Então $X(t)$ é $\mathcal{F}_t$-mensurável.

Dem: É suficiente provar o resultado para $\mathbb{T} = [0,T], T < \infty$ e arbitrário. Seja $\mathcal{H}$ a classe dos processos predizíveis com respeito a $\mathbb{F}^+$. Então $\mathcal{H}$ é um espaço vetorial fechado com relação a limites pontuais. Seja $\mathcal{E}_T$ como no Lema 3.3, definido com respeito a $\mathbb{F}^+$. Se $A \in \mathcal{E}_T$ é tal que $A = \{0\} \times B, B \in \mathcal{F}_{0+}$, então $\mathbb{1}_A(t,\cdot)$ é $\mathcal{F}_t$ mensurável pois $\mathbb{1}_A \equiv 0$. Se $A \in \mathcal{E}_T$ é tal que $A = (0,b] \times B, B \in \mathbb{F}_{0+}$ então $\mathbb{1}_A(t,\cdot)$ é $\mathcal{F}$-mensurável pois $B \in \mathcal{F}_{0+} \subset \mathcal{F}_t, t > 0$. Finalmente, se $A = (a,b] \times B, 0 < a < b \leq T, B \in \mathcal{F}_{a+}$, então $\mathbb{1}_A(t,\omega) = 0$ para $0 < t \leq a$ ou $t > b$ e $B \in \mathcal{F}_{a+} \subset \mathcal{F}_t$ $a < t \leq b$. Portanto $\mathbb{1}_A(t,\cdot)$ é $\mathcal{F}_t$ mensurável.

Segue pelo Lema 3.4 que $\mathcal{H}$ contém todos os processos mensuráveis com respeito a $\sigma(\mathcal{E}_T) = \mathcal{P}(\mathbb{F}^+)$.

# 2. Integrais estocásticas com respeito a martingales quadrados integraveis.

Consideremos $M = \{M(t) : t \in [0,T]\}$ um martingale contínuo a direita, quadrado integravel com $M(0) = 0$. Suponhamos também que a filtração associada é contínua a direita. Nosso objetivo é definir uma integral estocástica,

$$\int_0^T X(s)dM(s)$$

para uma certa classe de processos $X$.

Para tanto, consideremos primeiramente processos $X$ da forma

$$X(t,\omega) = \sum_{i=0}^{n-1} X_i(\omega)\mathbb{1}_{(t_u t_{u+1}]}(t),$$

onde $0 \leq t_0 < t_1 < \ldots < t_n \leq T, X_i$ $\mathcal{F}_{ti}$-mensurável e $E\int_0^T X^2(t)d\langle M\rangle(t) < \infty$, onde $\langle M\rangle$ é o processo de variação quadrática associado a $M$. Neste caso defina

$$\int_0^T X(t)dM(t) = \sum_{i=0}^{n-1} X_i\Big(M(t_{i+1}) - M(t_i)\Big) \tag{3.1}$$

e para $0 \leq t < T$

$$\int_0^t X(s)dM(s) = \int_0^T X(s)\mathbb{1}_{[0,t]}(s)dM(s).$$

Segue das definições da integral em (3.1) e do processo $\langle M\rangle$ que

$$E\Big(\int_0^T X(s)dM(0)\Big) = 0$$

e

$$E\left[\Big(\int_0^T X(s)dM(s)\Big)^2\right] = E\left[\int_0^T X^2(s)dM(s)\right].$$

Consideremos agora uma medida $m$ definida em $([0,T]\times\Omega P)$ pela seguinte relação:

$$\begin{aligned}\int_{[0,T]\times\Omega} X(t,\omega)dm(t,\omega) &= \int_\Omega\int_{[0,T]} X(t,\omega)d\langle M\rangle(t,\omega)\ dP(\omega)\\ &= E\Big(\int_0^T X(t)d\langle M\rangle(t)\Big)\end{aligned}$$

para todo processo predizível não negativo $X$.

Podemos agora definir a integral estocástica para todo processo $X \in L_2([0,T] \times \Omega, \mathcal{P}, m) = L_2(m)$. Para tanto, observemos que para $\mathcal{A}_T$ como no Lema 3.3 e $A \in \mathcal{P}$, existe $A_n, n = 1,2\ldots$ em $\mathcal{A}_T$ tal que

$$\int_0^T |\mathbb{1}_A - \mathbb{1}_{A_n}|^2 dm \leq \frac{1}{n}$$

pois $\mathcal{P} = \sigma(\mathcal{S}_T)$. Este fato demonstra o resultado seguinte.

**Proposição 3.6:** A classe $S$ de todos os processos predizíveis e contínuos a esquerda da forma

$$X(t,\omega) = X(0,\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=0}^{n-1} X(t_i,\omega)\mathbb{1}_{(t_i,t_{i+1}]}(t),$$

$0 = t_0 < t_1 < \ldots < t_n \leq T$, é densa em $L_2([0,T] \times \Omega, \mathcal{P}, m)$.

Seja $I : S \to L_2(\Omega, \mathcal{F}, P) = L_2(P)$, a aplicação definida por $I(X) = \int_0^T X(s)dM(s)$. Como para $X \in S \subset L_2([0,T] \times \Omega, P, m)$

$$\begin{aligned} ||T(X)||^2_{L_2(P)} &= E\Big[\Big(\int_0^T X(s)dM(s)\Big)^2\Big] = \\ &= E\Big[\int_0^T X^2(s)d\langle M\rangle(s)\Big] \\ &= \int_{[0,T]\times\Omega} X^2(s,\omega)dm(s,\omega) = ||X||_{L_2(m)}, \end{aligned}$$

$I$ é uma isometria de $S$ em $L_2(P)$ e como tal pode ser estendida univocamente a uma isometria de $L_2(m)$ em $L_2(P)$, pois $S$ é denso.

A aplicação $I(X)$ é chamada <u>integral estocástica de $X$ com respeito a $M$</u> e denotada por

$$
\begin{aligned}
I(X) &= \int_0^T X(s)dM(s) \\
&= \int_0^T XdM.
\end{aligned}
$$

Como consequência imediata da definição temos:

i) Se $X_1$ e $X_2 \in L_2(m)$ e $a \in \mathbb{R}$, então

$$\int_0^T (aX_1 + X_2)dM = a\int_0^T X_1 dM + \int_0^T X_2 dM.$$

ii) Se $X$, $X_n \in L_2(m)$ e $X_n \to X$, isto é $E(\int_0^T |X_n - X|^2 d\langle M\rangle) \to 0$ quando $n \to \infty$, então

$$\int_0^T X_n dM \to \int_0^T XdM$$

em $L_2(P)$.

Para $s, t \in [0,T], s < t$, definimos

$$\int_s^t X(u)dM(u) = \int_0^T X(u)\mathbb{1}_{(s,t]}(u)\, dM(u).$$

Outros processos estocásticos interessantes podem ser definidos através de integrais estocásticas pela relação

$$Y(t) = \int_0^t X(s)dM(s), \quad t \in [0,T].$$

**Proposição 3.7:** $Y$ é um martingale quadrado integravel que possui uma versão contínua a direita.

Dem: Claro $Y(t)$ é $\mathcal{F}_t$-mensurável e $E(Y^2(t)) < \infty$, para todo $t$. Para $X \in S$, isto é $X(t) = \sum_{i=1}^{n-1} X_i \mathbb{1}_{[t_i,t_{i+1}]}(t)$, podemos verificar diretamente que $Y$ é um martingale quadrado integravel. Para $X \in L_2(m)$, seja $X_n \in S$ com $X_n \to X$ e $s < t$

$$E\Big(|E\Big(\int_s^t XdM/\mathcal{F}_s\Big) - E\Big(\int_s^t X_n dM/\mathcal{F}_s\Big)|\Big)$$

$$\leq \Big[E\Big(\Big(\int_s^t (X - X_n)dM\Big)^2\Big)\Big]^{1/2} \leq \Big[E\int_0^T |X - X_n|^2 d\langle M\rangle\Big]^{1/2} \to 0.$$

Como $\int_0^t X_n dM$ é martingale, temos $E\Big(\int_s^t X_n dM/\mathcal{F}_s\Big) = 0$ para todo $n$ e portanto $E\Big(\int_s^t XdM/\mathcal{F}_s\Big) = 0$. Segue que $Y(t) = \int_0^t XdM$ é um martingale. Que $Y$ tem uma versão contínua a direita é uma consequência do Fato 2.2 pois $E(Y(t)) = 0$. □

**Proposiçã o 3.8:** Se $M$ é um martingale quadrado integral com trajetórias contínuas e $X \in L_2(m)$, então $Y(t) = \int_0^t XdM, t \in [0, T]$ tem uma versão contínua.

Dem: Seja $X_n \in S$, tal que para todo $n$

$$E\Big(\int_0^T |X - X_n|^2 d\langle M\rangle\Big) < \frac{1}{n^3}\ .$$

Segue que

$$E\Big(\sup_{0\leq t\leq T} |\int_0^T (X - X_n)dM|^2\Big) \leq 4E\int_0^T |X - X_n|^2 d\langle M\rangle$$

e

$$P\Big(\sup_{0\leq t\leq T}\Big|\int_0^t \big(X - X_n\big)dM\Big| > \frac{1}{n}\Big) \leq 4E\Big(\int_0^T \big|X - X_n\big|^2 dM\Big) < \frac{4}{n^2}\ .$$

Portanto,

$$\sum_n P\Big(\sup_{0\leq t\leq T}\Big|\int_0^t (X - X_n)dM\Big| > \frac{1}{n}\Big) < \infty$$

e pelo Lema de Borel Cantelli

$$P\Big(\overline{\lim_{n\to\infty}}\Big[\sup_{0\leq t\leq T} |\int_0^t (X - X_n)dM| > 1/n\Big]\Big) = 0.$$

Temos então que

$$\sup_{0\leq t\leq T}\left|\int_0^t (X - X_n)dM\right| \longrightarrow 0 \text{ q.c.}$$

quando $n \to \infty$ e como para todo $n$, $\int_0^t X_n dM$ é contínuo, o processo $\int_0^t XdM$ é de fato contínuo com probabilidade 1. □

# 3. Integral estocástica com respeito ao processo de Wiener. A integral de Itô.

Em $(\Omega, \mathcal{F}, P)$, seja $W = \{W(t) :\in [0,T]\}$ um processo de Wiener. Seja $\{\mathcal{F}_t^W : t \in [0,T]\}$ a filtração definida por $\mathcal{F}_t^W = \sigma(W(s) : 0 \leq s \leq t)$. Segue da definição de $W$ que se $s \leq t$,

$$E\left(W(t)/\mathcal{F}_s^W\right) = W(s) \quad \text{e}$$

$$E\left((W(t) - W(s))^2/\mathcal{F}_s^W\right) = t - s$$

Temos então que um processo de Wiener $W$ em $[0,T]$ é um martingale quadrado integral ($E(W(t)^2) = t < \infty$) com respeito a filtração $\{\mathcal{F}_t^W : t \in [0,T]\}$ com $\langle W\rangle(t) = t$ e trajetórias contínuas. Outras filtrações $\{\mathcal{F}_t : t \in [0,T]\}$ com respeito as quais $W$ é um martingale contínuo quadrada integral são possíveis de serem obtidas (ver por exemplo Lipster e Shiryayev [L - S]).

A definição de $\int_0^T XdW$ segue da seção anterior para integrandos $X$ predizíveis, satisfazendo

$$E\left(\int_0^T X^2(t)dt\right) < \infty .$$

No entanto, a integral estocástica na forma definida por Itô tem certos aspectos particulares que achamos conveniente desenvolver.

Consideremos um processo de Wiener $W = \{W(t) : t \in \mathbb{T}\}, \mathbb{T} = [0, T]$ ou $[0, \infty]$, que é um martigale com respeito a uma filtração $\{\mathcal{F}_t : t \in \mathbb{T}\}$. Seja $\mathcal{W}$ a classe de processos $X$ satisfazendo:

i) $X$ é mensurável;

ii) $X$ é adaptado a $\{\mathcal{F}_t : t \in \mathbb{T}\}$;

iii) $E \int_{\mathbb{T}} X^2(t)dt < \infty$.

Seja $\mathcal{S}_0 \subset \mathcal{W}$ a classe de processos simples, isto é processos de forma

$$X(t, \omega) = \sum_{i=0}^{n-1} X(t, \omega) \mathbb{1}_{[t_j, t_{j+1})}(t);$$

onde $0 = t_0 < \ldots < t_n$ é uma partição $\mathbb{T}$ e $X(t_j)$ é $\mathcal{F}_{t_j}$-mensurável. Para $X \in \mathcal{S}_0$ podemos definir $\int X dW$ como sendo a aplicação $I_W : \mathcal{S}_0 \to L_2(P)$, definida como

$$I_W(X) = \sum_{i=0}^{n-1} X(t_j) \Big[W(t_{j+1}) - W(t_j)\Big] .$$

Para outro processo $Y \in \mathcal{S}_0$, é fácil ver (exercício) que:

$$E\Big((I_W(X) - I_W(Y))^2\Big) = E\Big(\int_{\mathbb{T}} (X(t) - Y(t))^2 dt\Big).$$

**Proposição 3.9:** Seja $X \in \mathcal{W}$. Então existe um sequência $X_n, n = 1, 2 \ldots$ em $\mathcal{S}_0$ tal que:

$$E \int_{\mathbb{T}} |X(t) - X_n(t)|^2 dt \to 0 .$$

Dem: É suficiente provar o resultado para $X$ limitado em $[0, T]$ com $X(t, \omega) = 0, t \geq T$. Por conveniência, definimos $X(t, \omega) = 0, t < 0$. Como $\int_{-\infty}^{\infty} X^2(t, \omega)dt < \infty$ q.c., temos que

$$\int_{-\infty}^{\infty} |X(t + h) - X(t)|^2 dt \to 0 \text{ q.c. quando } h \to 0$$

e

$$\int_{-\infty}^{\infty} |X(t+h) - X(t))|^2 dt \leq 4 \int_{-\infty}^{\infty} X^2(t) dt \text{ q.c..}$$

Como $E\int_{-\infty}^{\infty} X^2(t)dt < \infty$

$$\int_{\Omega}\int_{-\infty}^{\infty} |X(t+h) - X(t)|^2 dt dP \to 0 \text{ quando } h \to 0.$$

Para $n = 1, 2, \ldots$, seja

$$a(n,t) = \sum_{j=-\infty}^{\infty} \left(j/2^n\right) \mathbb{1}_{[\frac{j}{2^n}, \frac{j+1}{2^n})}(t) .$$

Então,

$$\int_{\Omega}\int_{-\infty}^{\infty} |X(s + a(n,t)) - X(s+t)|^2 ds dP \to 0$$

quando $n \to \infty$ e como

$$\int_{\Omega}\int_{-\infty}^{\infty} |X(s + a(n,t)) - X(s+t)|^2 ds dP \leq 4 \int_{\Omega}\int_{-\infty}^{\infty} X^2(s) ds dP,$$

para todo $t$ temos

$$\int_{-1}^{T}\int_{\Omega}\int_{-\infty}^{\infty} |X(s + a(n,t)) - X(s+t)|^2 ds dP dt \to 0$$

quando $n \to \infty$ ou

$$\int_{-\infty}^{\infty}\int_{-1}^{T}\int_{\Omega} |X(s + a(n,t)) - X(s+t)|^2 dP dt ds \to 0.$$

Podemos portanto tomar uma subsequência tal que:

$$\int_{-1}^{T}\int_{\Omega} |X(s + a(n_i,t)) - X(s+t)|^2 dP dt \to 0$$

quando $n_i \to \infty$ para quase todo $s$. Escolhendo e fixando um $s_0 \in [0,1]$ com

$$\int_{-1}^{T}\int_{\Omega}|X(s_0+a(n_i,t)-X(s_0+t))|^2 dPdt \to 0$$

quando $n_i \to \infty$ e fazendo uma troca de variáveis obtemos

$$\int_{s_0-1}^{T+s_0}\int_{\Omega}|X(s_0+a(n_i,t-s_0))-X(t))|^2 dPdt \to 0.$$

Como $s_0 \in [0,1]$,

$$\int_{0}^{T}\int_{\Omega}|X(s_0,a(n_i,t-s_0))-X(t)|^2 dPdt \to 0.$$

Definindo $X_i(t) = X\Big(s_0, a(n_i, t-s_0)\Big)$, a proposição estava demonstrada. □

Segue que a aplicação $I_W$ de $S_0$ em $L_2(\Omega)$ pode ser estendida de maneira única a $\mathcal{W}$ definindo

$$\int XdW = I_W(X)$$

para todo $X \in \mathcal{W}$.

A construção da integral de Itô apresentada acima é mais geral que a integral estocástica para martingales quadrado integravel. Se $X \in \mathcal{W}$, então existe um processo $Y$ mensurável tal que $X = Y$ q.t.p. $dtdP$, pois a sequência construida na Proposição 3.9 pode ser tomada continua a esquerda e uma subsequencia convergindo q.t.p. $dtdP$ pode ser escolhida. Por outro lado, como foi mostrado por Courrége [Cor], integrais estocásticas $\int XdM$, onde $M$ não é um processo de Wiener não podem ser estendida para processos em $\mathcal{W}$.

Consideremos agora processos estocásticos definidos por

$$Y(t) = \int_0^t X(s)dW(s)$$

para $X \in \mathcal{W}$. O seguintes resultados serão necessários.

**Lema 3.10:** Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias integraveis tais que $E(X_n - X_k/X_1, \ldots X_n) = 0$ para $k = 1, 2, \ldots, n$. Então

i) $E\Big((\sup(X_1^+, \ldots, X_k^+))^\alpha\Big) \le \big(\frac{\alpha}{\alpha - 1}\big)^\alpha E\big[\big(X_n^+\big)^\alpha\big]$;

ii) $P\Big(\sup(X_1^+ \ldots, X_n^+) > a\Big) \le \frac{1}{a^2} E\big(X_n^+\big), a > 0$;

iii) $P\Big(\sup(|X_1 \ldots, |X_n||) > a\Big) \le \frac{EX_n^2}{a^2}, a > 0$;

iv) $E\Big((\sup(|X_1|, \ldots, |X_n|))^\alpha\Big) \le \big(\frac{\alpha}{\alpha - 1}\big)^\alpha E|X_n|^\alpha$.

**Proposição 3.11:** Seja $Y(t) = \int_0^t X(s)dW(s)$, $0 \le t \le T$, $X \in \mathcal{W}$. Então $Y = \{Y(t) : 0 \le t \le T\}$ é um martingale contínua e quadrado integravel.

Dem: Primeiramente suponhamos que $X$ é simples. Então

$$\int_0^t X(s)dW(s) = \sum_{i=1}^{n} X(t_i)\Big[W(t_{i+1} \wedge t) - W(t_i \wedge t)\Big]$$

e portanto $Y$ é contínuo com probabilidade 1. Sejam $\{s_{n,j} : j = 0, 1, \ldots, n\}, n = 1, 2 \ldots,$ partições de $[0, t]$ com $X(t)$ constante no intervalo $[s_{n,j}, s_{n,j+1})$. Como $\int_0^t X(s)dW(s)$ é contínuo,

$$\sup_{0 \le t \le T}\Big|\int_o^t X(s)dW(s)\Big|$$

é mensurável. Seja

$$Y_n = \sup_i \Big|\int_0^{s_{n,i}} X(s)dW(s)\Big| .$$

As variáveis aleatórias $\int_0^{s_{n,i}} X(s)dW(s)$ satisfazem as condições do Lema 3.10 e portanto

$$P\Big(Y_n > a\Big) \le \frac{1}{a^2}\int_0^T E\Big(X^2(s)\Big)ds$$

e

$$E\Big(Y_n^2\Big) \le 4\int_0^T E\Big(X^2(s)\Big)ds\ .$$

Se $\{s_{n,j}\} \subset \{s_{n+1,j}\}$ e $\bigcup_n\{s_{n,j}\}$ é denso em $[0,T]$, segue da continuidade de $\int_0^t X(s)dW(s)$, fazendo $n \to \infty$ nas desigualdades acima, que

$$P\Big(\sup_{0\le t\le T}\Big|\int_0^t X(s)dW(s)\Big| > a\Big) \le \frac{1}{a^2}\int_0^T E\Big(X^2(s)\Big)ds$$

e

$$E\Big(\sup_{0\le t\le T}\Big|\int_0^t X(s)dW(s)\Big|^2\Big) \le 4\int_0^T E\Big(X^2(s)\Big)ds.$$

Seja agora $X \in \mathcal{W}$. Então existe uma sequência de processos simples $X_n \in \mathcal{W}$ tal que $\int_0^T E(X(s) - X_n(s))^2 ds \to 0$ quando $n \to \infty$. Escolhendo uma subsequência $(n_k)$ tal que $\int_0^T E(X(s) - X_{n_k}(s))^2 ds \le 1/2^n$, como

$$\int_0^T E(X_{n_{k+1}}(s) - X_{n_k}(s))^2 ds \le 3/2^k,$$

obtemos a seguinte desigualdade:

$$P\Big(\sup_{0\le t\le T}\Big|\int_0^t X_{n_{k+1}}(s)dW(s) - \int_0^t X_{n_k}dW(s)\Big| > 1/k^2\Big) \le \frac{3k^4}{2^k}.$$

Pelo Lema de Borel-Cantelli, a série

$$\int_0^t X_{n_1}(s)dW(s) + \sum_{k=1}^{\infty}\left[\int_0^t X_{n_{k+1}}(s)dW(s) - \int_0^t X_{n_k}(s)dW(s)\right]$$

converge uniformemente em $t \in [0,T]$ q.c. para $Y(t) = \int_0^t X(s)dW(s)$. Este resultado juntamente com a Proposição 3.7 completa a demonstração. $\square$

Notemos que $\langle Y\rangle(t) = \int_0^t X^2(s)ds$ q.c.. De fato, tomando uma versão progressivamente mensurável temos de $X$ que $\int_0^t X^2(s)ds$ é $\mathcal{F}_t$ adaptado,

$$E(Y^2(t) - \int_0^t X^2(u)du/\mathcal{F}_s) = E\Big(\big(\int_s^t X(u)dW(u)\big)^2/\mathcal{F}_s\Big) - E\Big(\int_0^t X^2(u)du/\mathcal{F}_s\Big)$$

$$= E\Big(\big(\int_0^s X(u)dW(u)\big)^2/\mathcal{F}_s\Big) - E\Big(\int_0^s X^2(u)du/\mathcal{F}_s\Big)$$

$$+E\Big(\big(\int_s^t X(u)dW(u)\big)^2/\mathcal{F}_s\Big) - E\Big(\int_s^t X^2(u)du/\mathcal{F}_s\Big)$$

$$= E\Big(Y^2(s) - \int_0^s X^2(u)du/\mathcal{F}\Big)s \leq s.$$

Como $\int_0^t X^2(u)du$ é contínuo e crescente em $t$ q.c., ele é predizível. □

# 4. Integrais estocástica com respeito a martingales locais contínuos.

A extensão da teoria de integração para classe dos martigales locais contínuas requer um estudo adicional sobre integrais estocásticas com respeito a martigales quadrado integravel que apresentaremos a seguir na forma de lemas. Como as demonstrações destes resultados são extremamente tecnicas e longas nos referimos aos trabalhos de Kunita e Watanabe [K - W] e Meyer [M2] para a demonstração do primeiro lema e apresentaremos apenas uma demonstração resumida dos lemas subsequentes.

Nos lemas abaixo $M$ e $N$ são martingales quadrado integrável com respeito a uma mesma filtração e é assumido que $M(0) = N(0) = 0$.

**Lema 3.12:** i) Se $X$ é limitado e predizível, então

$$\Big|\int_0^T X(t)d|\langle M,N\rangle|(t)\Big| \leq \Big[\int_0^T X^2(t)d\langle M\rangle\langle N\rangle(t)\Big]^{1/2} \quad \text{q.c..}$$

ii) Se $X$ e $Y$ são mensuráveis, então

$$\int_0^T |X(t)Y(t)||d|\langle M,N\rangle|(t)| \leq \left[\int_0^T X^2(t)d\langle Mbig > (t)\int_0^T Y^2(t)d\langle N\rangle(t)\right]^{1/2} \text{ q.c.}$$

e

$$E\left[\int_0^T |X(t)Y(t)|d|\langle M,N\rangle|(t)\right]$$

$$\leq \left[E\int_0^T X^2(t)d\langle M\rangle(t)E\int_0^T X^2(t)d\langle N\rangle(t)\right]^{1/2}.$$

**Lema 3.13:** Se $X \in L_2(m)$, então

$$E\left[\int_0^t X(s)dM(s)\cdot N(s)\right] = E\left[\int_0^t X(s)d\langle M,N\rangle(s)\right]$$

para todo $t \in [0,T]$.

Dem: Se $X$ é simples o resultado pode ser verificado diretamente. Seja $X$ limitado e $(X_n)$ uma sequência de processos simples em $L_2(m)$ com

$$E\int_0^T |X(s)-X_n(s)|^2 d\langle M\rangle(s) \to 0.$$

Segue que

$$E\left|\left[\int_0^t X(s)dM(s) - \int_0^t X_n(s)dM(s)\right]N(t)\right| \leq$$

$$\leq \left[E\int_0^t \big(X(s)-X_n(s)\big)^2 d\langle M\rangle(s)E\langle N\rangle(t)\right]^{1/2} \to 0$$

E pelo Lema 3.12 i.

$$E\left|\int_0^t X_n(s)d\langle M,N\rangle(s) - \int_0^t X(s)d\langle M,N\rangle(s)\right|$$

$$\leq \left[E\int_0^t \Big(X(s)-X_n(s)\Big)^2 d\langle M\rangle(s)E\langle N\rangle(t)\right]^{1/2} \to 0.$$

Para $X \in L_2(m)$ arbitrário, a demonstração segue analogamente usando Lema 3.12 ii. □

**Lema 3.14:** Sejam $M$ e $N$ contínuos. Se $X \in L_2(m)$ e $Y(t) = \int_0^t X(s)dM(s)$, então

$$\langle Y, N\rangle(t) = \int_0^t X(s)d\langle M, N\rangle(s).$$

Dem: Seja $\tau$ um tempo de parada. $N^\tau(t) = N(\tau \wedge t)$ também é contínuo e quadrado integravel. Segue do Lema 3.13 que

$$E\left[Y(T)N^{(\tau)}(T) - \int_0^T X(s)d\langle M, N^\tau\rangle(s)\right] = 0.$$

Temos também

$$E\Big(Y(T)N^{(\tau)}(T)\Big) = E\Big(Y(T)N(\tau)\Big)$$

$$= E\Big(N(\tau)E(Y(T)/\mathcal{F}_\tau)\Big) = E\Big(N(\tau Y(\tau)\Big).$$

Como

$$\langle M, N^{(\tau)}\rangle(s) = \langle M, N\rangle(s \wedge \tau),$$

e

$$\int_0^T X(s)d\langle M, N^{(\tau)}\rangle(s) = \int_0^\tau X(s)d\langle M, N\rangle(s)\Big) = 0$$

e

$$E\Big((X(\tau)N(\tau) - \int_0^\tau X(s)d\langle M, N\rangle(s)\Big) = 0.$$

Portanto, se

$$Z(t) = Y(t)N(t) - \int_0^t X(s)d\langle M,N\rangle(s),$$

então $Z$ é adaptado e $E(Z(\tau)) = 0$ para todo tempo de parada $\tau$. Segue que $Z$ é um martingale uniformemente integravel e

$$\int_0^t X(s)d\langle M,N\rangle(s) = \langle Y,N\rangle(t). \qquad \square$$

**Lema 3.15:** Se $M$ é contínuo, $X \in L_2(m)$ e $Y(t) = \int_0^t X(s)dM(s)$, então

$$\langle Y\rangle(t) = \int_0^t X^2(s)d\langle M\rangle(s).$$

<u>Dem:</u> No Lema 3.14 tomemos $N = Y$. $\square$

**Lema 3.16:** Seja $X \in L_2(m)$, $\tau$ um tempo de parada e $Y(t) = \int_0^t X(s)dM(s)$. Então

$$Y^{(\tau)}(t) = \int_0^t X(s)\mathbb{1}_{[0,\tau]}dM(s).$$

<u>Dem:</u> $X\mathbb{1}_{[0,\tau]} \in L_2(m)$. Seja $Z(t) = \int_0^t X(s)\mathbb{1}_{[0,\tau]}(s)dM(s)$. Então

$$E\Big(Y^{(\tau)}(t) - Z(t)\Big)^2 = E\Big(\langle Y^{(\tau)} - Z\rangle(t)\Big) =$$

$$= E\langle Y^{\tau}\rangle(t) + E\langle Z\rangle(t) - 2E\langle Y^{(\tau)}, Z\rangle(t).$$

Mas

$$E\langle Y^{\tau}\rangle(t) = E\langle Y\rangle(\tau \wedge t) = E\int_0^{t\wedge\tau} X^2(s)d\langle M\rangle(s),$$

$$E\langle Z\rangle(t) = E\int_0^t X^2(s)\mathbb{1}_{[0,\tau]}(s)dM(s)$$

e

$$
\begin{aligned}
E\langle Y^{\tau}, Z\rangle(t) &= E\int_0^t X(s)1\!\!1_{[0,\tau]}(s)d\langle Y^{(\tau)}, M\rangle(s) \\
&= E\int_0^t X(s)1\!\!1_{[0,\tau]}(s)d\langle Y, M\rangle(\tau\wedge s) \\
&= E\int_0^{\tau\wedge s} X^2(s)1\!\!1_{[0,\tau]}(s)d\langle M\rangle(s) \\
&= E\int_0^t X^2(s)1\!\!1_{[0,\tau]}(s)d\langle M\rangle(s).
\end{aligned}
$$

Portanto $E\big(Y^{(\tau)} - Z\big)^2 = 0$ e a integral $\int_0^t X(s)1\!\!1_{[0,\tau]}(s)dM(s)$ pode ser escrita como $\int_0^{t\wedge\tau} X(s)dM(s)$. □

**Lema 3.17:** Seja $M$ contínuo e $X \in L_2(m)$. Se $\tau$ é um tempo de parada, então

$$
E\Big(\int_0^t X^2(t)d\langle M^{(\tau)}\rangle(t)\Big) < \infty
$$

(isto é, $X \in L_2(m^{(\tau)})$, onde $m^\tau$ é definida pela relação $\int_{[0,T]\times\Omega} X dm^{(\tau)} = \int_\Omega \int_{[0,t]} X d\langle M^{(\tau)}\rangle dP\Big)$ e

$$
\int_0^t X(s)dM^{(\tau)}(s) = \int_0^{t\wedge\tau} X(s)dM(s).
$$

<u>Dem:</u> Como $\langle M^{(\tau)}\rangle(t) = \langle M\rangle(\tau\wedge t)$, segue que $L_2(m) \subset L_2(m^{(\tau)})$. Dos Lemas anteriores temos que se $Y(t) = \int_0^t X(s)dM(s)$,

$$
E\Big(\int_0^{t\wedge\tau} X(s)dM(s) - \int_0^t X(s)dM^{(\tau)}(s)\Big)^2 =
$$

$$
= E\int_0^t X^2(s)1\!\!1_{[0,\tau]}(s)d\langle M\rangle(s) + E\int_0^t X^2(s)d\langle M^{(\tau)}\rangle(s)
$$

$$-2E\int_0^t X(s)d\langle M^{(\tau)}, Y^{(\tau)}\rangle(s) = 0 \qquad \square$$

**Colorário 3.18:** Seja $(\tau_n)$ uma sequência de tempos de parada tal que $\tau_n \nearrow T$ q.c.. Então para cada $t \in [0, T]$,

$$\int_0^t X(s)dM^{(\tau_n)}(s) = \int_0^{t\wedge\tau_n} X(s)dM(s) \xrightarrow[n\to\infty]{q.c.} \int_0^t X(s)dM(s).$$

Para $M$ um martingale local contínuo com $M(0) = 0$, seja $L_{2,\text{loc}}(\langle M\rangle)$ a classe de todo processo predizível $Y$ que satisfaz

$$\int_0^T X^2 d\langle M\rangle(s) < \infty \quad \text{q.c..}$$

Para $X \in L_{2,\text{loc}}(\langle M\rangle)$, seja $\tau_n$ a sequência de tempos de parada definida por:

$$\tau_n = \begin{cases} \inf\left\{t : |M(t)| \geq n \quad \text{ou} \quad \int_0^t X^2(s)d\langle M\rangle(s) \geq n\right\}, \\ T \text{ se o conjunto acima é vazio.} \end{cases}$$

Note que $\tau_n \nearrow$ e pela continuidade de $M$ e $\langle M\rangle, \tau_n \nearrow T$. Denotemos $M^{(\tau_n)}$ por $M^{(n)}$. Então $M^{(n)}$ são contínuos, quadrado integraveis e $\langle M^{(n)}\rangle(t) = \langle M\rangle(\tau_n \wedge t)$. Para $X \in L_2(m^{(n)})$, seja $Y^{(n)}$ uma versão contínua de

$$Y^{(n)}(t) = \int_0^t X(s)dM^{(n)}(s).$$

É fácil ver que o evento

$$\left(\bigcup_n \{\tau_n \nearrow T\}\right) \bigcap \left(\bigcap_n \{Y^{(n)} \text{ contínuo}\}\right) \bigcap \left(\bigcap_{n\geq m} \bigcap_m \{Y^{(m)}(t\wedge\tau_n) = Y^{(n)}(t)\}\right)$$

tem probabilidade 1. Portanto existe $n$ tal que, se $m \geq n$

$$Y^{(m)}(t) = Y^{(n)}(t) \quad \text{q.c.}$$

em $[0, T]$. Isto mostra que $Y^{(n)}$ converge q.c. e permite que integrais estocásticas de processos $X \in L_{2,\text{loc}}(\langle M \rangle)$ com respeito a martingales locais contínuos sejam definidas.

**Definição 3.19:** Seja $M$ um martingale local contínuo com $M(0) = 0$ e $X \in L_{2,\text{loc}}(\langle M \rangle)$. Então definimos

$$\int_0^t X(s)dM(s) \equiv \lim_{n \to \infty} \int_0^t X(s)dM^{(n)}(s) \quad \text{q.c.}$$

para todo $t \in [0, T]$.

Se $M$ é quadrado integrável, temos que $L_2(m) \subset L_{2,\text{loc}}(\langle M \rangle)$. O fato que para toda sequência de tempos de parada $\tau_n \nearrow T$

$$\int_0^t X(s)dM^{(\tau)}(s) \to \int_0^t X(s)dM(s)$$

mostra que as duas possíveis definições de integral neste caso coincidem.

Para mostrarmos a lineariedade da integral definida em 3.9 consideremos o seguinte lema.

**Lema 3.20:** Seja $\gamma_n$ uma sequência de tempos de parada $\gamma_n \nearrow T$ e $\gamma_n \leq \tau_n$ q.c.. Então

$$\int_0^t X(s)dM^{(\gamma_n)}(s) \to \int_0^t X(s)dM(s)$$

Dem: Seja $m \geq n$. Então $\gamma_n = \gamma_n \wedge \tau_n$ e pelo Lema 3.17.

$$\int_0^t X(s)dM^{\gamma_n}(s) = \int^{t\wedge\gamma_n} X(s)dM^{(\tau_m)}(s) = Y^{(m)}(t\wedge\gamma_n).$$

Pela construção de $Y^{(m)}$ existe $N$ tal que se $m \geq N, Y^{(m)}(t) = Y(t)$ em $[0,T]$. Portanto $Y^{(m)}(t\wedge\gamma_n) \to Y(t\wedge\gamma_n)$ q.c., quando $n\to\infty$ e $\int_0^t X(s)dM^{(\gamma_n)}(s) = Y(t\wedge\gamma_n)$. Como $Y$ é contínuo e $\gamma_n \nearrow T$, temos que $Y(t\wedge\gamma_n) = \int_0^t X(s)dM^{(\gamma_n)}(s) \to Y(t) = \int_0^t X(s)dM(s)$ q.c., quando $n\to\infty$. □

**Proposição 3.21:** Se $X_1$, $X_2 \in L_{2,\text{loc}}(\langle M\rangle)$ e $a \in I\!R$, então $\int_0^t(aX_1 + X_2)(s)dM(s) = a\int_0^t X_1(s)dM(s) + \int_0^t X_2(s)dM(s)$.

Dem: Seja $a = 1$ e sejam $(\tau_{1n})$, $(\tau_{2,n})$ e $(\tau_{2,n})$ as sequências de tempos de parada envolvidas nas definições das integrais de $X_1, X_2$ e $X_1 + X_2$ respectivamente. Seja $\tau_n = \tau_{1,n}\wedge\tau_{2,n}\wedge\tau_{3,n}$. Claro $\tau_n \nearrow T$ e $\tau_n \leq \tau_{i,n}$ $i = 1,2,3$. Segue do Lema 3.20 que

$$\int_0^t X_1(s)dM^{(\tau_n)} \longrightarrow \int_0^t X_1(s)dM(s)$$

$$\int_0^t X_2(s)dM^{(\tau_n)} \longrightarrow \int_0^t X_1(s)dM(s)$$

e

$$\int_0^t X_1(s) + X_2(s)dM^{(\tau_n)}(s) \longrightarrow \int_0^t (X_1+X_2)(s)(s)dM(s)$$

O resultado segue da linearidade de integrais estocásticas com respeito a martingales quadrado integraveis. Similarmente podemos provar o resultado para multiplicação por escalar. □

**Proposição 3.21:** Para $\varepsilon > 0$, $\eta > 0$ e $X \in L_{2,\text{loc}}(\langle M\rangle)$ temos:

$$P\left(\sup_{0\le t\le T}\left|\int_0^t X(s)dM(s)\right|\geq\varepsilon\right)\leq\frac{4\eta}{\varepsilon^2}+P\left(\int_0^t X(s)d\langle M\rangle(s)\geq\eta\right).$$

<u>Dem:</u> Seja $\gamma_m = \inf\{t : |M(t)| \geq m$ ou $\int_0^t X^2(s)d\langle M\rangle(s) \geq \eta\}$ ou $\gamma_m = T$ se o conjunto é vazio. Estão $\gamma_n$ é um tempo de parada e $\gamma_m \leq \tau_m$, onde $\tau_m$ é o tempo de parada usado na definição de $Y(t) = \int_0^t X(s)dM(s)$. $M^{(\gamma_n)}$ é contínuo e limitada. Seja $Z^{(m)}$ uma versão continua de $\int_0^t X(s)dM^{(\gamma_n)}(s)$. Temos então que:

$$P\Big(\sup_{0\le t\le T}|Y(t)|>\varepsilon\Big)\leq P\Big(\sup_{0\le t\le T}|Z^{(m)}(t)|\geq\varepsilon\Big)+P\Big(\sup_{0\le t\le T}|Y(t)-Z^{(m)}(t)|\neq 0\Big).$$

Usando desigualdades de martingales obtemos:

$$\begin{aligned}P\Big(\sup_{0\le t\le T}|Y(t)|>\varepsilon\Big) &\leq \frac{1}{\varepsilon^2}E\Big(\sup_{0\le t\le T}\int_0^t X(s)dM^{(\gamma_n)}(s)\Big)\\ &\leq \frac{4}{\varepsilon^2}E\int_0^T X^2(s)d\langle M^{(\gamma_n)}\rangle(s)\\ &\leq \frac{4\eta}{\varepsilon^2}.\end{aligned}$$

Pelo Lema 3.17,

$$\int_0^t X(s)dM^{(\gamma_n)}(s)=\int_0^{t\wedge\gamma_n}X(s)dM^{(\tau_m)}(s).$$

Portanto

$$Z^{(m)}(t)=Y(t\wedge\tau_m)\quad \text{q.c.}$$

para $t\in[0,T]$. Seja $B_m=\{\gamma_m=T\}$ e $A_m=\{\tau_m=T\}$. Então $B_m\subset A_m$. Sabemos que para quase todo $\omega\in A$,

$$Y^{(\gamma_m)}(t,\omega)=Y(t,\omega)$$

em $[0,T]$. Então para quase todo $\omega \in B_m$

$$Z^{(m)}(t,\omega) = Y(t,\omega)$$

em $[0,T]$. Portanto

$$P\Big(\sup_{0\le t\le T} |Z^{(m)}(t) - Y(t)| \neq 0\Big) \le P\Big(B_m^c\Big)$$

$$\le P\Big(\sup_{0\le t\le T} |M(t)| \ge m\Big) + P\Big(\int_0^T X^2(s)d\langle M\rangle(s) \ge \eta\Big).$$

Temos então que:

$$P\Big(\sup_{0\le t\le T} |Y(t)| > \varepsilon\Big) \le \frac{4\eta}{\varepsilon^2} + P\Big(\sup_{0\le t\le T} |M(t)| \ge m\Big) + P\Big(\int_0^T X^2(s)d < M > (s) \ge \eta\Big).$$

Como $m$ é arbitrário e $M$ é contínuo, o termo $P(\sup_{0\le t\le T} |M(t)| \ge m)$ converge a zero quando $m \to \infty$,o que prova a proposição. □

**Colorário 3.22:** Sejam $X_n$ e $X$ em $L_{2,\text{loc}}(\langle M\rangle)$ e suponhamos que

$$\int_0^T |X_n(s) - X(s)|^2 d\langle M\rangle(s) \to 0$$

em probabilidade, quando $n \to \infty$. Então,

$$\sup_{0\le t\le T} \Big|\int_0^t X_n(s)dM(s) - \int_0^t X(s)dM(s)\Big| \to 0$$

em probabilidade.

O Lema 3.17 permanece válido para martingales locais contínuos. Este fato nos permite mostrar que a definição da integral estocástica não depende da sequência de tempos de parada. Isto sera provado a seguir.

**Proposição 3.23:** Seja $X \in L_{2,\mathrm{loc}}(\langle M \rangle)$ e $\tau$ um tempo de parada. Então

$$\int_0^t X(s)dM^{(\tau)}(s) = \int_0^{t\wedge\tau} X(s)dM(s)$$

<u>Dem:</u> $M^{(\tau)}$ é um martingale local contínuo e $X \in L_{2,\mathrm{loc}}(\langle M^{(\tau)} \rangle)$, portanto

$$\int_0^t X(s)dM^{(\tau)}(s)$$

é definido. Como anteriormente, seja

$$Y(t) = \int_0^t X(s)dM(s)$$

e

$$Z(t) = \int_0^t X(s)dM^{(\tau)}(s).$$

Sejam $\tau_n$ e $\tau_n'$ os tempos de parada usados na Definição 3.19 de $Y(t)$ e $Z(t)$. Segue que $\tau_n < \tau_n'$ e para cada $t$

$$\int_0^t X(s)dM^{(\tau\wedge\tau_n)}(s) \to \int_0^t X(s)dM^{(\tau)}(s) = Z(t) \text{ q.c..}$$

Com probabilidade 1, $Y^{(\tau_n)}(t) = Y(t)$ para $n$ suficientemente grande e $Y^{(\tau_n)}(t\wedge\tau) = Y(t\wedge\tau)$. Portanto $Y^{(\tau_n)}(t\wedge\tau) \to Y(t\wedge\tau)$ q.c. e pelo Lema 3.17.

$$Y^{(\tau_n)}(t\wedge\tau) = \int_0^{t\wedge\tau} X(s)dM^{(\tau_n)}(s) = \int_0^t X(s)dM^{(\tau\wedge\tau_n)}.$$

Ou seja $Y(t\wedge\tau) = Z(t)$ q.c. para cada $t$. Como estamos considerando versões contínuas o resultado segue para todo $t \in [0,T]$ $\square$

**Proposição 3.24:** $Y(t) = \int_0^t X(s)dM(s)$ é um martingale local contínuo com variação quadrática

$$\langle Y\rangle(t) = \int_0^t X^2(s)d\langle M\rangle(s).$$

Dem: Da Proposição 3.23 temos que

$$Y^{(\tau_n)}(t) = \int_0^{t\wedge\tau_n} X(s)dM(s) = \int_0^t X(s)dM^{(\tau_n)}(s).$$

Portanto para cada $n, Y^{(\tau_n)}$ é um martingale contínuo e quadrado integrável com variação quadrática

$$\langle Y^{(\tau_n)}\rangle = \int_0^t X^2(s)d\langle M^{(\tau_n)}\rangle(s) = \int_0^{t\wedge\tau_n} X^2(s)d\langle M\rangle(s).$$

O resultado segue tomando $\tau_n \nearrow T$. □

**Proposição 3.25:** A definição de $Y(t) = \int_0^t X(s)M(s)$ não depende da escolha da sequência de tempos de parada $\tau_n \nearrow T$.

Dem: Seja $\gamma_n, n = 1, 2, \ldots$ uma sequência de tempos de parada, $\gamma_n \nearrow T$ e $X \in L_{2,\text{loc}}(\langle M\rangle)$. Então

$$\int_0^t X(s)dM^{(\gamma_n)}(s) = \int_0^{t\wedge\gamma_n} X(s)dM(s) = Y(t\wedge\gamma_n)$$

em $[0, T]$. Pela continuidade de $Y$

$$\sup_{0\leq t\leq T} \left|Y(t\wedge\tau_n) - Y(t)\right| \to 0 \quad \text{q.c..}$$

Portanto

$$\sup_{0\leq t\leq T} |\int_0^t X(s)dM^{(\tau_n)}(s) - Y(t)| \to 0 \text{ q.c.}$$ □

# 5. A Fórmula de Itô.

A fórmula de Itô, ou mudança de variáveis para integrais estocásticas sera apresentada inicialmente no caso de semimartingales contínuos $d$-dimensionais, isto é processos estocásticos da forma:

$$X(t) = \big(X_1(t), \dots, X_d(t)\big)$$

com

$$X(t) = X(0) + M(t) + B(t) \qquad t \geq 0,$$

onde

i) $E|X(0)| \leq \infty$, $M(0) = 0$ q.c. e $B(0) = 0$ q.c.

ii) $M = \{M(t) : t \geq 0\}$ é um martingale contínuo e quadrado integrável, isto é cada componente $M_i$ de $M$ é um martingale quadrado integravel.

iii) $B = \{B(t) : t \geq 0\}$ é contínuo e de variação limitada em intervalos finitos q.c. e $E(\mathrm{Var}_{[0,t]}B) < \infty$ para todo $t$, onde $\mathrm{Var}_{[0,t]}B$ denota a variação total de $B$ em $[0,t]$.

**Teorema 3.26:** Seja $X(t) = X(0) + M(t) + B(t)$ um semimartingale contínua unidimensional. Seja $F : \mathbb{R} \to \mathbb{R}$ limitada, continua e com derivadas de primeira e segunda ordem limitadas e continuas, isto é $F \in C_b^2(\mathbb{R})$. Então

$$\begin{aligned} F\big(X(t)\big) &= F\big(X(0)\big) + \int_0^t F'\big(X(s)\big)dM(s) + \int_0^t F'\big(X(s)\big)dB(s) \\ &+\frac{1}{2}\int_0^t F''\big((X(s)\big)d\langle M\rangle(s). \end{aligned}$$

<u>Dem:</u> Suponhamos que exista uma sequência $X^{(N)}$ de semimartigales contínuas $X^{(N)}(t) = X^{(N)}(0) + M^{(N)}(t) + B^{(N)}(t)$ tal que:

a) $E|X^{(N)}(0) - X(0)| \to 0$, $E|M^{(N)}(t) - M(t)|^2 \to 0$ e $E\int_0^T d\mathrm{Var}_{[0,t]}(B^{(N)} - B) \to 0$ quando $T < \infty$ e $N \to \infty$;

b) O resultado é válido para $X^{(N)}$.

Então o resultado é válido para $X$. De fato,

$$F\left(X^{(N)}(t)\right) = F\left(X^N(0)\right) \quad + \int_0^t F'\left(X^{(N)}(s)\right) dM^{(N)}(s) + \int_0^t F'\left(X^{(N)}(s)\right) dB^{(N)}(s)$$

$$+1/2 \int_0^t F''\left(X^{(N)}(s)\right) d\langle M^{(N)}\rangle(s).$$

$F\left(X^{(N)}(t)\right) \longrightarrow F\left(X(t)\right)$ e $F'\left(X^{(N)}(t)\right) \longrightarrow F'\left(X(t)\right)$ em probabilidade. Pela desigualdade de Doob e por a), segue que $\sup\limits_{0 \leq s \leq T} |M^{(N)}(s) - M(s)|^2 \to 0$ em probabilidade. Seja $C$ tal que $\max\left(|F(x)|, |F'(x)|\right) \leq C$. Seja

$$\begin{aligned} I_1 &= \int_0^t F'\left(X^{(N)}(s)\right) dM^{(N)}(s) - \int_0^t F'\left(X(s)\right) dM(s) = \\ &= \int_0^t \left(F'\left(X^{(N)}(s)\right) - F'\left(X(s)\right)\right) dM(s) + \int_0^t F'\left(X^{(N)}(s)\right) d\left(M^{(N)} - M\right)(s) = \\ &= I_2 + I_3. \end{aligned}$$

Pelo teorema da convergência dominada

$$E(I_2^2) = E \int_0^t \left|F'\left(X^{(N)}(s)\right) - F'\left(X^{(N)}(s)\right)\right|^2 d\langle M\rangle(s) \to 0,$$

pois

$$\int_0^t |F'(X^{(N)}(s) - F'(X(s))|^2 d\langle M\rangle(s) \leq 4C^2 \langle M\rangle(T)$$

e

$E\langle M\rangle(T) < \infty$.

$$E(I_3^2 \leq C^2 E(M^{(N)}(T) - M(T))^2 \longrightarrow 0$$

Portanto $I_1 \to 0$ em probabilidade. Seja

$$\begin{aligned}
I_4 &= \int_0^t F\Big(X^{(N)}(s)\Big)dB^{(N)}(s) - \int_0^t F'\Big(X(s)\Big)dB(s) = \\
&= \int_0^t \Big(F'(X^{(N)}(s)\Big) - F'\Big(X(s)\Big)dB(s) + \int_0^t F'\Big(X^{(N)}(s)\Big)d(B^{(N)} - B)(s) \\
&= I_5 + I_6.
\end{aligned}$$

$E|I_5| \leq E\int_0^t \ |F'\Big(X^{(N)}(s)\Big) - F'\Big(X(s)\Big)|d\mathrm{Var}_{[0,t]}B \to 0$ pelo teorema de convergência dominada. $E|I_6| \leq CE\int_0^t d\mathrm{Var}_{[0,t]}(B^{(N)} - B) \to 0$. Portanto $I_4 \to 0$ em probabilidade. Seja

$$\begin{aligned}
I_7 &= \int_0^t F''(X^{(N)}(s)\Big)d\langle M^{(N)}\rangle(s) - \int_0^t F''\Big(X(s)\Big)d\langle M\rangle(s) = \\
&= \int_0^{''} \Big(F''\Big(X^{(N)}(s)\Big) - F''\Big(X(s)\Big)d\langle M\rangle(s) \\
&+ \int_0^t F''\Big(X^{(N)}(s)\Big)d\Big(\langle M^{(N)}\rangle - \langle M\rangle\Big)(s) \\
&= I_8 + I_9.
\end{aligned}$$

Como no caso de $I_5$, podemos mostrar que $E|I_8| \to 0$. Agora

$$\begin{aligned}
|I_9| &\leq C\int_0^t d\mathrm{Var}_{[0,t]}\Big(\langle M^{(N)} + M, M^{(N)} - M\rangle\Big) \leq \\
&\leq C\Big(\langle M^{(N)} + M\rangle(t)\langle M^{(N)} - M\rangle(t)\Big)^{1/2}.
\end{aligned}$$

Segue que $E|I_9| \to 0$, pois

$$E\langle M^{(N)} - M\rangle(t) = E(M^{(N)}(t) - M(t)\Big)^2 \to 0 \ \text{e}$$

$$E\langle M^{(N)} + M\rangle(t) \leq 2(E(\langle M^{(N)}\rangle(t)\Big)^2 + E(M(t))^2) \to 4E(M(t))^2.$$

Portanto $I_7 \to 0$ em probabilidade. Temos então que nossa afirmação é verdadeira.

O próximo passo na demonstração do teorema requer que seja encontrada uma sequência $X^{(N)}$ de semimartingales contínuos satisfazendo (a) e (b) com $X^N(0), M^{(N)}(0)$ e $B^{(N)}(0)$ tomados limitados para cada $N$. Seja

$$\tau_N(\omega) = \begin{cases} \inf\left\{t : |M(t,\omega)| \geq N \text{ ou } \mathrm{Var}_{[0,t]}B \geq N\right\}, \\ T \text{ se o conjunto acima é vazio.} \end{cases}$$

Defina

$$\begin{aligned} &X^{(N)}(0) = X(0)\mathbb{1}(|X(0)| \leq N);\\ &M^{(N)}(t) = M^{(\tau_N)}(t) = M(t \wedge \tau_N);\\ &B^{(N)}(t) = B^{(\tau_N)}(t) = B(t \wedge \tau_N);\\ &X^{(N)}(t) = X^{(N)}(0) + M^{(N)} + B^{(N)}(t). \end{aligned}$$

Claro $E|X^{(N)}(0) - X(0)| \to 0$. Como $M^{(N)} - M$ é um martingale contínuo e quadrado integravel com $M^{(N)}(0) - M(0) = 0$, segue que

$$E(M^{(N)}(t) - M(t))^2 = E\langle M - M^{(N)}\rangle(t) =$$

$$= E\langle M\rangle(t) + E\langle M^{(N)}\rangle(t) - 2E\langle M, M^{(N)}\rangle(t)$$

$$= E\langle M\rangle(t) - E\langle M^{(N)}\rangle(t),$$

pois $\langle M, M^{(N)}\rangle(t) = \langle M\rangle(t \wedge \tau_N) = \langle M^{(N)}\rangle(t)$. Temos também que $\langle M^{(N)}\rangle(t,\omega) \to \langle M\rangle(t,\omega)$ q.c. e $\langle M^{(N)}\rangle(t,\omega) \leq \langle M\rangle(t,\omega)$ q.c. para todo $N$. Segue pelo teorema da convergência dominada que $E\langle M^{(N)}\rangle(t) \to E\langle M\rangle(t)$ e portanto $E|M^{(N)}(t) - M(t)|^2 \to 0$. Agora

$$\int_0^t d\mathrm{Var}_{[0,s]}(B - B^N) \leq \int_{\tau_N}^T d\mathrm{Var}_{[0,s]}B \to 0 \text{ q.c.}$$

e

$$\int_{\tau_N}^T d\mathrm{Var}_{[0,s]}B \leq \int_0^T d\mathrm{Var}_{[0,s]}B.$$

Novamente, como $E\int_0^T d\text{Var}_{[0,s]}B < \infty$, pelo teorema da convergência dominada segue que $E\int_0^t d\text{Var}_{[0,s]}(B - B^N) \to 0$. Portanto a sequência $X^{(N)}$ satisfaz (a) e (b).

Dado os resultados acima, o teorema estara provado se considerarmos o caso de um semimartingale $X$ limitada, isto é, $X(t) = X(0) + M(t) + B(t)$, onde $|X(0)| \le K$, $M(t)$ é uma martingale continua com $|M(t)| \le K$ para todo $t$ e $\int_0^T d\text{Var}_{[0,s]}B \le K$. Para tanto consideremos o processo tomando valores em $[-3K, 3K]$ e $C = \sup_{x\in[-3K,3K]} (|F'(x)|, |F''(x)|)$.

$$F(b) - F(a) = F'(a)(b-a) + \frac{1}{2}F''(a)(b-a) + R(a,b)$$

com $|R(a,b)| \le \varphi(|b-a|)(b-a)^2, \varphi$ uma função crescente convergindo para 0 em 0. Consideremos a seguinte partição aleatória de $[0,t]$. Seja $a > 0$, $t_0 = 0$ e

$$t_{i+1} = t \wedge (t_i + a) \wedge \inf\Big\{s > t_i : |M(s) - M(t_i)| > a, \text{ ou } |B(s) - B(t_i)| > a\Big\}.$$

Quando $a \to 0$, $\sup(t_{i+1} - t_i) \le a \to 0$ e $\sup|M(t_{i+1})| - M(t_i) \le a \to 0$. Seja

$$F\big(X(t)\big) - F\big(X(0)\big) = \sum_i \Big[F\big(X(t_{i+1}) - F\big(X(t_1)\big)\Big] =$$

$$= \sum_i F'\big(X(t_i)\big)\big(X(t_{i+1}) - X(t_i)\big) + \frac{1}{2}\sum_i F''\big(X(t_i)\big)\big(X(t_{i+1}) - X(t_i)\big)^2$$

$$+ \sum_i R\big(X(t_i), X\big(t_{i+1}\big)\big) =$$

$$= S_1 + \frac{1}{2}S_2 + S_3.$$

$$S_1 = \sum_i F'\big(X(t_i)\big)\big(M(t_{i+1}) - M(t_i)\big) + \sum_i F'\big(X(t_i)\big)\big(B(t_{i+1}) - B(t_i)\big) = S_1' + S_1''$$

$$E(S_1' - \int_0^t F'(X(s))dM(s))^2 = E\Big(\sum_i \int_{t_i}^{t_{1+i}} (F'(X(s)) - F'(X(t_i))^2 d\langle M\rangle(s)\Big)$$

$$\leq E\Big(\Big[\sup_i \sup_{t_i<s\leq t_{i+1}} \Big(F'(X(s)) - F'\big(X(t_i)\big)\Big)^2\Big]\langle m\rangle(t)\Big) \to 0,$$

pois $\langle M\rangle(t)$ é integravel e o supremo converge uniformemente a 0. Agora

$$\Big|S_1'' - \int_0^t F'\big(X(s)\big)dB(s)\Big| \leq \sum_i \int_{t_i}^{t_{i+1}} \big(F'\big(X(s)\big) - F'\big(X(t_i)\big)\big)d\mathrm{Var}_{[0,s]}B$$

$$\leq \Big\{\sup_i \sup_{t_i\leq s\leq t_{i+1}} \Big|F'\big(X(s)\big) - F'\big(X(t_i)\big)\Big|\Big\} \int_0^t d\mathrm{Var}_{[0,s]}B.$$

Portanto $E|S_1'' - \int_0^t F'(X(s))dB(s)| \to 0$. $S_2$ pode ser escrito da forma

$$S_2 = \sum_i F''\big(X(t_i)\big)\big(M(t_{i+1}) - M(t_i)\big)^2$$

$$+2\sum_i F''\big(X(t_i)\big)\big(M(t_{i+1}) - M(t_i)\big)^2 - M(t_i)\big)\big(B(t_{i+1}) - B(t_i)\big)+$$

$$+\sum_i F''\big(X(t_i)\big)\big(B(t_{i+1}) - B(t_i)\big)^2 =$$

$$S_2' + S_2'' + S_3'''.$$

Observemos primeiramente que $E\langle M\rangle^2(t) < \infty$ para todo $t \in [0,T]$ e

$$E\langle M\rangle^2(T) \quad = 2E\int_0^T \big(E(M(t)^2/\mathcal{F}_t) - M^2(t)\big)d\langle M\rangle(t)$$

$$\leq 2K^2 E\langle M\rangle(T) < \infty,$$

pois $E\langle M\rangle(t) = EM^2(t) \leq K^2$ para todo $t \in [0,T]$. Seja

$$\int_2^* = \sum_i F''(X(t_i))\big[\langle M\rangle(t_{i+1}) - \langle M\rangle(t_i)\big].$$

Segue que $E|S_2^* - \int_0^t F''(X(s))d\langle M\rangle(s)| \to 0$. Como

$$E[(\langle M\rangle(t_{i+1})-\langle M\rangle(t_i))-(\dot{M}(t_{i+1})-M(t_i))^2/\mathcal{F}_{t_i})=0$$

temos que:

$$E\Big(S_i^*-S_2'\Big)=\sum_i E\Bigg[\Big(F''(X(t_i))\Big)^2\Big\{\langle M\rangle(t_{i+1})-\langle \dot{M}(t_i)\rangle-\Big(M(t_{i+1})-M(t_i)\Big)^2\Big\}^2\Bigg]$$

$$\leq 2C^2\Big(E\Big[\sum_i\Big(\langle M\rangle(t_{i+1})-\langle M\rangle(t_i)\Big)^2+E\Big(\sum_i(M(t_{i+1})-M(t_i)\Big)^4\Big).$$

O primeiro valor esperado é limitado superiormente por

$$E\Big[\sup(\langle M\rangle(t_{i+1})-\langle M\rangle(t_i)\langle M\rangle(t)\Big].$$

O supremo converge a zero pois $\langle M\rangle(s)$ é uniformemente contínuo em $[0,t]$, dominado por $\langle M\rangle(t)$ e como foi mostrado $E\langle M\rangle^2(t)<\infty$. O segundo valor esperado é dominado por

$$E\Bigg[\sup_i\Big(M(t_{i+1})\quad -M(t_i)\Big)^2\sum_j\Big(M(t_{j+1})-M(t_j)\Big)\Bigg]\leq a^2E\Big(\sum_j(M(t_j)-M(t_i))\Big)^2$$

$$=a^2E(M^2(t))\to 0,$$

quando a $\to 0$, onde $|M(t_{i+1},\omega)-M(t_i,\omega)|\leq a$ pela maneira como a partição foi definida. Temos então que $E(S_2^*-S_2')^2\to 0$.

$S_2'''\to 0$ em $L_1$ pois $|S_2'''|\leq C\sup|B(t_{i+1})-B(t_i)|\int_0^t d\mathrm{Var}_{[0,s]}B$ e $\sup_i|B(t_{i+1})-B(t_i)|\leq a$. Analogamente $S_2''\to 0$ em $L_1$.

Finalmente temos que $S_3$ é majorado por

$$\sum_i\Big(X(t_{i+1})-X(t_i)\Big)^2\varepsilon\Big(|X(t_{i+1})-X(t_i)|\Big)\leq$$

$$\leq 2\varepsilon(2a)\sum_i\Big[\Big(B(t_{i+1})-B(t_i)\Big)^2+\Big(M(t_{i+1})-M(t_i)\Big)^2\Big].$$

Agora,

$$E\Big(\sum_i M(t_{i+1}) - M(t_i)\Big)^2 = E\Big(M^2(t)\Big),$$

$$\begin{aligned} E\Bigg[\sum_i \Big(B(t_{i+1}) - B(t_i)\Big)^2\Bigg] &\le aE\Bigg[\textstyle\sum_i |B(t_{i+1}) - B(t_i)|\Bigg] \\ &\le aE\int_0^t d\mathrm{Var}_{[0,s]}B, \end{aligned}$$

e $\varepsilon(2a) \to 0$ quando $a \to 0$. Portanto $E(S_3) \to 0$.

Coletando todos os resultados provados temos que

$$\begin{aligned} S_1 + \frac{1}{2}S_1 + S_3 \to \int_0^t F'(X(s))dM(s) &+ \int_0^t F'(X(s))dB(s) \\ &+ \tfrac{1}{2}\int_0^t F''(X(s))d\langle M\rangle(s) \end{aligned}$$

em probabilidade, quando $a \to 0$. □

A versão $d$-dimensional do Teorema 3.26 é apresentada a seguir. Exceto por problemas de notação a demonstração segue as mesmas linhas do caso unidimensional.

**Teorema 3.27:** Seja $X(t) = X(0) + M(t) + B(t)$ uma semimantingale contínua $d$-dimensional e seja $F : \mathbb{R}^d \to \mathbb{R}$ limitada, contínua com derivadas de maneira e segunda ordem contínuas e limitadas. Então

$$\begin{aligned} F\big(X(t)\big) = F\big(X(0)\big) &+ \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}\big(X(s)\big)dM_i(s) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}\big(X(s)\big)dB_i(s) + \\ &+ \tfrac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}\big(X(s)\big)d\langle M_i, M_j\rangle(s)\ . \end{aligned}$$

Uma outra versão de fórmula de Itô que também é provada da mesma maneira que os resultados anteriores é a seguinte:

**Teorema 3.28:** Seja $\{M(t) : t \in \mathbb{T}\}$ um martingale $d$-dimensional contínuo e quadrado integrável com respeito a uma filtração $\{\mathcal{F}_t : t \in \mathbb{T}\}, \{V(t) : t \in \mathbb{T}\}$ um processo contínuo, adaptado, com trajetórias q.c. de variação limitada em $[0,t], t \geq 0$, tal que $E\mathrm{Var}_{[0,t]}V < \infty$ e $F : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}, F(v, x_i \ldots x_d)$ limitada, contínua, com derivada parcial com respeito a $v$ contínua e limitada e derivadas de primeira e segunda ordem com respeito a $x_1, \ldots, x_d$. Então,

$$F\big(V(t), M(t)\big) = F\big(V(0), M(0)\big) + \int_0^t \frac{\partial F}{\partial v}\big(V(s), M(s)\big)\, dV(s)$$

$$+ \sum_{i=1}^d \int_0^1 \frac{\partial F}{\partial x_i}\big(V(s), M(s)\big)\, dM(s)$$

$$+ \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}\big(V(s), M(s)\big)\, d\langle M_i, M_j\rangle(s) \ .$$

O resultado dos Teorema 3.26 e 3.27 serão derivados agora para o caso de semimartingales locais contínuos sobre um intervalo finito $[0,T]$.

**Teorema 3.29:** Seja $X(t) = X(0) + M(t) + B(t)$ um semimartingale local, $d$-dimensional, contínuo e $F : \mathbb{R}^d \to \mathbb{R}$ uma função contínua com derivadas parciais de primeira e segunda ordem. Então

$$F\big(X(t)\big) = F\big(X(0)\big) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}\big(X(s)\big)\, dM_i(s) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}\big(X(s)\big)\, dB_i(s)$$

$$+ \frac{1}{2} \sum_{i=1}^d \sum_{i=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_i}\big(X(s)\big)\, d\langle M_i, M_i\rangle(s) \ .$$

Dem: Demonstraremos o resultado para $d = 1$.

Como $F(X(s))$, $F'(X(s))$ e $F''(X(s))$ são contínuas, as integrais $\int_0^T | F'(X(s))d \operatorname{Var}_{[0,s]}B$, $\int_0^T |F''(X(s))|d\langle M\rangle(s)$ e $\int_0^T (F'(X(s))^2 d\langle M\rangle(s)$ são finitas q.c.. Defina

$$\sigma_n = \inf\left\{t \leq T : |M(t)| \geq n \quad \text{ou} \quad \int_0^1 (F'(X(s))^2 d\langle M\rangle(s) \geq n \text{ ou } \operatorname{Var}_{[0,t]}B \geq n\right\}$$

ou $\sigma_n = T$ se o conjunto é vazio. Seja $\sigma_n' = T$ se $|X(0)| \leq n$ e $\sigma_n' = 0$ se $|X(0)| > n$. Seja $\tau_n = \sigma_n \wedge \sigma_n', X^{(n)}(0) = X(0)\mathbb{1}_{\{|X(0)|\leq n\}}$, $M^{(n)} = M(t \wedge \tau_n)$, $B^{(n)}(t) = B(t \wedge \tau_n)$ e $X^{(n)}(t) = X^{(n)}(0) + M^{(n)}(t) + B^{(n)}(t)$. Então $X^{(n)}(t)$ é um semimartingale contínuo e $X^{(n)}(t) \in [-3n, 3n]$. Considere $U$ e $V$ intervalos finitos e abertos tais que $[-3n, 3n] \subset V \subset \overline{V} \subset U$ e seja $\phi \in C_b^2(I\!R)$ tal que, $\phi(x) = 1$, se $x \in \overline{V}$ e $\phi(x) = 0$, se $x \in U^c$. Defina $H(x) = F(x)\phi(x)$. Claro $H \in C_b^2(I\!R)$ e em $[-3n, 3n]$ $H, H'$ e $H''$ coincidem com $F, F'$ e $F''$, respectivamente. Do Teorema 3.26 temos que

$$\begin{aligned}F\big(X^{(n)}(t)\big) = F\big(X^{(n)}(0)\big) \ &+ \int_0^t F'\big(X^{(n)}(s)\big)dM^{(n)}(s) + \int_0^t F'\big(X(s)\big)dB^{(n)}(s) \\ &+1/2\int_0^t F''\big(X^{(n)}(s)\big)d\langle M^{(n)}\rangle(s) \ .\end{aligned}$$

Como $X^{(n)} \to X(t)$ q.c. e $F'(X^{(n)}(t)) \to F'(X(t)), t \geq 0$, segue dos resultados de integrais estocásticas com respeito com respeito a martingales locais contínuas que

$$\int_0^t F'\big(X(s)\big)dM^{(n)}(s) \longrightarrow \int_0^t F'\big(X(s)\big)dM(s) \ .$$

em probabilidade quando $n \to \infty$.

Com probabilidade 1, temos que existe $N = N(\omega)$ tal que, se $n \geq N(\omega)$

$$\int_0^t F'\big(X^{(n)}(s)\big)dB^{(n)}(s) = \int_0^t F'\big(X(s)\big)dB(s).$$

Portanto,

$$\int_0^t F'\Big(X^{(n)}(s)\Big)dB^{(n)}(s) \longrightarrow \int_0^t F'\Big(X(s)\Big)dB(s) \qquad \text{q.c..}$$

Analogamente podemos mostrar que

$$\int_0^t F''\Big(X^{(n)}(s)\Big)d\langle M^{(n)}\rangle(s) \longrightarrow \int_0^t F''\Big(X(s)\Big)d\langle M\rangle(s) \quad \text{q.c..}$$

Fazendo $n \to \infty$ obtemos

$$F\Big(X^{(n)}(t)\Big) \longrightarrow F\Big(X(t)\Big).$$

Segue que

$$F\Big(X^{(n)}(0)\Big) + \int_0^t F'\Big(X^{(n)}(s)\Big)dM^{(n)}(s) + \int_0^t F'\Big(X^{(n)}(s)\Big)dB^{(n)}(s) +$$

$$\frac{1}{2}\int_0^T F''\Big(X^{(n)}(s)\Big)d\langle M^{(n)}\rangle(s)$$

converge para

$$F\Big(X(0)\Big) + \int_0^t F'\Big(X^{(n)}(s)\Big)dM^{(n)}(s)dM^{(n)}(s) + \int_0^t F'\Big(X(s)\Big)dB(s)$$

$$+\frac{1}{2}\int_0^t F''\Big(X(s)\Big)d\langle M\rangle(s). \qquad \square$$

# 4. Equações Diferenciais Estocásticas.

**Introdução:** Falando por alto, uma equação diferencial estocástica em $I\!R^d$, no sentido de Itô, é uma expressão do tipo:

$$dx = X(t,x)dt + \sum_{j=1}^{m} Y_j(t,x)dW^j \tag{I}$$

onde $X, Y_1 \ldots, Y_m$ são campos de vetores e $W = (W^1, ..., W^m)$ é um movimento Browniano canonico em $I\!R^m$. O campo $X$ é conhecido como a deriva ("drift") de (I) enquanto que os $Y_j' s$ são os coeficientes da parte difusa. Sendo uma equação diferencial, (I) só é de fato definida a partir do momento em que se específica o que se entende por suas soluções. Evidentemente, uma solução de (I) é um processo $\xi_t$ que satisfaz a equação integral correspondente. Por exemplo, se $\xi_t$ é definido no intervalo $[s, T]$, deve-se ter

$$\xi_t = \xi_s + \int_s^t X(u, \xi_u)du + \sum_{j=1}^{m} \int_s^t Y_j(u, \xi_u)\, dW_u^j$$

para $t\varepsilon\ [s, T]$ onde a primeira integral é uma integral usual sobre caminhos aleatórios e as outras são integrais estocásticas no sentido de Itô. No entanto, ao se definir de maneira precisa as soluções de (I), faz-se uma distinção – um tanto sutil – entre as soluções denominadas fortes e as soluções fracas. Esta distinção surge devido à arbitrariedade na escolha do movimento Brownianio: Enquanto nas soluções fortes este é fixado de antemão, para as soluções fracas, admite-se a possibilidade de construí-lo em algum espaço de probabilidade com filtração, isto é em alguma base estocástica. A distinção entre as soluções fortes e fracas dá origem a abordagens diferentes às questões de existência e unicidade. Em essencia, busca-se soluções fracas quando a preocupação é com as leis do processo enquanto que nas soluções fortes olha-se o processo propriamente dito.

Usualmente, a existência e unicidade de soluções fracas se obtém sob condições menos restritas que para as soluções fortes.

Neste capítulo, vamos desenvolver a teoria de existência e unicidade de soluções fortes para equações com coeficientes Lipschitz. Esta teoria é o análogo da teoria clássica de existência e unicidade para equações ordinárias que emerge do método das aproximações sucessivas (teorema do ponto fixo de Banach, veja p.ex. [Hö], [Di],...).

Na equação (I), as integrais estocásticas são tomadas em relação ao movimento Browniano.

Estas são as únicas equações que serão consideradas neste texto. Na literatura, no entanto, encontra-se equações sobre semi-martingales arbitrários. Referimos à [E], [I-W], para estas equações.

Outro comentário sobre (I), é que ao invés de integrais de Itô pode-se considerar também integrais de Stratonovich. Discutiremos este tipo de equação quando for feito o transporte da teoria em $\mathbb{R}^d$ para variedade diferenciáveis. Por se comportarem melhor sob mudanças de coordenadas, as equações de Strotonovich são definidas de maneira mais natural em variedades que as de Itô. As equações de Itô, porém, exigem menos regulariedade nos coeficientes, sendo possível para as mesmas resultados mais gerais de existência e unicidade de soluções.

## 4.1 Soluções fortes. Definições.

O domínio de (I) será tomado como sendo um aberto $U \subset [0,\infty)\times \mathbb{R}^d$. Os coeficientes serão então aplicações $X, Y_1, ..., Y_m : U \to \mathbb{R}^d$, que suporemos, frequentemente, que satisfazem uma condição de Lipschitz segundo a

**Definição 1.1:** Uma aplicação $F : V \subset [0, \infty) \times I\!R^d \to I\!R^d$ é dita Lipschitz em relação à $x$, se existe uma constante $L > 0$ t.q..

$$\Big|f(t,x) - f(t,y)\Big| \leq L\Big|x - y\Big|$$

para todo $(t,x)$, $(t,y) \in V$ ($|\cdot|$ denota uma norma em $I\!R^d$), $f$ é localmente Lipschitz se for Lipsichitz em alguma vizinhança de $(t,x)$ todo $(t,x)$. $f$ é Lipschitz global (abreviado L.g.) se $f$ é Lipschitz e $V$ é da forma $I \times I\!R^d$ com $I \subset [0, \infty)$ um intervalo.

Para definir as soluções fortes, fixamos de antemão um movimento Browniano em $I\!R^m$, que pode ser tomado sem perda de generalidade como sendo o movimento Browniano standart, que é definido sobre o espaço de Wiener $(\Omega, P, \mathcal{F}_t)$ onde $\Omega = \{\omega = [0, \infty) \to I\!R^m$: $\omega$ é contínua, $\omega(0) = 0\}$, $P$ é a medida de Wiener e $\mathcal{F}_t$ é o completamento da $\sigma$-algebra gerada por $\{W_s : 0 \leq s \leq t\}$. O processo $W_t$ é dado explicitamente por $W_t(\omega) = \omega(t)$, $\omega \in \Omega$, $t \in [0, \infty)$.

Para levar em conta soluções de (I), iniciadas em $s > 0$, consideraremos também a seguinte variação deste movimento Browniano: Para $0 \leq s \leq t$, seja $\mathcal{F}_{s,t}$ a $\sigma$-algebra completa gerada por $W_v - V_u, s \leq u \leq v \leq t$. A família $(\mathcal{F}_{s,t})_{s,t}$ é crescente em $t$ e decrescente em $s$. Além do mais, se $\theta_s$ é a translação $(\theta_s(\omega)(t) = \omega(t+s) - \omega(s)$ em $\Omega$, então $\mathcal{F}_{s,t} = \theta_s(\mathcal{F}_t)$ e daí que $t \to W_t - W_s$ é um $(\mathcal{F}_{s,t})$ – movimento Browniano, já que a medida de Wiener é $\theta$-invariante.

**Definição 1.2:** Tome (I) com domínio de definição $U \subset [0, \infty) \times I\!R^d$ e suponha que os coeficientes $X, Y_1 ..., Y_m$ sejam contínuos. Seja $(s,x) \in U$. Um processo contínuo $(t,\omega) \to \xi_{s,t}(x,\omega)$ definido para valores $(t,\omega)$ tais que $s \leq t \leq T(s,x,\omega)$ é uma solução (forte) de (I) com condição inicial $(s,x)$ se:

i) $\xi_{s,s}(x,\omega) = x$ e $(t, \xi_{s,t}(x,\omega)) \in U, q.s.$,

ii) $T(s,x,\cdot)$ é um $(\mathcal{F}_{s,t})$ – tempo de parada predizível,

iii) $\xi_{s,t}(x,\omega)$ é $(\mathcal{F}_{s,t})$ – adaptado e

iv) $\xi_{s,t}(x,\cdot) = x + \int_s^t X(u,\xi_{s,u}(x,\cdot)du + \sum_{j=1}^{m} \int_s^t Y_j\big(u,\xi_{s,u}(x,\cdot)\big)dW_u^i$

se $\leq t < T(s,x,\cdot)$. O tempo de parada $T(s,x,\cdot)$ é denominado tempo de explosão de solução.

**Observações 1.3:**

a) Como $\xi_{s,t}$ é um processo contínuo e adaptado, ele é predizível. Com os coeficientes contínuos, as integrais estocásticas em iv) fazem sentido. A fortiori, $\xi_{s,t}(x,\omega)$ é um $(\mathcal{F}_{s,t})$ semi-martingale.

b) Ao invés de $\xi_{s,t}(x,\omega)$ usaremos freqüêntemente as notações $\xi_{s,t}(x)$, $\xi_{s,t}(x)(\omega)$, $\xi_{s,t}(\omega)(x)$ ou $\xi_{s,t}(\omega)$. Estas últimas ressaltam a dependência $x \to \xi_{s,t}(\omega)(x)$ das soluções em relação à $x$.

c) É sempre conveniente salientar que (I) se reduz à uma equação ordinária no caso em que os coeficientes difusos $Y_1 \cdots, Y_m$ se anulam. Neste caso, cada variável aleatória $\xi_{s,t}(x)$ é constante e as soluções são determinísticas.

d) Outra redução de (I) é quando os coeficientes não dependem explicitamente de $t$. Estas equações são denominadas autonomas ou homogenêas no tempo. O domínio de definição é de forma $U = [0,\infty) \times V$ com $V$ aberto de $\mathbb{R}^d$.

**Definição 1.4:** Dadas duas soluções $\xi_{s,t}^1(x)$ e $\xi_{s,t}^2(x)$ com tempos de explosão $T^1(s,x)$ e $T^2(s,x)$, respectivamente, diremos que $\xi^1$ está contida em $\xi^2$ (notação $= \xi_{s,t}^1(x) \prec \xi_{s,t}^2(x)$) se $T^1(s,x) \leq T^2(s,x)$ e $\xi_{s,t}^1(x)$ e $\xi_{s,t}^2(x)$ são processos indistinguíveis até $T^1(s,x)$.

A relação $\prec$ define uma ordem parcial no conjunto das soluções. Uma solução é maximal se o for em relação a esta ordem.

**Observação 1.5:** A definição acima é analoga à de soluções maximais para equações diferenciais ordinárias. Lembremos que para estas equações mostra-se, independente de qualquer hipótese, por uma aplicação do lema de Zorn que toda solução está contida em uma solução maximal. Para equações estocásticas, mostraremos posteriormente, sob hipótese adicionais (Lipschitz) a existência de soluções maximais. Não é possível no entanto, construir soluções maximais a partir de alguma solução, de maneira totalmente abstrata via o lema de Zorn. A razão principal é que este procedimento exige que se tome limites de uma quantidade não enumerável de tempos de parada, o que pode dar origem a aplicações não mensuráveis.

**Exemplos 1.6:** a) Seja a equação $dx = dW$ em $\mathbb{R}^d$. A única solução (a menos de indistinguibilidade) com condição inicial $(s,x) \in [0,\infty) \times \mathbb{R}^d$ é dada pelo processo

$$\xi_{s,t}(x) = x + W_t - W_s$$

com tempo de explosão $T(s,x) = +\infty$. Este processo se escreve explicitamente

$$\xi_{s,t}(x,\omega) = x + \omega(t) - \omega(s) \ . \qquad (*)$$

As seguintes observações sobre estas soluções, antecipam os resultados gerais sobre dependência em relação às condições iniciais.

i) Por $(*)$, vê-se que $\forall\ \omega \in \Omega$, a aplicação $(s,t,x) \to \xi_{s,t}(x,\omega)$ é contínua (compare com teorema 4.8).

Em geral, esta continuidade é quase sempre $\beta$-Hölder em $s$ e $t$, todo $\beta < \frac{1}{2}$. Neste caso particular, esta propriedade é conseqüência da condição de Hölder dada por Levy para o movimento Browniano.

ii) $\forall\ \omega \in \Omega$, $s,t$ fixados, $x \in I\!R^d \to \xi_{s,t}(\omega)(x) \in I\!R^d$ é um difeomorfismo de classe $l^\infty$, por ser uma aplicação afim (compare com teorema 4.12 e corolário 5.2). Além do mais, se $s \leq u \leq t$

$$\begin{aligned}
\xi_{u,t}\Big(\xi_{s,u}(x)\Big)(\omega) &= \xi_{s,u}(x)(\omega) + \omega(t) - \omega(u) \\
&= x + \omega(t) - \omega(s) \\
&= \xi_{s,t}(x)(\omega)
\end{aligned}$$

todo $\omega \in \Omega$.

iii) Como $dx = dW$ é uma equação autonoma, $\xi_{s,t}$ depende apenas de $\xi_{0,t}$ (veja discussão ao final do § 4). De maneira precisa,

$$\begin{aligned}
\xi_{s,t}(\omega)(x) &= x + \omega(t) - \omega(s) \\
&= x + (\theta_s\omega)(t-s) \\
&= \xi_{0,t-s}(\theta_s(\omega))(x)
\end{aligned}$$

onde $(\theta_s(\omega))(t)$ é a translação em $\Omega$. Desta igualdade tem-se a equação de fluxo: Pondo $\phi_t(\omega) = \xi_{0,t}(\omega)$,

$$y_{t+s}(\omega) = y_t(\theta_s(\omega))_0 y_s(\omega).$$

b) Equações Lineares: Seja $I = (a,b) \subset [0,\infty)$ um intervalo e tome aplicações contínuas

$$A : I \to M_{dxd} \quad \text{e} \quad b_j : I \to \mathbb{R}^d, j = 1, \ldots, m$$

e defina

$$dx = A(t)x\, dt + \sum_{j=1}^{m} b_j(t)\, dW^j \tag{1.1}$$

Como para equações diferenciais ordinárias, a fórmula da variação das constantes também funciona: Seja $\Phi_{s,t}$ a solução fundamental de $x = A(t)x$. Então

$$\xi_{s,t}(x) = \Phi_{s,t}x + \Phi_{s,t} \sum_{j=1}^{m} \int_s^t \Phi_{u,s}\, b_j \;\; dW_u^j \tag{1.2}$$

é solução de (1.2). Isto se verifica diretamente por intermédio da fórmula de Itô. (1.2) é a única solução (a menos de indistinguibilidade) com condições iniciais $(s, x)$ como se conclui a partir do teorema 4.5. O tempo de explosão $T(s, x) = b$ é determinístico.

c) Equações Bilineares: Tomando $I$ como no exemplo anterior, sejam $A, B_1 \cdots B_m : I \to M_{dxd}$ contínuas e defina a equação "bilinear" em $x$ e $W$,

$$dx = A(t)x\, dt + \sum_{j=1}^{m} B_j(t)x\, dW^j, \; x \in \mathbb{R}^d \tag{1.3}$$

Pela fórmula de Itô, ve-se que se $g_{s,t}$ é solução de

$$dg = A(t)g\, dt + \sum_{j=1}^{m} B_j(t) g dW^i \quad g \in M_{d\times d} \tag{1.4}$$

com condição inicial $g_{s,s} = 1$ (= matriz identidade), então $\xi_{s,t}(x) = g_{s,t}\ x$ é solução de (1.4) com condição inicial $(s, x)$.

## 4.2 Desigualdades.

O método de demostração dos resultados sobre existência e unicidade de soluções e de dependência em relação a parâmetros, envolve uma série de majorações nas quais certos procedimentos e algumas desigualdades conhecidas aparecem reiteradamente. Por uma questão de conveniência na leitura e redação do texto, vamos catalogá-los aqui:

1. Para duas expressões $(\cdots)$ e $\{\cdots\}$ a desigualdade

$$(\cdots) \leq C\,\{\cdots\}$$

significa que existe uma constante $C > 0$ t.q. o lado direito majora o esquerdo. As vezes esta constante depende de alguma quantidade que aparece em $(\cdots)$ ou $\{\cdots\}$. Nesse caso colocaremos um subindice em $C$ para indicar esta dependência, se ela for relevante (p. ex. $C_\varepsilon$ na desigualdade $(cL)$ abaixo ou $C_p$ em (3.1)). Caso contrário, usaremos $C$ apenas mesmo que constantes diferentes sejam indicadas. Por exemplo, se $a_1$, $a_2$, $b > 0$ temos

$$C(a_1 + ba_2) \leq C(a_1 + a_2) \tag{C}$$

onde é claro, $C$ do lado esquerdo não é o mesmo que do lado direito. Esta desigualdade segue de

$a_1 + ba_2 \leq \max\{1, b\}(a_1 + a_2)$.

2. Desigualdades Algébricas

A.1) Para $x, y > 0$ e $p \in I\!R$,

$$(x + y)^p \leq C(x^p + y^p) \tag{A.1}$$

Aqui pode-se tomar $C = 2^p$. Para obter esta desigualdade, olhe primeiro o caso em que $x = 1$ analisando o gráfico de $y \in [0, \infty) \to (1 + y)^p/(1 + y^p)$. O caso geral saí daí colocando-se $x^p$ em evidência.

A.2) Para $x_1, \cdots, x_n$ reais positivos e $p \in I\!R$

$$(x_1 + \cdots + x_n)^p \leq C(x_1^p + \cdots + x_n^p) \tag{A.2}$$

Aqui pode-se tomar $C = n^p$. Esta desigualdade se obtém a partir de A.1 por indução sobre $n$.

A.3) Para $x > 0, \varepsilon > 0$ e $p \in I\!R$.

$$\left(\varepsilon + x\right)^p \leq C_\varepsilon \left(\varepsilon + x^2\right)^{p/2} \tag{A.3}$$

pois $(\varepsilon + x)^p/(\varepsilon + x^2)^{p/2}$ é limitada em $[0, \infty)$.

A.4) Para $(x_1, \cdots, x_d) \in I\!R^d$, $\varepsilon > 0$.

$$|x|,\ |x_i| \leq\ C\ (\varepsilon + |x|^2)^{1/2} \tag{A.4}$$

onde $|x|$ é uma norma qualquer em $I\!R^d$.

A.5) Para $a, b > 0$,

$$\left(1+\frac{1}{a}+\frac{1}{b}\right)\frac{a}{1+a}\,\frac{b}{1+b} = 1 - \frac{1}{1+a+b+ab} \leq 1 \tag{A.5}$$

A.6) Para $x, y \in \mathbb{R}^d$, $x \neq 0 \neq y$,

$$\frac{1}{1+|x|}\,\frac{1}{1+|y|}\,|x-y| \leq \left|\frac{x}{|x|^2} - \frac{y}{|y|^2}\right| \tag{A.6}$$

Com $|\cdot|$ a norma euclidiana em $\mathbb{R}^d$. Uma maneira de se obter esta desigualdade é a seguinte: Tome $x, z \in \mathbb{R}^d$ com $|x| = |z|$ e $r > 0$. Então pela simetria do triângulo isósceles de vértices $x, z$ e 0, tem-se $|x - rz| = |rx - z|$. De onde se tira que:

$$\frac{r}{1+r}\,\left|x - rz\right| \;\leq\; \left|rx - z\right| \cdot \tag{$*$}$$

Tomando agora $r = |y|/|x|$, (A.6) é equivalente à

$$\frac{|x|}{1+|x|}\,\frac{|x|}{1+r|x|}\,|x-y| \;\leq\; \left|x - \frac{1}{r}\,\frac{y}{r}\right|$$

que por sua vez é obtida de $(*)$ ao se escrever $z = \;\; y/r$ e usar o fato de que $|x|/(1+|x|) < 1$.

3. Uma consequência da desigualdade de Lipschitz:

Suponha que $f : I \times \mathbb{R}^d \to \mathbb{R}^d$ é Lipschitz global. Então dado $\varepsilon > 0$ existe $C_\varepsilon > 0$ dependendo de $\varepsilon$, t.q..

$$\left|f(t,x)\right| \leq\; C_\varepsilon\big(\varepsilon + |\,x\,|\big) \tag{cL}$$

$\leq C_\varepsilon\Big(\varepsilon + |\, x\,|^2\Big)^{1/2}$

De fato, $|\, f(t,x) - f(t,0)\,| \leq L\,|\,x\,|$ onde $L$ é a constante de Lipschitz. Daí que

$$\begin{aligned} \Big|f(t,x)\Big| &\leq |\, f\big(t,0\big)\,| + L\,|\,x\,| \\ &\leq \frac{|\, f\big(t,0\big)\,|}{\varepsilon}\,\varepsilon + L\,|\,x\,| \\ &\leq C_\varepsilon\Big(\varepsilon + |\,x\,|\Big) \qquad \text{por (C)} \\ &\leq C_\varepsilon\Big(\varepsilon + |\,x\,|^{1/2}\Big)^2 \qquad \text{por (A.3)} \end{aligned}$$

4. Desigualdade de Hölder: Para $r \geq 1$

$$\|\, f\,\|_r \leq \|\, f\,\|_p\, \|\, f\,\|_q \qquad \text{(H o r,p,q)}$$

com $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, onde $\|\, f\,\|_r = (\int |\, f\,|^r)^{1/r}$

5. Desigualdade de Burkhölder: Para $M_t$ martingale contínuo e $p \geq 2$,

$$E\Big[\,|\, M_t\,|^p\,\Big] \leq C_p\, E\Big[\langle M\rangle_t^{p/2}\Big] \qquad \text{(Bh)}$$

todo $t \in [0,T]$ se $E[\,|\, M_T\,|^p\,] < \infty$.

Para mostrar esta desigualdade, aplicamos a fórmula de Itô à $F(x) = |x|^p$ que é de classe $C^2$ se $p \geq 2$. Denotando por $\mathrm{sign}(x)$ a função sinal, temos

$$|M_t|^p = p\int_0^t |M_s|^{p-1}\mathrm{sign}(M_s)dM_s + \frac{1}{2}p(p-1)\int_0^t |M_s|^{p-2}d\langle M\rangle_s.$$

Tomando esperança nesta igualdade,

$$\begin{aligned} E|M_t|^p &\leq \tfrac{1}{2}p(p-1)E\left[\int_0^t |M_s|^{p-2}d\langle M\rangle_s\right] \\ &\leq \tfrac{1}{2}p(p-1)E\left[\sup_{0\leq s\leq t}|M_s|^{p-2}\langle M\rangle_t\right] \end{aligned}$$

que pela desigualdade de Hölder

$$\leq \frac{1}{2}p(p-1)E\left[\sup_{0\leq s\leq t}|M_s|^{p-2}\right]^{\frac{p-2}{p}} E\left[\langle M\rangle_t^{p/2}\right]^{2/p}$$

e pela desigualdade de Doob

$$\leq \frac{1}{2}p(p-1)q^{p-2}E\left[|M_t|^p\right]^{\frac{p-2}{p}} E\left[\langle M\rangle_t^{p/2}\right]^{2/p} \leq C_pE\left[\langle M\rangle_t^{p/2}\right]$$

pois $2/p \leq 1$.

6. Hölder + Burkhölder: Para $p \geq 0$ e $A_t$, $t \in [s,T]$ processo $p$-integrável tal que

$$E\left[\left|\int_s^T A_u\, dW_u\right|^p\right] < \infty$$

tem-se

$$E\left[\left|\int_s^t \mathcal{A}_u \, dW_u\right|^p\right] \leq C_p \left|t-s\right|^{p/2-1} \int_s^t \left[\left|\mathcal{A}_u\right|^p\right] du \qquad \text{(H+Bh)}$$

Pois $\int_s^t \mathcal{A}_u \, dW_u$ é um martingale cuja variação quadrática é $\int_s^t \mid \mathcal{A}_u \mid^2 \, du$. Por (Bh),

$$E\left[\left|\int_s^t \mathcal{A}_u dW_u\right|^p\right] \leq C_p \, E\left[\left|\int_s^t \left|\mathcal{A}_u\right|^2 du\right|^{p/2}\right],$$

e por $(H0 \; 1, p/2, \; p/p-2)$,

$$\left(\int_s^t \mid \mathcal{A}_u \mid^2 du\right)^{p/2} \leq \mid t-s \mid^{p/2-1} \int_s^t \mid \mathcal{A}_u \mid^p \; du$$

Destas duas desigualdades se obtém $(H+Bh)$.

7. Desigualdade de Gronwall: Seja $x : I \rightarrow \mathbb{R}^d$ contínua e suponha que $\forall \, t \in I$, $x$ tenha derivada à direita $x'(t)$ e satisfaça

$$\mid x'(t) \mid \leq \text{ a } \mid x(t) \mid + b$$

com $a, b$ constantes positivas. Fixando $t_0 \in I$, tem-se para $t > t_0$

$$\left|x(t)\right| \leq \left|x(t_0)\right| e^{a(t-t_0)} + b\frac{e^{a(t-t_0)}-1}{a}. \qquad \text{(Gr)}$$

## 4.3 Existência e unicidade de soluções.

Mostramos nesta secção a existência e unicidade de soluções de (I) assumindo condições de Lipschitz sobre seus coeficientes. Antes de entrar em detalhes, vamos fazer um paralelo com as equações ordinárias.

A idéia básica do método é aplicar o teorema do ponto fixo de Banach ao operador construido por intermédio da equação integral associada à (I): Se $\xi_t$ é um processo, define-se um novo processo $(\mathcal{A}\xi)_t$ por

$$(\mathcal{A}\xi)_t = x + \int_s^t X(u, \xi_u) + \sum_{j-1}^{m} \int_s^t Y_j(u, \xi_u) dW_u^j \tag{3.1}$$

A relação $\xi \to \mathcal{A}\xi$ é para ser vista como um operador em algum espaço (de processos) convenientemente escolhido. Se se verificar que $\mathcal{A}$ admite um único ponto fixo, este será a única solução de (I) com condição inicial $(s, x)$.

No caso clássico das equações ordinárias, se os coeficientes satisfazem uma condição de Lipschitz local, o que se faz é restringir a equação diferencial a retângulos (cilindros) de segurança ao redor de $(s, x)$ (veja por exemplo [Hö] cap. II, § 2, relação (4)). Com esta restrição, $\mathcal{A}$ torna-se uma contração de um espaço métrico completo, se encaixando portanto nas condições do teorema do ponto fixo de Banach.

Para equações estocásticas, este tipo de restrição não é possível sem uma consideração criteriosa de tempos de parada para os processos envolvidos. A razão disto é que enquanto as soluções de equações ordinárias são curvas bem localizadas, a natureza difusa dos processos estocásticos faz com que seus caminhos aleatórios sejam parados em tempos diferentes se se pretende localizá-los a conjuntos limitados. Um exemplo disto pode ser visto pelo movimento Browniano:

**Exemplo 3.1:** Seja a equação $dx = dW$ como no exemplo 1.6 a). Definindo o operador $\mathcal{A}$ com em (3.1), com $(s, x) = (0, 0)$, temos que $(\mathcal{A}\xi)_t = W_t$ se $\xi_t = 0$. Tome $M > 0$ e seja:

$$\tau_B = \inf\{t > 0 = W_t \notin (-M, M)\}$$

o primeiro tempo de saída de $(-M,M)$. Então $\forall\, t>0$, $p(\tau_B \le t) > 0$ (veja [P-S] prop. 2.6), isto é, $W_t$ sai de qualquer intervalo ao redor da origem com probabilidade $> 0$ antes do tempo $t$, qualquer $t>0$. Portanto não é possível restringir $A$ é um operador sobre um espaço de processos definidos em intervalos do do tipo $[0,\varepsilon)$ e a valores limitados.

Por razões desta natureza, ao invés de se empregar o procedimento clássico de se restringir o domínio da equação, dá-se um tratamento global à teoria, localizando posteriormente os resultados por intermédio de tempos de parada.

No que segue, usaremos frequentemente a seguinte

**Notação:** $Y_0 = X$ e $dW^0 = dt$. Assim (I) se escreve de maneira mais compacta como

$$dx = \sum_{j=0}^{m} Y_j(t,x)dW^j \tag{3.2}$$

**Teorema 3.2:** Suponha que os coeficientes de (I) sejam Lipschitz globais definidos em $U = I \times \mathbb{R}^d$ com $I = [a,b) \subset [0,\infty)$.

Então para todo $(s,x)\varepsilon U$ existe uma única solução maximal $\xi_{s,t}(x)$ com condição inicial $(s,x)$. Seu tempo de explosão é $T(s,x) = b$.

Além do mais, $\xi_{s,t}(x)$ é um processo em $L^p_c[s,\beta]$ todo $\beta \in [s,b)$ e $p \ge 2$.

A unicidade aqui, é a menos de indistinguibilidade.

<u>Dem:</u> Fixe $\beta \in [s,b)$ e um tempo de parada predizível $T$.

Tome processos $\xi,\eta \in L_c[s,\beta]$ tais que $\xi^T,\eta^T \in L^p_c[s,\beta], p \ge 2$ e defina

$$(A^T\xi)_t = x + \sum_{j=0}^{m} \int_s^{t\wedge T} Y_j(u,\xi_u)dW^j_u$$

O primeiro passo consiste em obter uma majoração para $\rho_{t,p}(\mathcal{A}^T\xi, \mathcal{A}^T\eta)$ onde $\rho_{t,p}$ é à distância em $L^p_c[s,t]$, $t \in [s,\beta]$.

Por A.2),

$$\left|\mathcal{A}^T\xi - \mathcal{A}^T\eta\right|^p \leq C\sum_{j=0}^{m}\left|\int_s^{t\wedge T}\Big(Y_j(u,\xi_u) - Y_j(u,\eta_u)\Big)dW_u^j\right|^p$$

e portanto

$$\rho_{t,p}(\mathcal{A}^T\xi, \mathcal{A}^T\eta) \leq C\sum_{j=0}^{m}E\left[\sup_{s\leq u\leq t}\left|\int_s^{u\wedge T}\Big(Y_j(v,\xi_u) - Y_j(v_1,\eta_v)\Big)dW_v^j\right|^p\right] \tag{3.3}$$

Pela condição de Lipschitz e por $(Hol,p,q)$,

$$\sup_{s\leq u\leq t}\left|\int_s^{u\wedge T}(X(v,\xi_v) - X(v,\eta_v))dv\right|^p \leq C\int_s^{t\wedge T}\left|\xi_v - \eta_v\right|^p dv$$

e portanto, o termo correspondente à $j=0$ em (3.3) é majorado por

$$CE\left[\int_s^{t\wedge T}\left|\xi_v - \eta_v\right|^p dv\right] \tag{$*$}$$

No que diz respeito aos outros termos, seja

$M_u^j = \int_s^{t\wedge T}(Y_j(v,\xi_v) - Y_j(v,\eta_v))dv$. Então $M_u^j$ é um martingale e portanto pela desigualdade de Doob, $E[\sup_{s\leq u\leq t} \mid M_u^j \mid^p] \leq CE[\mid M_t^j \mid^p]$. Daí que pela desigualdade de Lipschitz e por $(Ho+Bh)$, cada um dos termos correspondentes é $j \geq 1$ é majorado por (*). Temos então, por (C), que

$$\rho_{t,p}\left(\mathcal{A}^T\xi, \mathcal{A}^T\eta\right) \leq CE\left[\int_s^{t\wedge T}|\xi_u - \eta_u|^p du \tag{3.4}$$

Passemos agora à verificação de que $\mathcal{A}$, definido em (3.1) define numa contração. No caso específico em que $T = \infty$ (e portanto $\xi, \eta \,\varepsilon\, L^p_c\,[s,\beta]$), (3.4) se escreve

$$\rho_{t,p}\left(\mathcal{A}\xi, \mathcal{A}\eta\right) \leq CE\left[\int_s^t\left|\xi_u - \eta_u\right|^p du\right] \tag{3.5}$$

O que mostra em particular, que $\mathcal{A}(L^p_c) C L^p_c$. Aplicando o teorema de Fubini no 2º membro de (3.5),

$$\rho_{t,p}\left(\mathcal{A}\xi, \mathcal{A}\eta\right) \leq\ C\int_s^t\ \rho_{u,p}(\xi,\eta)du$$

Iterando-se esta desigualdade, temos:

$$\begin{aligned}\rho_{t,p}(\mathcal{A}^n\xi,\ \mathcal{A}^n\eta)\ &\leq\ C^n\rho_{t,p}(\xi,\eta)\int_s^t\int_s^{t1}\cdots\int_s^{tn-1}dt_{n-1}\cdots dt_1 du\\ &\leq \frac{C^n(\beta-S)^n}{n!}\ \rho_{\beta,p}\ (\xi,\eta)\end{aligned}$$

**Por esta desigualdade, vê-se que $\mathcal{A}^n$ é uma contração de $L^p_c\ [s,\beta]$ todo $\beta \in [s,b]$ e $p \geq 2$ algum $n$. Pelo teorema do ponto fixo de Banach (veja corolário II 1.2 em [Hõ]), $\mathcal{A}$ tem um único ponto fixo em $L^p_c[s,\beta)$ e portanto (I) tem uma única solução em $L^p_c[s,\beta)$ todo $p \geq 2$.**

É claro que as soluções nos, diferentes $L_c^p$ coincidem pois $L_c^p \subset L_c^q$ se $p \geq q$. Isto no entanto não é suficiente para mostrar a unicidade das soluções.

É necessário verificar que não existem soluções que não sejam quadraticamente integráveis. Para isto, seja $\xi$ a única solução em $L_c^2[s,\beta]$ e suponha que $\eta \in L_c[s,\beta]$ seja outra solução. Para cada inteiro positivo $n$, defina

$$T_n = \inf\{t > s : |\xi_t| \geq n \quad \text{ou} \quad |\eta_t| \geq n\}$$

Colocando $T_n = \infty$ se o conjunto é vazio. Então $T_n$ é tempo de parada pois $\xi$ e $\eta$ são processos adaptados.

A desigualdade (3.4) se aplica à $\xi$ e $\eta$ pois $\xi^{T_n}$ e $\eta^{T_n}$ são limitados e portanto quadráticamente integráveis.

Usando a notação $\tau_{t,n} = E\Big[\sup\limits_{s \leq u \leq t} |\xi^{T_n} - \eta^{T_n}|^2\Big]$, (3.4) (e o fato que $\xi$ e $\eta$ são soluções) fornece

$$\tau_{t,n} \leq C \int_s^t \tau_{u,n}\, du$$

Por $(Gr)$, $\tau_{t,n} \leq e^{c(t-s)}\tau_s$ e como $\tau_s = 0, \tau_{t,n} = 0$ e daí que $\xi^{T_n} = \eta^{T_n}$. Como $T_n \to \infty$, $\xi = \eta$ e daí a unicidade da solução em $L_c[s,\beta]$.

Por fim, se $s \leq \beta_1 \leq \beta_2 < b$ a restrição à $[s,\beta_1]$ da solução definida em $[s,\beta_2]$ também é solução o que mostra, por extensão, a existência de uma única solução em $[s,b)$. Seu tempo de explosão é $T(s,x) = b$ e ela é evidentemente maximal.

Para relaxar as condições do teorema acima e obter existência e unicidade para equações localmente Lipschitz, usamos o

**Lema 3.3:** Seja $U \subset [0,\infty) \times \mathbb{R}^d$ aberto e $f : U \to \mathbb{R}^d$ contínua e localmente Lipschitz em relação à $x$. Então, se $K \subset U$ é compacto, existe $\tilde{f} = [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ com $\tilde{f} = f$ em $K$ e $\tilde{f}$ continua e (globalmente) Lipschitz. Se $f$ de classe $C^k$ o mesmo ocorre com $\tilde{f}$.

<u>Dem:</u> Escolha $K'$ compacto com $K \subset \text{int } K' \subset K' \subset U$ e $\phi = \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ $C^\infty$, com suporte em $K'$ e tal que $\phi \equiv 1$ em $K$. Então

$$\tilde{f}(t,x) = \begin{cases} \phi(t,x)f(f,x) & \text{se } (t,x) \in U \\ 0 & \text{se } (t,x) \notin U \end{cases}$$

Satisfaz o que se pede.

**Teorema 3.4:** Suponha que os coeficientes de (I) sejam definidos em $U \subset [0,\infty) \times \mathbb{R}^d$ e sejam contínuos e localmente Lipschitz em relação à $x$.

Então para todo $(s,x) \in U$ existe uma única solução maximal $\xi_{s,t}(x)$ com condição inicial $(s,x)$.

Seja $T(s,x)$ o seu tempo de explosão. Então se $T(s,x,\omega) < \infty$, $\xi_{s,t}(x,\omega)$ se acumula na fronteira de $U$ quando $t \to T(s,x,\omega)$.

<u>Dem:</u> Seja $K_n \subset U$ uma sequência de compactos com interior não vazio tal que $K_n \subset \text{int } K_{n+1}$ e $\bigcup_n K_n = U$. Para cada $n$, seja $Y_j^n, j = 0, \cdots, m$ a extensão globalmente Lipschitz que coincide com $Y_j$ em $K_n$, garantida pelo lema anterior.

Defina

$$dx = \sum_{j=o}^{m} Y_j^n\,(t,x)\,dW_j \tag{$I_n$}$$

e denote por $\xi^n_{s,t}(x)$ suas soluções iniciadas em $(s,x)$. Defina

$$T_n(s,x) = \begin{cases} \inf\left\{t \geq s : \left(t, \xi^n_{s,t}(x)\right) \notin K_n\right\} \\ \infty \quad \text{se } \left\{\cdots\right\} = \phi \end{cases}$$

Cada $T_n(s,x)$ é um $(\mathcal{F}_{s,t})$ – tempo de parada e com o as soluções $\xi^n_{s,t}(x)$ são contínuas e $K_n \subset \text{int } K_{n+1}$, tem-se que $T_n(s,x,\omega) < T_{n+1}(s,x,\omega)$ se $T_n(s,x,\omega) < \infty$. Além do mais, se $t < T_n(s,x)$ então

$$\xi^n_{s,t}(x) = x + \sum_{j=0}^{m} \int_s^t Y_j\left(u, \xi^n_{s,u}(x)\right) dW^j_u \tag{3.6}$$

pois no interior de $K_n, Y^n_j = Y_j$. Seja $T(s,x) = \lim_{n\to\infty} T_n(s,x) = \sup_m T_m(s,x)$. Então $T(s,x)$ é um tempo de parada predizível pois no caso em que $T(s,x)(\omega) < \infty$, $T_n(s,x)(\omega) < T_{n+1}(s,x)(\omega)$ e portanto $T_n(s,x)(\omega) < T(s,x)(\omega)$ todo $n$, isto é, os tempos de parada $T_n$ anunciam $T$.

Defina agora $\xi_{s,t}(x) = \xi^n_{s,t}(x)$ se $s \leq t < T_n(s,x)$. Então $\xi_{s,t}(x)$ é bem definido pois $\xi^n_{s,t}(x) = \xi^{n+1}_{s,t}(x)$, pela unicidade da solução de $(I_{n+1})$. Por (3.6), $\xi_{s,t}(x)$ é uma solução local de (I) com tempo de explosão $T(s,x)$.

Se $T(s,x)(\omega) < \infty$ então $T_n(s,x)(\omega) < \infty$ todo $n$ e $\xi_{s,T_n(s,x)}(x)$ $(\omega)$ está na fronteira de $K_n$ e como $UK_n = U$, tem-se que os pontos de acumulação de $\xi_{s,t}(x)(\omega)$ se situam na fronteira de $U$. Esta observação mostra a última parte do enunciado e mostra também que $\xi_{s,t}(x)$ é não prolongável, isto é, é solução maximal.

Por fim, a unicidade da solução maximal segue do fato que qualquer solução maximal quando restrita à $K_n$, isto é, para valores de $t < T_n(s,x)$, coincide com $\xi^n_{s,t}(x)$ e portanto com $\xi_{s,t}(x)$.

## 4.4 Dependência contínua em relação às condições iniciais.

Vamos estudar a dependência contínua das soluções de (I) em relação às condições iniciais, ainda sob a hipótese básica de que os coeficientes satisfazem uma condição de Lipschitz. A inspiração continua sendo a teoria clássica das equações ordinárias. Nesse caso, o que se mostra, sob condições de Lipschitz é que as soluções de $\dot{x} = f(t,x)$ dependem não apenas continuamente das condições iniciais $(s,x)$ mas satisfazem também uma condição de Lipschitz em relação à estas variáveis (na verdade, as soluções são diferenciais em relação à $s$). O que se obtém para equações estocásticas, difere do caso clássico nos seguintes pontos:

Em primeiro lugar, a unicidade das soluções maximais, que foi mostrada no § anterior é a menos de indistinguibilidade. Como o conjunto das condições iniciais é não enumerável, é necessário escolher as soluções com as diferentes condisões iniciais de maneira consistente, para evitar, por exemplo, que o conjunto $\Omega'$ em que $\xi_{s,t}(x,\omega)$ esteja definido para todo $s,t,x$ e $\omega \in \Omega'$ seja de medida nula. Em outras palavras, é necessário tomar uma modificação das soluções $\xi_{s,t}(x,\omega)$ de tal forma que $\xi_{s,t}(x,\cdot)$ esteja definido num conjunto de probabilidade total que para toda condição inicial $(s,x)$ e $t \geq s$. Isto será feito com o auxílio do teorema de Kolmogorov sobre modificações contínuas de processos estocásticos e campos aleatórios. Neste sentido, obtemos resultados do tipo: para $\omega \in \Omega'$ com $P(\Omega') = 1$, a aplicação $(s,t,x) \to \xi_{s,t}(x,\omega)$ está definida e é contínua.

Outra diferença em relação ao caso clássico esta no fato de que apesar de ser possível obter continuidade Lipschitz de $\xi_{s,t}(x)$ em relação à $x$ não se tem diferenciabilidade ou mesmo continuidade Lipschitz em relação à $s$ e $t$ (as soluções do exemplo 3.1 a) são dadas pelo movimento Browniano, cujas trajetórias não são Lipschitz). Tem-se no entanto, em relação à $s$ e $t$, continuidade $\beta$-Hölder com $\beta < 1/2$ arbitrário.

Como no caso do teorema de existência e unicidade, consideramos primeiramente o caso Lipschitz global (L.g.), localizando posteriormente os resultados via tempos de parada.

**Lema 4.1:** Suponha que (I) seja L.g.. Então para todo $p \in \mathbb{R}$ e $\beta \in [a,b)$,

$$E\left[\left(\varepsilon + \left|\xi_{s,t}(x) - \xi_{s,t}(y)\right|^2\right)^p\right] \leq C_p\left(E + \left|x-y\right|^2\right)^p \tag{4.1}$$

para todo $s \leq t \leq \beta$, $x, y \in \mathbb{R}^d$ e $\varepsilon \geq 0$. $C_p$ não depende de $\varepsilon$.

Dem: É suficiente considerem o caso em que $\varepsilon > 0$.

Para $\varepsilon = 0$, toma-se limite com $\varepsilon \to 0$, que é possível pois $C_p$ não depende de $\varepsilon$.

Como $\xi_{s,t}(x)$ e $\xi_{s,t}(y)$ são soluções de (I), estes processos são semi-martingales continuos em relação à filtração $(\mathcal{F}_{s,t})_{t \leq s}$. Portanto $t \to \eta_t = \xi_{s,t}(x) - \xi_{s,t}(y)$ também é um semi-martingale. Fixando $p \in \mathbb{R}$ e $\varepsilon > 0$ defina:

$$f(x) = \varepsilon + |x|^2 \qquad \text{e} \qquad F(x) = \left(\varepsilon + |x|^2\right)^p$$

Então $F : \mathbb{R}^d \to \mathbb{R}$ é $C^\infty$ e portanto $F(\eta_t)$ é um processo ao qual se aplica a fórmula de Itô.

Temos:

$$F(\eta_t) = F(\eta_s) + \int_s^t A_u du + \sum_{j=1}^m \int_s^t B_u^j dW_u^j$$

Com $A, B^1 \ldots, B^m$ processos. Como a esperança de uma integral estocástica se anula e $\eta_s = x - y$, tomando esperanças na igualdade acima fica

$$E\left[F(\eta_t)\right] = F(x-y) + E\left[\int_s^t A_u du\right] \tag{4.2}$$

Calcula-se $A_u$ por intermédio da fórmula de Itô.

Para isto, escreva os coeficientes de (I) em coordenadas como $X = (X^1, \dots, X^d)$, $Y_j = (Y_j^1, \dots, Y_j^d), j = 1, \dots, m$.

Então

$$A_u = \sum_{i=1}^{d} \frac{\partial F}{\partial x^i}(\eta_u)\alpha_u^i + \frac{1}{2}\sum_{i,k}^{d} \frac{\partial^2 F}{\partial x^i \partial x^k}(\eta_u)\sum_{j=1}^{m}\left(\beta_j^i\, \beta_j^k\right)_u \tag{4.3}$$

onde

$$\alpha_u^i = X^i\left(u, \xi_{s,u}(x)\right) - X^i\left(u, \xi_{s,u}(y)\right)$$

$$\beta_j^i = Y_j^i\left(u, \xi_{s,u}(x)\right) - Y_j^i\left(u, \xi_{s,u}(y)\right)$$

Da expressão de $F$, $\dfrac{\partial F}{\partial x^i}(\eta) = 2p\, f(\eta)^{p-1}\eta^i$ e $\dfrac{\partial^2 F}{\partial x^i \partial x^j}(\eta) = 4p\,(p-1)\, f(\eta)^{p-2}\eta^i\eta^j$, onde $\eta^i, i = 1, \dots, d$ são as coordenadas de $\eta$ em $\mathbb{R}^d$.

Para majorar $|A_u|$, use as desigualdades

a) Por Lipschitz

$$|\alpha^i| \leq C|\eta| \quad , \quad |\beta_j^i| \leq C\,|\eta|$$

b) Por A.4

$$|\eta| \leq C\, f(\eta)^{1/2} \quad ; \quad |\eta^i| \leq C\, f(\eta)^{1/2}$$

Por estas desigualdade, ao majorar os termos da 1ª soma em (4.3), $|\alpha^i|$ colabora com $f(\eta)^{1/2}$ e $|\partial F/\partial x^i \partial x^j|$ com $f(\eta)^{p-1/2}$. Estes termos são portanto majorados por $C\, f(\eta)^p$. Na segunda soma, $|\partial^2 F/\partial x^i \partial x^k|$ colabora com $f(\eta)^{p-1}$ e $\beta_j^i$ com $f(\eta)^{1/2}$. De onde se tira que

$$|A_u| \leq C\, f(\eta_u)^p\,.$$

Voltando a (4.2), concluimos que

$$E\Big[F(\eta_t)\Big] \leq F(x-y) + C\int_s^t E\Big[F(\eta_u)\Big]du$$

como $C$ não dependendo de $\varepsilon$. Por $(Gr)$

$$E\Big[F(\eta_t)\Big] \leq \Big(e^{C(t-s)} + \frac{e^{C(t-s)}+1}{C}\Big)F(x-y)$$

$\leq Cf(x-y)$ o que mostra (4.1).

A desigualdade do lema anterior permite mostrar através do teorema de Kolmogorov a existência de modificação contínua de $(x,\omega) \to \xi_{s,t}(x,\omega)$, $s$ e $t$ fixados.

**Teorema 4.2:** Com (I) Lipschitz global e $s$ se $t$ fixados, existe $\Omega' \subset \Omega$ com $P(\Omega') = 1$, tal que para todo $\omega \in \Omega'$

$$\xi_{s,t}(x) \quad \text{e} \quad \int_s^t Y_j\Big(u, \xi_{s,t}(x)\Big)dW^j \ , \ j = 0, \ldots, m$$

são contínuas como função de $x$. A continuidade é $\beta$-Hölder para qualquer $\beta < 1$. Tem-se além do mais

$$\xi_{s,t}(x) = x + \sum_{j=0}^{m}\int_0^t Y_j\Big(u, \xi_{s,u}(x)\Big)dW^j \tag{4.4}$$

todo $\omega \in \Omega'$, $x \in \mathbb{R}^d$.

obs: A sutileza desta última afirmação está no fato de que cada $x \in I\!R^d$ (4.1) vale q.s pois $\xi_{s,t}(x)$ é modificação de solução de (I). No entanto como $I\!R^d$ é não enumerável, em princípio pode não se ter a igualdade para todo $x$ e q.t. $\omega$.

Dem: A desigualdade (4.1) com $\varepsilon = 0$, garante que o campo aleatório $(x,\omega) \to \xi_{s,t}(x,\omega)$ satisfaz as condições do teorema de Kolmolgorov com expoentes $\alpha_i = \gamma = p$. Como resultado, temos que existe uma modificação contínua de $\xi_{s,t}(x,\omega)$, denotada da mesma maneira, tal que para $\beta > 0$ com $\beta < \frac{p-d}{p}$ se tenha

$$\left|\xi_{s,t}(\omega)(x) - \xi_{s,t}(\omega)(y)\right| \leq C_\omega \sum_{j=1}^{d} \left|x_j - y_j\right|^p$$

se $x, y \in I\!R^d$ são suficientemente próximos (dependendo de $\omega$). Em outras palavras, $\xi_{s,t}(x,\omega)$ admite uma modificação $\beta$-Hölder contínua. Como $p \geq 2$ é arbitrário, $\beta$ com $0 < \beta < 1$ também é arbitrário.

Quanto as integrais se $j \geq 1$ $(H + Bh)$ fornece, para $p \geq 2$

$$E\left[\left|\int_s^t \Big(Y_j(u, \xi_{s,u}(x)\big) - Y_j\big(u, \xi_{s,u}(y)\Big) dW_u^j\right|^p\right] \leq \tag{4.5}$$

$$\leq C\left|t-s\right|^{p/2-1} \int_s^t E\left[\left|Y_j\big(u, \xi_{s,u}(x)\big) - Y_j\big(u, \xi_{s,u}(y)\big)\right|^p\right] du ,$$

e se $j = 0$, usando (Ho 1, p,q.) como na demonstração de (H+Bh) no § 2, obtém-se a mesma densigualdade. Aplicando a desigualdade de Lipschitz no segundo membro de (4.5) e posteriormente (4.1) com $\varepsilon = 0$, o primeiro membro de (4.5) fica majorado por

$$C_p|t-s|^p \ |x-y|^p \leq C|x-y|^p$$

De onde se ve que as integrais satisfazem a condição de Kolmogorov com expoentes $\gamma = \alpha_i = p$, admitindo portanto modificação $\beta$-Hölder contínua com $0 < \beta < 1$.

Como realizamos um número finito de modificações, existe de fato $\Omega'$ como no enunciado.

Para ver a segunda afirmação use A.2) juntamente com as majorações obtidas acima, para concluir que:

$$D_t(x) = \xi_{s,t}(x) - x - \sum_{j=0}^{m} \int_s^t Y_j\big(u, \xi_{s,u}(x)\big) dW_u^j$$

admite uma modificação contínua. O teorema de Kolmogorov garante então a existência de $\Omega'$ com $P(\Omega') = 1$ t.q. $D_t(x)(\omega) = 0$ todo $\omega \in \Omega'$.

**Colrolário 4.3:** Tome $s_0 < s < t$ e escolhas contínuas, garantidas pelo teorema anterior, de $\xi_{s_0,t}(x)$, $\xi_{s_0,s}(x)$ e $\xi_{s,t}(x)$. Então

$$\xi_{s_0,t}(x) = \xi_{s,t}\left(\xi_{s_0,s}(x)\right)$$

Dem: Defina

$$\tilde{\xi}_{s,t}(x) = \begin{cases} \xi_{s_0,t}(x) & \text{se} \quad t \leq s \\ \\ \xi_{s,t}(\xi_{s_0 s}(x)) & \text{se} \quad t > s \end{cases}$$

Então $t \to \tilde{\xi}_{s_0,t}(x)$ é um processo $(\mathcal{F}_{s_0,t})$ – mensurável e contínuo.
Tem-se $\tilde{\xi}_{s_0,s_0}(x) = x$ e $\tilde{\xi}_{s_0,t}(x)$ satisfaz (I) para $t \leq s$.
Para $t > s$, substituindo $\xi_{s_0,s}(x)$ no lugar de $x$ em (4.4), fica

$$\begin{aligned}\tilde{\xi}_{s_0,s}(x) &= \xi_{s,t}\left(\xi_{s_0,s}(x)\right) = \xi_{s_0,s}(x) + \textstyle\sum_{j=0}^{m}\int_s^t Y_j\left(u, \tilde{\xi}_{s_0,u}(x)\right)dW_u^j \\ &= x + \textstyle\sum_{j=0}^{m}\int_s^t Y_j\left(u, \tilde{\xi}_{s_0,u}(x)\right)dW_u^j\end{aligned}$$

Portanto $\tilde{\xi}_{s_0,t}(x)$ é solução com condições inicial $(s_0, x)$, e pela unicidade se tira o corolário.

O próximo objetivo é mostrar a continuidade das soluções em relação às variáveis $s, t$ e $x$ conjuntamente, mostrando a existência de uma modificação contínua do campo aleatório $(s, t, x) \to \xi_{s,t}(x)$. Desta forma uma extenção do teorema 4.2 será obtida. Para essa extensão, será utilizado o teorema 4.2 por intermediário do corolário 4.3.

**Desigualdades 4.4:** Para as desigualdades a seguir, assumimos que (I) é L.g. com domínio $[a, b) \times \mathbb{R}^d$.

Os elementos de $[a, b)$, $t, s, s'$ etc..., que aparecem estarão sempre em algum intervalo compacto $[a, \beta], \beta < b$, fixado. Os elementos de $\mathbb{R}^d$ são arbitrários. As soluções $\xi_{s,t}(x)$ são escolhidas como no teorema 4.2.

1. para $p \in \mathbb{R}$ e $\varepsilon > 0$,

$$E\left[\left(\varepsilon + \left|\xi_{s,t}(x)\right|^2\right)^p\right] \leq C_\varepsilon^p\left(\varepsilon + |x|^2\right)^p \tag{4.6}$$

obs: A constante $C_\varepsilon^p$ depende de $\varepsilon$. Veja comentário no decorrer da demonstração.

2. Para $p \geq 2$ e $s < t < t'$,

$$E\left[\left|\int_t^{t'} Y_j\left(u, \xi_{s,u}(x)\right)dW_u^j\right|^p\right] \leq C|t' - t|^{p/2}\left(1 + |x|^p\right) \tag{4.7}$$

para $j = 0, \ldots, m$.

3. Para $p \geq 2$,

$$E\left[\left|\xi_{s,t}(x) - y\right|^p\right] \leq C\Big(|x-y|^p + |t-s|^{p/2}(1+|x|^p)\Big) \tag{4.8}$$

4. Para $p \geq 2$ e $s < s' < t$,

$$E\left[\left|\xi_{s,t}(x) - \xi_{s',t'},(y)\right|^p\right] \leq C\Big(|x-y|^p + |s-s'|^{p/2}(1+|x|^p)\Big) \tag{4.9}$$

5. Para $p \geq 2$ e $s < s' < t$,

$$E\left[\left|\int_{s'}^{t}\Big(Y_j\big(u, \xi_{s,u}(x)\big) - Y_j\big(u, \xi_{s',u}(y)\big)\,dW_u^j\right|^p\right] \tag{4.10}$$

$$\leq C\Big(|x-y|^p + |s-s'|^{p/2}(1+|x|^p)\Big)$$

para $j = 0, \ldots, m$.

6. Para $p \geq 2$ e $\max(s, s') < \min(t, t')$

$$E\left[\left|\xi_{s,t}(x) - \xi_{s',t'}(y)\right|^p\right] \leq \tag{4.11}$$

$$\leq C\Big(|x-y|^p + (1+|y|^p+|x|^p)(|t-t'|^{p/2} + |s-s'|^{p/2})\Big)$$

obs: A partir desta última desigualdade e do teorema de Kolmogorov, a existência de modificação contínua de $(s,t,x) \to \xi_{s,t}(x)$ é quase que imediata. As outras desigualdades são necessárias na prova de (4.11).

Dem:

1. É semelhante à do lema 4.1. Sejam $f(x) = \varepsilon + |x|^2$ e $F(x) = f(x)^p$ e aplique a fórmula de Itô para o $(\mathcal{F}_{s,t})$ – semi-martingale $F(\xi_{s,t}(x)), s, x$ fixados. Usando a notação $\xi_t = \xi_{s,t}(x)$, temos

$$d\Big(F(\xi_t)\Big) = \mathcal{A}_t\, dt + \sum_{j=1}^{m} B_t^j\, dW_t^j$$

com

$$\mathcal{A}_t = \sum_{i=1}^{d} X^i(t,\xi_t)\frac{\partial F}{\partial x^i}(\xi_t) + \frac{1}{2}\sum_{i,k=1}^{d}\frac{\partial^2 F}{\partial x^i \partial x^k}(\xi_t)\sum_{j=1}^{m} Y_j^i(t,\xi_t)Y_j^k(t,\xi_t)\ ,$$

onde os superindices indicam coordenadas em $\mathbb{R}^d$. Como a esperança de uma integral estocástica (no sentido de Itô) se anula,

$$E\Big[F(\xi_t)\Big] = F(x) + \int_s^t E\Big[\mathcal{A}_u\Big]\, du\ .$$

Para majorar $E[\mathcal{A}_u]$, use (cL) para encontrar uma constante $C^\varepsilon$ t.q. $|Y_j^i(t,x)| \leq C^\varepsilon f(x)^{1/2} j = 0, 1, \cdots, m$ (é aqui que surge a diferença com o lema 4.1 lá $|Y_j^i(t,x) - Y_j^i(t,y)|$, é majorado por algo que independe de $\varepsilon$). Fazendo majorações como no lema 4.1, concluimos que $E[\mathcal{A}_t] \leq C_p^\varepsilon\, E[F(\xi_t)]$ e daí que

$$E\Big[F(\xi_t)\Big] \leq F(x) + \int_s^t E\Big[F(\xi_u)\Big]du$$

de onde sai (4.6) por uma aplicação da desigualdade de Gronwall (Gr).

2. Ponha $\mathcal{A}_u = Y_j(u, \xi_{s,u}(x))$. Por (H+Bh) se $j \geq 1$ e $(H_0 1, p, q)$ se $j = 0$,

$$E\left[\left|\int_t^{t'} \mathcal{A}_u dW_u^j\right|^p\right] \;\leq\; C_p|t' - t|^{p/2-1}\int_t^{t'} E\left[\left|\mathcal{A}_u\right|^p\right]du$$

(compare com (4.5)). Por (cL) com $\varepsilon = 1$, $|Y_j(u, y)| \leq C(1 + |y|)$ e por A3)

$$|\mathcal{A}_u|^p \leq C\Big(1 + |\xi_{s,u}(x)|^2\Big)^{p/2}$$

e por (4.6),

$$E\Big[|\mathcal{A}_u|^p\Big] \leq C\Big(1 + |x|^2\Big)^{p/2} .$$

Daí que

$$\begin{aligned} E\left[\left|\int_t^{t'} \mathcal{A}_u dW_u^j\right|^p\right] &\leq C_p|t - t'|^{p/2-1}\int_t^{t'} (1 + |x|^2)^{p/2}du \\ &\leq C_p|t - t'|^{p/2}\Big(1 + |x|\Big)^p \qquad \text{por (A.1)} , \end{aligned}$$

isto é, tem-se (4.7).

3. Temos

$$\xi_{s,t}(x) - y = x - y + \sum_{j=0}^m \int_s^t Y_j\Big(u, \xi_{s,u}(x)\Big)dW_u^j$$

e portanto por (A.2),

$$E\left[\left|\xi_{s,t}(x)-y\right|^p\right] \leq \; C\Big(|x-y|^p+\sum_{j=0}^{m}\Big|\int_s^t Y_j(u,\xi_{s,u}(x))dW^j\Big|^p\Big)$$

$$\leq C\Big(|x-y|^p+|t-s|^{p/2}(1+|x|^p)\Big)$$

por (4.7) e (C) concluindo a demonstração de (4.8).

4. Com $z, z' \varepsilon I\!\!R^d$ arbitrários, $\xi_{s',t}(z)$ e $\xi_{s,s'}(z')$ são variáveis aleatórias independentes já que $\xi_{s',t}(z)$ é $\mathcal{F}_{s',t}$ – mensurável e estas $\sigma$-algebras são independentes. Portanto, levando em conta que $\xi_{s,t}(x) = \xi_{s',t}(\xi_{s,s'}(x))$ (corolário 4.3), tem-se a desintegração regular

$$E\left[\left|\xi_{s,t}(x)-\xi_{s',t}(y)\right|^p\right] = \int E\left[\left|\xi_{s',t}(z)-\xi_{s',t}(y)\right|^p\right] P\Big(\xi_{s,s'}(x)\in dz\Big)$$

e por (4.1), com $\varepsilon = 0$,

$$\leq C\int |x-y|^p \; P\Big(\xi_{s,s'}(x)\in dz\Big)$$

e pela definição de $P(\xi_{s,s'}(x)\in dz)$

$$\leq CE\left[\left|\xi_{s,s'}(x)-y\right|^p\right]$$

e por (4.8),

$$\leq C\Big(|x-y|^p+|s'-s|^{p/2}(1+|x|^p)\Big)$$

que é (4.10).

5. Por (H+Bh) se $j \geq 1$ e (Ho 1, p,q) se $j = 0$, o lado esquerdo de (4.10) é majorado por

$$C \int_{s'}^{t} E\left[\left|Y_j\Big(u, \xi_{s,u}(x)\Big) - Y_j\Big(u, \xi_{s',u}(y)\right|^p\right] du$$

que e pela condição de Lipschitz

$$\leq C \int_{s'}^{t} E\left[\left|\xi_{s,u}(x) - \xi_{s',u}(y)\right|^p\right] du$$

que por (4.9)

$$\leq C\Big(|x-y|^p + |s-s'|^{p/2}(1+|x|^p)\Big)$$

6. Vamos supor $s < s' < t < t'$. Os outros casos são análogos.
Pelo corolário 4.3, $\xi_{s,t}(x) = \xi_{s',t}(\xi_{s,s'}(x))$, portanto

$$\xi_{s,t}(x) = \xi_{s,s'}(x) + \sum_{j=0}^{m} \int_{s'}^{t} Y_j\Big(u, \xi_{s,u}(x)\Big) dW_u^j$$

e como

$$\xi_{s',t'}(y) = y + \sum_{j=0}^{m} \int_{s'}^{t} Y_i\Big(u, \xi_{s',u}(y)\Big) dW_u^j + \sum_{j=0}^{m} \int_{t}^{t'} Y_j\Big(u, \xi_{s',u}(y)\Big) dW_u^j$$

por (A.2), temos que

$$\left|\xi_{st}(x) - \xi_{s',t}(y)\right|^p \leq C\Big(\left|\xi_{s,s'}(x) - y\right|^p + \sum_{j=0}^{m} \left|\int_{t}^{t'} Y_j\Big(u, \xi_{s',u}(y)\Big) dW_u^j\right|^p$$

$$+ \sum_{j=0}^{m} \left|\int_{s^1}^{t} \Big(Y_j\Big(u, \xi_{s,x}(x)\Big) - Y_j\Big(u, \xi_{s',u}(y)\Big)\Big) dW_u^j\right|^p .$$

Usando, respectivamente (4.8), (4.7) e (4.10) no 1º, 2º e 3º termos do segundo membro, temos

$$E\left[\left|\xi_{s,t}(x)-\xi_{s',t}(y)\right|^p\right] \leq C\Big(|x-y|^p+|s-s'|^{p/2}(1+|x|^p\Big)+|t-t'|^{1/2}(1+|y|^p)\Big)$$

$$\leq C\Big(|x-y|^p+(1+|x|^p+|y|^p\Big)\Big(|s-s'|^{p/2}+|t-t'|^{1/2}\Big)$$

que é (4.11).

**Teorema 4.5:** Suponha que (I) seja L.g. com domínio de definição $[a,b)\times I\!R^d$ e denote por $\xi_{s,t}(x)$ suas soluções maximais.
Então existem modificações contínuas dos campos aleatórios

$$(s,t,x) \longrightarrow \xi_{s,t}(x)$$

$$(s,t,x) \longrightarrow \int_s^t Y_j\big(u,\xi_{s,u}(x)\big)dW_u^j \quad j=0,1,\dots,m$$

tal que

$$\xi_{s,t}(x)=x+\sum_{j=0}^{m}\int_s^t Y_j\big(u,\xi_{s,u}(x)\big)dW_u^j \quad \text{todo } (s,t,x) \text{ q.s.}$$

O domínio de definição dos campos aleatórios é

$$\Big\{(s,t,x)\in[a,b)\times[a,b)\times I\!R^d : s\leq t\Big\}.$$

Além do mais, $(s,t,x) \to \xi_{s,t}(x)$ é $(\beta,\beta,\alpha)$ - Hölder continua todo $\beta < \frac{1}{2}$ e $\alpha < 1$. Para esta modificação, tem-se $\xi_{s,t}(x) = \xi_{u,t}(\xi_{s,u}(x))$, todo $x$ e $s < u < t$ q.s..

Dem: É suficiente tomar soluções definidas em intervalos compactos $[a,\beta] \subset [a,b)$. Tomando uma sequência $\beta_u \to b$, teremos uma quantidade enumerável de modificações o que acarreta uma modificação contínua.
Para soluções definidas em $[a,\beta]$, tomamos $(s,t,x) \to \xi_{s,t}(x)$ com $x$ restrito é conjuntos limitados, isto é, com o campo aleatório definido um $\{(s,t,x) : s \leq t \text{ e } |x| < M\}$. Então (4.11) se $s \leq t$ e (4.8) se $s = t$ garante que para $p \geq 2$

$$E\left[\left|\xi_{s,t}(x) - \xi_{s',t'}(y)\right|^p\right] \leq C\left[|x-y|^p + |t-t'|^{p/2} + |s-s'|^{p/2}\right]$$

Daí que o campo aleatório $(s,t,x) \to \xi_{s,t}(x)$ satisfaz a condição de Kolmogorov com expoentes $\gamma = p, \alpha_1 = \alpha_2 = p/2$ e $\alpha_3 = p$, qualquer $p \geq 2$. Portanto, $\xi_{s,t}(x)$ com $x$ restrito a conjuntos limitados de $\mathbb{R}^d$ admite uma modificação $(\beta;\beta,\alpha)$-Hölder continua com $\beta < (p/2-d)/p$, $\alpha < (p-2d)/p$ e $p \geq 2$ arbitrários. Pode-se tomar então qualquer $\beta < \frac{1}{2}$ e $\alpha < 1$, pois $(p/2-d)/p \to 1/2$ e $(p-2d)/p \to 1$ quando $p \to \infty$.

Como $\mathbb{R}^d$ é união enumerável de conjuntos limitados, tem-se uma modificação contínua de $\xi_{st}(x)$, $x \in \mathbb{R}^d$.

Quanto à modifiação dos campos aleatórios definidos pelas integrais, suponha $s \leq s^1$ e $t < t^1$ (os outros casos são similares) e escreva

$$\int_s^t Y_j\big(u, \xi_{s,u}(x)\big)dW_u^j - \int_{s^1}^{t^1}\big(u, \xi_{s^1,u}(y)\big)dW_u^j =$$

$$= \int_s^{s^1} Y_j\big(u, \xi_{s,u}(x)\big)dW^j + \int_t^{t^1} Y_j\big(u, \xi_{s^1,u}(y)\big)dW_j$$

$$+\int_{s^1}^{t}\left(u,\xi_{s,u}(x)\right)-Y_j\left(u,\xi_{s^1,u}(y)\right)dW_u^j$$

Usando a desigualdade (A.2) e posteriormente (4.7) no primeiro e segundo termos e (4.10) no terceiro termo do segundo membro, obtém-se a condição de Kolmogorov com expoentes $\gamma = p, \alpha_1 = \alpha_2 = p/2$ e $\alpha_3 = p, p \geq 2$ arbitrário.

As integrais tem portanto o mesmo tipo de continuidade que $\xi_{s,t}(x)$. Por fim, a igualdade

$$\xi_{s,t}(x) = x + \sum_{j=0}^{m}\int_s^1 Y_j\left((u,\xi_{s,u}(x)\right)dW_u^j \quad \text{q.s.}$$

Segue por um raciocínio análogo ao da demonstração do teorema 4.2.

Antes de extender o resultado acima para equações localmente Lipschitz é conveniente (não apenas conveniente mas essencial) desenvolver a propriedade de fluxo das soluções ainda com a condição de Lipschitz global sobre os coeficientes.

Lembrando o caso das equações ordinárias: Suponha que $f(t,x)$ é Lipschitz global, então a solução única $\xi_{s,t}(x)$ de $\cdot x = f(t,x)$ com condição inicial $(s,x)$ satisfaz $\xi_{s,t}(x) = \xi_{u,t}(\xi_{s,u}(x))$ todo $s \leq u \leq t$. Além do mais para $s,t$ fixados, a aplicação $\xi_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ é um homeomorfismo.

Para equações estoásticas, já foi mostrada a propriedade de composição desde que se faça uma escolha conveniente. O que falta verificar, para um paralelismo completo com as equações ordinárias é a propriedade de homeomorfismo de $\xi_{s,t}$. Aqui o que se busca é algo da forma: Para $\omega \in \Omega' \subset \Omega$ com $P(\Omega') = 1$, $\xi_{s,t}(\omega) : \mathbb{R}^d \to \mathbb{R}^d$ é um homeomorfismo. Vamos mostrar esta propriedade verificando primeiramente que $\xi_{s,t}(\omega)$ é 1-1. Para isto precisamos do

Lema 4.6: Assuma (I) L.g. e defina para $x, y \in \mathbb{R}^d$, $s < t$, $s,t \in [a,b)$,

$$\eta_{s,t}(x,y) = \frac{1}{\xi_{s,t}(x) - \xi_{s,t}(y)}$$

($= +\infty$ se o denominador se anula). Tome $s > 0$ então para $p \geq 2, \max(s,s') < \min(t,t'), |x-y| \geq \delta$ e $|x'-y'| \geq \delta$,

$$E\left[\left|\eta_{s,t}(x,y) - \eta_{s',t'}(x',y')\right|\right]^p \tag{4.12}$$

$$C\delta^{-p}\Big(|x-x'|^p + |y-y'|^p + (1 + |x|^p + |y|^p + |x'|^p + |y'|^p|t-t'|^{p/2} + |s-s'|^{p/2}\Big).$$

<u>obs:</u> Nesta desigualdade, assumimos $\infty - \infty = 0$.

<u>Dem:</u> Por uma manipulação simples e por (A.2) (com dois termos), temos

$$\left|\eta_{s,t}(x,y) - \eta_{s',t'}(x',y')\right|^p \leq$$

$$\leq 2^p|\eta_{s,t}(x,y)|^p|\eta_{s',t'}(x',y')|^p\Big(|\xi_{s,t}(x) - \xi_{s',t'}(x')|^p + |\xi_{s,t}(y) - \xi_{s',t'}(y')|^p\Big)$$

Aplicando (Ho 1, 2, 2) duas vezes e a desigualdade triangular (para a soma no último termo) temos que o lado esquerdo da desigualdade acima é majorado por

$$CE\left[\left|\eta_{s,t}(x,y)\right|^{4p}\right]^{1/4} E\left[\left|\eta_{s',t'}(x',y')\right|^{4p}\right]^{1/4}.$$

$$\cdot\left(E\left[\left|\xi_{s,t}(x) - \xi_{s',t'}(x')\right|^{2p}\right]^{1/2} + E\left[\left|\xi_{s,t}(y) - \xi_{s',t'}(y')\right|^{2p}\right]^{1/2}\right)$$

Por (4.1) com $\varepsilon = 0$ e $-4p$ no lugar de $p$,

$$E\left[\left|\eta_{s,t}(x,y)\right|^{4p}\right]^{1/4} \leq C\left|x-y\right|^{-p} \tag{$*$}$$

Com uma desigualdade semelhante para $\eta_{s',t'}(x',y')$.

Com (4.11) (com expoente $2p$) seguido de (A.2) (com expoente 1/2),

$$E\left[\left|\xi_{s,t}(x) - \xi_{s',t'}(x')\right|^{2p}\right]^{1/2} \leq \tag{$**$}$$

$$\leq C\Big(|x-x'|^p + \big(1+|x|^p+|y|^p\big)\big(|t-t'|^{p/2} + |s-s'|^{p/2}\big)\Big)$$

Juntando ($*$) e ($**$) e usando a fato de que $|x-y|^{-p} \leq \delta^{-p}$ se $|x-y| \leq \delta$ ($p \geq 2$) chega-se a (4.12).

Da desigualdade (4.12) juntamente com o teorema de Kolmogorov, se obtem a injetividade q.s. de $\xi_{s,t}(x)$. De fato, (4.12), com $p$ suficientemente grande, garante que $\eta_{s,t}(x,y)$ satisfaz a condição de Kolmogorov com $x, y$ restritos a conjuntos limitados e $\mathbb{R}^d$ e satisfazendo $|x-y| > \delta$. Por um argumento análogo ao da demonstração do teorema 4.5, concluimos que $\eta_{s,t}(x,y)$ admite uma modificação contínua no domínio $\{(s,t,x,y) = s < t, \ |x-y| > \delta\}$. Como $\delta$ em (4.12) é arbitrário, o campo aleatório $\eta_{s,t}(x,y)$ definido em $\{(s,t,x,y) = s < t, \ x \neq y\}$ admite também uma modificação contínua e portanto

$$P\big(\omega : |\eta_{s,t}(x,y)| < \infty \ \text{ todo } \ s < t, \ x \neq y\big) = 1$$

Em outras palavras, $\{\omega : \xi_{s,t}(\omega)$ é injetiva todo $s < t\}$ é de probabilidade total, isto é, $\xi_{s,t}(\omega)$ é injetiva todo $s \leq t$ q.s..

Sobre esta injetividade, vale o seguinte comentário na mesma linha das observações feitas no início do § : A desigualdade (4.1) com $\varepsilon = 0$ e $p = -1$ se escreve:

$$E\left[\left|\xi_{s,t}(x) - \xi_{s,t}(y)\right|^{-1}\right] \leq C\,|x-y|^{-1}$$

de onde se tira que com $s < t$ e $x \neq y$ fixados, $P(\omega : \xi_{s,t}(x) \neq \xi_{s,t}(y)) = 1$. Isto no entanto não mostra a injetividade no sentido exposto acima, já que os conjuntos excepcionais, em que se tem $\xi_{s,t}(\omega)(x) = \xi_{s,t}(\omega)(y)$ podem, em princípio serem a um conjunto com probabilidade positiva (ou mesmo não mensurável) pois $\mathbb{R}^d \times \mathbb{R}^d$ é não enumerável.

O objetivo agora é mostrar a sobrejetividade q.s. de $\xi_{s,t}(\omega)$. A estratégia a ser seguida, é de se extender $\xi_{s,t}(\omega)$ a transformações na esfera $S^d$ por intermédio da compactificação de $\mathbb{R}^d$ por um ponto. A partir daí, usar a sobrejetividade usando o fato de que $\xi_{s,t}(\omega)$ é homotópica à $\xi_{s,s}(\omega) =$ identidade.

Seja então $\tilde{\mathbb{R}}^d = \mathbb{R}^d \cup \{\infty\}$ $(\approx S^d)$ a compactificação por um ponto de $\mathbb{R}^d$ e defina o campo aleatório $\tilde{\xi}_{s,t}(x)$, extensão de $\xi_{s,t}(x)$, definido para $x \in \tilde{\mathbb{R}}^d$ e a valores em $\tilde{\mathbb{R}}^d$, por

$$\tilde{\xi}_{s,t}(x) = \begin{cases} \xi_{s,t}(x) & \text{se} \quad x \in \mathbb{R}^d \\ & \\ \infty & \text{se} \quad x = \infty \end{cases} \qquad s < t.$$

Este campo aleatório é mensurável. Vamos mostrar a seguir, com o auxílio do teorema de Kolmogorov, que $\tilde{\xi}_{s,t}(x)$ admite uma modificação contínua.
Para isto, vamos definir $\theta : \tilde{\mathbb{R}}^d \to \tilde{\mathbb{R}}^d$ por

$$\theta(x) = \begin{cases} x/|x|^2 & \text{se} \quad x \in I\!R^d - \{0\} \\ \infty & \text{se} \quad x = 0 \\ 0 & \text{se} \quad x = \infty \end{cases}$$

Identificando $\tilde{I\!R}^d$ com $S^d = \{(y_1, \ldots, y_{d+1}) \in I\!R^{d+1} : |y| = 1\}$ via a projeção estereográfica

$$x = (x_1, \ldots, x_d) \in I\!R^d \to \frac{2}{1+|x|^2}\Big((x_1, \ldots, x_d, \frac{-1+|x|^2}{2}\Big) \, ,$$

$\theta$ fica sendo transformação:

$$\theta\Big(y_1, \ldots, y_d, \, y_{d+1}\Big) = \Big(y_1, \ldots, y_d, -y_{d+1}\Big)$$

De onde se vê que $\theta$ é um homeomorfismo de $I\!R^d$ e $\theta^2 =$ identidade.

**Lema 4.7:** Para $\varphi : I\!R^d \to I\!R^d$ continua, seja $\tilde{\varphi} : \tilde{I\!R}^d \to \tilde{I\!R}^d$ sua extensão com $\tilde{\varphi}(\infty) = \infty$. Então $\tilde{\varphi}$ é contínua se $\Psi = \tilde{I\!R}^d \to I\!R$ dada por

$$\Psi(x) = \begin{cases} \dfrac{1}{1+|\tilde{\varphi}(\theta(x))|} & \text{se} \quad x \in \tilde{I\!R}^d - \{0\} \\ \\ 0 & \text{se} \quad x = 0 \end{cases}$$

for contínua em 0.

obs: Note que $0 < \Psi(x) < \infty$.

Dem: Como $\lim\limits_{x \to 0} \Psi(x) = 0$, $\quad |\varphi(\theta(x))| \to \infty$ se $x \to 0$. Por ser $\theta$ um homeomofismo, tem-se então que $|\varphi(x)| \to \infty$ se $x \to \infty$, o que mostra a continuidade de $\tilde{\varphi}$.

**Lema 4.8:** Com (I) L.g e as notações acima, defina

$$\Psi_{s,t}(x) = \begin{cases} \dfrac{1}{1+|\tilde{\xi}_{s,t}(\theta(x))|} & \text{se } x \in \tilde{I\!R}^d - \{0\} \\ 0 & \text{se } x = 0 \end{cases}$$

Então para $p \geq 2$, $s < t$, $s' < t'$ e $x, y \in I\!R^d$

$$E\left[\left|\Psi_{s,t}(x) - \Psi_{s',t'}(y)\right|^p\right] \leq C\left(|x-y|^p + |t-t'|^{p/2} + |s-s'|^{p/2}\right) \tag{4.13}$$

Dem: Suponhamos inicialmente que $x$ e $y$ não são nulos. Então $\theta(x)$ e $\theta(y) \in I\!R^d$ e as majorações são feitas em $I\!R^d$. Como

$$|\Psi_{s,t}(x) - \Psi_{s',t'}(y)|^p = |\Psi_{s,t}(x)|^p \, |\Psi_{s',t'}(y)|^p \left(|\xi_{s,t}(\theta(x))| - |\xi_{s',t'}(\theta(y))|^p\right)$$

$$\leq \ |\Psi_{s,t}(x)|^p \, |\Psi_{s',t'}(y)|^p \, |\xi_{s,t}(\theta(x)) - \xi_{s',t'}(\theta(y))|^p$$

temos ao aplicar (Ho 1, 2, 2) duas vezes que

$$E\left[\left|\Psi_{s,t}(x) - \Psi_{s',t'}(y)\right|^p\right] \leq$$

$$\leq E\left[\left|\Psi_{s,t}(x)\right|^{4p}\right]^{1/4} E\left[\left|\Psi_{s',t'}(y)\right|^{4p}\right]^{1/4} E\left[\left|\xi_{s,t}(\theta(x)) - \xi_{s',t'}(\theta(y))\right|^{2p}\right]^{1/2} .$$

Majorando os dois primeiros termos do 2º membro por (4.6) (com $\varepsilon = 1$ e expoente - $4p$) e o último termo por 4.11, o 1º membro fica majorado por:

$$C\left(1+|\theta(x)|\right)^{-p}\left(1+|\theta(y)|\right)^{-p}\left(|\theta(x)-\theta(y)|^{p}+\left(1+|\theta(x)|+|\theta(y)|\right)^{p}\left(|t-t'|^{p/2}+|s-s'|^{p/2}\right)\right.$$

Como $\theta(x), \theta(y) \in I\!R^d$, por (A.5)

$$\left(1+|\theta(x)\right)^{-p}\left(1+|\theta(y)\right)^{-p}\left(1+|\theta(x)|+|\theta(y)||\right)^{p}\leq 1$$

e por (A.6)

$$\left(1+|\theta(x)|\right)^{-p}\left(1+|\theta(y)|\right)^{-p}|\theta(x)-\theta(y)|^{p}\leq\left|\frac{\theta(x)}{|\theta(x)|^{2}}-\frac{\theta(y)}{|\theta(y))|^{2}}\right|^{p}=|x-y|^{p}\ .$$

Portanto, $(*)$ é majorado por

$$C\left(|x-y|^{p}+|t-t'|^{p/2}+|s-s'|^{p/2}\right)\ . \qquad (**)$$

No caso em que $x=y=0$ não há nada a verificar pois $\Psi_{s,t}(x)=\Psi_{s',t'}(y)=0$. Por fim se, por exemplo $x\neq 0$ e $y=0$, aplicando (A.2) (com expoente 2) e posteriormente (4.6) (com $\varepsilon=1$ e expoente $-p$),

$$E\left[\left|\Psi_{s,t}(x)\right|^{2p}\right]\leq C\left(1+|\theta(x)|\right)^{-p} \qquad (***)$$

$$\leq C|x|^{2p}. \qquad \text{(pois } p>0\text{)}$$

$(**)$ junto com $(***)$ é o mesmo que (4.13).

O lema 4.7, a desigualdade (4.13) (com $p$ suficientemente grande) e o teorema de Kolmogorov mostram que $\xi_{s,t}(x)$ admite uma modificação contínua. Em outras palavras, existe $\Omega' \subset \Omega$ com $P(\Omega') = 1$ tal que para todo $(s, t \in [a,b), s < t, \tilde{\xi}_{s,t}(\omega) = S^d \to S^d$ é contínua, todo $\omega \in \Omega'$.

Como $\tilde{\xi}_{s,t}(\omega)$ é contínua também como função de $t$, $u \to \tilde{\xi}_{s,u}(\omega)$ define uma homotopia entre $\tilde{\xi}_{s,s}(\omega) =$ identidade e $\tilde{\xi}_{s,t}(\omega)$. Isto é suficiente para mostrar a sobrejetividade de $\tilde{\xi}_{s,t}(\omega)$. De fato, como $S^d$ é não contráctil, a identidade em $S^d$ é uma aplicação inessencial (não homotópica à uma constante) e como a homotopia é uma relação de equivalência, $\tilde{\xi}_{s,t}(\omega)$ também é inessencial. Por esta razão $\tilde{\xi}_{s,t}(\omega)$ é sobrejetora, pois aplicações inessenciais de $S^d$ são sobrejetoras (veja p.ex. [H-Y] teo. 4.13). Daí segue que $\xi_{s,t}(\omega)$ é sobre pois $\tilde{\xi}_{s,t}(\omega)(\infty) = \infty$.

Juntando isto com o comentado anteriormente, temos que

$$P\Big(\omega : \tilde{\xi}_{s,t}(\omega) \text{ é bijeção todo } s \leq t\Big) = 1.$$

Isto é, $\tilde{\xi}_{s,t}(\omega)$ é q.s inversível e como $\tilde{\mathbb{R}}^d$ é compacto, $\tilde{\xi}_{s,t}(\omega)$ é aplicação fechada e portanto sua inversa $\xi_{s,t}(\omega)^{-1}$ é contínua e daí que $\tilde{\xi}_{s,t}(\omega)$ é homeomorfismo de $S^d$. Novamente, por ser $\tilde{\xi}_{s,t}(\omega)(\infty) = \infty$ $\xi_{s,t}(\omega) : \mathbb{R}^d \to \mathbb{R}^d$ é homeomorfismo q.s.
Toda essa discussão se cristaliza no

**Teorema 4.9:** Suponha que (I) é L.g. com domínio $I \times \mathbb{R}^d, I = [a,b)$. Então as soluções maximais $\xi_{s,t}(x,\omega) = \xi_{s,t}(\omega)(x) = \xi_{s,t}(x)(\omega)$ podem ser escolhidas de tal forma que:

i) $\omega \to \xi_{s,t}(x)(\omega)$ é $\mathcal{F}_{s,t}$ – mensurável todo $(s,t,x) \in I \times I \times \mathbb{R}^d$, $s \leq t$.

ii) Existe $\Omega' \subset \Omega$, mensurável com $P(\Omega') = 1$ t.q. $\forall\, \omega \in \Omega'$.

$$(s,t,x) \to \xi_{s,t}(\omega)(x)$$

é $(\beta,\beta,\alpha)$ - Hölder contínua todo $\beta < \frac{1}{2}$ e $\alpha < 1$.

iii) Com o mesmo $\Omega'$ que em ii), $\forall\omega \in \Omega'$ e $\forall(s,t) \in I \times I$ com $s \leq t$,

$$\xi_{s,t}(\omega) = \mathbb{R}^d \to \mathbb{R}^d$$

é um homeomorfismo.

iv) Para $s \leq u \leq t$, $x \in \mathbb{R}^d$ e $\omega \in \Omega'$

$$\xi_{s,t}(x)(\omega) = \xi_{u,t}\Big(\xi_{s,u}(x)\Big)(\omega)$$

obs: Como o número de teoremas, lemas, proposições, etc... que podem ser escritos num texto é geralmente finito (e este é o caso deste texto), a quantidade de modificações realizadas nas soluções para que todas as propriedades acima fossem verificadas, é finita. Existe portanto uma modificação que faz com que as mesmas sejam satisfestas simultâneamente.

Antes de extender o teorema acima a equações localmente Lipschitz, indiquemos sua extensão à equações L.g. dependentes de parâmetros de dimensão finita.

**Exercício 4.10:** Tome os coeficientes $Y_i, j = 0,\ldots,m$ dependentes do parâmetro $\lambda \in \mathbb{R}^k$ e suponha que

$$|Y_j\big(t,x,\lambda\big) - Y_j\big(t,y,\mu\big)| \leq L|x-y| + K|\lambda-\mu|^\alpha, \alpha > 0$$

Mostra a existência de uma modificação que além de satisfazer as propriedades anteriores é $\alpha$-Holder contínua em relação. à $\lambda$. (Sugestão: seja $\xi^{\lambda}_{s,t}(x)$ a solução com parâmetro $\lambda$. Escreva

$$|\xi^{\lambda}_{s,t}(x) - \xi^{\mu}_{s^1,t^1}(y)|^p \leq |\xi^{\lambda}_{s,t}(x) - \xi^{\lambda}_{s^1,t^1}(y)|^p + |\xi^{\lambda}_{s^1,t^1}(y) - \xi^{\lambda}_{s^1,t^1}(y)|^p$$

e use o 2º termo do 2º membro para incorporar um termo do tipo $|\lambda - \mu|^{\alpha p}$ em (4.11).

O objetivo agora é obter o correspondente do teorema anterior para equações localmente Lipsichitz. Voltemos então à situação do teorema 3.4 denotando, como lá, por $T(s,x)$ o tempo de explosão da solução maximal $\xi_{s,t}(x)$ com condição inicial $(s,x)$. A possibilidade da escolha de uma versão contínua para o campo aleatório $\xi_{s,t}(x), (s,x) \in U,\ s < t < T(s,x)$, pode ser mostrada quase que imediatamente a partir da construção da solução maximal (no teorema 3.4) e do teorema 4.12.

Lembrando que para a construção da solução maximal $\xi_{s,t}(x)$, tornam sequências de compactos $K_n \neq \phi$ com $K_n \subset \text{int } K_{n+1}$ e $\bigcup_n K_n = U$ e de equações diferenciais L.g. cujas soluções denotadas por $\xi^n_{s,t}(x)$ coincidem com $\xi_{s,t}(x)$ se $T < t_n(s,x)$ onde $T_n(s,x)$ é o primeiro tempo em que $\xi^n_{s,t}(x)$ atinge a fronteira de $K_n$.

Cada campo aleatório $\xi^n_{s,t}(x)$, sendo solução maximal de uma L.g., admite modificação contínua. Como a quantidade das $\xi_n' s$ é enumerável, é possível toma-las simultaneamente continuas, isto é, existe $\Omega' \subset \Omega$ com $P(\Omega') = 1$ t.q. $(s,t,x) \to \xi^n_{s,t}(x)(\omega)$ é contínua todo $n$ e todo $\omega \in \Omega'$. Daqui para frente tomaremos sempre uma escolha como esta. Feito isto, consideremos o domínio de definição de $\xi_{s,t}(\omega), \omega \in \Omega'$. Este é dado por

$$D_{s,t}(\omega) = \left\{x \in I\!\!R^d : T(s,x)(\omega) > t\right\},$$

que é um aberto de $I\!\!R^d$. De fato, $T(s,x) = \sup_n\ T_n(s,x)$. Portanto se definirmos

$$D^n_{s,t} = \left\{x \in I\!\!R^d : T_n(s,x) > t\right\}$$

temos que $D_{s,t} = \bigcup_n D^n_{s,t}$. Por outro lado, $D^n_{s,t}$ é aberto, pois se $T_n(s,x) > t$, pela definição de $T_n$, temos que $\xi^n_{s,t}(x) \in \text{int } K_n$, $s \leq u \leq t$. E daí que $\xi^n_{s,u}(y) \in \text{int } K_n$ para $y$ numa vizinhança de $x$ e $s \leq u \leq t$, isto é, $T_n(s,y) > t, y \in D^n_{s,t}$ e $D^n_{s,t}$ é aberto.
Isto posto, temos a aplicação $\xi_{s,t}(\omega) : D_{s,t}(\omega) \to I\!R^d$, que é contínua pois $\xi_{s,t}(\omega)(x) = \xi^n_{s,t}(\omega)(x)$ algum $n$.

**Teorema 4.11:** Suponha (I) localmente Lipschitz e seja $\Omega'$, com $P(\Omega') = 1$ como acima. Para $s \leq t$ e $\omega \in \Omega'$ seja

$$I_{s,t}(\omega) = \left\{ y \in I\!R^d : y = \xi_{s,t}(\omega)(x), x \in D_{s,t}(\omega) \right\}$$

a imagem de $\xi_{s,t}(\omega)$. Então

i) $I_{s,t}(\omega)$ é aberto e $\xi_{s,t}(\omega) : D_{s,t}(\omega) \to I_{s,t}(\omega)$ é homeomorfismo

ii) Para $s \leq u \leq t$, $D_{s,t} \subset D_{s,u}, \{\xi_{s,u}(x) : x \in D_{s,t}\} \subset D_{u,t}$ e

$$\xi_{s,t} = \xi_{u,t} \ o \ \xi_{s,u} \tag{4.14}$$

<u>Dem:</u> Como $D_{s,t} = \bigcup D^n_{s,t}, \xi_{s,t} = \xi^n_{s,t}$ em $D^n_{s,t}$ e pelo teorema 4.9 $\xi^n_{s,t}$ é um homeomorfismo, temos que $\xi_{s,t}(\omega) : D_{s,t}(\omega) \to T_{s,t}(\omega)$ é um homeomorfismo. Segue daí e do teorema da invariança do domínio ([H-Y] teo. 6.53) que $I_{s,t}(\omega)$ é aberto.

As propriedades em ii) seguem imediatamente do teorema 4.9 e do fato que $\xi_{s,t}(x) = \xi^n_{s,t}(x)$ algum $n$.

Para concluir esta seção, vamos discutir as equações autonomas (ou homogeneas no tempo), isto é, em que os coeficientes não dependem de $t$. O domínio de definição $U$

destas equações é da forma $U = [0,\infty)\times V$, $V \subset I\!R^d$ aberto e como no caso das equações ordinárias, todas as soluções serão discritas em termos das soluções com condições iniciais $(0,x), x \in V$.

De fato, tomemos o campo aleatório $(t,x) \to \varphi_t(x) = \xi_{0,t}(x)$, que, pelo teorema 4.11, é contínuo se os coeficientes são Lipschitz. Mostremos que para $s > 0$, se tem

$$\xi_{s,t}(\omega) = \varphi_{t-s}\big(\theta_s(\omega)\big) \ , \ t \geq s$$

onde $(\theta_s(\omega))(t) = \omega(t+s) - \omega(s)$ é a translação no espaço de Wiener.

Defina $\eta_t = \varphi_{t-s}(\theta)$. Então $\eta_t$ é $\mathcal{F}_{s,t}$ mensurável (pois $\varphi_t$ é $\mathcal{F}_{0,t-s}$ - mensurável e $\theta_s(\mathcal{F}_{0,t-s}) = \mathcal{F}_{s,t}$) e tem-se

$$\begin{aligned}\left(\int_s^t X(\eta_u)du\right)(\omega) &= \int_s^t X\big(\eta_u(\omega)\big)du \\ &= \int_0^{t-s} X\big(\eta_{u+s}(\omega)\big)du \\ &= \left(\int_0^{t-s} X(\varphi_u)du\right)\big(\theta_s(\omega)\big)\end{aligned}$$

pois $\eta_{u+s}(\omega) = \varphi_u(\theta_s(\omega))$. E também,

$$\begin{aligned}\Big(\int_s^t Y_j(\eta_u)dW_u^j\Big)(\omega) &= \Big(\int_0^{t-s} y_j(\eta_{u+s}dW^j\Big)(\theta_s(\omega)) \\ &= \Big(\int_0^{t-s} Y_j(\varphi_u)dW_u^j\Big)(\theta_s(\omega))\end{aligned}$$

pela propriedade de translação das integrais estocásticas sobre movimentos Brownianos. Como $\varphi_y$ é solução de (I), temos que $\eta_t, t \geq s$ é solução de (I) com condição inicial $(s,x)$. Pela unicidade temos então que $\xi_{s,t}(\omega) = \eta_t(\omega) = \varphi_{t-s}(\theta_s(\omega))$ q.s..

Esta igualdade, juntamente com (4.14) fornece a equação de fluxo para soluções de equações estocásticas autonomas:

$$\varphi_{t+s}(\omega) = \varphi_t(\theta_s(\omega)) \circ \varphi_s(\omega) \tag{4.15}$$

## 4.5 Dependência diferenciável.

Os resultados sobre dependência diferenciavel das soluções de uma equação ordinária em relação as condições iniciais também tem um análogo para equações estecásticas. Lembramos que se $\dot{x} = f(t,x)$ é uma equação diferencial ordinária, a solução $\xi_{s,t}(x)$ com condição inicial $(s,x)$ é diferenciável de classe $C^k$ em relação à $x$ se o mesmo ocorre com $f(t,x)$. Além do mais, sua diferencial em $x, d(\xi_{s,t})_x$ satifaz a equação adjunta

$$\dot{g} = \frac{\partial f}{\partial x}(t, \xi_{s,t}(x))g \qquad g_0 = \text{identidade} \tag{5.1}$$

que é uma equação linear com coeficientes variáveis (continuos) no espaço das matrizes quadradas. Nesta equação, $\frac{\partial f}{\partial x}(t,y)$ denota o diferencial da aplicação $y \to f(t,y)$ no ponto $y$.

Para equações estocásticas, tem-se também a diferenciabilidade das soluções em relação à $x$.

Existe, no entanto, uma diferença no que diz respeito à hipótese que se faz sobre os coeficientes:
Continuidade apenas de suas $n$-ésimas derivadas parciais não basta. É nescessário que se assuma algum tipo de regularidade nessa continuidade para que se possa utilizar certas desigualdades. Por exemplo em [K1], mostra-se que $\xi_{s,t}(x)$ é q.s. $\beta$-Hölder contínua, todo $\beta < \alpha$, se os coeficientes são de classe $C^k$ e as $k$-ésimas derivadas parciais são $\alpha$-Hölder contínuas, $\alpha > 0$. Isto é feito à custa da prova de uma desigualdade extra que

evitaremos aqui. Por questão de simplicidade na exposição vamos tomar $\alpha = 1$, isto é, vamos assumir que as $k$ – ésimas derivadas parciais são Lipschitz contínuas. Esta condição é garantida no caso em que os coeficientes são de classe $C^{k+1}$.

O correspondente à (5.1) para equações estocásticas é o par de equações

$$\begin{aligned} dx &= X\Big(t,x\Big)dt + \textstyle\sum_{j=0}^{m} Y_i(t,x)dW^j \\ dg &= \ X'(t,x)gdt + \textstyle\sum_{j=0}^{m}\ Y_j^1(t,x)gdW^j, \qquad x \in I\!R^d, gEM_{d\times d} \end{aligned} \tag{5.2}$$

onde $X'(t,x)$, $Y_j'(t,x)$ denota a diferencial de $X, Y_j$ em relação a $x$, avaliada em $(t,x)$. Note que como em (5.1) a segunda equação é linear em $g$ e seus coeficientes são obtidos por substituição em $Y_j'(t,x)$ das soluções da equação em $x$. Como estas soluções são processos estocásticos, essa equação linear não é do tipo que viemos tratando até agora. Temos no entanto, que as duas equações de(5.2) tomadas conjuntamente é do tipo (I) em $I\!R^d \times M_{d\times d}$.

**Teorema 5.1:** Suponha que (I) é L.g. e que seus coeficientes sejam diferenciaveis em relação à $x$ e que as diferenciais parciais $Y_j^1, j = 0, \ldots, m$, sejam contínuas, globalmente Lipschitz em relação à $x$ e limitadas.

Então existe uma modificação contínua de $\xi_{s,t}(x)$ t.q. $x \to \xi_{s,t}(x)$ é q.s. de classe $C^1$ e o par $\Big(\xi_{s,t}(x),\ d(\xi_{s,t})_x\Big)$ é a única solução de (5.2) com condições iniciais $(s, x, 1)$ ($1$ = matriz identidade).

<u>Dem:</u> Antes de mais nada, com as hipóteses sobre os coeficientes, (5.2) é L.g., já que as diferenciais parciais são L.g. (que se usa para verificar a condição de Lipschitz da segunda equação em relação à $x$) e globalmente limitadas (de onde se obtém a condição de Lipschitz em relação à $g$). Portanto o teorema de existência e unicidade e de continuidade em relação às condições iniciais funciona para (5.2)

Defina, para $x, y \in I\!R^d$

$$\tilde{Y}_j\Big(t,x,y\Big) = \int_0^1 Y_j'\Big((t, x + u(y-x)\Big)du \ \in \ M_{d\times d}$$

$j = 0, \dots, m$, e considere a equação

$$\begin{aligned} dx &= \textstyle\sum_{j=0}^m \ Y_j(t,x)\ dW^j \\ dy &= \textstyle\sum_{j=0}^m \ Y_j(t,y)\ dW^j \\ d\eta &= \textstyle\sum_{j=0}^m \ \tilde{Y}_j(t,x,y)\ \eta dW^j \end{aligned} \tag{5.3}$$

em $I\!R^{3d}$. Esta equação é L.g. como pode-se verificar a partir das hipóteses sobre os diferenciais parciais e a definição de $\tilde{Y}_j$.
Para $x, v \in I\!R^d$, $\lambda \in I\!R$, $\lambda \neq 0$, seja

$$\eta_{s,t}(x,\lambda) = \frac{1}{\lambda}\Big(\xi_{s,t}(x+\lambda v) - \xi_{s,t}(x)\Big)$$

onde $\xi_{s,t}(x)$ é solução de (I).

Utilizando o teorema fundamental do cálculo, verifica-se facilmente que a tripla $(\xi_{s,t}(x), \xi_{s,t}(x+\lambda v), \eta_{s,t}(x,\lambda))$ satisfaz (5.3) com condições iniciais $(x, x+\lambda v, v)$. Pela continuidade em relação às condições iniciais, $\lim\limits_{\lambda\to 0} \eta_{s,t}(x,\lambda)$ existe q.s. e coincide com a solução com condições iniciais $(x,x,v)$, isto é, $\xi_{s,t}(x)$ admite derivadas direcionais em todas as direções e estas são contínuas, existindo portanto a diferencial $(d\xi_{s,t})_x$, $x \in I\!R^d$, q.s.. Como a terceira equação de (5.3) se reduz à segunda de (5.2) quando $x = y$ (olhando as colunas de $g$ em (5.2)), tem-se também que a diferencial satisfaz (5.2).

**Corolário 5.2:** Nas condições do teorema 5.1, $d(\xi_{s,t})_x$ é inversível todo $s, t, x$, q.s.. Além do mais det $(d(\xi_{s,t})_x) > 0$ e $\xi_{s,t}$ é q.s. difeomorfismo de classe $C^1$.

<u>Dem:</u> Este corolário é conseqüência imediata dos teoremas de suporte veja [I-W] e da invertibilidade da diferencial do fluxo associado à uma equação ordinária. Por esta razão, vamos apenas indicar aqui uma prova direta.
Defina em $\mathbb{R}^d \times M_{d\times d}$

$$\begin{aligned} dx &= \textstyle\sum_{j=0}^{m} Y_j(t,x)dW^j \\ dh &= -hX'(t,x)dt - \textstyle\sum_{j=1}^{m} hY_j'(t,x)^2 dW^j \end{aligned} \tag{5.4}$$

As soluções $h_{s,t}$ de (5.4) com $h_{s,s} = 1$ são as inversas das soluções $g_{s,t}$ de (5.2) com $g_{s,s} = 1$. Para ver isto, aplique a fórmula de Itô para concluir que $d(h_{s,t}g_{s,t}) = 0$. Este procedimento deixará claro o porque do quabrado das diferenciais dos elementos difusos em (5.4) (sugerimos que se compare com a equação correspondente ao caso das equações ordinárias). Por aí se ve que $d(\xi_{s,t})_x$ é inversível, e pela continuidade das trajetórias tem determinante $> 0$.

A última afirmação segue do fato que um homeomorfismo diferenciável cujo diferencial é inversível é um difeomorfismo.

**Corolário 5.3:** Suponha que os coeficientes de(I) sejam de classe $C^k$ em relação à $x, x \geq 1$, as derivadas parciais de ordem $\leq k$ sejam contínuas (em relação à $(t,x)$) e as derivadas parciais de ordem $k$ sejam globalmente Lipschitz e limitadas.

Então existe uma versão de $\xi_{s,t}(x)$ tal que $x \to \xi_{s,t}(x)$ é difeomorfismo de classe $C^k$ q.s..

Dem: É simplesmente uma questão de diferenciar (5.2) sucessivamente.

**Observação 5.4:** Como comentado acima a condição de Lipschitz nas $k$-ésimas derivadas parciais pode ser substituida por $\alpha$-Hölder continuidade, qualquer $\alpha > 0$ (veja [K1]).

**Teorema 5.5:** Com a situação e notações como no teorema 4.11, suponha além do mais que os coeficientes tenham $k$-ésimas derivadas parciais em relação à $x$ e estas sejam Lipschitz contínuas.
Então para $\omega \in \Omega'$.

$$\xi_{s,t}(\omega) : D_{s,t}(\omega) \longrightarrow I_{s,t}(\omega)$$

é um difeomorfismo $C^k$.

Dem: Procede-se como no teorema 4.11: $\xi_{s,t}(\omega)$ coincide para $t < T_n(s,k)$ com a solução de uma equação global. Como estas são difeomorfismo $C^k$, $\xi_{s,t}(\omega)$ também é.

## 4.6 Equações de Stratonovich equações em variedades.

Ao invés de (I), consideremos equações de Stratonovich

$$dx = X(t,x)dt + \sum_{j=1}^{m} Y_j(t,x) \circ dW^j \quad , \quad x \in \mathbb{R}^d \tag{S}$$

Suas soluções são definidas como as de (I), com a diferença, é claro, que na equação integral correspondente as integrais são tomadas no sentido de Stratonovich:

$$\xi_t = x + \int_s^t X(u,\xi_u)du + \sum_{j=1}^{m} \int_s^t Y_j(u,\xi_u) \circ dW_u^j \tag{6.1}$$

Como os integrandos de integrais de Strotonovich são semi-martingales, é necessário que se faça hipóteses adicionais sobre os coeficientes de (S) além da continuidade (como em (I)), para que a integrais estocásticas em (6.1) possam ser definidas. As seguintes condições são suficientes para dar sentido a (6.1) e também para se mostrar a existência e unicidade de soluções de (S):

i) $X$ é contínua, tem derivadas parciais em relação à $x$ e estas são contínuas em relação à $(t,x)$.

ii) $Y_j, j = 1,\ldots,m$, é de classe $C^1$, tem derivadas parciais de 2ª ordem e estas são contínuas como função de $(t,x)$.

Com estas hipóteses sobre os coeficientes, (S) é equivalente à uma equação de Itô com coeficientes de classe $C^1$. Para ver isto, seja a equação de Itô

$$dx = \tilde{X}(t,x)dt + \sum_{j=1}^{m} Y_j(t,x)\, dW^j \tag{6.2}$$

com

$$\tilde{X}(t,x) = X(t,x) + \frac{1}{2}\sum_{j=1}^{m}(D_{Y_j}Y_j)(t,x) \tag{6.3}$$

onde $D_{Y_j}Y_j$ denota a derivada direcional de $Y_j$ na direção dela mesma. Com (H), (6.2) é $C^1$ e portanto localmente Lipschitz e daí que admite soluções únicas, modificações contínuas etc...

Seja $\xi_t$ solução de (6.2). Como $Y_j$ é $C^2$ em relação à $x$, $C^1$ em relação à $t$ e $t$ é processo de variação limitada a fórmula de Itô se aplica à $Y_j(t,\xi_t)$ mostrando que $Y_j(t,\xi_t)$ é semi-martingale e portanto as integrais $\int_s^t Y_j(u,\xi_u)_0 dW_u^j$ fazem sentido.

Usando a fórmula de Itô para calcular a variação quadrática $< Y_j(\cdot,\xi), W^j >$, temos que

$$\int_s^t Y_j(u,\xi_u)_0 dW_u^j = \int_s^t Y_j(u,\xi_u) dW_u^j + \frac{1}{2}\left(\left\langle Y_j, W^j\right\rangle_t - \left\langle Y_j, W^i\right\rangle_s\right)$$

$$= \int_s^t Y_j(u,\xi_u) dW_u^j + \frac{1}{2}\sum_{j=1}^m \int_s^t \left(D_{Y_j} Y_j\right)(u,\xi_u) du$$

Conseqüentemente, soluções de (6.2) são soluções de (S).

Em vista da teoria desenvolvida anteriormente, temos então o

**Teorema 6.1:** Seja (S) definida em $U \subset [0,\infty) \times \mathbb{R}^d$ e suponha que seus coeficientes satisfaçam (H). Então as conclusões do teorema 4.11 valem para (S).

Se além do mais $X$ é $C^{k+1}$ e $Y_j$ é $C^{k+2}$ em relação à $x$ e $C^{k+1}$ em relação à $t$, a propriedade de difeomorfismo $C^k$, como no teorema 5.5, vale para (S).

**Observação:** Com as condições (H), $\tilde{X}$ em (6.2) é de classe $C^1$ e $Y_j$ de classe $C^2$. Isto é um tanto exagerado para se ter modificações contínuas de soluções de (6.2) já que pelo teorema 4.11, Lipschitz é suficiente. A hipótese (H) é no entanto exigida para se garantir a existência das integrais da Stratonovich. O que pode ser enfraquecido na formulação acima é a hipótese sobre $X$. Ao invés de $C^1$, Lipschitz basta.

Apesar das equações de Stratonovich exigirem uma regulariedade maior nos coeficientes para se garantir a existência e unicidade de soluções, (S) se comporta "melhor" que (I) sob mudanças de coordenadas.

Isso é expresso pela fórmula de Itô aplicada a soluções de (I) ou (S). Para explicitá-la, vamos simplificar as expressões e supor em (I) ou (S) que $m = 1$. (S) então será escrita como $dx = Xdt + Y_0 dW$.

Vamos usar também a notação $XF$ para indicar a derivada de $F$ na direção do vetor $X$. Desta forma, o campo $X$ define um operador diferencial (de 1ª ordem) $F \to XF$.

**Teorema 6.2:** i) Seja $\xi_t$ solução de $(I)$ e $F : \mathbb{R}^d \to \mathbb{R}$ de classe $C^2$. Então:

$$d\big(F(\xi_t)\big) = \Big(XF + \frac{1}{2}\sum_{i,k} \frac{\partial^2 F}{\partial x^i \partial x^k} Y^i Y^k(t,\xi_t) + YF(t,\xi_t)\ dW \tag{6.4}$$

onde $Y^i$ denota a $i$-ésima componente de $Y$.

ii) Seja $\xi_t$ solução de (S) e $F : \mathbb{R}^d \to \mathbb{R}$ de classe $C^3$. Então

$$\begin{aligned} d\big(F(\xi_t)\big) &= XF(t,\xi_t)dt + YF(t,\xi_t)_0\ dW \\ &= (LF)(t,\xi_t)dt + (YF)(t,\xi_t)\ dW \end{aligned} \tag{6.5}$$

onde $L$ é o operador diferencial de 2ª ordem

$$L = X + \frac{1}{2}Y^2$$

e $Y^2 F = Y(YF)$.

Dem: i) (6.2) é exatamente a fórmula de Itô. Observe que se a solução for escrita em coordenadas como $\xi = (\xi^1, \ldots, \xi^d)$ então $d\xi^i = X^i dt + Y dW$ e portanto a variação quadrática que aparece na fórmula de Itô é

$$d\langle \xi^i, \xi^j \rangle = Y^i Y^j dt$$

ii) A primeira igualdade em (6.5) é a fórmula de Itô em termos de diferenciais de Strolonovich. Lembre que esta só vale para $F$ de classe $C^3$. A segunda igualdade sai ou de primeira, calculando a variação quadrática $\left\langle Y(\cdot, \xi\cdot), W \right\rangle$ ou por (6.2) e (6.4). Por este último caminho, escrevemos

$$\begin{aligned} d\big(F(\xi)\big) &= \Big(\tilde{X}F + \tfrac{1}{2}\sum_{i,k} \frac{\partial^2 F}{\partial x^i \partial x^k} Y^i Y^k\Big) dt + YFdW \\ &= \Big(XF + \tfrac{1}{2}(D_Y Y)F + \tfrac{1}{2}\sum_{i,k} \frac{\partial^2 F}{\partial x^i \partial x^k} Y^i Y^k\Big) dt + YFdW \end{aligned}$$

(6.5) então se obtém a partir de

$$\begin{aligned} Y^2 &= \textstyle\sum_{i,k} Y^i Y^k \dfrac{\partial^2}{\partial x^k \partial x^k} \quad + \textstyle\sum_{i,k} Y^i \dfrac{\partial Y^k}{\partial x^i} \dfrac{\partial}{\partial x^k} \\ &= \textstyle\sum_{i,k} Y^i Y^k \dfrac{\partial^2}{\partial x^i \partial x^k} \quad + D_Y Y \end{aligned}$$

que pode ser obtido diretamente, bastando que se escreva o operador de primeira ordem $Y$ em coordenadas como

$$Y = Y^1 \frac{\partial}{\partial x^1} + \cdots + Y^d \frac{\partial}{\partial x^d}$$

**Observações 6.3:** i) O correspondente à (6.4) e (6.5) para $m > 1$ é obtido por soma em $j = 1, \cdots, m$ das expressões onde $Y$ aparece.

ii) As fórmulas (6.4) e (6.5) esclarecem o que queriamos dizer por melhor comportamento, sob mudanças de coordenadas, das equações de Stratonovich em relação às de Itô. Em (6.4), os símbolos $Y^i, Y^k$ e $\dfrac{\partial^2 F}{\partial x^i \partial x^k}$ só puderam ser escritos porque estava implícito que tinhamos feito uma escolha prévia de um sistema de coordenadas. Por outro lado, as expressões em (6.5) têm sentido intrínseco, isto é, são escritos sem menc,ão explicita (ou implícita) a nenhum sistema de coordenadas. Como veremos, (6.4) se altera ao se tomar diferentes sistemas de coordenadas enquanto que (6.5) permanece com a mesma cara. Por razões dessa ordem é que as equações diferenciais estocásticas em variedades são as equações de Strotonovich.

Definimos a seguir as equações de Strotonovich em variedades diferenciáveis, e comentaremos posteriormente a maneira de se introduzir equações de Itô.

**Definição 6.4:** Seja $M$ uma variedade diferencial de classe $C^k$, $k \geq 3$, que suporemos paracompacta, Hausdorff e conexa.

Uma equação diferencial de Stratonovich em $M$ é uma expressão como $(S)$ com $X(t,x) = Y_0(t,x), Y_1(1,x), \ldots, Y_m(t,x)$ campos de vetores em $M$ isto é, $Y_i : [a,b) \times M \to TM$ com $[a,b) \subset [0,\infty)$. Supomos que estes campos satisfazem $(H)$.

Uma solução de $(S)$ com condição inicial $(s,x)$ e tempo de explosão $T(s,x)$ é um processo $\xi_{s,t}(x)$ a valores em $M$ que satisfaz i) - iii) da definição 1.2 e

iv) Para $F : M \to \mathbb{R}$ de classe $C^3$, $F(\xi_{s,t}(x))$ é um semi-martingale e

$$F\Big(\xi_{s,t}(x)\Big) = F(x) + \int_s^t XF\Big(u, \xi_{s,u}(x)\Big)du + \sum_{j=}^{m} \int_s^t Y_j F(u, \xi_{s,u})_0 \, dW_u^j \tag{6.6}$$

para $s < t < T(s,x)$.

Soluções maximais são definidas como em (1.3).

**Observação:** Por (6.6), as soluções de (S) em $M$ são semi-martingales a valores em $M$ no sentido de L. Schwartz [Sch].

A construção de soluções de $(S)$ em $M$ é feita – naturalmente – através do teorema 6.1, tomando sistemas de coordenadas (cartas) de $M$:

Seja $\varphi : U \subset M \to Y \subset \mathbb{R}^d, U, V$ abertos, um sistema de coordenadas em $M$. Então $\varphi$ é um difeomorfismo de classe $C^k (k \geq 3)$ e daí que se denotarmos por $\varphi_*$ sua diferencial, os campos $Y_i$ em $M$ definem campos $\tilde{Y}_j = \varphi_* Y_j$ em $V \subset \mathbb{R}^d$. Tomando a equação diferencial

$$dy = \sum_{j=0}^{m} \tilde{Y}_j(t, y)_0 \, dW^j \tag{6.7}$$

em $V$, o teorema 6.1 se aplica pois $\tilde{Y}_j$ satisfaz (H) da mesma forma que $Y_j$. Portanto, se $y = \varphi(x) \in V$, (6.7) admite solução $\tilde{\xi}_{s,t}(y)$ com tempo de explosão $T_V(s, x)$. Voltando à $M$, temos que $\xi_{s,t}(x) = \varphi^{-1}\left(\tilde{\xi}_{s,t}(y)\right)$ é um processo em $U$ que satisfaz i)-iii) da definição de solução. A condição inicial é $x = \varphi^{-1}(y)$. Quanto à iv), seja $F = M \to \mathbb{R}$ de classe $C^3$. Restringindo à $U$ e aplicando a fórmula de Itô (6.5),

$$F_0\varphi^{-1}\left(\tilde{\xi}_{s,t}(y)\right) = F_0\varphi^{-1}(y) + \sum_{j=o}^{m} \int_s^t \tilde{Y}_j(F_0\varphi^{-1})\left(u, \tilde{\xi}_{s,u}(y)\right)_0 dW_u^j$$

onde usamos como anteriormente $X = Y_0$ e $dW^0 = dt$.

Como $\tilde{Y}_j = \varphi_* Y_j$, $\tilde{Y}_j(F_0\varphi^{-1}) = (Y_j F)_0 \varphi$ e daí que

$$F\left(\xi_{s,t}(x)\right) = F(x) + \sum_{j=0}^{m} \int_s^t (Y_j F)\left(u, \xi_{s,u}(x)\right)_0 dW_u^j$$

e iv) é satisfeita.

Desta forma, temos um procedimento para construir soluções de $(S)$ em $M$. Verifiquemos que este procedimento fornece a mesma solução se tomarmos outro sistema de coordenadas $\varphi_1 : U_1 \to V_1$, com $x \in U_1 \bigcap U$. Seja como acima $\xi_{s,t}(x)$ solução de (S) em $U_1 \bigcap U$ construido via a equação em $\varphi(U_1 \bigcap U)$. Usando a notação $\psi = \varphi_1 \; o \; \varphi^{-1}$, temos por (6.5), aplicada à cada coordenada de $\psi$ que

$$\left(\tilde{\xi}_{s,t}(y)\right) = \psi(y) + \sum_{j=0}^{m} \int_s^t \left(\tilde{Y}_j\psi\right)\left(u, \tilde{\xi}_{s,u}(y)\right)_0 dW_u^j$$

Usando agora o fato de que $\tilde{Y}_j\psi(u, \tilde{\xi}_{s,u}(y)) = (\varphi_{1*}Y)(u, \varphi_1\xi_{s,u}(x))$, que pode ser verificado imediatamente, concluimos que $\psi(\tilde{\xi}_{s,t}(y))$ é solução de (6.7) em $V_1$. Pela unicidade das soluções, os dois sistemas de coordenadas fornecem a mesma solução em $U \bigcap U_1$.

Para verificar a existência e unicidade de soluções maximais, modificações contínuas, etc..., para (S) em $M$, faremos uso dos teoremas de imersão de Whitney (veja [N] para uma abordagem direta).

Lembramos que uma imersão de uma variedade $M$ de dimensão $d$ em $\mathbb{R}^n$, $n \geq d$, é uma aplicação $i : M \to \mathbb{R}^n$ cujo diferencial $i_*$ tem posto máximo em todo os pontos, isto é, é injetiva. Se além do mais, $i : M \to i(M)$ é um homeomorfismo (com a topologia de $i(M)$ a induzida por $\mathbb{R}^n$), $i$ é denominado um mergulho e $M$ é dita mergulhada em $\mathbb{R}^n$. Neste caso, toda estrutura topológica e diferenciável de $M$ é obtida da estrutura induzida sobre o subconjunto $i(M)$, sendo possível não distinguir nem mesmo na notação, $M$ de $i(M)$.

De acordo com Whitney, toda variedade paracompacta e conexa $d$-dimensional, admite um mergulho em $\mathbb{R}^n$ com $n \geq 2d + 1$ (veja p.ex. [N] cap.2). Este resultado nos permite considerar $M$ como subvariedade de $\mathbb{R}^n$.

Para construir soluções maximais, necessitamos ainda do seguinte resultado de extensão: Tomando $M$ mergulhada em $\mathbb{R}^n$, seja $X(t, x)$ um campo de vetores em

$M$. $X$ pode ser visto como uma aplicação $X = [a,b) \times M \to \mathbb{R}^n$. Então existe $X^e = [a,b) \times U^e \to \mathbb{R}^n$ com $U^e$ aberto e $M \subset U^e$ tal que $X^e(t,x) = X(t,x)$ se $x \in M$. $X^e$ pode ser tomado com a mesma regulariedade que $X$. Por exemplo se $X$ é $C^k$, o mesmo ocorre com $X^e$. Além do mais, $U^e$ pode ser tomado de tal forma que $M$ é fechado em $U^e$ (veja [ ] prop. 2-5-14).

Isto posto, partindo de $(S)$ em $M$ construimos

$$dx = \sum_{j=0}^{m} Y_j^e(t,x) \circ dW^j \qquad (S^e)$$

em $U^e$. Assumindo que $(S)$ satisfaz $(H)$, o memo ocorre com $(S^e)$ e portanto o teorema 6.1 se aplica.

Seja $\xi_{s,t}^e(x)$ na solução maximal com a escolha contínua em relação às condições iniciais. Afirmamos que as soluções iniciadas em $x \in M$ não saem de $M$ q.s., i.e., para $\omega \in \Omega'$ com $P(\Omega') = 1$, $\xi_{s,t}^e(\omega)(x) \in M$ se $x \in M$. Esta afirmação é conseqüência imediata dos teoremas de suporte, do resultado correspondente para equações ordinárias e do fato que $M$ é fechada em $U^e$. O esquema para uma verificação direta é o seguinte:

Tome $(\varphi_n, U_n)$ uma seqüência de sistemas de coordenadas em $M$ com $M = \bigcup_n U_n$ (que existe pois $M$ é paracompacta). Em $U_n$ construa a solução $\xi_{s,t}^n(x)$ como acima, com tempo de explosão $T_n(s,x)$. Como $(S^e)$ é extensão de $(S)$, $\xi_{s,t}^n(x)$ é solução de $(S^e)$ e portanto pela unicidade, $\xi_{s,t}^n(x) = \xi_{s,t}^e(x)$ q.s., $s \le t < T_n(s,x)$.
Além do mais,

$$T_n(s,x) = \inf \left\{ t > s = \xi_{s,t}^e(x) \notin U_n \right\}$$

Pelo fato de que a quantidade dos $U_n's$ é enumerável, existe $\Omega'$ com $P(\Omega') = 1$ t.q. $\xi_{s,t}^e(y)(\omega) = \xi_{s,t}^n(y)(\omega)$, todo $n$, $s \le t < T_n(s,x,\omega), y \in U_n$, $\omega \in \Omega'$. Usando agora que $\xi_{s,t}(x) = \xi_{u,t}(\xi_{s,u}(x))$ com $u = T_n(s,x)$, tem-se que $\xi^e_{s,t}(x)(\omega)$ não saí de $M$ se $\omega \in \Omega'$, $x \in M$.

Conseqüentemente, $\xi^e_{s,t}(x)$ é solução de $(S)$ e como é solução maximal de $(S^e)$, é também solução maximal de $(S)$ e como tal será denotada por $\xi_{s,t}(x)$.

Pelo teorema 3.4, $(t, \xi_{s,t}(x))$ se acumula em fronteira de $[a,b)\times U^e$ quanto $t \to T(s,x)$ se $T(s,x) < b$. Como $M \subset U^e$, temos que se $T(s,x) < b$, $\lim\limits_{t \to T(s,x)} \xi_{s,t}(x) = \infty$, isto é, $\xi_{s,t}(x)$ sai de qualquer compacto de $M$ conforme $t \to T(s,x)$ (o $\infty$ neste limite é o da compactificação de $M$ por um ponto). Em particular, se $M$ é compacto, $T(s,x) = b$ todo $s, x$, q.s..

O domínio de $\xi_{s,t}(\omega)$ é

$$D_{s,t}(\omega) = \left\{x \in M: \ t < T(s,x)\right\}$$

Como $\xi^e_{s,t}(\omega)$ é um homeomorfismo sobre sua imagem em $U^e$ e $M$ é subvariedade mergulhada, temos que $D_{s,t}(\omega) = M \bigcap \{x \in U^e : t < T(s,x)\}$ e por restrição $\xi_{s,t}(\omega) = D_{s,t}(\omega) \to I_{s,t}(\omega)$ é um homeoformismo com $I_{s,t}(\omega)$ aberto de $M$.

Da mesma forma, a propriedade de difeomorfismo de $\xi_{s,t}(\omega)$ é conseqüência do fato de que a restrição de difeomorfismos a subvariedades mergulhadas são difeomorfismos entre abertos da subvariedade.

Em resumo, temos o

**Teorema 6.5:** Seja $(S)$ definida em $[a,b) \times M$ e satisfazendo $(H)$. Então tem-se existência e unicidade de soluções, modificações contínuas, diferenciáveis, etc..., como no teorema 6.1.

Se $M$ é compacta para cada $(s,x), T(s,x) = b$ q.s.

Vamos definir agora equações de Itô em variedades diferenciáveis. A primeira observação a se fazer é que como a fórmula de Itô (6.4) para soluções de equações de Itô depende da escolha de uma sistema de coordenadas, o drift de uma equação de Itô numa variedade não pode ser considerado como um campo de vetores. Isto porque ao se fazer uma mudança de coordenadas, o drift de $(I)$ não muda como um campo de vetores. De fato, suponhá dados dois sistemas de coordenadas $x = (x^1, \ldots, x^d) : U \subset M \to \mathbb{R}^d$ e $y = (y^1, \ldots, y^d) : U \to \mathbb{R}^d$ ligados mutuamente por $y = \psi(x)$. Denote por $\partial/\partial x^i$ e $\frac{\partial}{\partial y^i}$ respectivamente os campos canônicos associados à $x$ e $y$ (quando escrito no sistema de coordenadas $x$, $\frac{\partial}{\partial x^i}$ é o campo constante na direção do $i$-ésimo vetor considerado; o mesmo com $\partial/\partial y^i$ na coordenada $y$). Seja $(I)$ escrito em coordenadas $(x^1, \ldots, x^d)$ por

$$dx = X\big(t, x\big)dt + Y\big(t, x\big)dw \qquad (I_x)$$

Escrevendo $X$ e $Y$ em coordenadas como $X = (X^1, \ldots, X^d), Y = (Y^1, \ldots, Y^d)$, seja $\xi_t$ solução de $(I_x)$. Por (6.4),

$$d\big(\varphi(\xi_t)\big) = \left(X + \frac{1}{2}\sum_{i,k} \frac{\partial^2 \varphi}{\partial x^i \partial x^k} Y^i Y^k\right)(t, \xi_t)\ dt + Y(t, \xi_t) dW$$

onde $\frac{\partial^2 \varphi}{\partial x^i \partial x^k}$ é derivada parcial coordenada a coordenada.

Daí que $(I_n)$ quando escrito no sistema de coordenadas $y$ é a equação de Itô

$$dy = \tilde{X}(t, y)dt + \tilde{Y}(t, y)dW \qquad (\mathrm{I_y})$$

onde, se escrermos $\tilde{X} = (X^1, \ldots, X^d); \tilde{Y} = (Y^1, \ldots, Y^d)$ em coordenadas em relação à $y$, temos

$$\tilde{Y}^i = \sum_{k=1}^{d} \frac{\partial y^i}{\partial x^k} Y^k$$

$$X^i = \sum_{k=1}^{d} \frac{\partial y^i}{\partial x^k} Y^k + \frac{1}{2} \sum_{k,l=1}^{d} \frac{\partial^2 y^i}{\partial x^k \partial x^l} Y^k Y^l.$$

Estas espressões mostram que enquanto os coeficientes difusos de uma equação de Itô se comportam como campos de vetores (isto é: mudam como campos de vetores ao se mudar o sistema de coordenadas), a deriva não é verdadeiramente um campo de vetores mas sim um objeto cuja fórmula de mudança de coordenadas envolve derivadas de 2ª ordem da aplicação que realiza a mudança de coordenadas.

Para introduzir equações de Itô em $M$, é necessário considerar não apenas sua estrutura diferenciável (como no caso das equações de Stratronovich) mas introduzir além do mais uma conexão afim em $M$. Lembremos que uma conexão afim em $M$ é uma aplicação que associa a pares de campos de vetores $(X, Y)$ em $M$ um campo de vetores $\nabla_X Y$. Esta aplicação é bilinear quando se considera multiplicações (escalares) por funções, a menos da relação

$$\nabla_X(fY) = f \nabla_X Y + (Xf)Y \tag{C}$$

com $f : M \to \mathbb{R}$ função e $X, Y$ campos de vetores ($Xf$ é derivada direcional) (veja [K-N], [C], para maiores detalhes). O exemplo mais simples e intuitivo de conexão é dado pela conexão plana, isto é, a derivada direcional em $\mathbb{R}^d$: Seja $\mathbb{R}^d$ com coordenadas $x = (x^1, \ldots, x^d)$. Então a expressão

$$\nabla_{\partial/\partial x^i} \partial/\partial x^i = 0 \tag{6.8}$$

define uma conexão em $\mathbb{R}^d$ que, para campos

$$X = X^1 \frac{\partial}{\partial x^1} + \cdots + X^d \frac{\partial}{\partial x^d} \quad \text{e} \quad Y = Y^1 \frac{\partial}{\partial x^1} + \cdots + Y^d \frac{\partial}{\partial x^d}$$

arbitrários é dada – em virtude de $(C)$ – por

$$\nabla_X Y = \sum_{i,j}\Big(X^i \frac{\partial y^j}{\partial x^i}\Big)\frac{\partial}{\partial x^j} = \sum_j \Big(XY^j\Big)\frac{\partial}{\partial x^j} = D_X Y. \tag{6.9}$$

Tendo uma conexão, o drift de (I) em $M$ será o operador de 2ª ordem

$$L = X + \sum_{j=1}^{m} Y_j^2 - \frac{1}{2}\ \sum_{j=1}^{m} \nabla_{Y_j} Y_j \tag{6.10}$$

onde, como anteriormente, $Y_j^2 f = Y_j(Y_j f)$ é um operador de 2ª ordem. De maneira especifica,

**Definição 6.6:** Seja $M$ de classe $C^k, k \geq 3$. Uma equação de Itô em $M$ se constitui de campos de vetores $X, Y_1, \ldots, Y_m, m \geq 0$ em $M$ e uma conexão afim $\nabla$. Uma solução desta equação é um processo $\xi_t$ em $M$ satisfazendo i) - iii) da definição 1.2 e

iv) para $F : M \to \mathbb{R}$ de classe $C^2, F(\xi_t)$ é um semi-martingale e

$$F(\xi_t) = F(x) + \int_s^t (LF)(u, \xi_u)du + \sum_{j=1}^{m} \int_s^t Y_j(u, \xi_u)dW_u^j \tag{6.11}$$

onde $L$ é o operador definido em (6.10).

Para esclarecer a relação entre esta definição e os comentários anteriores, fazemos as seguintes observações (nas quais tomamos $m = 1$ por simplicidade de notação):

a) (6.11) é uma maneira de se escrever (6.4) independentemente de sistemas de coordenadas. No caso em que $M = \mathbb{R}^d$ e $\nabla$ é a conexão plena definida em (6.9), (6.11) se reduz à (6.4). Isto porque se $Y = \sum Y^i \dfrac{\partial}{\partial x^i}$ então

$$Y^2 = \sum_{i,k} Y^i Y^k \frac{\partial^2}{\partial x^i \partial x^k} + \sum_{i,k} Y^i \frac{\partial y^k}{\partial x^i} \frac{\partial}{\partial x^k} = \sum_{i,k} Y^i Y^k \frac{\partial^2}{\partial x^i \partial x^k} + D_Y Y \qquad (*)$$

e daí que

$$L = X + \frac{1}{2} \sum_{i,k} Y^i Y^k \frac{\partial^2}{\partial x^i \partial x^k}$$

pois neste caso $\nabla_Y Y = D_Y Y$.

b) Fixando um sistema de coordenadas $x = (x^1, \cdots, x^d)$, por (6.11) a equação de Itô em $M$ define em $\mathbb{R}^d$ a equação $dx = Z(x)dt + Y(x)dW$ onde $Z = \mathbb{R}^d \to \mathbb{R}^d$ é dada por $Z = (z^1, \cdots, z^d)$ com $z^i = Lx^i$ e $Y$ é o próprio campo, coeficiente difuso. A fórmula de Itô (6,4) fornece – para $F : \mathbb{R}^d \to \mathbb{R}, C^2$,

$$d(F(x)) = \left( ZF + \sum_{i,k} \frac{\partial^2 F}{\partial x^i \partial x^k} Y^i Y^k \right) dt + YFdW$$

Verefiquemos que o drift desta expressão coincide com LF. Pela igualdade $(*)$ na observação a) acima,

$$z^i = Lx^i \;= \left(X - \tfrac{1}{2}\nabla_Y Y\right)x^i + \tfrac{1}{2}Y^2 x^i$$

$$+\left(X - \tfrac{1}{2}\nabla_Y Y\right)x^i + \tfrac{1}{2}\textstyle\sum_k Y^k \dfrac{\partial y^i}{\partial x^k}$$

pois $\dfrac{\partial^2 x^i}{\partial x^k \partial x^l} = 0$. Daí que (indicando coordenadas por superíndices).

$$\begin{aligned}
F &= \sum_i Z^i \frac{\partial F}{\partial x^i} \\
&= \sum_i \Big(X - \frac{1}{2}\nabla_Y Y\Big)^i \frac{\partial F}{\partial x^i} + \frac{1}{2} \sum_{i,k} Y^k \frac{\partial Y^i}{\partial x^k} \frac{\partial F}{\partial x^i} \\
&= XF - \tfrac{1}{2}\Big(\nabla_Y Y\Big)F + \tfrac{1}{2}\textstyle\sum_{i,k} Y^k \dfrac{\partial Y^i}{\partial x^k} \dfrac{\partial F}{\partial x^i}\ .
\end{aligned}$$

Por $(*)$ novamente, temos então que

$$\begin{aligned}
ZF + \frac{1}{2}\sum_{i,k} \frac{\partial^2 F}{\partial x^i \partial x^k}\, Y^i Y^k &= XF- \\
\frac{1}{2}\Big(\nabla_Y Y\Big)F + \frac{1}{2}\, Y^2\, F & \\
&= LF\ .
\end{aligned}$$

Usando esse fato com $F = y^i$ onde $y = (y^1, \cdots, y^d)$ é outro sistema de coordenadas, vemos que (6.11) define em diferentes sistemas de coordenadas em $M$ equações de Itô equivalentes.

Em virtude dessa última observação, os resultados de existência e unicidade de (I) em $M$ seguem de maneira similar que para as equações de Stratonovich. Deixamos os detalhes disto para o leitor. Deixemos também a verificação de que, uma vez fixado $\nabla$,

$$dx = Xdt + Y \circ dW$$

é equivalente à equação de Itô com drift $\tilde{X} = X - \frac{1}{2}\sum_{j=1}^{m} \nabla_{Y_j} Y_j$ e mesmos coeficientes de difusão.

**Exemplos 6.7:** a) Os sistemas lineares do exemplo 1.6 b) são escritos também em forma de Strotonovich

$$dx = A(t)xdt + \sum_{j=1}^{m} b_j(t) \circ dW^j$$

A equação de Itô correspondente tem os mesmos coeficientes já que $D_{Y_j} Y_j = 0$ pois $Y_j(t,x) = b_j(t)$ não depende de $x$.

b) As equações bilineares de Stratonovich são da forma

$$dx = A(t)xdt, + \sum_{j=1}^{m} B_j(t)x \circ dW^j$$

A equação de Itô correspondete é dada por

$$dx = \Big(A(t) + \frac{1}{2} \sum_{j=1}^{m} B_j(t)^2\Big)xdt + \sum_{j=1}^{m} B_j(t)x\, dW^j$$

Como no exemplo 1.6 c), os sistemas bilineares também admitem soluções fundamentais, isto é, as soluções com condições iniciais $(s,x)$, $x \in \mathbb{R}^d$ são dadas pela solução com condição inicial $(s,1)$ da equação associada em $M_{d\times d}$. Este exemplo é um caso particular dos exemplos seguintes.

c) Seja $G$ um grupo de Lie e $X, Y_1, \cdots, Y_m$ campos invariantes à direita em $G$ (por simplicidade, assumimos que são independentes de $t$). Defina

$$dg = X(g)dt + \sum_{j=1}^{m} Y_j(g) \circ dW^j \quad g \in G \tag{6.12}$$

Como os campos $X, Y_j$ são $C^\infty$ (na verdade analíticos), soluções maximais existem. Denote por $1$ a identidade em $G$ seja $g_t$ a solução com $g_0 = 1$. Tome $h \in G$ arbitrário e $F : G \to I\!R$ de classe $C^3$. Seja $R_h = G \to G$ a translação à direita $R_h(g) = gh$. Pela fórmula de Itô (6.3),

$$\begin{aligned} F(g_t h) &= F o R_h(g_t) + \sum_{j=0}^{m} \int_0^t \Big(Y_j(F o R_h)\Big)(g_s) \circ dW_s^j \\ &= F(h) + \sum_{j=0}^{m} \int_0^t (Y_j F)(g_s h) \circ dW_s^j \end{aligned}$$

pois $Y_j$ é invariante à direita e portanto $Y_j(F o R_h) = (Y_j F) o R_h$. Daí que a solução $\varphi_t(h)$ com condição inicial $(0, h)$ é $\varphi_t(h) = g_t h = L_{g_t}(h)$ onde $L_{g_t}$ é a translação à esquerda. O difeomorfismo $\varphi_t$ se confunde portanto com $g_t$.

Seja $T$ o tempo de explosão de $g_t$. Então $T = \infty$ q.s. De fato, fixe uma vizinhança $U$ de identidade de $G$ e defina por indução a seqüência de tempos de parada.

$$\begin{aligned} T_1 &= \inf \Big\{t > 0 : g_t \notin U\Big\} \\ T_n &= \inf \Big\{t > t_{n-1} : g_t g_{T_{n-1}}^{-1} \notin U\Big\} \end{aligned}$$

com $T_n = \infty$ se o conjunto é vazio. Então $T \geq T_n$ todo $n$ pois $g_t$ está definido para $t < T_n$. Mostremos que $T_n \to \infty$ q.s.. Seja $\theta_{T_n}$ o shift $\theta_{T_n}(\omega)(t) = \omega(t + T_n(\omega)) - \omega(T_n(\omega))$. Pela fórmula do fluxo (4.15, $g_{t+T_{n-1}}(\omega) = g_t(\theta_{T_{n-1}}(\omega))g_{T_{n-1}}(\omega)$, de onde se vê que $T_n - T_{n-1} = T_1 \cdot \theta_{T_{n-1}}$. Segue então da propriedade de Markov forte ([I-W]) teo. II.6.4) que a sequencia de variáveis aleatórias $T_1, T_2 - T_1, \ldots, T_n - T_{n-1}, \ldots$ é independente e identicamente distribuida. Agora $E[T_1] > 0$ pois $T_1 > 0$ q.s., portanto pela lei forte dos grandes números, $T_n = (T_n - T_{n-1}) + \ldots + (T_2 - T_1) + T_1 \to \infty$ q.s.. Consequentemente $T = \infty$ q.s. e pelo que foi comentado acima, todas as soluções têm tempo de explosão $\infty$.

d) Seja $G$ um grupo de Lie $M$ uma variedade diferenciável na qual $G$ atua por difeomorfismos $x \to gx, x \varepsilon M, g \varepsilon G$. Seja $X$ um campo de vetores invariante à direita em $G$. Denotando por $e^X$ a aplicação exponecial em $G$, a expressão

$$\tilde{X}(x) = \frac{d}{dt}(e^{tX}x)_{/t=0}$$

define um campo de vetores em $M$. A trajetória do campo $X$ que passa pela identidade de $G$ (i.e. $\varphi^{tX}$) é uma solução fundamental para o fluxo $\varphi_t$ associado ao campo $\tilde{X}$, isto é, $\varphi_t(x) = e^{tX}$.

Da mesma forma, a solução $g_t$ do exemplo anterior é uma solução fundamental de

$$dx = \tilde{X}(x)dt + \sum_{j=1}^{m} \tilde{Y}_j(x) \circ dW^j \tag{6.13}$$

em $M$, onde $X, Y_1, \cdots, Y_m$ são como em (6.11). Isto porque se $F : M \to R$ é de classe $C^3$ então para $x \in M$, a função $F_x : G \to F(gx)$ é de classe $C^3$ e

$$F(g_t x) = F(x) + \sum_{j=0}^{m} \int_0^t (Y_j F_x)(g_s) \; o \; dW_s^j$$

Mas pelo fato de que o fluxo de $\tilde{X}$ é $e^{tX}$, $Y_j F_x = (\tilde{Y}_j F)_x$ portanto

$$F(g_t x) = F(x) + \sum_{j=0}^{m} \int_0^t \tilde{Y}_j F(g_s x) \; o \; dW_S^j$$

isto é, soluções de (6.13) são da forma $g_t x$ com $g_t$ como no exemplo anterior. Em particular, seus tempos de explosão são iguais a $\infty$.

Casos particulares de (6.13), são os sistemas bilineares em que $G = Gl(d, \mathbb{R})$ e $M = R^d$ e os sistemas lineares em que $G$ é o grupo das transformações afins de $R^d$ e $M = R^d$.

e) Ao invés de uma ação global como no exemplo anterior, seja $\pi = U \subset G \times M \to M$ uma ação local. Isto é, $U$ é um aberto de $G \times M$ com $\{1\} \times M \subset U$ e tal que as seções $U_x = g \varepsilon G = (g, x) \varepsilon U$ sejam conexos e $\pi$ é diferenciável e satisfaz

$$\pi(gh, x) = \pi\big(g, \pi(h, x)\big) \qquad \text{(quando estiver definido)}$$

$$\pi(1, x) = x.$$

As aplicações $x \varepsilon M \to gx = \pi(g, x) \varepsilon M, g \varepsilon G$ definem difeomorfismos entre abertos de $M$ e tem-se $(gh)x = g(hx)$ quando ambos os lados estiverem definidos.

Da mesma forma que em d), (6.13) define uma equação em $M$ tendo $g_t$ como solução fundamental. A diferença é que o tempo de explosão de $g_t x$ é o primeiro tempo em que $g_t$ atinge a fronteira de $U_x$.

Um exemplo desta situação é dado por

f) Equações de Ricatti: Seja $G = Gl(2, \mathbb{R})$, $M = \mathbb{R}$ e seja $\pi$ definida por

$$\phi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, x\right) = \frac{a\,x + b}{c\,x + d}$$

com $U = \left\{\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) x\right) : cx + d > 0\right\}$. Os campos invariantes à direita são da forma $X(g) = Ag$ com $A$ matriz $2 \times 2$. Pela definição de $\tilde{X}$, temos de maneira equivalente que

$$\tilde{X}(x) = \left(d\theta_x\right)_1(A)$$

onde $\theta_x : g \varepsilon G \rightarrow gx \;\varepsilon\; M$ e $(d\theta_x)_1$ indica sua diferencial na identidade. Daí que, deferenciando (6.11) em relação à $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ temos que se $X(g) = Ag$ com $A = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \alpha \end{smallmatrix}\right)$ então $\tilde{X}$ é o campo associado à equação de Ricatti

$$\dot{x} = (\alpha - \delta)x + \beta \; - \gamma\, x^2 \; .$$

(6.13) se escreve então como

$$dx = \left((\alpha - \delta)x + \beta - \gamma x^2\right)dt + \sum_{j=1}^{m} \left((\alpha_j - \delta_j)x + \beta_j - \gamma_j x^2\right) \circ dW^j$$

com $X(g) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ g e $Y_j(g) = \begin{pmatrix} \alpha_j & \gamma_j \\ \beta_j & \delta_j \end{pmatrix} g$

O tempo de explosão da solução $g_t x$ é

$$\inf\Big\{t > 0 : c_t x + d_t = 0\Big\}$$

onde escrevemos $g_t = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix}$

## 4.7 Fluxos estocásticos.

Vamos olhar aqui, com mais detalhes, os fluxos gerados por equações diferenciais estocásticas. A idéia de fluxo ficou implicita, por exemplo, no teorema 4.11. Iniciamos definindo formalmente o que se entende por fluxo estocástico em uma variedade $M$, a qual supomos que satisfaz as condições usuais. Como anteriormente, o espaço de probabilidade $(\Omega, \mathcal{F}, P)$ é o espaço de Wiener.

**Definição 7.1:** Um campo aleatório $\xi_{s,t}(x,\omega)(= \xi_{s,t}(\omega)(x)$ etc...) $0 \leq s \leq t < \infty, x \,\varepsilon M$ a valores em $M$ é um fluxo estocástico Browniano se

i) $\xi_{s,t}(x,\omega)$ é contínuo em $(s,t,x)$ q.s.

ii) $\xi_{s,s}$ = identidade q.s.

iii) $\xi_{s,t}(x) = \xi_{u,t}(\xi_{s,u}(x))$, todo $x, s \leq u \leq t$ q.s.

iv) Para $0 \leq t_0 < t_1 \ldots < t_k, x_0, \ldots, x_{k-1} \,\varepsilon M$

$$\xi_{t_0,t_1}(x_0), \ldots, \xi_{t_{k-1},t_k}(x_{k-1})$$

são independentes.

v) $\xi_{s,t}(\omega)$ é um homeomorfismo (global) de $M$ q.s., isto é, $P(\omega : \xi_{s,t}(\omega)$ é homeomorfismo todo $s \leq t) = 1$.

O fluxo é de classe $C^k$ se

vi) $P(\omega : \xi_{s,t}(\omega)$ é difeomorfismo $C^k$ todo $s \leq t) = 1$

**Observações:** a) Um fluxo propriamente dito deve satisfazer i), iii) e v) (ou vi)). O termo Browniano se refere á iv) (veja [K2])

b) No caso em que $\xi_{s,t}(\omega)$ é constante como função de $\omega$, um fluxo estocástico se reduz a um fluxo determinístico.

c) Ao invés de em $[0, \infty)$ um fluxo pode ser definido também em $[0, b)$, $b > 0$.

d) Com $S$ fixo, $\omega \to \xi_{s,t}(\omega)$ define um processo a valores no grupo dos homeomorfismos de $M$ (difeomorfismos se se tem vi)).

**Exemplo 7.2:** Na situação do exemplo c) do § 6, seja $g_t$ o processo a valores em $G$, que é um grupo de homeomorfismos de $M$. Colocando $\xi_{s,t}(\omega) = g_t(\theta_s(\omega)), \xi_{s,t}$ é um fluxo estocástico Browniano em $M$ pois $g_t$ é solução de uma equação de Stratonovich. A propriedade iv) é consequência da $\mathcal{F}_t$ - mensurabilidade de $g_t$. Como foi verificado, $g_t$ está definido para todo $t \geq 0$ e daí que $\xi_{s,t}$ é um fluxo em $[0, \infty)$.

Sobre a ligaçãó entre fluxos estocásticos e equações estocáticas, uma das questões é saber se um fluxo estocástico pode ser obtido de uma equação estocástica, no sentido do teorema 4.11 ou do exemplo acima. No caso determinístico, quando se tem defeenciabilidade, o fluxo é solução de uma equação deferencial $\dot{x} = X(t, x)$ que é obtida por diferenciação do fluxo $X(t,x) = \frac{d}{ds}(\xi_{t,s}(x))\big|_{s=0}$. No caso estocástico, um fluxo nem sempre é obtido de uma equação do tipo $(S)$ (ou $(I)$ em $\mathbb{R}^d$). Isto porque viemos considerando um caso específico de equações estocásticas. O que é possível mostrar (e para isto nos referimos a [Bx], [LJ] ou [K2] é que todo o fluxo é obtido de uma equação estocástica se ao invés de $(S)$, que é uma equação sobre $m$ movimentos Brownianos, tomarmos equações sobre uma quantidade infinita de movimentos Brownianos, isto é, movimentos Brownianos em espaços funcionais.

A reciproca desta questão, que é a de se garantir que as soluções de $(S)$ ou $(I)$ são fluxos estocásticos, vai nos ocupar pelo resto desta seção. Em vista do teorema 4.11 (6.1, 6.5), as soluções de $(S)$ geram um fluxo, como na definição acima, no caso em que o domínio $D_{s,t}(\omega)$ e a imagem $I_{s,t}(\omega)$ de $\xi_{s,t}(\omega)$ coincidem com $M$ q.s. . Em contraposição ao caso determinístico, pode ocorrer que $D_{s,t}(\omega) = M$ todo $s,t$ e não se ter no entanto a sobrejetividade de $\xi_{s,t}(\omega)$. A discussão detalhada disto, passa antes por uma análise de equação que rege o fluxo inverso. A equação das inversas de um fluxo determinístico é essencialmente mesma que a que rege o fluxo. Por exemplo, no caso antonomo, se $\varphi_t$ é gerado por $\dot{x} = X(x)$ então $\varphi_t^{-1} = \varphi_{-t}$ e $\varphi_t^{-1}$ é gerado por $x = -X(x)$, isto é, as trajetórias de $-X$ simplesmente alteram o sentido em que se percorre as trajetórias de $X$. Se $X$ é campo completo, isto é se $\varphi_t$ é globalmente definido todo $t \varepsilon R, \varphi_t^{-1}$ também é globalmente definido.

Para equações estocásticas, a equação que rege as inversas das soluções, também é obtida tomando-se os campos que a definem em sentido contrário, sendo necessário, no entanto, um cuidado na escolha do caminho aleatório específico das soluções da equação de retorno que fornece a inversa $\xi_{s,t}^{-1}(\omega)$ da solução $\xi_{s,t}(\omega)$ de $(S)$. Os seguintes argumentos heuristicos devem esclarecer o que está em jogo nesta escolha.

Começando com

$$dx = X(x)dt + Y(x) \circ dW \tag{S}$$

dividimos esta equação por $dt$. Como os caminhos aleatórios de $W$ não são (com probabilidade 1) diferenciáveis, $dW/dt$ não faz sentido como derivada de cada caminho $W_t(\omega)$. Escrevemos no entanto, formalmente $\circ\, dW/dt = u(t)$ e obtemos a família de equações deferencíaveis ordinárias

$$\dot{x} = X(x) + u(t)Y(x) \qquad t \geq 0 \tag{SC}$$

$(SC)$ é o que se denomina um sistema de controle. Cada <u>controle</u> $u : [0, \infty) \to \mathbb{R}$ (que seja p. ex. localmente integrável para se garantir a existência de soluções), define uma solução $\varphi_t(u)(x)$ da equação escrita em $(SC)$. As trajetórias de $(SC)$ são as soluções $\varphi_t, t \geq 0$ para os diferentes controles. Estas trajetórias aproximam os caminhos aleatórios das soluções $\varphi_t(\omega)$ de $(S)$ (veja os teoremas de aproximação e suporte em [I-W] [K3]). Fixando um controle $u, \varphi_t(u)$ é um difeomorfismo (local) em $M$. Para se obter sua inversa $\varphi_t^{-1}(u)$, como solução de uma equação diferencial é necessário percorrer as trajetórias $t \to \varphi_t(u)(x)$ em sentido contrário. Por isso, $\varphi_t^{-1}(u)$ satisfaz

$$\dot{x} = -X(x) + v(s)\big(-Y(x)\big)$$

onde $v(s) = u(t-s), 0 \leq s \leq t$. Dito de outra maneira, o sistema de controle associado aos campos $-X$ e $-Y$ fornecem as inversas dos fluxos associados à $(SC)$ desde que os controles sejam tomados em sentido contrário no tempo.

Da mesma forma, a equação que rege as inversas do fluxo $\xi_{s,t}(\omega)$ associado à $(S)$ será dada pelos campos $-X, -Y$, desde que se tome $\omega$ em "sentido contrário no tempo".

A formalização deste "sentido contrário no tempo" se faz por intermédio do movimento Browniano de retorno e as correspondentes integrais estocásticas e equações diferenciais estocásticas de retorno. Este movimento Browniano é construido da seguinte maneira: Sendo $(\Omega, P)$ o espaço de Wiener, seja como anteriormente $\mathcal{F}_{s,t} = \sigma\{W_t - W_s\}$ $0 \leq s \leq t$. $\mathcal{F}_{s,t}$ é decrescente com $s$ ou, dito de outra maneira, $u\varepsilon[0,t] \to \mathcal{F}_{t-u,t}$ cresce com $u$. Deixamos ao leitor a verificação de que $u\varepsilon[0,t] \to \widehat{W}_u = W_t - W_{t-u}$ é um $\mathcal{F}_{t-u,t}$ movimento Browniano (use a caracterização do movimento Browniano por martingales e o fato de que $E[W_s - W_{s_1}/\mathcal{F}_{s,t}] = 0$ se $s_1 < s < t$). Uma vez verificado isto, para processos $u\varepsilon[o,t] \to f_u$ para os quais a integral de Itô faz sentido, vamos denotar $\int_0^u f_\nu d\widehat{W}_\nu$ por $\int_s^t f_\nu d\widehat{W}_\nu$ onde $s = t - u$. O mesmo para integrais de Stratonovich. Estas são as integrais de retorno. Elas são, como as consideradas anteriormente, integrais sobre movimentos Brownianos. O uso da realização $\widehat{W}_u$ para o movimento Browniano é conveniente no momento em que se vai escrever a equação satisfeita pelo fluxo inverso.

Com esta terminologia, vejamos o exemplo do fluxo inverso de $dx = dW$. Como no exemplo 1.6a) (tomando $\varphi_t = \xi_{0,t}$),

$$\varphi_t(\omega)(x) = x + \omega(t)$$

Para traçar a trajetória $s\ \varepsilon[0,t] \to x + \omega(s)$ em sentido contrário, colocamos $u\varepsilon[0,t] \to x + \omega(t) - \Big(\omega(t) - \omega(t-u)\Big) = x + W_t(\omega) - \widehat{W}_u(\omega)$. Daí que se $\widehat{\varphi}_u(\omega)(x)$ é solução de $dx = -d\widehat{W}$, $\varphi_u(\omega)\ o\varphi_t(\omega)(x) = \varphi_{t-u}(\omega)(x) 0 \leq u \leq t$, isto é, $\varphi_t(\omega) = \varphi_t^{-1}(\omega)$ (note que são ambos avaliados no mesmo $\omega\varepsilon\Omega$).

**Teorema 7.3:** Tomando $(S)$ como no teorema 6.5 (ou 6.1) com domínio $U = [0,\infty) \times M$, a inversa $\xi_{s,t}^{-1}(\omega)$ de $\xi_{s,t}(\omega), s \leq t$, satisfaz para todo $F$ de classe $C^3$,

$$F\Big(\xi_{s,t}^{-1}(x)\Big) = F(x) - \sum_{j=0}^{m} \int_s^t Y_j F\Big(u, \xi_{u,t}^{-1}(x)\Big)\ o\ d\widehat{W}_u \tag{7.1}$$

isto é, $\xi_{s,t}^{-1}(x)$ satisfaz a equação de retorno

$$dx = -\sum_{j=0}^{m} Y_j(t,x) \; o \; d\widehat{W}^j \tag{7.2}$$

<u>dem</u>: Sejam os campos aleatórios

$$\eta_j(y) = \int_s^t Y_j F\Big(u, \xi_{s,u}(y)\Big) \; o \; dW^j, j = 0, \cdots, m \; .$$

Mostremos que para $x$ fixado,

$$\eta_j\Big(\xi_{s,t}^{-1}(x)\Big) = \int_s^t Y_j F\Big(u, \xi_{u,t}^{-1}(x)\Big) \; o \; d\big(\widehat{W}\big)_u^j$$

onde $o \; d\big(\widehat{W}\big)^0 = dt$. Se $j = 0$, $\int_s^t XF(u, \xi_{s,u}(y))du$ é aproximada por somas do tipo

$$\sum_{i=0}^{k} XF\Big(t_i, \xi_{s,t_i}(u)\Big)\Big(t_{i+1} - t_i\Big)$$

com $t_0, \ldots, t_k$ partição de $[s,t]$. Avaliando em $y = \xi_{s,t}^{-1}(x)$, temos que

$$XF\Big(t_i, \xi_{s,t_i}\Big(\xi_{s,t}^{-1}(x)\Big) = XF\Big(t_i, \xi_{t_i,t}^{-1}(x)\Big)$$

pois

$$\xi_{s,t}^{-1} = \xi_{s,t_i}^{-1} \; o \; \xi_{t_i,t}^{-1}.$$

Portanto $\eta_j(\xi_{s,t}^{-1}(x))$ é aproximada por somas do tipo

$$\sum_{i=0}^{k} XF\Big(t_i, \xi_{t_i,t}^{-1}(x)\Big)$$

e daí a igualdade para $j = 0$. Para $j \geq 1$, $\int_s^t Y_j F\Big(u, \xi_{s,u}(y)\Big)$ $o$ $dW^j$ é aproximada por somas do tipo

$$\frac{1}{2}\sum_{i=0}^{k}\Big(Y_jF(t_i, \xi_{s,t_i}(y)\Big) + Y_jF\Big(t_{i+1}, \xi_{s,t_i+1}(y)\Big)\Big(W_{t_i+1}^j - W_{t_i}^j\Big)$$

Portanto $\eta_j\Big(\xi_{s,t}^{-1}(x)\Big)$ é aproximado por

$$\frac{1}{2}\sum_{i=0}^{k}\Big(Y_jF(t_{i+1}, \xi_{t_i+1}^{-1}(x)\Big) + Y_jF\Big(t_i, \xi_{t_i,t}^{-1}(x))\Big)\Big(\widehat{W}_{t-t_i}^j - \widehat{W}_{t-t_{i+1}}^j\Big)$$

Daí que

$$\eta_j\Big(\xi_{s,t}^{-1}(x)\Big) = \int_s^t Y_jF\Big(u, \xi_{u,t}^{-1}(x)\Big)\ o\ d(\widehat{W})_u^j\ .$$

Devido a estas igualdades, pondo $y = \xi_{s,t}^{-1}(x)$ em

$$F\Big(\xi_{s,t}(y)\Big) = F(y) + \sum_{j=0}^{m}\int_s^t Y_jF\Big(u, \xi_{s,u}(y)\Big)\ o\ d(W)_u^j$$

temos

$$F(x) = F\Big(\xi_{s,t}^{-1}(x)\Big) + \sum_{j=0}^{m} \int_s^t Y_j F\Big(u, \xi_{u,t}^{-1}(x)\Big) \; o \; d(\tilde{W})_u^j$$

e portanto o teorema.

**Observação:** Para tomar a equação de retorno em forma de Itô, passa-se à Stratonvich, faz-se retorno e volta à Itô. Por exemplo, em $\mathbb{R}^d$ com $m = 1$, $Xdt + YdW$ é equivalente à $(X - \frac{1}{2}\, D_Y Y)dt + Y \; o \; dW$ cuja equação de retorno é $(-X + \frac{1}{2}\, D_Y Y)dt - Y \; o \; d(\widehat{W})$. Voltando à Itô, a equação de retorno fica $(-X + D_Y Y)dt - Y d(\widehat{W})$.

A questão da globalização dos homeomorfismo $\xi_{s,t}(\omega)$ está ligada à completude da equação e da correspondente equação de retorno, isto é, à possibilidade de se extender $\xi_{s,t}$ à $t \varepsilon [s, \infty)$. Por isto, introduzimos a

**Definições 7.4:** a) Uma equação estocástica é dita conservativa (ou completa) se para todo $(s, x), T(s, x) = \infty$ q.s.

b) A equação é estritamente conservativa se $P\Big(T(s, x) = \infty \text{ todo } (s, x)\Big) = 1$

**Observações:** a) Nem sempre um sistema conservativo é estritamente conservativo (veja exemplo a seguir). A razão é a de sempre: existe uma quantidade não enumerável de conjuntos excepcionais. Em dimensão 1, os dois conceitos são equivalentes; comentaremos sobre isso adiante

b) A completude estrita é equivalente à que os domínios

$$D_{s,t}(\omega) = \Big\{x\varepsilon M : t < T(s,x)\Big\}$$

coincidam com $M$.

**Exemplo 7.5:** Seja $dx = dW$ em $\mathbb{R}^2 - 0$ (ou $R^d - 0$ com $d \geq 2$). Suas soluções são dadas pela restrição de

$$\xi_{s,t}(\omega)(x) = x + \omega(t) - \omega(s)$$

à $\mathbb{R}^2 - \{0\}$. O tempo de explosão é então

$$T(s,x) = \inf\Big\{t > s = x + \omega(t) - \omega(s) = 0\Big\}.$$

Isto é $T(s,x)$ é o primeiro tempo em que o movimento Browniano em $\mathbb{R}^2$ atinge o conjunto $\{0\}$. Como conjuntos unitários em $\mathbb{R}^2 - \{0\}$ são polares para movimento Browniano (veja [P-S]) para todo $s, x$, $T(s,x) = \infty$ q.s. e a equação é conservativa. Por outro lado, dado $\omega$ exitem $x$ e $s$ t.q. $x + \omega(t) - \omega(s) = 0$ algum $t < \infty$ (por exemplo $s = 1, t = 2$ e $x = \omega(2) - \omega(1)$). Daí que

$$\Big\{\omega : T(s,x)(\omega) = \infty \ \text{ todo } \ s,x\Big\} = \phi$$

e $dx = dW$ não é estritamente conservativa em $\mathbb{R}^2 - \{0\}$.

Em virtude do Teorema 7.3 e com os conceitos acima, temos:

**Teorema 7.6:** As soluções de $(S)$ em $M$ definem um fluxo estocástico Browniano (como na definição 7.1) se e só se $(S)$ e

$$dx = -\sum_{j=0}^{m} Y_i\Big(t,x\Big) \circ dW^j \qquad \text{(SR)}$$

São estritamente conservativos.

<u>dem:</u> É só observar que (7.2) é equivalente à (SR) já que $\widehat{W}$ é também um movimento Browniano.

Pelo teorema 7.3, a imagem $I_{s,t}(\omega)$ de $\xi_{s,t}(\omega)$ coincide com $M$ no caso que (SR) é estritamente conservativo.

Alguns casos em que se tem fluxos estocásticos (globais) gerados por equações estocásticos são:

1. Equações globalmente Lipschitz em $\mathbb{R}^d$. Pelo Teorema 4.9, tais equações são estritamente conservativas. A equação de retorno de uma L.g. também é L.g.

2. Equações invariantes à direita (ou à esquerda) em grupos de Lie. Como foi visto no exemplo 6.7c, as soluções são da forma $\varphi_t(\omega)(g) = g_t(\omega)g$ com $g_t(\omega) = \varphi_t(\omega)(1)$. Como o tempo de explosão é $\infty$, o sistema é estritamente conservativo; (SR) também é invariante à direita.

   O mesmo ocorre com equações obtidas por ações globais de $G$, como no exemplo 6.7d).

3. Como enunciado no teorema 6.5, se $M$ é compacta, $(S)$ é conservativa (assim como $(SR)$). Pelo método descrito para construção de soluções em M-via mergulhos – $(S)$ em $M$ é a restrição de uma equação em $\mathbb{R}^N$. Por ser $M$ compacta, essa última pode ser tomada L.g. e portanto equações em variedades compactas são estritamente conservativas.

4. Se $M$ é de dimensão 1, equações conservativas são estritamente conservativas. Neste caso, $M \approx S^1$ ou $\mathbb{R}^1$. No primeiro caso tem-se compacidade. Quando $M \approx \mathbb{R}^1$, $\xi_{s,t}(\omega) : \quad D_{s,t}(\omega) \to T_{s,t}(\omega)$ é estritamente crescente pois é um homermorfismo homotópico à $\xi_{s,s}(\omega) =$ identidade. Por tanto, fixando $s$ e tomando

$$\Omega' = \bigcap_{n \in \mathbb{Z}} \Big\{\omega : T(s,n)(\omega) = \infty\Big\}$$

vemos que para $\omega \in \Omega'$, $D_{s,t}(\omega) = \mathbb{R}^1$ e tem-se portanto completude estrita.

## 4.8 Processos de difusão.

Antes de introduzir formalmente os processos de difusão, voltemos à fórmula de Itô (6.5) que descreve as soluções de uma equação de Stratonovich em uma variedade $M$ em forma de integral de Itô. Supondo uma situação homogenea no tempo (como faremos ao longo de toda esta secção), temos para $f : M \to \mathbb{R}$,

$$\Big(\varphi_t(x)\Big) = f(x) + \int_0^t Lf\Big(\varphi_s(x)\Big)ds + \int_0^t Yf\Big(\varphi_s(x)\Big)dW_s$$

onde $L$ é um operador de segunda ordem. Tomando esperança em relação à medida de Wiener,

$$E\Big[f(\varphi_t(x))\Big] = \int_0^t E\Big[Lf\Big(\varphi_s(x)\Big)\Big]ds$$

e daí que se escrevermos

$$\Big(T_t f\Big) = E\Big[f\Big(\varphi_t(x)\Big)\Big],$$

$t \longrightarrow T_t$ define um semi-grupo (em algum espaço funcional convenientemente escolhido) e como temos

$$\frac{\partial}{\partial t}\Big(T_t f\Big)(x) = T_t(LF)(x), \tag{8.1}$$

o gerador infinitesimal de $T_t$ é o operador $L$, isto é $T_t$ é o semi-grupo associado à equação de evolução

$$\frac{du}{dt} = Lu$$

$$u_0 = f \ .$$

Estes comentários mostram a existência de uma ligação estreita entre as equações estocásticas e os operadores diferenciais de segunda ordem do tipo parabólico (como é o caso do operador $L$). Com o propósito de detalhar um pouco mais essa ligação, começamos por observar que para obter (8.1) não foi necessário o conhecimento completo das soluções da equação estocástica mas apenas de média de variáveis aleatórias do tipo $f(\varphi_t(x))$. A igualdade (8.1) depende então não das soluções como um todo mas apenas das leis (probabilidades em $M$) das aplicações aleatórias $\omega \to \varphi_t(x)(\omega)$. O conjunto destas leis é o que se entende por processo de difusão associado à equação estocástica.

Nesta seção vamos definir de maneira precisa o que são processos de difusão e construir o processo definido por uma equação diferencial. Em contraposição aos processos que são soluções de $(S)$, o processo de difusão depende apenas do operador $L$. Como pode ser visto a partir dos comentários acima, os processos de difusão são de interesse no estudo dos operadores diferenciais parabólicos (e em particular elipticos). Não vamos nos aprofundar aqui nesta direção. Para isto nos referimos à [Dy], [B-G] e [I-W]. Vamos, no entanto, mostrar como construir equações estocásticas (e portanto processos de difusão) associados aos operadores de Laplace-Beltrami em variedades Riemannianas.

Como preliminar à definição dos processos de difusão, tomemos $M$ uma variedade diferenciável. Sejam $\overline{M} = M \bigcup \{\infty\}$ sua compactificação por um ponto e $C$ o conjunto das curvas $\omega : [0, \infty) \to \overline{M}$. Fixando de antemão um tempo de explosão $T : C \to [0, \infty]$, denotemos que $W_T(M)$ o conjunto das curvas $\omega \in C$ tais que

a) $\omega(t) = \infty$ se $t \geq T(\omega)$

b) $\omega(t) \in M$ se $t \in [0, T(\omega))$ e $\omega : [0, T(\omega)) \to M$ é contínua.

Um conjunto cilindrico em $W_T(M)$ é um conjunto da forma

$$\Big\{\omega : \omega(t_1) \in A_1, \ldots, \omega(t_n) \in A_n\Big\}$$

onde $A_1, \ldots, A_n$ são barelianos em $\overline{M}$ e $0 \leq t_1 < t_2 < \ldots < t_n$. Denotemos por $C(M)$ a $\sigma$-algebra em $W_T(M)$ gerada pelos conjuntos cilindricos e por $C_t(M)$ a $\sigma$-algebra gerada por aqueles conjuntos cilindricos em que $t_n \leq t$.

**Definição 8.1:** Com as notações acima, um processo de Markov com tempo de explosão $T$ é uma família de probabilidades $P_x$, $x \in \overline{M}$ em $(W_T(M), C(M))$ satisfazendo

i) $P_x(\omega : \omega(0) = x) = 1$ todo $x \in \overline{M}$,

ii) Para $A \in C(M)$, a aplicação

$$x \in M \to P_x(A) \in [0, 1]$$

é Borel mensurável,

iii) Para $x \in M$, $0 \leq s < t$, $A \in C(M)$ e $B \subset \overline{M}$ boreliano,

$$P_x\Big(A \bigcap \{\omega : \omega(t) \in B\}\Big) = \int_A P_{\omega'(s)}\Big(\omega : \omega(t-s) \in B\Big) P_x(d\omega') \ . \qquad (8.2)$$

O processo é dito conservativo no caso em que $P_x(\omega : T(\omega) = \infty) = 1$ todo $x \in M$.

A respeito desta definição, fazemos as seguintes

**Observações 8.2:** a) Esta definição cobre apenas os processos de Markov homogêneos no tempo; que são iniciados em $t = 0$.
b) Para $x \in M$ e $t \geq 0$, defina para cada boreliano $B \subset M$,

$$P\Big(t, x, B\Big) = P_x\Big(\omega : \omega(t) \in B\Big) \ . \qquad (8.3)$$

A aplicação $B \to P(t, x, B)$ define uma medida sobre os borelianos de $M$. A menos que o processo seja conservativo, $P(t, x, \cdot)$ não é uma probabilidade em $M$. Mesmo assim, as medidas $P(t, x, \cdot)$ são denominadas probabilidades de transição do processo de Markov. Por aplicações sucessivas de (8.2), temos que se $0 \leq t_1 < \ldots < t_n$ então

$$P_x\Big(\omega : \omega(t_1) \in B_1, \ldots, \omega(t_n) \in B_n\Big) =$$

$$= \int_{B_1} P\Big(t, x, dx_1\Big) \int_{B_2} P\Big(t_2 - t_1, dx_2\Big) \ldots \int_{B_n} P\Big(t_n - t_{n-1}, x_{n-1}, dx_n\Big)$$

com $B_1, \ldots, B_n$ borelianos de $M$. De onde se vê que as probabilidade de transição caracterizam o processo de Markov.

c) Um processo de Markov não é exatamente um processo estocástico no sentido do capítulo I, mas uma família de processos cada um dos quais iniciados em $x \in M$.

A igualdade (8.2) é o que se denomina propriedade de Markov. Os processos de difusão são processos de Markov que satisfazem (8.2) não apenas para tempos determinísticos mas também para tempos de parada, isto é, satisfazem a propriedade forte de Markov. Aqui, tempos de parada são tomados em relação à seguinte filtração: Seja $D_t$ o completamento de $C_t(M)$ em relação à $P_x$, $x \in M$, e defina

$$C_{t+}(M) = \bigcap_{s>t} D_t \, .$$

A filtração $C_{t+}$ é contínua à direita. Como é usual, denotarmos por $C_\infty(M)$ a $\sigma$-algebra gerada por $C_{t+}, t \geq 0$ e para $T$ um $C_{t+}$–tempo de parada, colocamos

$$C_T = \Big\{A \in C_\infty(M) : A \bigcap \big\{\omega : T(\omega) \leq t\big\} \in C_{t+}(M) \ \text{ todo } t \geq 0\Big\} \, .$$

**Definição 8.3:** Um processo de difusão em $M$ é um processo de Markov que satisfaz a seguinte propriedade forte de Markov.

Para $T$ um $C_{t+}$-tempo de parada, sejam $A \in C_T$ e $B$ boreliano em $M$. Então

$$\begin{aligned} & P_x\Big(A \bigcap \big\{\omega : \omega(t + T(\omega)) \in B\big\}\Big) \\ &= \int_A P_{\omega'(T(\omega'))}\Big(\omega : \omega(t) \in B\Big) P_x(d\omega') \, . \end{aligned} \tag{8.4}$$

**Definição 8.4:** Seja $L$ um operador no espaço das funções contínuas em $M$ com domínio $D(L)$. Um processo de difusão gerado por $L$ é um processo tal que

$$f\big(\omega(t)\big) - f\big(\omega(0)\big) - \int_0^t (Lf)\big(\omega(s)\big)\, ds$$

é um mantingale em relação à filtração $C_{t+}$ e $P_x$ todo $x \in M$.

Como exemplo de um processo de difusão gerado por um operador diferencial, citemos o movimento Browniano em $\mathbb{R}^d$ que é gerado pelo operador de Laplace $\Delta = \frac{\partial^2}{\partial x_j^2} + \cdots + \frac{\partial^2}{\partial x_d^2}$ que é um operador cujo domínio é o espaço das funções de classe $C^2$.

Mais geralmente, dada uma equação diferencial estocástica em $M$

$$dx = X(x)dt + \sum_{j=1}^{m} Y_j(x) \circ dW^j, \tag{S}$$

Seja $L = X + \frac{1}{2}\sum_{j=1}^{m} Y_j^2$ o operador de segunda ordem que aparece na fórmula de Itô (6.5) considerado como um operador cujo domínio é o espaço das funções de classe $C^3$. As soluções de $(S)$ definem uma difusão gerada por $L$ da seguinte maneira:

Como anteriormente, denotamos por $\varphi_t(\omega)(x)$ a solução com condição inicial $(0, x)$ (essas soluções são suficientes para se ter todas as soluções, pela homogeneidade de $(S)$ no tempo). Sendo $T$ o tempo de explosão das soluções, $t \to \varphi_t(x)(\omega)$ define para cada $\omega \in \Omega$ ($\Omega$ = espaço de Wiener) um elemento de $W_T(M)$. Desta maneira, temos uma aplicação $\Omega \to W_T(M)$ que é $\mathcal{F}_t, C_{t+}$-mensurável. Fixando $x \in M$, a imagem da medida de Wiener $P$ por esta aplicação define uma medida $P_x$ em $(W_T(M), C_t)$. Deixamos ao leitor a verrificação de que a família $P_x$ define um processo de difusão em $M$ (a propriedade i) da definição 8.1 é conseqüência da dependência contínua em relação às condições iniciais e a propriedade de Markov forte segue da mesma propriedade para o processo de Wiener juntamente com a mensurabilidade de $\varphi_t(x)$ e a propriedade de fluxo). Pela fórmula de Itô (6.5), este é um processo de difusão gerado pelo operador $L = X + \frac{1}{2}\sum_{j=1}^{m} Y_j^2$. As probabilidades de transição deste processo são dadas por

$$P\Big(t, x, B\Big) = P\Big(\omega : \varphi_t(\omega)(x) \in B\Big) , \tag{8.5}$$

isto é, $P(t,x,\cdot)$ é a lei (sob P) da aplicação aleatória $\omega \to \varphi_t(x)(\omega)$.

Convém repetir o comentário feito acima de que o processo definido pelas probabilidades de transição (8.5) depende de $\varphi_t(\omega)$ apenas através de suas leis e não da solução como um todo. Neste sentido, observamos que um operador do tipo parabólico $L$ que pode ser escrito como $L = X + \frac{1}{2}\sum Y_j^2$, admite – em geral – mais de uma expressão deste tipo. Dito de outra maneira, o processo definido a partir das soluções de uma equação estocástica pode ser definido por mais de uma equação diferencial (veja exemplos abaixo). Não é surpreendente portanto a afirmação de que o fluxo $\varphi_t(\omega)$ não pode ser construido a partir do processo. O que pode ser feito, nesta linha, é a reconstrução do fluxo $\varphi_t(\omega)$, a partir do processo associado à equação estocástica

$$dx = X(x)dt + \sum Y_j(x) \circ dW^j$$

$$dy = X(y)dt + \sum Y_j(y) \circ dW^j$$

em $M \times M$ (veja [B] , [K2]).

Como comentamos anteriormente, esta construção permite um estudo dos operadores $L$ via métodos de cálculo estocástico. É natural então se perguntar sobre o procedimento inverso, isto é, o de se obter equações estocásticas (e processos de difusão) geradas por um operador pré-determinado. No que segue vamos fazer uma construção deste tipo para os operadores de Laplace-Beltrami em variedades Riemannianas. Esta construção abre a possibilidade de uma abordagem probabilística a uma série de questões em geometria diferencial e em topologia das variedades. Ao leitor interessado nestas aplicações do cálculo estocástico, deixamos as referências [Bi 1 e 2], [I-W], [SG],...

Seja $M$ uma variedade Riemanniana cuja métrica é denotada por $<,>$. Para $f : M \to \mathbb{R}$ seja grad$f$ o seu gradiente, isto é, o único campo de vetores em $M$ que satisfaz

$df = <$ graf $f, \cdot >$ onde $df$ é a diferencial da $f$. Dado um campo de vetores $X$ em $M$, denotamos por $X$ o seu divergente. div $X$ é uma função $M \to I\!R$ e é definido por $(\text{div } X)(x) = tr(A_x)$ onde $A_x : T_xM \to T_xM$ é a transformação linear

$$v \in T_xM \to A_xv = \nabla_v X \in T_xM$$

com $\nabla$ a conexão Riemanniana de $<,>$. O Laplaciano (ou operador de Laplace-Beltrami) em $M$ é o operador de segunda ordem definido, para $f$ de classe $C^2$, por

$$\Delta_f = \text{div}(\text{grad } f)$$

(Veja [K-N] vol.2 nota 14)., Alguns exemplos destes operadores são:
a) Em $I\!R^d$ com a métrica Euclidiana e coordenadas $(x^1, \ldots, x^d)$,

$$\text{grad } f = \frac{\partial f}{\partial x^1} \frac{\partial}{\partial x^1} + \cdots + \frac{\partial f}{\partial x^d} \frac{\partial}{\partial x^d}$$

$$\text{e como} \quad \nabla_{\partial/\partial x^i} \partial/\partial xj = 0, \quad \text{div} \left( a^1 \frac{\partial}{\partial x^1} + \cdots + a^d \frac{\partial}{\partial x^d} \right) = \frac{\partial a^1}{\partial x^1} + \cdots + \frac{\partial a^d}{\partial x^d}$$

Daí que

$$\Delta f = \frac{\partial^2}{\partial (x^1)^2} + \cdots + \frac{\partial^2}{\partial (x^d)^2}$$

e $\Delta$ é o Laplaciano clássico.

b) No circulo $S^1$, parametrizado pelo angulo $\theta$, seja a métrica dada por $\langle \frac{d}{d\theta}, \frac{d}{d\theta} \rangle = 1$. Como em a), o Laplaciano é $\Delta = \frac{d^2}{d\theta^2}$.

Para se construir difusões em $M$ geradas pelo Laplaciano $\Delta$, existe um procedimento geral e intrínseco (Eells e Elworthy [E-E]) com o qual se obtém a difusão por integração de uma equação estocástica não exatamente em $M$ mas no conjunto (fibrado) dos referenciais ortonormais de $M$. A idéia aí, é que a equação estocástica e o fluxo gerado por ela no fibrado dos referenciais desempenhe o mesmo papel que o fluxo geodésico gerado por $M$, que é o fluxo gerado por um campo de vetores no fibrado dos referenciais.

Antes de desenvolver essa construção intrínseca, vamos ver um procedimento relativamente mais simples (do ponto de vista formal) de se obter difusões geradas por $\Delta$, no caso em que $M$ é uma subvariedade isométricamente mergulhada em $\mathbb{R}^n$. Essas duas construções dão origem a fluxos estocásticos completamente diferentes. Para ilustrar essa diferença, vejamos primeiramente o caso unidimensional.

A) <u>$\dim M = 1$.</u> Neste caso $M \approx S^1$ ou $\mathbb{R}^1$. Como sempre é possível parametrizar por comprimento de arco, a métrica é a canonica, isto é, como nos exemplos b) e a) acima. Em ambos os casos, parametrizando por $\theta$, $\Delta = \frac{d^2}{d\theta^2}$ e a equação estocástica $d\theta = o\ dW$ define um processo de difusão cujo gerador é $\frac{1}{2}\Delta$. Em $\mathbb{R}^1$, o processo é o próprio movimento Browniano (processo de Wiener). No caso de $S^1$, as soluções da equação são obtidas pela imagem do movimento Browniano em $\mathbb{R}^1$ via a aplicação de recobrimento $\theta \to e^{i\theta}$. Isto pode ser visto pela fórmula de Itô.

Por imersão em $\mathbb{R}^n$, outras equações estocásticas em $M$, tendo $\Delta$ como gerador infinitesimal podem ser obtidas. Seja $x(t)$ uma curva diferenciável em $\mathbb{R}^n$ e suponha que $|x'(\theta)| = 1$ todo $\theta$.

Então $x$ define uma imersão de $\mathbb{R}^1$ ou $S^1$ se a curva for fechada. Tome uma base $\{e_1, \dots e_n\}$ de $\mathbb{R}^n$ e defina para cada $i = 1, \dots, n$ o campo de vetores $X_i$ em $\mathbb{R}^1$ (ou $S^1$) por

$X_i(\theta)$ = projeção ortogonal de $e_i$ sobre a reta tangente à curva em $x(\theta)$.

Como $x$ é parametrizada por comprimento de arco, temos

$$X_i(\theta) = \langle e_i, x'(\theta)\rangle \frac{d}{d\theta} .$$

Seja a equação estocástica em $\mathbb{R}^1$ (ou $S^1$) dada por

$$d\theta = \sum_{j=1}^{n} X_j(\theta) \; o \; dW^j \qquad \text{(BE)} .$$

Como para um campo $X(s) = a(s)\frac{d}{ds}$ tem-se

$$X^2 = aa' \frac{d}{ds} + a^2 \frac{d^2}{ds^2}$$

o gerador infinitesimal de (BE) é

$$\begin{aligned} L &= \frac{1}{2}\left(\sum_{j=1}^{n} \langle e_j, x'(\theta)\rangle^2\right)\frac{d^2}{d\theta^2} + \frac{1}{2}\left(\sum_{j=1}^{n} \langle e_j, x'(\theta)\rangle \langle e_j, x''(\theta)\rangle\right)\frac{d}{d\theta} \\ &= \frac{1}{2}|x'(\theta)|^2 \frac{d^2}{d\theta^2} + \langle x'(\theta), x''(\theta)\rangle \frac{d}{d\theta} \\ &= \frac{1}{2}\frac{d^2}{d\theta^2} = \frac{1}{2}\Delta . \end{aligned}$$

Um caso específico desta situação é dado pela imersão $\theta \to (\cos\theta, \operatorname{sen}\theta)$ de $S^1$ em $\mathbb{R}^2$. Temos

$$X_1(\theta) = -\operatorname{sen}\theta \frac{d}{d\theta} \quad \text{e} \quad X_2(\theta) = \cos\theta \frac{d}{d\theta}$$

e (BE) é

$$d\theta = -\operatorname{sen}\theta \; o \; dW^1 + \cos\theta \; o \; dW^2$$

que evidentemente não coincide com $d\theta = odW$. Para ver a diferença entre os fluxos gerados por estas equações, vamos interpretá-las como equações em espaços homogeneos da mesma forma que no exemplo 6.7 d). Por simplicidade, ao invés de $X_1$ e $X_2$ como acima, tomamos $Y_1(\theta) = -\text{sen}2\theta\frac{d}{d\theta}$ e $Y_2(\theta) = \cos 2\theta\frac{d}{d\theta}$ (que são obtidas de $X_1$ e $X_2$ pela mudança de coordenadas $\sigma = \theta/2$). Seja o grupo de Lie $Sl(2,\mathbb{R})$ das matrizes $2\times 2$ com determinante 1. Campos invariantes à direita em $Sl(2,\mathbb{R})$ são da forma $Y(g) = Ag$ com $A$ matriz $2\times 2$. Os elementos de $Sl(2,\mathbb{R})$ são transformações lineares inversiveis em $R^2$. Interpretando $S^1$ como o conjunto dos raios em $\mathbb{R}^2$ partindo da origem, obtém-se uma ação de $Sl(2,\mathbb{R})$ em $S^1$. Sejam $Y_1(g) = B_1g, Y_2 = B_2g$ e $X(g) = Ag$ campos invariantes à direita com

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Denote por $\tilde{Y}_1, \tilde{Y}_2$ e $\tilde{X}$ os campos induzidos em $S^1$. Então escrevendo os campos lineares em $\mathbb{R}^2$ associados à $B_1, B_2$ e $A$ em coordenadas polares e tomando a coordenada $\theta$, vemos que $\tilde{Y}_1(\theta) = -\text{sen }2\theta\frac{d}{d\theta}$ $\tilde{Y}(\theta) = \cos 2\theta\ \frac{d}{d\theta}$ e $\tilde{X}(\theta) = \frac{d}{d\theta}$. Daí se vê que o fluxo definido por $d\theta = o\, dW$ é um fluxo de rotações em $S^1$ (já que $A$ é anti-simétrica) enquanto que o fluxo definido por (BE) envolve outras transformações de $Sl(2,\mathbb{R})$ em $S^1$ que não apenas as rotações.

O que este caso tem de característico do caso geral é que $d\theta = o\, dW$ é obtido via a construção intrínseca enquanto que (BE) requer que a variedade esteja imersa em $\mathbb{R}^n$ (BE=Browniano Extrínseco).

**B)** Variedades imersas: Seja $M$ uma subvariedade $d$-dimensional de $\mathbb{R}^n$ com a métrica induzida pela imersão. Sendo $\{e_1, \ldots, e_n\}$ uma base ortonormal de $\mathbb{R}^n$, sejam $X_j, j = 1, \ldots, n$ os campos de vetores em $M$ definidos por $X_j(x)$ =projeção ortogonal de $e_j$

sobre o espaço tangente à $M$ em $x$. $X_j$ é o gradiente de $x \in M \to \langle e_j, x\rangle \in \mathbb{R}$. Seja a equação

$$dx = \sum_{j=1}^{m} X_j(x) \circ dW^j \tag{BE}$$

em $M$. Vamos verificar que $L = \sum_{j=1}^{m} X_j^2$ coincide com o Laplaciano em $M$, e daí que as soluções de $(BE)$ fornecem um processo de difusão tendo $\frac{1}{2}\Delta$ como gerador infinitesimal.

Antes de mais nada, verifiquemos que o operador $L$ não depende de $\{e_1, \ldots, e_n\}$. Seja $\{f_1, \ldots, f_n\}$ outra base ortonormal e defina $Y_1, \ldots, Y_n$ por projeção ortogonal dos $f_j's$. Como $f_i = \sum a_{ji} e_i$, temos que $Y_i = \sum_j a_{ji} X_j$ e daí que

$$\begin{aligned} Y_i^2 &= \Big(\sum_j a_{ji} X_j\Big)\Big(\sum_k a_{ki} X_k\Big) \\ &= \sum_{j,k} a_{ji} a_{ki} X_j X_k \end{aligned}$$

e

$$\begin{aligned} \sum Y_i^2 &= \sum_{j,k} \sum_i a_{ji} a_{ki} X_j X_k \\ &= \sum_j X_j^2 \end{aligned}$$

pois $(a_{ji})$ é matriz ortogonal. Tomando agora $h : M \to \mathbb{R}$, temos $X_i h = \langle X_i, \text{grad } h\rangle$ e

$$\begin{aligned} X_i^2 h &= X_i \langle X_i, \text{grad } h\rangle \\ &= \langle \nabla_{X_i}, \text{grad } h\rangle + \langle X_i, \nabla_{X_i} \text{grad } h\rangle \end{aligned}$$

onde $\nabla$ é a conexão de Levi-Civita em $M$. Fixando $x \in M$, como $L$ é independente da base, podemos supor que $e_i$ é ou tangente ou ortogonal à $M$ em $x$. Neste caso somando em $i$, o último termo da expressão acima coincide com div (grad h), isto é , com $\Delta h$. Por outro lado, o primeiro termo do segundo membro se anula. Para ver isto, note primeiramente que se $e_i$ é normal à $M$ em $x$ então $X_i(x) = 0$ e daí que $(\nabla_{X_i} X_i)(x) = 0$. E se $e_i$ é tangente à $M$ então $(\nabla_u X_i)(x) = 0$ todo $u \in T_x M$. De fato, denotando por $D_u X$ a derivada (em $\mathbb{R}^n$) de $X$ na direção de $u, \nabla_u X$ é a componente tangencial de $D_u X$. Tomando campos normais $Y_1, \ldots, Y_{n-d}$ à $M$, $X_i = e_i - \sum \langle e_i, Y_j \rangle Y_j$. Daí que

$$\begin{aligned} D_u X_i &= -\sum_j D_u\Big(\langle e_i, Y_j\rangle Y_j\Big) \\ &= \sum_j \langle e_i, D_u Y_j\rangle Y_j - \sum_j \langle e_i, Y_j\rangle D_u Y_j \end{aligned}$$

Como a componente tangencial só é dada por este último termo e $\langle e_i, Y_j\rangle = 0$ (se $e_i$ é tangente), temos que $\nabla_u X_i = 0$. Concluimos então que

$$L = \sum_{i=1}^{n} X_i^2 = \Delta$$

e $(BE)$ define uma difusão gerada por $\frac{1}{2}\Delta$.

C) Construção Intrínseca: Antes de exibir a equação diferencial, vamos rever alguns fatos sobre fibrados principais e conexões sobre os mesmos. Veja [K-N] para detalhes.

Seja $OM$ o conjunto das bases ortonormais de $M$, isto é, para cada $x \in M$, $O_x M$ é o conjunto das bases ortonormais de $T_x M$ (em relação ao produto interno $\langle , \rangle_x$) e $OM = \bigcup_{x \in M} O_x M$. Uma maneira conveniente de representar bases em $T_x M$ é via os referenciais, isto é, aplicações lineares inversíveis $p : \mathbb{R}^d \to T_x M$. Sendo $\{e_1, \ldots, e_d\}$ base de $R^d$, $\{p(e_i)\}$ é base de $T_x M$. Um referencial é ortonormal se $\langle pu, pv\rangle_x = \langle u, v\rangle$ todo $u, v \in \mathbb{R}^d$ ($o, \langle , \rangle$ no segundo membro desta igualdade é o produto interno em $\mathbb{R}^d$). Vamos denotar por $\pi$ a projeção canônica $\pi : p \in O_x M \to x \in M$.

Seja $O(d)$ o grupo das matrizes ortogonais $d \times d$. Ele atua em $OM$ por $R_a(p) = pa = poa$, $p \in OM, a \in O(d)$. A ação é à direita pois $R_{ab} = R_b o R_a$. É uma ação livre (isto é $R_a(p) = p$ algum $p \Rightarrow a =$ identidade) e transitiva sobre as fibras $O_xM$ (isto é, dados $p, q \in O_xM$, existe $a \in O(d)$ com $q = pa$). Fixando $p \in O_xM$, todo $q \in O_xM$ se escreve de maneira única como $q = pa, a \in O(d)$.

Alguns exemplos de fibrados de referenciais são:

**Exemplos 8.5:** a) Seja $M = \mathbb{R}^d$ com a métrica euclidiana.

Os campos $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^d}$ formam em cada $x \in \mathbb{R}^d$ uma base ortonormal. Denotado por $p(x)$ o referencial associado, qualquer outro referencial ortonormal é da forma $q = p(x)a$, um único $a \in O(d)$. A aplicação $p(x)a \to (x, a)$ identifica $OM$ com $\mathbb{R}^d \times O(d)$. Por esta identificação $R_a(x, b) = (x, ba)$. Em particular, $O\mathbb{R}^1 \approx \mathbb{R}^1 \times O(1)$ e $O\mathbb{R}^1$ se constitui de duas cópias de $\mathbb{R}^1$: os referenciais $\frac{d}{dx}$ e $-\frac{d}{dx}$.

b) Seja $M = S^d = \{x \in \mathbb{R}^{d+1} : |x| = 1\}$ com a métrica induzida por essa imersão. O espaço tangente $T_xS^d$ se identifica com o hiperplano em $\mathbb{R}^{d+1}$ ortogonal à $x$ e uma base ortonormal deste hiperplano é uma base ortonormal de $T_xS^d$. Esta base juntamente com $x$ forma uma base ortonormal de $\mathbb{R}^{d+1}$. Daí que $OS^d$ se identifica como o conjunto das bases ortonormais em $\mathbb{R}^{d+1}$. Fixando a base canônica $\{e_1, \ldots, e_{d+1}\}$, toda base ortonormal se escreve como $\{ae_1, \ldots, ae_d\}$, um único $a \in O(d+1)$. Daí que $OS^d$ se identifica com $O(d+1)$.

Em particular, se $d = 1, OS^1 \approx O(2)$ se constitui de duas cópias de $S^1$: os referenciais que apontam no sentido horário e os de sentido contrário.

Como os elementos de $OM$ são referenciais dos espaços tangentes à $M$, tensores em $M$ se representam de maneira bastante conveniente como funções em $OM$. Por exemplo, cada vetor tangente $v \in T_xM$ define uma aplicação $F_v : O_xM \to \mathbb{R}^d$ por $f_v(p) = p^{-1}v$

(isto é, $f_v(p)$ são as coordenadas de $v$ na base $p$). A função $F_v$ não é arbitrária. De fato, se $a \in O(d)$, $f_v(pa) = (poa)^{-1}v = a^{-1}op^{-1}v = a^{-1}f_v(P)$. Reciprocamente, toda aplicação $f : O_xM \to I\!R^d$ que satisfaça $f(pa) = a^{-1}f(p)$ define um vetor $v \in T_xM$ por $v = pf(p)$. Da mesma forma, um campo de vetores $X : M \to TM$ define $f : OM \to I\!R^d$, $f(p) = p^{-1}X(\pi(p))$, que satisfaz

$$f(pa) = a^{-1}f(p)$$

e reciprocamente, toda aplicação $f : OM \to I\!R^d$ que satisfaz esta igualdade define um campo de vetores em $M$.

Para construir o movimento Brawniano intrínseco em $M$ é necessário descrever a conexão de Levi-Civita associada à sua métrica no contexto de conexões em fibrados principais.

Uma conexão em $OM$ é uma distribuição (no sentido de Fronbenius) horizontal e invariante à direita. Isto significa o seguinte: Uma distribuição em $OM$ é uma aplicação $p \in OM \to H_p$ com $H_p$ subespaço de $T_pOM$. A distribuição é horizontal quando a diferencial $\pi_*$ de $\pi$ restrita à $H_p$ é um isomorfismo entre $H_p$ e $T_{\pi(p)}M$. Em outras palavras $H_p$ é horizontal se dim $H_p = dimM$ e $H_p$ tem intersecção trivial com os subespaços verticais, isto é, com os subespaços tangentes à $O_xM$. A distribuição é invariante à direita se $R_aH_p = H_{pa}$.

**Exemplos 8.6:** a) Na situação do exemplo 8.5 a), com a identificação $OI\!R^d \approx I\!R^d \times O(d)$, a distribuição

$$(x,b) \to H_{(x,b)} = \Big\{(v,0) \in T_{(x,b)}\big(I\!R^d \times O(d)\big) : v \in I\!R^d\Big\}$$

é uma conexão. É imediato verificar que $H_{(x,b)}$ é horizontal. A invariança segue do fato que $R_a(x,b) = (x, ba)$ e portanto a diferencial $R_{a*}$, expressa em coordenadas tem matriz da forma

$$\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$$

com blocos quadrados de dimensões $d(= \dim \mathbb{R}^d)$ e $\frac{d(d-1)}{2}(= \dim O(d))$.

Outras conexões em $\mathbb{R}^d \times O(d)$ podem ser obtidas definindo subespaços horizontais $H_{(x,1)}$ em $(x,1), x \in \mathbb{R}^d, 1 =$ identidade de $O(d)$ e colocando $H_{(x,b)} = R_{b*}H_{(x,1)}$.

Em particular, a única conexão no caso $d = 1$ é dada pelo próprio espaço tangente à $\mathbb{R}^1 \times O(1)$ (já que $O(1) = \{1, -1\}$ é um grupo discreto e portanto seu espaço tangente é trivial).

b) Um exemplo de conexão na situação do exemplo 8.5 b), por ser construido por invariança: O espaço tangente na identidade de $O(d+1)$ (visto como subvariedade do espaço das matrizes) é o espaço $SO(d+1)$ das matrizes anti-simétricas. Este subespaço se decompõe em $SO(d+1) = H \oplus SO(d)$ onde $H$ é o subespaço das matrizes do tipo

$$\begin{pmatrix} 0 & \beta \\ -\beta^t & 0 \end{pmatrix}$$

com $\beta$ matriz linha $1 \times d$ e $SO(d)$ é

$$\begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}$$

Com $A$ $d \times d$ anti-simétrica. Ponha $H_1 = H$ e $H_g = L_{g*}(H)$ onde $L_g : O(d+1) \to O(d+1)$, $g \in O(d+1)$ é a translação à esquerda $L_g(h) = gh$. A distribuição $g \to H_g$ define uma conexão em $OS^d \approx O(d+1)$. Para detalhes, nos referimos à [K-N] (veja Teorema de Wang).

Quando $d = 1$, como em $\mathbb{R}^1$, existe uma única conexão em $OS^1$.

A relação entre estas conexões em $OM$ e as derivadas covariantes $\nabla_X Y$ se obtém da seguinte forma: Dada uma conexão (subespaços horizontais em $OM$) e $X$ e $Y$ campos de vetores, fixe $x \in M$ e $p \in O_x M$. Como $H_p$ está em bijeção com $T_x M$, existe um único $v \in H_p$ t.q. $\pi_* v = X(x)$($v$ é o levantamento horizontal de $X(x)$). Tomando $f_Y : OM \to \mathbb{R}^d$, $f_Y(q) = q^{-1} Y(\pi(q))$, define-se

$$\left(\nabla_X Y\right)(x) = \left(v f_Y\right)(p)$$

com $v f_Y$ derivada direcional usual. Verifica-se que esta expressão independe de $p$ e que variando $x$ em $M, \nabla_X Y$ satisfaz $(C)$ do § 6. $\nabla$ é a derivada covariante associada à conexão. A construção reciproca, isto é, dos subespaços horizontais a partir de $\nabla$ também é possível, em geral, desde que se considere o conjunto de todos os referenciais e não apenas os ortonormais como fizemos até aqui. Subespaços horizontais em $OM$ podem ser tomados desde que $\nabla$ seja adaptada á métrica, isto é, se tenha para $X, Y, Z$ campos de vetores,

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \ .$$

**Exemplo 8.6:** Na situação dos exemplos a), os subespaços horizontais $H_{(x,b)} = \{(v, 0)\}$ definem a derivação covariante em que

$$\nabla_{\partial/\partial x^i} \partial/\partial x_j = 0$$

isto é, a derivada direcional em $\mathbb{R}^d$. Isto porque as coordenadas de $\partial/\partial x_i$ na base $\{\partial/\partial xi, \ldots, \partial/\partial x^d\}$ são dadas por $e_j = (0 \ldots, 1, \ldots, 0)$ e portanto $f_{\partial/\partial x^j}(x, b) = b^{-1} e_j$

que é constante como função de $x$ e daí que sua derivada na direção de vetores horizontais se anula.

Uma conexão em $OM$ é dita sem torsão no caso em que sua derivada convariante satisfaz

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

para campos de vetores $X, Y$ quaisquer. O que se mostra é que em $OM$ existe uma única conexão sem torsão. Esta é a conexão de Levi-Civita ou Riemanniana associada à métrica em $M$.

Vamos por fim introduzir os campos de vetores com os quais se define a equação estocástica em $OM$, de cuja soluções se obtém um movimento Browniano em $M$.

Fixamos uma conexão em $OM$ e $v \in \mathbb{R}^d$, o campo de vetores standart $B_v$ associado à $v$ é o único campo horizontal (i.e., $B_v(p) \in H_p$ todo $p$) cuja projeção sobre $M$ tem coordenadas $v$ em relação à base $p$:

$$\pi_*\Big(B_v(p)\Big) = pv \qquad \text{todo } p$$

Algumas das propriedades dos campos standard são:

i) $R_{a*}(B_v) = B_{a^{-1}v}$.

ii) $B_v(p) \neq 0$ todo $p$ se $v \neq 0$.

iii) $[B_v, B_w]$ é vertical se a conexão é sem torsão

e

iv) Seja $(B_v)_t$ o fluxo gerado por $B_v$. Então $(B_v)_t(p)$ se projeta na (única) geodésica da conexão que começa em $x = \pi(p)$ e é tangente (em $x$) à $pv$.

Esta última propriedade, aliás, pode se tomada como a definição das geodésicas de uma conexão. Ela diz em essencia que o fluxo das geodésicas (o fluxo glodésico) da conexão é obtido pelo fluxo do campo de vetores $B_v$ (a equação das geodésicos em $M$ é uma equação de segunda ordem. Para obte-la como equação de primeira ordem é necessário levantá-la ao fibrado, considerando mais variáveis).

Observamos que os campos $B_v$ não se projetam a campos de vetores em $M$. Isto porque $\pi_*(B_v(pa)) = pav \neq pv = \pi_*(B_v(p))$.

Exemplos 8.7: a) Na situação dos exemplos a), a conexão plana $H_{(x,b)} = \{(v, O)\}$ é a conexão de Levi-Civita da métrica euclidiana em $I\!R^d$ (pois $\nabla_{\partial/\partial x^j} = 0$. Nos referenciais da forma $p = (x,1)$, $x \in I\!R^d, B_v(p) = v$. Para outros referenciais obtém-se $B_v(p)$ por translação à direita. Note que apesar dos campos $B_v$ não serem projetaveis em $I\!R^d \times \{1\} \approx I\!R^d$, restrito à esta subvariedade, $B_v$ é um campo constante, $(B_v)_t(x) = x + tv$ e as geodésias são – é claro – as retas em $I\!R^d$.

b) Na situação dos exemplos b), a conexão introduzida em 8.6, por invariança é a conexão Riemanniana da métrica em $S^d$. Os campos standard em $OS^d$ são os campos invariantes à esquerda associados às matrizes em $H_1$.

No caso específico em que $d = 1$, $O(1)$ tem duas componentes conexas, ambas identicas à $S^1$. Dado $v = 1 \in I\!R^1$, em uma delas $B_v = \frac{d}{d\theta}$ e em outra $B_v = -\frac{d}{d\theta}$. Como no caso de $I\!R^d$, apesar de $B_v$ não se projetar, sua restrição à uma das componentes conexas pode ser considerada como um campo em $S^1$ propriamente dito.

Seja agora $\{e_1, \ldots e_d\}$ uma base ortogonal de $I\!R^d$ e defina o operador de segunda ordem

$$\Delta_{o_M} = B_{e_1}^2 + \cdots + B_{e_d}^2 \,.$$

Este operador é o levantamento horizontal do Laplaciano $\Delta$ em $M$, no seguinte sentido

**Proposição 8.8:** Para $f : M \to \mathbb{R}$,

$$\Delta_{o_M}\Big(f \; o \; \pi\Big) = \Big(\Delta f\Big) \; o \; \pi$$

<u>Dem:</u> Ponha $Y = \text{grad} f$ e seja $F_Y(p) = p^{-1}Y(\pi(p))$ sua expressão equivariante. Se $x = \pi(p)$,

$$\begin{aligned} Y(x) &= \langle Y, pe_1\rangle + \cdots + \langle Y, pe_d\rangle pe_d \\ &= \Big(pe_1.f\Big)pe_1 + \cdots + \Big(pe_d.f\Big)pe_d \; . \end{aligned}$$

Daí que $F_Y(p) = (pe_1.f, \ldots, pe_d.f)$. Usando a expressão da derivada covariante em termos da derivada de $F_Y$, temos

$$\Big(\text{div } Y\Big)(x) = tr\Big(v \to B_v F_Y\Big)$$

e portanto

$$\Big(\Delta f\Big)\Big(\pi(p)\Big) = \sum_i \langle \Big(B_{e_i} F_Y\Big)(p), e_i\rangle \; .$$

Por outro lado, $B_{e_i}(f o \pi)(p) = (pe_i.f)(x)$, pois $\pi_* B_{e_i} = pe_i$ e daí que

$$\begin{aligned} \Delta_{OM}\Big(f \; o \; \pi\Big) &= \textstyle\sum_i B^2_{e_i}\Big(f \; o \; \pi\Big) = \sum_i \langle B_{e_i} F_Y(p), e_i\rangle \\ &= \Big(\Delta f\Big) \; o \; \pi \; . \end{aligned}$$

Em vista desta proposição, se $\varphi_t(p)$ é o fluxo associado à

$$dp = \sum_{j=1}^{d} B_{e_j}(p) \circ dW^j,$$

então $\pi(\varphi_t(p))$ é um processo de difusão em $M$ gerado por $\Delta$.

# 5. Bibliografia


[Ba] Baxandale, P.: "Brownian motions in the diffeomorphism group I". Composito Mathematica 53, pp. 19-50 (1984).

[B] Billingsley, P.: Probability and Measure" Wiley (1979).

[B-G] Blumenthal, R.M. e Getoor, R.K.: "Markov Processes and Potential Theory". Ac. Press (1968).

[Bi] Bismut, J-M.: "The Atiyah-Singer Theorems: a probabilistic approach. I the index theorem. II the Lefchetz fixed point formulas. J. Func. Anal., 57 pp. 56-99 e 329-348 (1984).

[C] Carmo, M.P.: "Geometria Riemanniana". Projeto Euclides (1979).

[Ca] Cambanis, S.: "The measurability of a stochastic processes of a second ordem and its linear space". Proc. Amer. Math. Soc. 47, 467-475 (1975).

[Co] Cohn, D.K.: "Measurable choice of limit points and existence of separable and measurable processes. Z.W.u.G. 22, 161-165, (1972).

[Cor] Courrège, P.: "Integrales stochastiques et martingales de carre integrable". Semin Brelot, Choquet-Deny, 7-e annee (1962/63).

[C-D] Chung, K.L. e Doob, J.L.: "Fields, optionality and mensurability". Amer. J.Math. 87, 397-424 (1965).

[Di] Dieudonné, J.: "Foudations of Modern Analysis". Ac. Press (1960).

[D] Doob, T.L.: "Stochastic Processes", Wiley (1953).

[Dy] Dynkin, E.B.: "Markov Processes". 2 vol. Springer (1965).

[E] Elliot, R.J.: "Stochastic Calculus and Applications", Springer (1982).

[El] Elworthy, K.D.: "Stochastic Differential Equations on Manifolds", Lec. Notes Cambridge University Press, (1982).

[G-S] Gihman I.I. e Skorohod, A.V.: The theory of stochastic processes". 3 vol., Springer (1974).

[H] Halmos, P.R.: "Measure Theory". Van Nonstrand (1954).

[H-J] Hoffmann - Jorgensen: Existence of measurable modification of stochastic processes. Z.W., 25, 205-207, (1973).

[Hõ] Hõnig, C.S.: "Aplicações da Topologia à Analise". Projeto Euclides (1976).

[H-Y] Hocking, J.G. e Young, G.S.: "Topology". Addison Wesley (1964).

[I-W] Ikeda, N. e Watanabe, S.: "Stochastic Differential Equations and Diffusion Processes". Norh-Holland, 2ª ed., (1989).

[K-N] Kobayashi, S. e Nomizu, K.: "Foundations of Differential Geometry". Interscience (1963).

[K1] Kunita, M.: "Stochastic Differential Equations and Stochastic Flow of Diffeomorphisms". Ecole d'Eté de Saint-Flour XII. LNM - Springer 1097, pp.143-303 (1982).

[K2] Kunita, H.: "Stochastic Flows and Applications". Tata Inst. Research - Springer (1986).

[K3] Kunita, H.: "Convergence of Stochastic Flows Connected with Stochastic Ordinary Differential Equations".

[K4] Kunita, H.: "On Backward Stochastic Differential Equations", Stochastics, vol.6 pp.293-313 (1982).

[K-W] Kunita, H. e Watanabe, S.: "On Square Integrable Martingales". Nagoya Math. J. vol. 30 pp.209-245 (1967).

[Kuo] Kuo, H.H.: "Gaussian Measures in Banach Spaces", LNM nº 463 Springer (1975).

[L-S] Liptser, R.S. e Shiryayev, A.N.: "Statistics of Radom Processes". vol. 1 Springer (1974).

[M1] Meyer, P.A.: "Probability and Potentials". Blaisdell-Walthan, (1968).

[M2] Meyer, PA.: Un Cours sur les intégrales stochastique. LN.N 511. Springer (1976).

[N] Narasimhan, R.: "Analysis on Real and Complex Manifolds". North-Holland (1968).

[P-S] Port, S.C. e Stone, C.J.; "Brownian Motion and Classical Potential Theory". Ac. Press (1978).

[R] Royden, H.L.: "Real Analysis". Macmillam (1968).

[S] Shiryayev, A.N.: "Probability". Springer (1984).

[Sch] Schwartz, L.: "Semi-Martingales a Valeurs sur des Varietés et Martingales Conformes sur des Varietés Analytiques Complexes", LNM nº 780 - Springer (1980).

[S-G] "Seminaire de probabilité XVI. Supplément: Gémetrie Differentiell Stochastique". LNM - Springer (1982).