Cogent Mathematics ›
Volume 2, 2015 - Issue 1

Research Article

Response surface designs using the generalized variance inflation factors

Diarmuid O'Driscoll & Donald E. Ramirez | Guohua Zou (Reviewing Editor)

Article: 1053728 | Received 22 Dec 2014, Accepted 15 May 2015, Published online: 14 Jun 2015

Cite this article https://doi.org/10.1080/23311835.2015.1053728 Check for updates

Abstract

We study response surface designs using the generalized variance inflation factors for subsets as an extension of the variance inflation factors.

Keywords: variance inflation factors generalized variance inflation factors response surface designs

Public

Respons ... inflation
facto ... ession
mod ... e inflation
factor G ... nding a
linear re ... rms. The

using GVIF, the H310 design can be improved for the standard global optimality criteria of A, D, and E.

1. Introduction

We consider a linear regression $Y=X\beta+\varepsilon$ with X a full rank $n\times p$ matrix and $L(\varepsilon)=N(0,\sigma^2 I_n)$. The variance inflation factor VIF, Belsley (1986), measures the penalty for adding one non-orthogonal additional explanatory variable to a linear regression model, and they can be computed as a ratio of determinants. The extension of VIF to a measure of the penalty for adding a subset of variables to a model is the generalized variance inflation factor GVIF of Fox and Monette (1992), which will be used to study response surface designs, in particular, as the penalty for adding the quadratic terms to the model.

2. Variance inflation factors

For our linear model $Y=X\beta+\varepsilon$, let D_X be the diagonal matrix with entries on the diagonal $D_X[i,i]=(X'X)_{i,i}^{-1/2}$. When the design has been standardized $X\to XD_X$, the VIFs are the diagonal entries of the inverse of $S_X=D_X(X'X)D_X$. That is, the VIFs are the ratios of the actual variances for the explanatory variables to the “ideal” variances had the columns of X been orthogonal. Note that we follow Stewart (1987) and do not necessarily center the explanatory variables.

For our linear model $Y=X\beta+\varepsilon$, view $X=[X_{[p]},x_p]$ with x_p the pth column of X and $X_{[p]}$ the matrix formed by the remaining columns. The variance inflation factor VIF_p measures the effect of adding column x_p to $X_{[p]}$. For notational convenience, we demonst … onal to the previous … w and pth column …

The VIFs … te that the inverse …

$$VIF_p = [S_X^{-1}]_{p,p} = [D_X^{-1}(X'X)^{-1}D_X^{-1}]_{p,p} = (x_p'x_p)^{1/2}\frac{\det(C_{p,p})}{\det(X'X)}(x_p'x_p)^{1/2} = \frac{\det(M_p)}{\det(X'X)} \quad (1)$$

the ratio of the determinant of the idealized moment matrix M_p to the determinant of the moment matrix $X'X$. This definition extends naturally to subsets and is discussed in the next section.

For an alternate view of the how collinearities in the explanatory variables inflate the model variances of the regression coefficients when compared to a fictitious orthogonal reference design, consider the formula for the model variance

$$Var_M(\hat{\beta}_j) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2}\frac{1}{1-R_j^2}$$

where R_j^2 is the square of the multiple correlation from the regression of the jth column of $X=[x_{ij}]$ on the remaining columns as in Liao and Valliant (2012). The first term $\sigma^2/\sum(x_{ij}-\bar{x}_j)^2$ is the model variance for $\hat{\beta}_j$ had the jth explanatory variable been orthogonal to the remaining variables. The second term $1/(1-R_j^2)$ is a standard definition of the $j$thVIF as in Thiel (1971).

---

3. Generalized variance inflation factors

In this section, we introduce the GVIFs as an extension of the classical variance inflation factors VIF from Equation 1. For the linear model $Y=X\beta+\varepsilon$, view $X=[X_1,X_2]$ partitioned with X_1 of dimension $n\times r$ usually consisting of the lower order terms and X_2 of dimension $n\times s$ usually consisting of the higher order terms. The idealized moment matrix for the (r,s) partitioning of X is

Followin… that is for $X_2|X_1$…

(2)

as in Equation 10 of Fox and Monette (1992), who compared the sizes of the joint confidence regions for β for partitioned designs and noted when X=[X[p],xp] that GVIF[xp|X[p]]=VIFp. Equation 2 is in the spirit of the efficiency comparisons in linear inferences introduced in Theorems 4 and 5 of Jensen and Ramirez (1993). A similar measure of collinearity is mentioned in Note 2 in Wichers (1975), Theorem 1 of Berk (1977), and Garcia, Garcia, and Soto (2011). For the simple linear regression model with p=2, Equation 2 gives VIF=11-ρ2 with ρ the correlation coefficient as required. Fox and Monette (1992) suggested that X1 contains the variables which are of “simultaneous interest,” while X2 contains additional variables selected by the investigator. We will set X1 for the constant and main effects and set X2 the (optional) quadratic terms with values from X1.

Willan and Watts (1978) measured the effect of collinearity using the ratio of the volume of the actual joint confidence region for β^ to the volume of the joint confidence region in the fictitious orthogonal reference design. Their ratio is in the spirit of GVIF as det(X′X) is inversely proportional to the square of the volume of the joint confidence region for β^. They also introduced a measure of relative predictability and they note: “The existence of near linear relations in the independent variables of the actual data reduces the overall predictive efficiency by this factor.” For a simple case study, consider the simple linear regression model with n=4, x1=[-2,-1,1,2]′, and y=[4,1,1,4]′. The 95% prediction interval for x1=0 is 2.5±10.20. If the model also includes x2=[-2.001,-1.001,1.001,2.001]′, then the 95% prediction interval for (x1,x2)=(0,0) is 2.5±46.02 demonstrating the loss of predictive efficiency due to the collinearity introduced by x2.

For the (r,s) partition of X=[X1,X2] with X1 of dimension n×r and X2 of dimension n×s, set

and den

R 1-1/2Is×s;
(3)

with det

$$\det(R)=\frac{\det(X'X)}{\det(X_1'X_1)\det(X_2'X_2)}=\frac{1}{GVIF(X_2|X_1)};$$

equivalently,

$$\det(R)=\det(I_{r\times r}-B_{r\times s}B_{s\times r}')=\det(I_{s\times s}-B_{s\times r}'B_{r\times s})$$

where $B_{r\times s}=X_1'X_1^{-1/2}(X_1'X_2)X_2'X_2^{-1/2}$.

In the case $\{r=p-1,\ s=1\}$, $X_2=x_p$ is a $n\times 1$ vector and the partitioned design $X=[X_1,x_p]$ has $\det(R)=1-[x_p'X_1X_1'X_1^{-1}X_1'x_p]/x_p'x_p$. From standard facts for the inverse of a partitioned matrix, for example, Myers (1990, p. 459), $VIF_p=[R^{-1}]_{p,p}=[D_{(p-1,1)}^{-1}(X'X)^{-1}D_{(p-1,1)}^{-1}]_{p,p}$ can be computed directly as

$$x_p'x_p^{1/2}(X'X)_{p,p}^{-1}x_p'x_p^{1/2}=\frac{x_p'x_p}{x_p'x_p-x_p'X_1X_1'X_1^{-1}X_1'x_p}=\frac{1}{1-[x_p'X_1X_1'X_1^{-1}X_1'x_p]/x_p'x_p}=\frac{1}{\det(R)}=GVIF(X_2|X_1).$$

Table 1. CCD with parameter a, canonical index γX_2, and GVIF

Display Table

We study the eigenvalue structure of $M(r,s)$ in Appendix 1. Let $\{\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_{\min(r,s)}\geq 0\}$ be the non-negative singular values of $X_1'X_1^{-1/2}(X_1'X_2)X_2'X_2^{-1/2}$. It is shown in Appendix 1 that an alternative formulation for GVIF is

$$GVIF(X_2|X_1)=\prod_{i=1}^{\min(r,s)}(1-\lambda_i^2)^{-1}. \quad (4)$$

4. Qua

For the r … as the

ident … $1'X_2X_2'X_2^{-}$

1/2. Fo … we have

From Equation 4, $GVIF(x_2|1,x)=(1-\lambda^2)^{-1}$ where $\lambda=\rho_1^2+\rho_2^2$ is the unique positive singular value of $[\rho_1,\rho_2]'$. Denote

$$\gamma_{X_2}=\rho_1^2+\rho_2^2$$

as the canonical index with $GVIF(x_2|1,x)=\frac{1}{1-\gamma_{X_2}}=\frac{1}{\det(R)}$. Surprisingly, many higher order designs also have the off-diagonal entry of the canonical moment matrix with a unique positive singular value with $GVIF(X_2|X_1)=\frac{1}{1-\gamma_{X_2}}$ with the collinearity between the lower order terms and the upper order terms as a function of the canonical index γ_{X_2}.

---

5. Central composite and factorial designs for quadratic models ($p=6$)

In this section, we compare the central composite design (CCD) X of Box and Wilson (1951) and the factorial design Z. The design points are shown in Table A1 of Appendix 2. Both designs are 9×6 and use the quadratic response model

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{11}x_1^2+\beta_{22}x_2^2+\beta_{12}x_1x_2+\varepsilon.$$

The CCD traditionally uses the value $a=2$ in four entries, while the factorial design uses the value $a=1$. To study the difference in the designs with these different values, we computed the GVIF to compare the "orthogonality" between the lower order terms X_1 of dimension 9×3 and the higher order quadratic terms X_2 of dimension 9×3. The off-diagonal $B_{3\times3}$ entry of R from Equation 3 in Section 3 has the form

$$B_{3\times3}=\begin{bmatrix}\rho_1 & \rho_2 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

with $\rho_1=$ … $1)=\frac{1}{1-\gamma_{X_2}}$
as in the … , if $a=1$,
then $\rho_1=$ … l choice of
$a=2$ … ing the
greate… O'Driscoll
and Ram…

6. Larger designs (p=10)

We consider the quadratic response surface designs for

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_{11}x_1^2+\beta_{22}x_2^2+\beta_{33}x_3^2+\beta_{12}x_1x_2+\beta_{13}x_1x_3+\beta_{23}x_2x_3+\varepsilon \quad (5)$$

with n responses and with X partitioned into $X_1|X_2$ with X_1 the four lower order terms (r=4) and X_2 the six quadratic terms (s=6). Four popular designs are given in Appendix 2. They are the hybrid designs (H310 and H311B) of Roquemore (1976) Tables A2 and A3, the Box and Behnken (1960) (BBD) design Table A4, and the CCD of Box and Wilson (1951) Table A5.

For each design, we compute the 10×10 canonical moment matrix. It is striking that, for all these designs, the off-diagonal 4×6 array in R has only one non-zero singular value with its square the canonical index γ_{X2}. It follows that $GVIF(X_2|X_1)=1\frac{1}{1-\gamma_{X2}}$.

Table 2. Hybrid designs H310, H311B, Box and Behnken BBD,and CCD

Display Table

Table 3. Singular values for off-diagonal array of R for BDD and SCD with GVIF

Display Table

Table 2 r … o the GVIF with the … rms.

7. Mor…

Let X be … Table A6

for the quadratic response surface model ($r=4$ and $s=6$) as in Equation (5). Let $\alpha=\{\alpha_1\geq\ldots\geq\alpha_r\geq 0\}$ and $\beta=\{\beta_1\geq\ldots\geq\beta_r\geq 0\}$ be the non-negative singular values of the off-diagonal array for R_X and R_Z, respectively. As $\alpha_i\leq\beta_i (1\leq i\leq r)$ in Table 3, it follows that $GVIF(X_2|X_1)\leq GVIF(Z_2|Z_1)$ showing less collinearity between the lower and higher order terms for the BDD design.

---

8. An improved H310 design

When the diagonal matrix $\Lambda_{r\times s}$ in Equation 6 in Appendix 1 has only one non-zero entry, we have denoted the square of this value the canonical index. We extend this definition to the case when $X_1'X_1^{-1/2}(X_1'X_2)X_2'X_2^{-1/2}$ has multiple positive singular values. The Frobenious norm for a rectangular matrix $A_{r\times s}$ is defined by $||A||_F^2=\sum_{i=1}^{r}\sum_{j=1}^{s}a_{ij}^2=\mathrm{trace}(A'A)$. For a design matrix X, we extend the definition of the canonical index with $\gamma_X^2=||\Lambda_{r\times s}||_F^2$. Alternatively, $\gamma_X^2=\mathrm{trace}(X_2'X_2^{-1}(X_2'X_1)X_1'X_1^{-1}(X_1'X_2))$ as in Equation 7.

We examine, in detail, the H310 design matrix $X_{11\times 10}$, Table A2 in Appendix 2, with our attention to the value of -0.1360 in row 2 for x_3. In succession, we will replace the values {1.1736,0.6386,-0.9273,1.0000,1.2906,-0.1360} by a free parameter and use γ_X^2 to determine an optimal value. For example, replacing the four entries which are 1.1736 with c_1, we calculate the minimum value for $\gamma_X^2=0.8199$ with $c_1=1.1768$ denoted c_{min} in Table 4. These values are within the four digit accuracy of the data. We performed a similar calculation with c_2 using the four entries which are 0.6386; with c_3 with the four entries which are -0.9273; with c_4 with the eight entries which are 1; and with c_5 with the single entry 1.2906. The original design has $\gamma_X^2=0.8199$. The entries in the H310 design are given to four significant digits. With this precision, the original design is nearly optimal with respect to the canonical index γ_X^2 for the first five entries in Table ... with c_{min}=-0...

Ta...

Display...

Denote the “improved” H310 design as the H310 design with the value of c6=-0.01264. The “improved” H310 also has a unique positive singular value for the off-diagonal of R with its square the canonical index γX2. All of the standard design criteria favor the “improved” H310 design over the H310 design, which was originally constructed based on the rotatability criterion to maintain equal variances for predicted responses for points that have the same distance from the design center. As usual A(X)=tr((X′X)-1), D(X)=det((X′X)-1), and E(X)=max{eigenvalues of (X′X)-1}. The small relative changes Δ in the design criteria are shown in Table 5 in Column 4.

Table 5. Design criteria for the “improved” H310 with Δ the relative change

Display Table

The abnormality of the second row in H310 has been noted in Jensen (1998) who showed that the design is least sensitive to the second row of X, the row containing the value c6=-0.1360.

9. Conclusions

The VIF measure the penalty for adding a non-orthogonal variable to a linear regression. The VIF can be computed as a ratio of determinant as in Equation 1. A similar ratio criterion was studied by Fox and Monette (1992) to measure the effect of adding a subset of new variables to a design and they dubbed it the generalized variance inflation factor GVIF, Equation 2. We have noted the relationship between GVIF and the ... atrix and have use... esign of Roquem... n “improv... g the stan...

Addit...

Funding

The authors received no direct funding for this research.

Notes on contributors

Diarmuid O’Driscoll

Diarmuid O’Driscoll is the head of the Mathematics and Computer Studies Department at Mary Immaculate College, Limerick. He was awarded a Travelling Studentship for his MSc at University College Cork in 1977. He has taught at University College Cork, Cork Institute of Technology, University of Virginia, and Frostburg State University. His research interests are in mathematical education, errors in variables regression, and design criteria. In 2014, he was awarded a Teaching Heroes Award by the National Forum for the Enhancement of Teaching and Learning (Ireland).

Donald E. Ramirez

Donald E Ramirez is a full professor in the Department of Mathematics at the University of Virginia in Charlottesville, Virginia. He received his PhD in Mathematics from Tulane University in New Orleans, Louisiana. His research is in harmonic analysis and mathematical statistics. His current research interests are in statistical outliers and ridge regression.

Refer

1. Belsle ... ssing
condit ... 3).
Ca
Goog

2. Berk, ... nal of the

Web of Science ® | Google Scholar

3. Box, G. E. P., & Behnken, D. W. (1960). Some new three-level designs for the study of quantitative variables. Technometrics, 2, 455–475.
Google Scholar

4. Box, M. J., & Draper, N. R. (1974). On minimum-point second order design. Technometrics, 16, 613–616.
Web of Science ® | Google Scholar

5. Box, G. E. P., & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, Series B, 13, 1–45.
Web of Science ® | Google Scholar

6. Eaton, M. L. (1983). Multivariate statistics. New York, NY: Wiley.
Google Scholar

7. Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. Journal of the American Statistical Association, 87, 178–183.
Web of Science ® | Google Scholar

8. Garcia, C. B., Garcia, J., & Soto, J. (2011). The raise method: An alternative procedure to estimate the parameters in presence of collinearity. Quality and Quantity, 45, 403–423.
Web of Science ® | Google Scholar

9. Hartle surfaces.
Biome

10. Je rder designs.
Biome

11. Jensen, D. R., & Ramirez, D. E. (1993). Efficiency comparisons in linear inference. Journal of Statistical Planning and Inference, 37, 51–68.
Web of Science ® | Google Scholar

12. Liao, D., & Valliant, R. (2012). Variance inflation in the analysis of complex survey data. Survey Methodology, 38, 53–62.
Web of Science ® | Google Scholar

13. Myers, R. (1990). Classical and modern regression with applications (2nd ed.). Boston, MA: PWS-Kent.
Google Scholar

14. O'Driscoll, D., & Ramirez, D. E. (in press). Revisiting some design criteria (under review).
Google Scholar

15. Roquemorer, K. G. (1976). Hybrid designs for quadratic response surfaces. Technometrics, 18, 419–423.
Web of Science ® | Google Scholar

16. Stewart, G. W. (1987). Collinearity and least squares regression. Statistical Science, 2, 68–84.
Google Scholar

17. Thiel, H. (1971). Principles of econometrics. New York, NY: Wiley.
Google Scholar

18. Wiche ... of
Econo

19. Wil ... s.
Techn

Appendix 1

We study the eigenvalue structure of $M(r,s)$. Let $\{\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\min(r,s)} \geq 0\}$ be the non-negative singular values of $X_1'X_1^{-1/2}(X_1'X_2)X_2'X_2^{-1/2}$.

As with the canonical correlation coefficients Eaton (1983), write the off-diagonal rectangular array $B_{r\times s}$ of R as $P\Lambda Q'$ with P and Q orthogonal matrices and $\Lambda_{r\times s}$ the rectangular diagonal matrix with the non-negative singular values down the diagonal. Set

$$L = \begin{bmatrix} P_{r\times r} & 0_{r\times s} \\ 0_{s\times r} & Q_{s\times s} \end{bmatrix}.$$

For notational convenience, we assume $r \leq s$. The matrix L is orthogonal and transforms $R \to L'RL$ into diagonal matrices:

$$\begin{bmatrix} I_r & \Lambda_{r\times s} \\ \Lambda_{s\times r}' & I_s \end{bmatrix} = \begin{bmatrix} I_r & [SV_{r\times r}|0_{r\times(s-r)}] \\ [SV_{r\times r}|0_{r\times(s-r)}]' & I_s \end{bmatrix} \quad \text{(A1)}$$

with $\Lambda_{r\times s} = [SV_{r\times r}|0_{r\times(s-r)}]$ where $SV_{r\times r}$ is the diagonal matrix of the non-negative singular values. Since L is orthogonal, this transformation has not changed the eigenvalues. To compute the determinant of R, convert the matrix in Equation 6 into an upper diagonal matrix by Gauss Elimination on $\Lambda_{s\times r}'$. This changes r of the 1′s on the diagonal in rows $r+1$ to $r+r$ into $1-\lambda_i^2$, and thus $\det(R) = \prod_{i=1}^{\min(r,s)}(1-\lambda_i^2)$ with

$$GVIF(X_2|X_1) = \prod_{i=1}^{\min(r,s)} \frac{1}{1-\lambda_i^2}.$$

The singular values of $R_{12} = X_1'X_1^{-1/2}(X_1'X_2)X_2'X_2^{-1/2}$ are the non-negative square roots of the eigenvalues of $\Lambda'\Lambda$ denoted by

eigva … 1(X2′X1)).
(A2)

If the tra … ote that

IrΛ … ×r′)-1(Is-

Appendix 2

Table A1. The lower order matrix for the CCD with center run with a=2, n=9 and the lower order matrix for the factorial design with center run n=9

Display Table

Table A2. The lower order matrix for the hybrid (H310) design of Roquemore (1976) with center run, n=11

Display Table

Table A3. The lower order matrix for the hybrid (H311B) design of Roquemore (1976) with center run, n=11

Display Table

Table A4. The lower order matrix for the Box and Behnken (1960) design (BBD) with center run, n=13

Display Table

Table

C

1)

Display

Typesetting math: 100%

Table A6. The lower order matrix for the Box and Draper (1974) minimal design (BDD) with center run, n=11

Display Table

Table A7. The Lower order matrix for the small composite design of Hartley (1959) (SCD) for $\alpha=1.732$ with center run, n=11

Display Table