Artículo de investigación

Analysis of exprimental designs with outliers

Análisis de diseños experimentales con valores atípicos

Análise de desenhos experimentais com outliers

Recibido: 16 de enero de 2019. Aceptado: 06 de febrero de 2019

Written by:

Muhammad Salman Shabbir (Corresponding Author)[18]

Ahmed F. Siddiqi[19]

Normalini Md Kassim[20]

Faisal Mustafa[21]

Mazhar Abbas[22]

Abstract

Primary purpose of the article is to develop outlier robust designs. As a matter of fact, negative effect of outliers in any experimental settings is established where the outliers at any specific design point can destroy the features of the design for which it is being developed. It is attempted here in this article to develop a version of robustness for central composite designs which may protect it for outliers by introducing the idea of minimax outlying effect. This involves the calculation of the degree of outlying effect(s) outlier(s) may produce and then minimize the maximum of these outlying effects in an attempt to equalize the influence of all design points. On comparison, these outlier robust designs are proved to be more optimal, on the scales of A, D, and E optimalities, against existing conventional rotatable, orthogonal, and other such designs. The outlier robust designs, developed here, are suitable for settings prone to outliers where conventional designs fail to represent and analyze the processes and systems.

Keywords: Central Composite Designs, Robust Designs, Outliers, Minimax.

Resumen

El objetivo principal del artículo es desarrollar diseños robustos atípicos. De hecho, el efecto negativo de los valores atípicos en cualquier configuración experimental se establece donde los valores atípicos en cualquier punto de diseño específico pueden destruir las características del diseño para el que se está desarrollando. En este artículo se intenta desarrollar una versión de robustez para los diseños compuestos centrales que pueden protegerlo de los valores atípicos mediante la introducción de la idea del efecto periférico minimax. Esto implica el cálculo del grado de efecto (s) externo (s) que puede producir un valor atípico y luego minimizar el máximo de estos efectos externos en un intento de igualar la influencia de todos los puntos de diseño. En comparación, se demuestra que estos diseños robustos atípicos son más óptimos, en las escalas de las optimidades A, D y E, frente a los diseños convencionales existentes, ortogonales, rotativos y otros similares. Los diseños robustos atípicos, desarrollados aquí, son adecuados para configuraciones propensas a los valores atípicos en los que los diseños convencionales no representan ni analizan los procesos y sistemas.

Palabras claves diseños compuestos centrales, diseños robustos, valores atípicos, Minimax.

[18] Postdoctoral Fellow. School of Management. Universiti Sains Malaysia, Penang, Malaysia Email: salman@usm.my
[19] UCP Business School. University of Central Punjab. Lahore, Pakistan; Email: dr.ahmedsiddiqi@ucp.edu.pk
[20] School of Management. Universiti Sains Malaysia, Penang, Malaysia Email: normalini_mk@yahoo.com
[21] UCP Business School. University of Central Punjab, Lahore Pakistan Email: faisal.mustafa@ucp.edu.pk
[22] Department of Management Sciences. COMSATS University Islamabad, Vehari Campus Email: mazharabbas@ciitvehari.edu.pk

Resumo

Objetivo principal do artigo é desenvolver projetos robustos outlier. De fato, o efeito negativo de outliers em qualquer ambiente experimental é estabelecido onde os outliers em qualquer ponto de design específico podem destruir os recursos do design para o qual ele está sendo desenvolvido. Neste artigo, tenta-se desenvolver uma versão de robustez para projetos compostos centrais que possam protegê-lo de outliers, introduzindo a ideia de efeito periférico minimax. Isso envolve o cálculo do grau de efeito (s) outlier (s) outlier (s) pode produzir e, em seguida, minimizar o máximo desses efeitos periféricos em uma tentativa de equalizar a influência de todos os pontos do projeto. Em comparação, esses designs robustos discrepantes são comprovadamente mais otimizados, nas escalas de otimalidades A, D e E, contra os designs convencionais rotacionais, ortogonais e outros existentes. Os designs robustos outlier, desenvolvidos aqui, são adequados para configurações propensas a outliers em que projetos convencionais não representam e analisam os processos e sistemas.

Palavras-chave: projetos compostos centrais, projetos robustos, outliers, minimax.

Introduction

In layman's terms, using the language of noted physicist Stephen Hawkins, an outlier is "an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism". Outliers are everywhere. Experience shows that, in a typical experimenting setting, 1 to 10% of all observations are surprising in one way or the other and should be termed as outliers. It may be a result of keypunch error, misplaced decimal point, recording or transmission error, unusual happening such as earthquake, fire, or members of a different population slipping into the sample by mistake or unknowingly. Each result in outliers. A usual approach is to discard such suspicious observations by labelling them erroneous. However, as Festing and Altman (2002) put it, such observations should not be discarded unless there is independent evidence that the observation is incorrect. Further a seemingly unusual observation is not necessarily being erroneous all the time. It may be an indication of something unusual, hitherto unknown dimension. The current paper is written under the same assumption that no data point should be discarded. While techniques and methods should be devised to accommodate such surprising observations. It is attempted here to develop a technique which is robust enough to anticipate outliers in the domain of experimental designs.

Irrespective of the nature of these surprising observations, data analysis has always suffered because of these. There are numerous examples available in the academic literature showing how outliers ruin the classical statistical analysis. Whole complexion of the data, and of its interpretation, changes because of these outliers (Siddiqi, 2003). This drives the analyst to think seriously about these observations. The statistical literature seems to be divided between (i) rejecting, and (ii) accommodating outliers.

Peirce (1852), Peirce (1877) was probably the first who put forward some statistical reasoning for idea of deleting observation on the name of wrongly recorded observations. He proposes to delete observations when the probability of the system of errors obtained by retaining them is lesser than that of rejecting them multi- plied by the probability of making so many, and no more, wrongly recorded observations. This opens the gate of providing statistical excuses for such deletions; Gould (1855), Winlock (1856), Chauvenet (1960), Saunder (1903), Irwin (1925), Thompson (1935), Student (1927), Dixon (1950), Grubbs (1950), Chang et al. (2005) just a few names in the list of authors who advocate a rejection of outliers by labelling them to be erroneous observations. Beckman and Cook (1983), Barnett and Lewis (1964), Knorr et al. (2000), Bay and Schwabacher (2003), Cousineau and Chartier (2015) present an excellent commentary on different techniques to identify, detect and then delete these outliers. Broadly speaking, deleting outliers means the rejection of original observations.

The other school of thought, which attempts to accommodate outliers by devising different techniques. Probably, it was Glaisher (1873), Glaisher (1874) where the concept of accommodating outliers was discussed with acceptable statistical details. He expounded the idea of differing distributions for the observations coming out from a single experiment by assuming that each observation came from a normal distribution with unknown means and unequal variances. Using a scheme of iterative re-weighted least squares, obtained an estimate for the mean. Irrespective of the fact that the assumptions are too big to meet, this idea of iterative re- weighted least squares is the precursor of many techniques used today. The also gives birth to the idea of robust regression where attempts are made to develop models that can automatically and intrinsically off-set the negative effects of outliers, if there exist any. These models work on the methodology of down weighing an out weigh observation so that all observations have identical weights. Newcomb (1886) puts the iterative weighting of observations to sounder basis by suggesting a mixed system of observation of changing variances in which the mean value is estimated by weighting the observation. Then comes the Edgeworth (1887) who hypothesizes three models for outliers and illustrates each using Monte Carlo simulation methods. Daniell (1920) modified the base of weights by using sampling distribution with weights decreasing with the size of deviation from the mean for a heavy tailed normal distribution.

The idea of iterative re-weighted least squares culminates into four broad types of robust estimators in the domain of robust regression. (i) L estimators, which involves weighing of order statistics. It is primarily non-parametric in nature and very close to distance based upon ranks of observations. (ii) M estimators, which involves the weighing of residuals, instead of observations. (iii) Bayesian estimators which makes use of priori and posteriori probability distribution instead of weights, and (iv) a hybrid method consisting of outright rejection followed by estimation (weights = 0 or 1) with the emphasis on the performance of the estimator.

Box and Draper (1975) introduced their version of robust design by minimizing the discrepancy caused by the outlier(s). Despite computational complexity, their version of the robust design has conceptual appeal. Siddiqi (2003) has compared these four estimators in the domain of designed experiments for a better, or at least optimum estimator in a particular situation. Siddiqi (2008), Siddiqi (2011) developed designs robust to outliers using minimax criterion of Akhtar and Prescott (1986). The designs show many desirable properties besides robustness and alphabetic optimality but, unfortunately, their robustness was limited to only one outlier.

The current article discusses the design robustness for centre composite designs (CCD) as illustrated in Figure a. These designs are developed by Box and Hunter (1957) for analyzing response surfaces and are special designs having three distinct types of points, (i) factorial points, (f), coming primarily from some known full, or fractional, balanced or unbalanced, orthogonal or non-orthogonal, factorial design, (ii) axial points, (a), at a distance α from the centre, number of which are a multiple of factors, and (iii) a few central points, (c). This makes a CCD(f,a,c) design. Curious reader should be referred to Box and Draper (1987), Montgomery and Myers (1995) for more details on response surfaces and central composite designs. Figure a shows a typical CCD(23, a,1) where a 22 full factorial design is augmented with $2\times2 = 4$ axial and a single central point. The design is a cuboid, developed by 3 factors, whose two levels are shown by dots at the edges of the cuboid, the four axial points are the axes, while the single central point is shown at the centre. The following sections shows how this design is made robust for outliers.

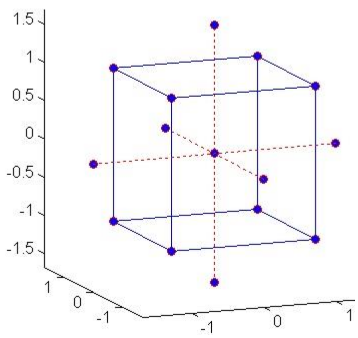


Figure a:A Typical Central Composite Design Developed with 23 Complete factorial Design & Augmented with Central, and Axial and Points

The very rationale of CCD are the optimal designs in different scenarios. The flexible axial distance α, the option of replication, the inclusion of design points at the centre of the cube, etc. all these characteristic features of the design are making them the most attractive choice in any manufacturing setting. However, the outliers are capable enough to mar these attractive features. All the different versions of these designs, like Rotatable, Orthogonal, Optimal, etc. which are developed to give results even in compromised situations seem to fail in the presence of outliers. It is attempted here, in the article, to develop outlier robust design, based on the idea of minimax, which can retain the attractive features of the CCD designs.

Outlier Robust Minimax Designs

The designs are based upon the philosophy of robust regression. As has been discussed earlier in previous section, the designs attempt to accommodate outliers instead of deleting them. In simple language, these designs are developed by minimizing the maximum outlying effect. These designs would ensure that no design point has an effect large enough to disturb the design.

More specifically, the minimax designs attempt to

1. calculate outlying effects of all design points,
2. sort these outlying effects in decreasing order of magnitudes,
3. lower these larger effects, by intelligently changing the design, up to the extent where all the design points have identical outlying effects.

The philosophy of minimax is not unique to outliers. Akhtar and Prescott (1986) has used this philosophy to develop design robust to missing observations. Sitter (1992) has used this philosophy to develop designs robust to poor estimates. Heo et al. (2001), Mukerjee and Huda (1985), Wiens (1990), Wiens (1992) use this criterion to develop restricted designs for linear regression. The list of such authors is very long, each using this philosophy to develop design to circumvent some un-wanted.

- Developing Outlier Robust Minimax Designs. Consider a classical response surface model of the form

$$y = X\beta + \varepsilon$$

Consider the additive setup, the y is an $(n\times 1)$ vector of response variable, X is the $(n\times p)$ known design matrix of the rank p $(p \leq n)$, β is a $(p \times 1)$ vector of unknown regression parameters, while ε is an $(n \times 1)$ vector of residuals assumed to be normally distributed with zero mean and σ2 variance.

For a CCD(f, a, c), the X matrix is composed of three types of points; the factorial, the axial whose numbers are simply a multiple of factorial points, and the central. Further, there may exists replications of either factorial, axial or both. So,

$$\begin{aligned} n &= (f \times r_f) + (a \times r_a) + c \\ &= (f \times r_f) + (2 \times f \times r_a) + c \\ &= (r_f + 2r_a)f + c \end{aligned}$$

where rf and ra denotes the replications for factorial and axial parts respectively. Similarly,

$$\begin{aligned} p &= 1 + \sum_{i=1}^{k} f + \binom{f}{2} + --- + \binom{f}{f} \\ &= 1 + kf + \binom{f}{2} + --- + \binom{f}{f} \end{aligned}$$

with k denotes the degree of the polynomial used in the response surface, defined above.

Provided the type and nature of the regression function is correct, the response variable y is influenced only by the given independent variables in X. The information matrix, X'X (due to Fisher (1922), developed by these independent variables, is a repository of this influence (Chatterjee and Hadi, 1986, Cook and Weisberg, 1980). Atkinson et al. (2014), Gao and Yang (2015), Hoaglin and Welsch (1978), Jauffret (2007) among others, use different variants, mostly partial and based upon some subset, of this information matrix for extracting partial information about the subset of the whole data. Siddiqi (2008), Siddiqi (2011) use the information matrix to develop outlier robust designs by introducing minimax criteria to develop minimax designs. If Oj denotes the outlying effect of an outlier, happens to be the jth design point, its value, as calculated by Siddiqi (2008), is given by

$$\begin{aligned} O_j &= \frac{|X'X| - |X'_{(j)}X_{(j)}|}{|X'X|} \\ &= 1 - \frac{|X'_{(j)}X_{(j)}|}{|X'X|} \end{aligned}$$

Using Rao and Toutenburg (1995)

$$\begin{aligned} |X'_{(a)}X_{(a)}| &= |X'X - x'_a x_a| \\ &= |X'X||1 - x_a(X'X)^{-1}x'_a| \end{aligned}$$

This implies, the outlying effect of a single outlier, at the jth design point, is given by

$$\begin{aligned} O_j &= 1 - \frac{|X'X||1 - x_j(X'X)^{-1}x'_j|}{|X'X|} \\ &= h_j \end{aligned}$$

with hj as the jth element in the main diagonal of the HAT matrix, defined by Hoaglin and Welsch (1978) and given by X'(X'X)-1 X. Continue with the idea, if Ojk denotes the outlying effect of jth and kth two design points, its value in terms on information matrix is given by

$$\begin{aligned} O_{jk} &= \frac{|X'X| - |X'_{(jk)}X_{(jk)}|}{|X'X|} \\ &= 1 - \frac{|X'_{(jk)}X_{(jk)}|}{|X'X|} \end{aligned}$$

Using Rao and Toutenburg (1995)

$$\begin{aligned} O_{jk} &= 1 - \frac{|X'_{(j)}X_{(j)} - x'_{(k)}x_{(k)}|}{|X'X|} \\ &= 1 - \frac{|X'_{(j)}X_{(j)}|\left|1 - x'_{(k)}\left(X'_{(j)}X_{(j)}\right)^{-1}x_{(k)}\right|}{|X'X|} \\ &= 1 - \frac{|X'X||1 - x_j(X'X)^{-1}x'_j|}{|X'X|}\left|1 - x'_{(k)}\left(X'_{(j)}X_{(j)}\right)^{-1}x_{(k)}\right| \\ &= 1 - |1 - x_j(X'X)^{-1}x'_j|\left|1 - x'_{(k)}\left(X'_{(j)}X_{(j)}\right)^{-1}x_{(k)}\right| \end{aligned}$$

Using Rao and Mitra (1973), Rao and Rao (1998)

$$\left(X'_{(a)}X_{(a)}\right)^{-1} = (X'X)^{-1} + \frac{(X'X)^{-1}x'_a x_a(X'X)^{-1}}{1 - x_a(X'X)^{-1}x'_a}$$

This implies to

$$
\begin{aligned}
O_{jk} &= 1-\left|1-x_j(X'X)^{-1}x'_j\right|\left|1-x'_{(k)}\left\{(X'X)^{-1}+\frac{(X'X)^{-1}x'_jx_j(X'X)^{-1}}{1-x_j(X'X)^{-1}x'_j}\right\}x_{(k)}\right| \\
&= 1-\left|1-x_j(X'X)^{-1}x'_j\right|\left|1-x'_{(k)}(X'X)^{-1}x_{(k)}-\frac{x_k(X'X)^{-1}x'_jx_j(X'X)^{-1}x'_{(k)}}{1-x_j(X'X)^{-1}x'_j}\right|
\end{aligned}
$$

Using the symbiology of Hoaglin and Welsch (1978)

$$H = X'(X'X)^{-1}X$$

which is usually termed as HAT matrix in statistical literature. It is a positive definite, symmetric matrix (Eubank, 1984) with

$$
\begin{aligned}
h_{aa} &= x_a(X'X)^{-1}x'_a \\
h_{ab} &= x_a(X'X)^{-1}x'_b
\end{aligned}
$$

as being diagonal and off-diagonal components of the HAT matrix, respectively,

$$
\begin{aligned}
O_{jk} &= 1-\left(1-h_{jj}\right)\left\{1-h_{kk}-\frac{h_{jk}^2}{1-h_{jj}}\right\} \\
&= 1-\left(1-h_{jj}\right)(1-h_{kk})+h_{jk}^2
\end{aligned}
$$

which gives the combined outlying effect of two design points for a CCD(f,a,c). As per the philosophy of minimax designs, this is to be calculated for each pair of the design points to find an axial distance for which it is identical for all these pairs.

- Some Properties of Ojk

1. It is positive for all values of j and k.
2. $0 \le O_{jk} \le 1$
3. Upper bounds for Ojk depends upon the replication of the design; more the replication lesser is the upper bounds.
4. Ojk remains the same for all replications of factorial points (Siddiqi, 2008), for all replications of axial points, and for all center points, especially when the factorial design in the CCD is complete. However, for fractional factorials, these effects may be different.

Consider a central composite design developed for a 22 complete factorial design and augmented with 2 × 2 = 4 axial point and 2 (say) central points where each of the axial point lies at a distance a from the center, i.e., CCD (22,a,2). The design matrix, X, for such a design, using Box and Hunter (1957), is given by

$$
X=\begin{pmatrix}
1 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & a & 0 & a^2 & 0 & 0 \\
1 & -a & 0 & a^2 & 0 & 0 \\
1 & 0 & a & 0 & a^2 & 0 \\
1 & 0 & -a & 0 & a^2 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

The actual realization of the design matrix depends upon the axial distance, a. The outlying effect of the design points may be calculated by using the expressions developed earlier in this section.

As a matter of fact, these outlying effects are calculated on the basis of the HAT matrix; the outlying effect of any single design point happens to be the respective diagonal term, while the outlying effect of two design point is a linear combination of diagonal and off-diagonal terms. There is a Mathematica code available in Appendix A to facilitate the calculations involved in the development of this hat matrix. For the design CCD (22,a,2), the hat matrix in this case is a (10 × 10) matrix, with each component requires the realization of the axial distance a, is given by

$$H=\begin{pmatrix} \frac{5a^6+6a^4-40a^2+96}{12a^6-8a^4-16a^2+96} & -\frac{a^6}{12a^4-32a^2+48} & \cdots & \frac{a^2(a^2-2)}{6a^4-16a^2+24} \\ -\frac{a^6}{12a^4-32a^2+48} & \frac{5a^6+6a^4-40a^2+96}{12a^6-8a^4-16a^2+96} & \cdots & \frac{a^2(a^2-2)}{6a^4-16a^2+24} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a^2(a^2-2)}{6a^4-16a^2+24} & \frac{a^2(a^2-2)}{6a^4-16a^2+24} & - & \frac{a^4+4}{6a^4-16a^2+24} \end{pmatrix}$$

As a matter of interest, each component of this matrix is associated with some design point of the CCD(22, a,2); the diagonal components with a single design point, while the off-diagonal components with two design points. More specifically, first four diagonal elements are representing factorial part, second four the axial part, and the last two diagonal elements are representing the two replications of central part of the CCD(22, a,2). Owing to the symmetric nature of the HAT matrix lower off-diagonal points are similar to upper off-diagonal components. Further, due to a peculiar structure of the design, as stated earlier, the diagonal elements corresponding to factorial points are identical, the diagonal elements corresponding to axial points are identical to each other, and the diagonal elements corresponding to central points are also identical to each other. This results in these six different types of outlying effects,

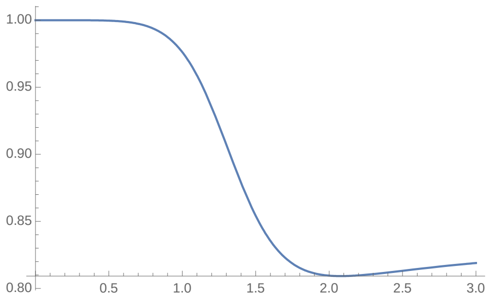
- ff : Ojk is calculated for pairs of design points where both points belong to the factorial portion of the CCD,
- aa : Ojk is calculated for pairs of design points where both points belong to the axial portion of the CCD,
- cc : Ojk is calculated for pairs of design points where both points belong to the central portion of the CCD,
- fa : Ojk is calculated for pairs of design points where points belong to factorial and axial portions of the CCD,
- fc : Ojk is calculated for pairs of design points where points belong to factorial and central portions of the CCD,
- ac : Ojk is calculated for pairs of design points where points belong to axial and central portions of the CCD,

The outlier robust minimax design attempts to minimize the maximum outlying effects to make them identical. This would not let any design point to be outlying enough to destroy the properties of the design. As a matter of interest, Ojk would not necessarily be the same for all factorial design points. As a matter of fact, for full factorial designs it would be the same while for fractional factorial, confounded, split plot designs it would not be the same. Similarly, it would not be the same for all axial or central design points. This requires to investigate the maximas for each of the type of design effects. The Mathematica code, in Appendix A, also helps in digging out these maximas. Table 1 shows these maximas for different combinations of design points for CCD (22, a,2).

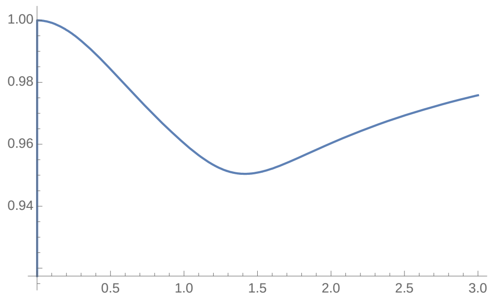
Table 1. Expressions Showing the Maximas for Different Types of Design Effects

Design Effect	Maxima
ff	$\frac{5a^6 - 4a^4 - 8a^2 + 48}{6a^6 - 4a^4 - 8a^2 + 48}$
fa	$\frac{6a^8 + 6a^6 + a^5 - 14a^4 + 2a^3 + 24a^2 + 96}{2(a^2 + 2)^2(3a^4 - 8a^2 + 12)}$
fc	$\frac{3a^6 + 14a^4 - 72a^2 + 192}{8(3a^6 - 2a^4 - 4a^2 + 24)}$
aa	$\frac{3a^8 + 4a^6 - 11a^4 + 22a^2 + 32}{(a^2 + 2)^2(3a^4 - 8a^2 + 12)}$
ac	$\frac{6a^6 - 9a^4 + 7a^2 + 26}{6a^6 - 4a^4 - 8a^2 + 48}$
cc	$\frac{a^4 + 4}{3a^4 - 8a^2 + 12}$

All these expressions in Table I are driven by the axial distance, a, as all of these expressions are some polynomial of a. Fig. 2 is the visual demonstration of these maximas. Each of the subfigure, from 2(a) to 2(f), is drawn for one of the maxima shown in Table I where the horizontal axis measures the axial distance, a while the vertical axis shows how the outlying effect changes with a. It is quite interesting to note that each of these sub-figures are non-linear in nature and shows parabolic curves. This implies that the outlying effect is neither increasing nor decreasing continuously with changing a and their exist a maxima, or minima, for each of these effects.

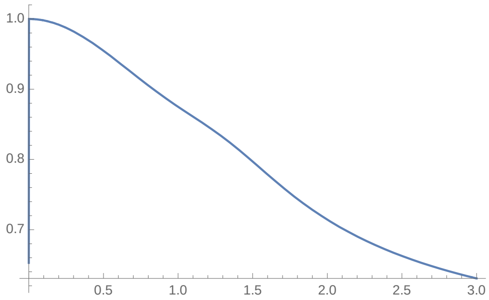


(a) Both Factorial Points

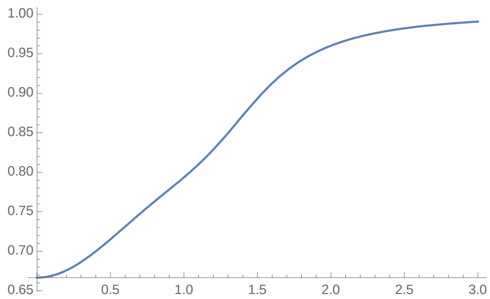


(b) Factorial & Axial Points

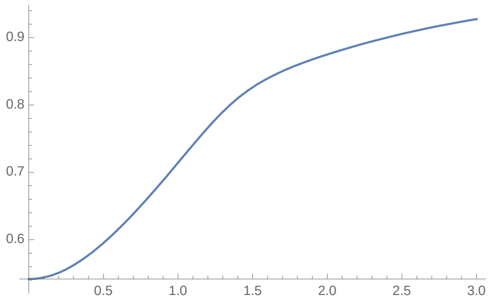
 ISSN 2322- 6307



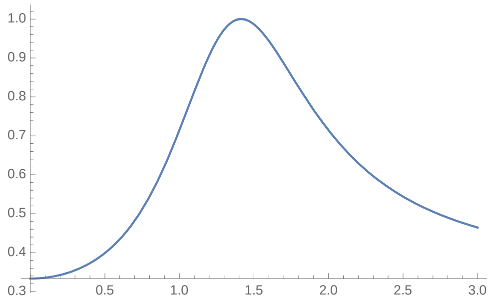
(c) Factorial & Central Points



(d) Both Axial Points



(e) Axial & Central Points



(f) Both Central Points

Figure 2. Outlying Effects for Different Combinations of Design Points As Varied with Axial Distance

$\sum_j^n \sum_{k\neq j}^n O_{jk} = f(a,p)$ due to Cook and Weisberg (1982) who discussed many properties of the hat matrix. In other words, for a constant p, the only source of variation in $\sum_j^n \sum_{k\neq j}^n O_{jk} = f(a,p)$ is a. This implies that any change in Ojk for some values of j and k would be offset by a similar but opposite change in Ojk for other values of j and k, i.e., lowering one effect implicitly off-set by an increase in some other effect. It is also explicit in Fig. 3 which is showing all these curves on a single graph.

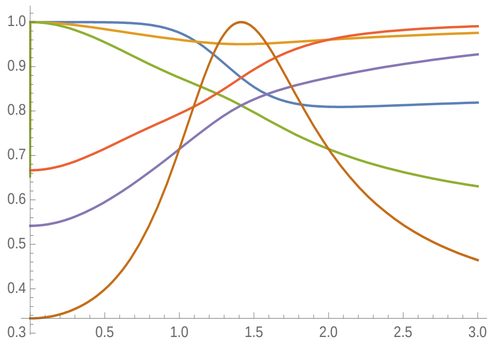


Figure 3. Outlying Effects for Different Combinations of Design Points As varied with Axial Distance

The minimax design requires a value for a at which either

1. all the outlying effects are equal, or
2. as many as possible outlying effects are identical while all remaining are lesser.

As a matter of fact, the situation 1 is rarely possible and one has to dig out for the second scenario. This second situation calls for a very complex non-linear programming modeling. Table 2 shows all such combinations for the CCD(22,a,2).

Table 2. Minimax Design

Design Points		Axial Distance	Outlying Effect	, both
ff	fa	1.11382	0.95598	0.744405,0.812556,0.827658,0.859052,0.95598,0.95598
ff	fc			
ff	aa	1.41421	0.875001	0.8125,0.8125,0.875,0.875,0.950444,1.
ff	ac	1.58563	0.837759	0.781149,0.837758,0.837758,0.910095,0.951258,0.951921
ff	cc	1.81761	0.814383	0.740872,0.814383,0.814383,0.860926,0.94386,0.956451
		1.22733	0.927861	0.773111,0.83424,0.842872,0.927862,0.927862,0.952652
fa	fc			
fa	aa	2	0.96032	0.714286,0.714286,0.809524,0.875,0.960317,0.960317
fa	ac			
fa	cc	1.58434	0.837759	0.781386,0.837601,0.83797,0.90986,0.951902,0.951902
fa	cc	1.26235	0.951902	0.781386,0.837601,0.841547,0.918042,0.951902,0.951902
fc	aa	1.2515	0.839253	0.778858,0.839254,0.839254,0.921123,0.944901,0.952118
fc	ac	1.41421	0.812501	0.8125,0.8125,0.875,0.875,0.950444,1.
fc	cc	2	0.71429	0.714286,0.714286,0.809524,0.875,0.960317,0.960317
fc	cc	2.14229	0.696754	0.751808,0.817698,0.855077,0.855077,0.949601,0.955032
aa	ac			
aa	cc	1.64328	0.920068	0.770626,0.829331,0.844382,0.920068,0.920068,0.952899
aa	cc	1.09522	0.809299	0.739524,0.809299,0.809299,0.861639,0.956633,0.959867
ac	cc	1.75086	0.855079	0.751808,0.818514,0.855077,0.855077,0.935829,0.955032
ac	cc	1	0.71429	0.714286,0.714286,0.793651,0.875,0.960317,0.97619

The table is made up of values as produced by the Mathematica code, given in Appendix A. The first two columns shows the design point's combination while the third shows a sorted list, in increasing order of magnitudes, of all outlying effects at an axial distance where the design point's in first two columns interact. Since the list is sorted, the last value shows the maximum outlying effect. And if last two outlying effects happens to be identical then this shows that the two maximas intercept at an a value for which all the other maximas are lesser. And this is the axial distance for which minimax designs are defined. For the CCD(22,a,2);

- there does not exist any single a for which all the outlying effects happens to be identical,
- there does exist many a for which at least two outlying effects are identical, like at a 1.11382, both ff and fa are identical, at a = 1.41421 ff and aa are identical, etc. The Table 2 lists all such values of a where some outlying effects are identical.

However, the minimax designs are looking for an a, as per second situation discussed above, for which, as many as possible, outlying effects are identical while all remainings are lesser. This occurs at a = 1.58563 for a CCD(22,a,2). So CCD(22,1.58563,2) is an outlier robust design with 2 factors.

A similar exercise can be made for other choices of factors, and of their replications. The Mathematica code given in Appendix A is capable enough to develop outlier robust designs for almost all possible choices of number of factors, design structures like full or fractional, central points, and their replications. Table 3 shows these designs for a few such options.

Comparisons

An attempt is made here to compare these minimax designs with a few already exists established designs, like

1. Box and Draper (1975) Designs,
2. Box and Hunter (1957) Rotatable Designs
3. Geramita et al. (1976) Orthogonal Designs

The comparison is made on the basis of design's alphabetic optimalities like

- A which focuses in minimizing the average variance of the estimates of the response surface model's coefficients,

- D which focuses in maximizing the differential Shannon (1948) information content of the parameter estimates including both of its versions;

- main effect only,
- squared effect only calculations, and
- E which focuses on maximizing the minimum eigenvalue of the information matrix.

More curious reader should refer to Atkinson et al. (2007), Kiefer (1985) for detailed and better understanding of these design alphabetic optimalities. Table 3 shows the results of this comparison where the outlier robust designs are compared with these designs. The Mathematica code given in Appendix A is used to generate all these outlier robust designs and the corresponding comparisons.

Table 3. Comparison of Minimax Designs with Other Designs

Design		Rotatable	Orthogonal	BD	MinMax
CCD(22, a, 2)	a	1.41421	0.413812	1.41421	1.58434
Factorial=4 x 1	A-optimality	1.4375	18.183	1.4375	1.25805
Axial=4 x 1	D-optimality	0.00001526	0.0052729	0.00001526	5.6329x10-6
Central=2	Dmain	0.015625	0.0530303	0.015625	0.0122902
Total=10	Dsqre	0.0078125	2.1159	0.0078125	0.00385188
	E-optimality	0.772022	17.0513	0.772022	0.67263
CCD(2A,a,2)	a	1.68179	0.332478	1.79999	2.0688
Factorial=8 x 1	A-optimality	1.41557	82.7971	1.34838	1.13038
Axial=6 x 1	D-optimality	3.7221 x 10-11	0.000032403	1.6054 xlO-11	2.172 x 10-12
Central=2	Dmain	0.000392598	0.00179975	0.000329384	0.000220206
Total=16	Dsqre	0.0000976563	69.6923	0.0000504226	0.0000122875
	E-optimality	0.677562	40.9184	0.654114	0.506229
CCD(23,a,2)	a	2.	0.330012	1.41935	1.38795
Factorial=8 x 2	A-optimality	0.916667	84.8812	1.14273	1.1457
Axial=6 x 1	D-optimality	8.9826 xl0-14	2.7979x 10-7	3.2180x 10-12	3.8340 xl0“12
Central=2	Dmain	0.000072338	0.000234435	0.000124455	0.000127801
Total=24	Dsqre	0.000012207	37.0039	0.00027047	0.000327538
	E-optimality	0.531874	42.1553	0.546279	0.524726
CCD(23,a,2)	a	1.68179	0.274548	2.10957	1.95542
Factorial=8 x 1	A-optimality	1.28345	88.882	0.973953	1.09769
Axial=6 x 2	D-optimality	2.3441 xl0“12	0.000020543	6.6191 x 10-14	2.4201 xl0“13
Central=2	Dmain	0.000138804	0.00174795	0.0000582218	0.0000791102
Total=22	Dsqre	0.0000174386	80.595	1.5437x10-6	3.5449x10“®
	E-optimality	0.677552	44.0013	0.449302	0.549965
CCD(23,a,2)	a	2.	0.26814	2.03969	2.1638
Factorial=8 x 2	A-optimality	0.805556	97.2141	0.767485	0.662901
Axial=6 x 2	D-optimality	6.3159x 10-15	2.0495x 10-7	4.5674 xl0“15	1.6756x 10-15
Central=2	Dmain	0.0000305176	0.000231435	0.0000287536	0.0000238759
Total=30	Dsqre	2.1798x10-6	48.7032	1.7795x10-6	9.5856x 10-7
	E-optimality	0.48591	48.3607	0.452294	0.360068
CCD(23“1,a, 2)	a	1.41421	0.336822	1.54038	1.34512
Factorial=4 x 1	A-optimality	3.16667	105.154	2.90189	3.3442
Axial=3 x 2	D-optimality	7.9472x 10-8	22.2544	2.3217 x 10-8	1.5880x 10-7
Central=2	Dmain	0.00195312	0.0132414	0.001495	0.00226132
Total=12	Dsqre	0.00078125	125.495	0.000339083	0.0012577
	E-optimality	0.654508	38.848	0.680328	0.704644
CCD(24,a,2)	a	2.	0.279624	2.	2.31839
Factorial=16 x 1	A-optimality	1.27083	246.123	1.27083	1.06249
Axial=8 x 1	D-optimality	2.8555x 10_2°	7.7160x 10_1°	2.8555x 10_2°	1.9848x 10-21
Central=2	Dmain	3.0140x10-®	0.0000146765	3.0140x10-®	1.9530x10-®
Total=26	Dsqre	3.1789x 10-7	8545.98	3.1789x 10-7	4.2569x 10-8
	E-optimality	0.627111	81.7853	0.627111	0.479169

CCD(2b, a, 2)	a	2.37841	0.244901	2.02905	1.46112
Factorial=32 x 1	A-optimality	1.05392	556.566	1.0983	1.13595
Axial=10 x 1	D-optimality	7.3012x 10-34	5.1530x 10-18	1.4572x 10-32	2.3734 x10-30
Central=2	Dmain	6.559x 10-9	2.9249x 10-8	9.4847 x 10-9	1.5932x 10-8
Total=44	Dsqre	2.6609x 10_1°	2.3328x10-9	3.9048x 10-9	8.5649x 10-7
	E-optimality	0.559727	138.997	0.539294	0.241989

Different choices of the number of factors, the design structure; either full or fractional, their replications are explored, just to save the space and show the optimalities, the newly generated outlier robust designs have managed to attain. While for each design, an a is calculated as per the description of the design. And, then values are calculated for different optimalities criteria. Lesser the value of these optimality criteria better the design be.

As per the literature, and the expressions given in Atkinson et al. (2007) a design with smaller optimality coefficient is considered to be more optimal. A closer look in these values, reveal the optimality of outlier robust designs in almost every configuration and structure. As a matter of fact, the outlier robust designs are 30% to 50% more efficient to these designs.

Concluding Remarks

Outliers have always been the problem set of points in almost all domains of statistical analysis. Different approaches are available in academic statistical literature to tackle their negative effects. Typically, analysts use to delete these observations by labeling them as erroneous or typo. However, such an approach is not welcoming, especially in the domain of experimental design, where each and every observation corresponds to a specific design point. Another approach is to accommodate such points but after down-weighing their outlying effects. This assures the presence of these points in the analysis, while minimizing their negative effects.

The same accommodative approach is used in this paper by introducing a new type of designs, that are robust in nature, and attempt to minimize the maximum outlying effect of design points. It derives the expressions to calculate these outlying effects by using a version Fisher's information matrix. Then it employs a non linear programming scheme to minimize maximum outlying effects. It also compares these Minimax designs with Box and Draper (1975) Designs, Box and Hunter (1957) Rotatable Designs, and Geramita et al. (1976) Orthogonal Designs, by using alphabetic optimalities like A, D, and E. It turns out that Minimax designs are more optimal as compared to these aforementioned designs for almost all choices of factors, design structures; either full or fractional, and their replications.

The Minimax designs developed here in this paper are capable enough to accomodate a pair of outliers only. However the methodology may be extended for more outliers. They are robust, that is, outliers cannot disturb or belittle their desirable design properties, like optimalitity, etc. The paper uses a Mathematica code, given in Appendix A to facilitate different derivations and calculations involved in the development of these Minimax designs. Even a high performance computer (Intel's 6th Generation i7 with 16 GB RAM) takes lengthy durations for the development of these designs. However better programming skills may reduce these lengthy durations.

Reference

Akhtar, M., & Prescott, P. (1986). Response surface designs robust to missing observations. Communications in Statistics-Simulation and Computation, 15(2), 345-363.

Atkinson, A., Donev, A., & Tobias, R. (2007). Optimum experimental designs, with SAS (Vol. 34). Oxford University Press.

Atkinson, A. C., Fedorov, V. V., Herzberg, A. M., & Zhang, R. (2014). Elemental information matrices and optimal experimental design for generalized regression models. Journal of Statistical Planning and Inference, 144, 81-91.

Barnett, V., & Lewis, T. (1964). Outliers in statistical data, Chichester: John Wiley, 1995. 584p.

Bay, S. D., & Schwabacher, M. (2003, August). Mining distance-based outliers in near linear time with randomization and a simple pruning rule. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 29-38). ACM.

Beckman, R. J., & Cook, R. D. (1983). Outlier………. s. Technometrics, 25(2), 119-149.

Box, G. E., & Draper, N. R. (1975). Robust designs. Biometrika, 347-352.
Box, G. E., & Draper, N. R. (1987). Empirical model-building and response surfaces. John Wiley & Sons.
Box, G. E., & Hunter, J. S. (1957). Multi-factor experimental designs for exploring response surfaces. The Annals of Mathematical Statistics, 28(1), 195-241.
Chang, L. C., Jones, D. K., & Pierpaoli, C. (2005). RESTORE: robust estimation of tensors by outlier rejection. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 53(5), 1088-1095.
Chatterjee, S., & Hadi, A. S. (1986). Influential observations, high leverage points, and outliers in linear regression. Statistical Science, 379-393.
Chauvenet, W. (1960). A Manual of Spherical and Practical Autonomy (Vol. II.
Cook, R. D., & Weisberg, S. (1980). Characterizations of an empirical influence function for detecting influential cases in regression. Technometrics, 22(4), 495-508.
Cook, R. D., & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall.
Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment: a review. International Journal of Psychological Research, 3(1), 58-67.
Daniell, P. J. (1920). Observations weighted according to order. American Journal of Mathematics, 42(4), 222-236.
Dixon, W. J. (1950). Analysis of extreme values. The Annals of Mathematical Statistics, 21(4), 488-506.
Edgeworth, F. Y. (1887). Xli. on discordant observations. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 23(143), 364-375.
Eubank, R. L. (1984). The hat matrix for smoothing splines. Statistics & probability letters, 2(1), 9-14.
Festing, M. F., & Altman, D. G. (2002). Guidelines for the design and statistical analysis of experiments using laboratory animals. ILAR journal, 43(4), 244-258.
RA Fisher, M. A. (1922). On the mathematical foundations of theoretical statistics. Phil. Trans. R. Soc. Lond. A, 222(594-604), 309-368.
Gao, G., & Yang, J. (2015, June). Matrix Based Regression with Local Position-Patch and Nonlocal Similarity for Face Hallucination. In International Conference on Intelligent Science and Big Data Engineering (pp. 110-117). Springer, Cham.
Geramita, A. V., Geramita, J. M., & Wallis, J. S. (1976). Orthogonal desingns. Linear and Multilinear Algebra, 3(4), 281-306.
Glaisher, J. W. L. (1873). On the rejection of discordant observations. Monthly Notices of the Royal Astronomical Society, 33, 391-402.
GLAISHER, J. 1874. Note on a Paper by Mr. Stone; On the Rejection of Discordanr Observations. Monthly Notices of the Royal Astronomical Society, 24, 251.
Gould, B. A. (1855). On Peirce's Criterion for the Rejection of Doubtful Observations, with tables for facilitating its application. The Astronomical Journal, 4, 81-87.
Grubbs, F. E. (1950). Sample criteria for testing outlying observations. The Annals of Mathematical Statistics, 21(1), 27-58.
Heo, G., Schmuland, B., & Wiens, D. P. (2001). Restricted minimax robust designs for misspecified regression models. Canadian Journal of Statistics, 29(1), 117-128.
Hoaglin, D. C., & Welsch, R. E. (1978). The hat matrix in regression and ANOVA. The American Statistician, 32(1), 17-22.
Irwin, J. O. (1925). On a criterion for the rejection of outlying observations. Biometrika, 238-250.
Jauffret, C. (2007). Observability and Fisher information matrix in nonlinear regression. IEEE Transactions on Aerospace and Electronic Systems, 43(2), 756-759.
Kiefer, J. (1985). Collected Papers III: design of experiments. Springer-Verlag.
Knorr, E. M., Ng, R. T., & Tucakov, V. (2000). Distance-based outliers: algorithms and applications. The VLDB Journal—The International Journal on Very Large Data Bases, 8(3-4), 237-253.
Montgomery, D. C., & Myers, R. H. (1995). Response surface methodology: process and product optimization using designed experiments, Raymond H. Meyers, Douglas C. Montgomery, A Wiley-Interscience Publications.
Mukerjee, R., & Huda, S. (1985). Minimax second-and third-order designs to estimate the slope of a response surface. Biometrika, 72(1), 173-178.
Newcomb, S. (1886). A generalized theory of the combination of observations so as to obtain the best result. American journal of Mathematics, 343-366.
Peirce, B. (1852). Criterion for the rejection of doubtful observations. The Astronomical Journal, 2, 161-163.
Peirce, B. (1877, May). On Peirce's criterion. In Proceedings of the American Academy of Arts

and Sciences (Vol. 13, pp. 348-351). American Academy of Arts & Sciences.
Rao, C. R., & Mitra, S. K. (1973). Theory and application of constrained inverse of matrices. SIAM Journal on Applied Mathematics, 24(4), 473-488.
Rao, C. R., & Rao, M. B. (1998). Matrix algebra and its applications to statistics and econometrics. World Scientific.
Rao, C. R., & Toutenburg, H. (1995). Linear models. In Linear models (pp. 3-18). Springer, New York, NY.
Saunder, S. A. (1903). Note on the use of Peirce's criterion for the rejection of doubtful observations. Monthly Notices of the Royal Astronomical Society, 63, 432-436.
Shannon, C. E. (1948). A mathematical theory of communication. Bell system technical journal, 27(3), 379-423.
Siddiqi, A. F. (2003). Outliers in Designed Experiments; Classical Robust & Resistant Methods. Journal of Statistics, 10, 49-65.
Siddiqi, A. F. (2008). Outlier robust designs. International Journal of Applied Mathematics and Statistics™, 13(M08), 64-77.
Siddiqi, A. F. (2010). Outlier Robust Draper & Lin Designs. Pakistan Journal of Statistics and Operation Research, 7(1), 87-99.
Sitter, R. R. (1992). Robust designs for binary data. Biometrics, 1145-1155.
STUDENT 1927. Errors of routine analysis. Biometrika, 151-164.
Thompson, W. R. (1935). On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation. The Annals of Mathematical Statistics, 6(4), 214-219.
Wiens, D. P. (1990). Robust minimax designs for multiple linear regression. Linear algebra and its applications, 127, 327-340.
Wiens, D. P. (1992). Minimax designs for approximately linear regression. Journal of Statistical Planning and Inference, 31(3), 353-371.
Winlock, J. (1856). On Professor Airy's objections to Peirce's criterion. The Astronomical Journal, 4, 145-147.

Mathematica Code

```
Clear["Global'*"]
{fact,rns,rf,ra,cent}={2,4,1,1,2};
d={"1","a","b","ab","c","ac","bc","abc"};
Print["Outlier Robust Center Composite Designs ; Single Outlier Case"]
Print["Basic Design = 2^"<>ToString[fact]<>" in "<>ToString[rns]<>" runs"]
Print["Design Pts. =",d]
Print["2nd order RSCCD = factorial x "<>ToString[rf]<>", axial x "<>ToString[ra]<>", center pts. x
"<>ToString[cent]]
{w,ax,des,rows,col}={1,2 fact,rns+2 fact+cent,rns rf+2 fact ra+cent,(fact+1) (fact+2)/2};
{t,q}={2 fact ra+cent,((rns rf+t)^.5-(rns rf)^.5)^2};
x=Array[ds,{des,col}];
Do[ds[i,j]=0,{j,col},{i,rows}]
Do[ds[i,1]=1,{i,rows}]
Do[rn=d[[i]];
k=1;
Do[If[StringTake[rn,{k}]==FromCharacterCode[95+j],ds[i,j]=1;
k++;
If[StringLength[rn]<k,rn=rn<>"z"],ds[i,j]=-1],{j,2,fact+1}],{i,rns}]
Do[w++;Do[If[j==w,ds[i,j]=a;
ds[i+1,j]=-a,ds[i,j]=0;
ds[i+1,j]=0],{j,2,fact+1}],{i,rns+1,rns+ax,2}]
Do[ds[i,j]=ds[i,j-fact]^2,{j,fact+2,ax+1},{i,des}]
Do[t1=2;
t2=t1+1;
Do[ds[i,j]=ds[i,t1] ds[i,t2];
t2++;
If[t2>fact+1,t1++;t2=t1+1],{j,ax+2,col}],{i,des}] Do[Do[f=x[[i]];
AppendTo[x,Flatten[f]],{i,rns}],{j,2,rf}]
Do[Do[f=x[[rns+i]];
AppendTo[x,Flatten[f]],{i,ax}],{j,2,ra}]
Print[MatrixForm[x]]
xxi=Inverse[Transpose[x].x];
hat=x.xxi.Transpose[x];
wgt2[dp1_,dp2_]:=1-(1-hat[[dp1,dp1]]) (1-hat[[dp2,dp2]])+hat[[dp1,dp2]]^2;
tm=0;Do[t=wgt2[1,i]/.a->1;
if[t>tm,tm=t;ff=i],{i,1,rns}];
tm=0;Do[t=wgt2[1,i]/.a->1;
if[t>tm,tm=t;fa=i],{i,rns+1,rns+ax}];
tm=0;Do[t=wgt2[1,i]/.a->1;
if[t>tm,tm=t;fc=i],{i,rns+ax+1,des}];
tm=0;Do[t=wgt2[(rns+1),i]/.a->1;
if[t>tm,tm=t;aa=i],{i,rns+2,rns+ax}];
tm=0;Do[t=wgt2[(rns+1),i]/.a->1;
if[t>tm,tm=t;ac=i],{i,rns+ax+1,des}];
tm=0;Do[t=wgt2[(rns+ax+1),i]/.a->1;
if[t>tm,tm=t;cc=i],{i,rns+ax+2,des}];
Plot[{wgt2[1,ff],wgt2[1,fa],wgt2[1,fc],wgt2[rns+1,aa],wgt2[rns+1,ac],wgt2[rns+ax+1,cc]},{a,0,3},PlotLegends-
>LineLegend["Expressions"]]
{desp[1],desp[2],desp[3],desp[4],desp[5],desp[6]}={wgt2[1,ff],wgt2[1,fa],wgt2[1,fc],wgt2[rns+1,aa],wgt2[rns+1
,ac],wgt2[rns+ax+1,cc]};
{ord2,oeff}={0,10};
Do[Do[solv=a/.NSolve[desp[i]-desp[j]==0,a];
Do[If[Re[solv[[k]]]==solv[[k]]&&solv[[k]]>0,tsol=Sort[Table[desp[u]/.a->solv[[k]],{u,6}]];
If[tsol[[5]]==tsol[[6]]&&tsol[[5]]<oeff,ord2=solv[[k]];
oeff=tsol[[6]]]]],{k,Length[solv]}],{j,(i+1),6}],{i,5}]
Print["ORD2 is observed at a = ",ord2," where the outlying effect is ",oeff," which is minimum."]
Print[MatrixForm[x/.a->ord2]]
var=Variance[Flatten[Table[wgt2[i,j],{i,1,(rows-1)},{j,(i+1),rows}]]];
bd2=a/.Last[NMinimize[var,a>0,a]]
Print["BD2 is observed at a = ",bd2," where the variance is ",var/.a->bd2," which is minimum."]
Plot[var,{a,0,3}]
orth=N[(q*rns*rf/4)^0.25];
```

```
Print["Orthogonal Design is observed at a =",orth]
rot=N[(rns*rf)^0.25];
Print["Rotatable Design is observed at a =",rot]
aop=Sum[xxi[[i,i]],{i,col}];
dop=Det[xxi];
dm=Det[Inverse[Transpose[x][[Range[2,fact+1]]].Transpose[Transpose[x][[Range[2,fact+1]]]]]];
ds=Det[Inverse[Transpose[x][[Range[2+fact,2*fact+1]]].Transpose[Transpose[x][[Range[2+fact,2*fact+1]]]]]];
eopt=Max[Eigenvalues[xxi]];
t={{" ","Rotatable","Orthogonal","BD","MinMax"},{"alpha",rot,orth,bd2,ord2},{"A-optimality",aop/.a->rot,aop/.a->orth,aop/.a->bd2,aop/.a->ord2},{"D-optimality",dop/.a->rot,dop/.a->orth,dop/.a->bd2,dop/.a->ord2},{" Dmain",dm/.a->rot,dm/.a->orth,dm/.a->bd2,dm/.a->ord2},{" Dsqre",ds/.a->rot,ds/.a->orth,ds/.a->bd2,ds/.a->ord2},{"E-optimality",eopt/.a->rot,eopt/.a->orth,eopt/.a->bd2,eopt/.a->ord2}};
Print["The Analysis"]
Print[TableForm[t]]
TeXForm[TableForm[t]]
```