Vdgraph: A Package for Creating Variance Dispersion Graphs

by John Lawson

Abstract This article introduces the package **Vdgraph** that is used for making variance dispersion graphs of response surface designs. The package includes functions that make the variance dispersion graph of one design or compare variance dispersion graphs of two designs, which are stored in data frames or matrices. The package also contains several minimum run response surface designs (stored as matrices) that are not available in other R packages.

Introduction

Response surface methods consist of (1) experimental designs for collecting data to fit an approximate relationship between the factors and the response, (2) regression analyses for fitting the model and (3) graphical and numerical techniques for examining the fitted model to identify the optimum. The model normally used in response surface analysis is a second order polynomial, as shown in Equation (1).

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}^{k} \beta_{ij} x_i x_j \quad (1)$$

The fitted equation is examined in order to predict the factor coordinates of the maximum or minimum response within the experimental region, or to simply explore the relationship between the factors and response.

Since it is not known in advance what neighborhood will be of most interest in the design space, a desirable response surface design will be one that makes the variance of a predicted value as uniform as possible throughout the experimental region. Standard response surface designs, such as the uniform precision central composite design, are constructed so that the variance of a predicted value will be near constant within a coded radius of one from the center of the design.

One way to visualize the uniformity of the variance of a predicted value for designs with more than two factors is to use the variance dispersion graph proposed by Myers et al. (1992).

Variance of a predicted value

The variance of a predicted value at a point $(x_1, \cdots, x_k)$ in the experimental region is given by Equation (2)

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x} \quad (2)$$

where $\mathbf{X}$ is the design matrix for the quadratic model in Equation (1), σ^2 is the variance of the experimental error, and

$$\mathbf{x} = [1, x_1, \cdots, x_k, x_1^2, \cdots, x_k^2, x_1 x_2, \cdots]$$

is a vector valued function of the coordinates (of the point in the experimental region) whose elements correspond to the columns of the design matrix $\mathbf{X}$.

Run	x_1	x_2
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5	-1	0
6	1	0
7	0	-1
8	0	1
9	0	0

Table 1: Face Center Cube Design or 3^2 Design

For the face-centered cube design, or 3^2 design shown in Table 1, Figure 1 is a contour plot of the scaled variance of a predicted value in the range of $-1.5 \le x_1 \le 1.5, -1.5 \le x_2 \le 1.5$. The scaled variance of a predicted value is $N\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$, where N is the number of points in the experimental design.

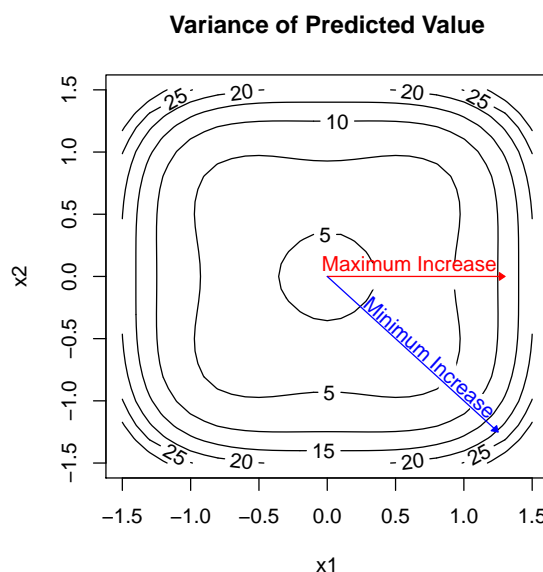**Variance of Predicted Value**

Figure 1: Contour plot of $N\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$

As seen in Figure 1, the variance of a predicted value increases faster along the line $x_2 = 0$ than along

the line $x_2 = -x_1$. This is easy to visualize in two dimensions, but would be more difficult to see in higher dimensions.

Variance dispersion graphs

A variance dispersion graph allows one to visualize the uniformity of the scaled variance of a predicted value in multidimensional space. It consists of three curves: the maximum, the minimum and the average scaled variance of a predicted value on a hypersphere. Each value is plotted against the radius of the hypersphere. Figure 2 shows the variance dispersion graph of the design shown in Table 1.

In this figure it can be seen that the maximum scaled variance of a predicted value is near 14 at a radius of 1.4 in coded units, while the minimum scaled variance is less than 10 at the same radius. This is the same phenomenon that can be seen in the contour plot of the scaled variance of a predicted value shown in Figure 1. The path of the maximum and minimum variance through the design space will be determined by the design and may not follow straight lines as shown in the specific example in this contour plot.
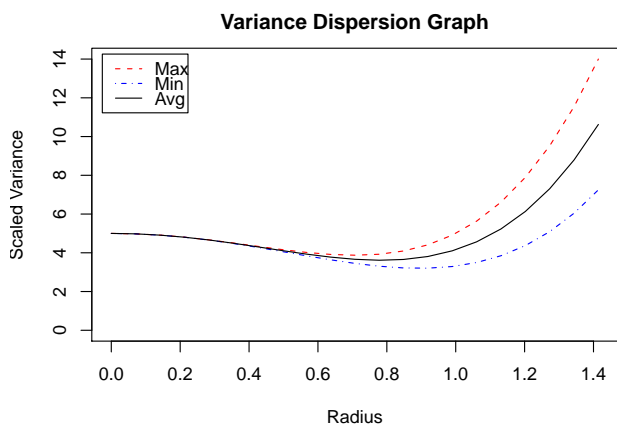

Figure 2: Variance dispersion graph for the design in Table 1.

Unlike the contour plot of the scaled prediction variance, the variance dispersion graph has the same format for a k dimensional response surface design as for a two dimensional design like Table 1.

Recent textbooks such as Montgomery (2005), Myers et al. (2009), and Lawson (2010) illustrate variance dispersion graphs as a tool for judging the merits of a response surface design. These graphs can be produced in commercial software such as **SAS ADX** (see SAS Institute Inc., 2010) and **Minitab** (see Minitab Inc., 2010) by using a downloadable macro (see Santiago, 2009).

The Vdgraph package

Vining (1993a), and Vining (1993b) published FOR-TRAN code for creating variance dispersion graphs. Vining's code obtains the maximum and minimum prediction variance on hyperspheres using a combination of a grid search and Nelder-Mead search as described by Cook and Nachtsheim (1980). The package **Vdgraph** (Lawson, 2011) incorporates this code in R functions that make the graphs.

The package includes the function `Vdgraph` for making a variance dispersion graph of one design and the function `Compare2Vdg` for comparing the variance dispersion graphs of two designs on the same plot. The package also includes several minimum run response surface designs stored as matrices. These include Hartley's small composite design for 2 to 6 factors, Draper and Lin's small composite design for 5 factors, the hexagonal rotatable design for 2 factors and Roquemore's hybrid designs for 3 to 6 factors.

Examples

The first example shown below illustrates the use of the R function `Vdgraph` to make variance dispersion graphs of a three factor Box-Behnken design created by the `bbd` function in the R package **rsm** (see Lenth, 2009).

```
> library(rsm)
> BB.des3 <- bbd(3)
> Vdgraph(BB.des3)
number of design points= 16
number of factors= 3
          Radius    Maximum    Minimum    Average
 [1,] 0.00000000   4.000000   4.000000    4.00000
 [2,] 0.08660254   3.990100   3.990067    3.99008
 [3,] 0.17320508   3.961600   3.961067    3.96128
 [4,] 0.25980762   3.918100   3.915400    3.91648
 [5,] 0.34641016   3.865600   3.857067    3.86048
 [6,] 0.43301270   3.812500   3.791667    3.80000
 [7,] 0.51961524   3.769600   3.726400    3.74368
 [8,] 0.60621778   3.750100   3.670067    3.70208
 [9,] 0.69282032   3.769600   3.633067    3.68768
[10,] 0.77942286   3.846100   3.627400    3.71488
[11,] 0.86602540   4.000000   3.666667    3.80000
[12,] 0.95262794   4.254106   3.766067    3.96128
[13,] 1.03923048   4.633600   3.942400    4.21888
[14,] 1.12583302   5.166116   4.214067    4.59488
[15,] 1.21243557   5.881600   4.601067    5.11328
[16,] 1.29903811   6.812500   5.125000    5.80000
[17,] 1.38564065   7.993638   5.809067    6.68288
[18,] 1.47224319   9.462100   6.678067    7.79168
[19,] 1.55884573  11.257600   7.758400    9.15808
[20,] 1.64544827  13.422175   9.078067   10.81568
[21,] 1.73205081  16.000000  10.666667   12.80000
```
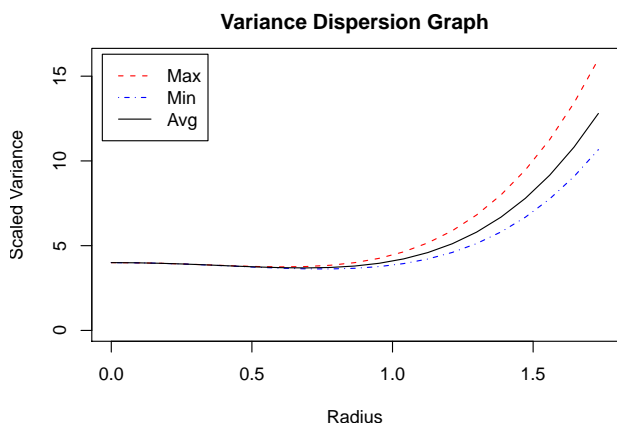
Figure 3: Variance dispersion graph for `BB.des3`.

The result from the first example (shown above) includes a listing of the coordinates of the plot and the graph shown in Figure 3.

The second example illustrates the use of `Compare2Vdg` by comparing the variance dispersion graph of Draper and Lin's small composite design for 5 factors (`SCDDL5`) (Draper and Lin, 1990) with Hartley's Small Composite Design (`SCDH5`) (Hartley, 1959). Hartley's small composite design requires only 28 runs by utilizing a $\frac{1}{2}$ fraction of the factorial portion of the design.

Run	x_1	x_2	x_3	x_4	x_5
1	1	-1	1	1	1
2	1	1	-1	-1	-1
3	-1	1	1	-1	1
4	1	-1	1	-1	1
5	1	1	-1	1	1
6	1	1	1	-1	-1
7	-1	1	1	1	-1
8	-1	-1	1	1	-1
9	-1	-1	-1	-1	1
10	1	-1	-1	1	-1
11	-1	1	-1	1	1
12	-1	-1	-1	-1	-1
13	$-\alpha$	0	0	0	0
14	α	0	0	0	0
15	0	$-\alpha$	0	0	0
16	0	α	0	0	0
17	0	0	$-\alpha$	0	0
18	0	0	α	0	0
19	0	0	0	$-\alpha$	0
20	0	0	0	α	0
21	0	0	0	0	$-\alpha$
22	0	0	0	0	α
23	0	0	0	0	0

Table 2: Draper and Lin"s Small Composite Design for 5 Factors

Although Draper and Lin's design (shown in Table 2 with $\alpha = 1.86121$) further reduces the number of

runs to 23, by substituting a 12 run Plackett-Burman design in the factorial portion, its variance dispersion graph reveals that the variance of a predicted value is not nearly as uniform as it is for the Hartley's design.

```
> data(SCDH5)
> data(SCDDL5)
> Compare2Vdg("Hartley's Small Composite-5 fac",SCDH5,
> +"Draper and Lin's Small Composite-5 fac",SCDDL5)
```
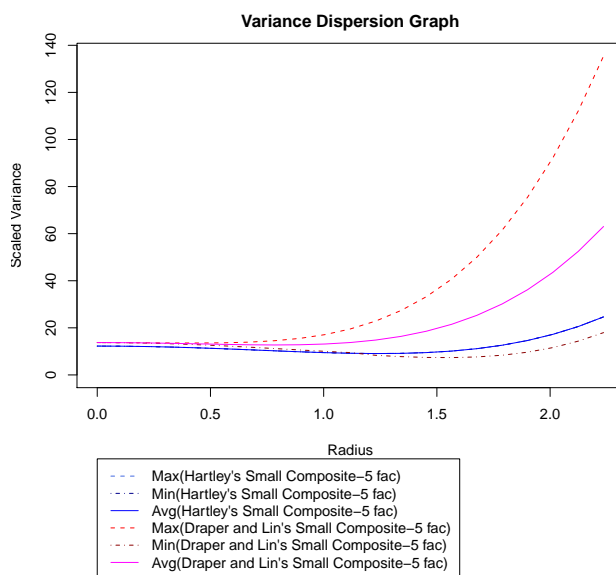

Figure 4: Comparison of Two Variance Dispersion Graphs.

As seen in Figure 4, Hartley's small composite design for 5 factors is rotatable since the three blue curves for the max, min and average scaled prediction variance coincide. The scaled variance of a predicted value for Hartley's design is near the minimum scaled variance of a predicted value for Draper and Lin's design throughout the experimental region.

Acknowledgement

Bibliography

R. D. Cook and C. J. Nachtsheim. A comparison of algorithms for constructing d-optimal designs. *Technometrics*, 22:315–324, 1980.

N. R. Draper and D. K. J. Lin. Small response surface designs. *Technometrics*, 32:187–194, 1990.

H. O. Hartley. Smallest composite design for quadratic response surfaces. *Biometrics*, 15:611–624, 1959.

J. Lawson. *Vdgraph: This package creates variance dispersion graphs for response surface designs*, 2011. URL http://CRAN.R-project.org/package=Vdgraph. R package version 1.0-1.

J. S. Lawson. *Design and Analysis of Experiments with SAS*. CRC Press, Boca Raton, 2010.

R. V. Lenth. Response surface methods in R, using rsm. *Journal of Statistical Software*, 32(7):1–17, 2009.

Minitab Inc. Minitab software for quality improvement, 2010. URL http://www.minitab.com.

D. C. Montgomery. *Design and Analysis of Experiments*. John Wiley & Sons, New York, sixth edition, 2005.

R. H. Myers, G. Vining, A. Giovannitti-Jensen, and S. L. Myers. Variance dispersion properties of second order response surface designs. *Journal of Quality Technology*, 24:1–11, 1992.

R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. John Wiley & Sons, New York, 2009.

E. Santiago. Macro: VDG.MAC, 2009. URL www.minitab.com/support/macros/default.aspx.

SAS Institute Inc. Getting started with the SAS 9.2 ADX interface for design of experiments, 2010. URL http://support.sas.com/documentation/cdl/en/adxgs/60376/PDF/default/adxgs.pdf.

G. G. Vining. A computer program for generating variance dispersion graphs. *Journal of Quality Technology*, 25:45–58, 1993a.

G. G. Vining. Corrigenda: A computer program for generating variance dispersion graphs. *Journal of Quality Technology*, 25:333–335, 1993b.

John Lawson
Department of Statistics
Brigham Young University
Provo, UT (USA)
lawson@byu.edu