Confidence regions for the stationary point of a quadratic response surface based on the asymptotic distribution of its MLE

Valeria Sambucini
Dipartimento di Statistica, Probabilità e Statistiche Applicate,
Sapienza Università di Roma
valeria.sambucini@uniroma1.it

Abstract

Response-surface methodology aims at finding the combination of factors levels which optimizes a response variable. A second order polynomial model is typically employed to make inference on the stationary point of the true response function. A suitable reparametrization of the polynomial model, where the coordinates of the stationary point appear as the parameter of interest, is used to derive unconstrained confidence regions for the stationary point. These regions are based on the asymptotic normal approximation to the sampling distribution of the maximum likelihood estimator of the stationary point. A simulation study is performed to evaluate the coverage probabilities of the proposed confidence regions. Some comparisons with the standard confidence regions due to Box and Hunter are also showed.

KEY WORDS: asymptotic distribution, response surface methodology, second-order polynomial model, unconstrained optimization.

1 Introduction

In response-surface methodology (RSM) a second order polynomial model is typically used to find the optimal setting for k predictor variables that maximizes or minimizes a response variable of interest, Y. RSM is actually a combination of statistical and mathematical techniques originally proposed by Box and Wilson (1951) in the field of chemical industry. Subsequently, this set of procedures has been further developed and refined with applications in many scientific areas. For a comprehensive and detailed description about this methodology see, for instance, the textbooks by Khuri and Cornell (1996), Box and Draper (2007) and Myers *et al.* (2009).

In actual applications, it is common practice to code the original input variables to get dimensionless factors, $X_1, ..., X_k$, having zero mean. The standard quadratic model can be written in matrix notation as

$$y = \beta_0 + \boldsymbol{x}^T\boldsymbol{\beta} + \boldsymbol{x}^T\mathbf{B}\boldsymbol{x} + \varepsilon, \tag{1}$$

where $\boldsymbol{x}$ is a fixed combination of the levels of the k input variables, β_0 is the intercept term,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \cdots & \frac{1}{2}\beta_{1k} \\ \frac{1}{2}\beta_{12} & \beta_{22} & \cdots & \frac{1}{2}\beta_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\beta_{1k} & \frac{1}{2}\beta_{2k} & \cdots & \beta_{kk} \end{bmatrix}$$

and ε is the random error having a normal distribution with zero mean and unknown variance σ^2. Typically the adequacy of the quadratic model is investigated by using an F test for the significance of the contribution of the second order terms (including interactions) to a model that already contains the linear terms. If model (1) is found to be adequate, the interest is focused on the stationary point of the true response function, denoted by $\boldsymbol{\xi} = (\xi_1, ..., \xi_k)$. When $n > k$ observations of the response variable are available, an estimate of $\boldsymbol{\xi}$ is obtained as $\hat{\boldsymbol{\xi}} = -\frac{1}{2}\hat{\mathbf{B}}^{-1}\hat{\boldsymbol{\beta}}$, where $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{B}}$ are the maximum likelihood estimates (MLEs) of $\boldsymbol{\beta}$ and $\mathbf{B}$ respectively. The nature of $\hat{\boldsymbol{\xi}}$ is determined by the eigenvalues of matrix $\hat{\mathbf{B}}$: if they are all positive, the fitted quadratic surface has a minimum at the stationary point $\hat{\boldsymbol{\xi}}$ and, if they are all negative, $\hat{\boldsymbol{\xi}}$ is a maximum. When instead the eigenvalues have mixed signs, the estimated stationary point results to be a saddle point. Many methods for constructing confidence intervals on the eigenvalues of matrix $\mathbf{B}$ have been discussed in the literature (see, for instance, Carter *et al.*, 1986; Carter *et al.*, 1990; Peterson, 1993; Bisgaard and Ankenman, 1996): if any of these intervals cover the value of zero, the true surface may have some type of rising behavior. The MLE $\hat{\boldsymbol{\xi}}$ is simply a point estimate of the true stationary point and the construction of confidence regions on the location of the true parameter is needed to assess how accurate is our estimate and how much flexibility is available in choosing optimum conditions. The standard unconstrained confidence regions for the stationary point are due to Box and Hunter (1954). Denoting by $\hat{\boldsymbol{d}}(\mathbf{x})$ the first derivative vector $\partial\hat{y}/\partial\boldsymbol{x}$, the $100 * (1-\alpha)\%$ Box and Hunter confidence region consists of all the points $\boldsymbol{x}$ such that

$$\hat{\boldsymbol{d}}(\boldsymbol{x})^T \, \hat{\mathbf{V}}_{\boldsymbol{d}}^{-1} \, \hat{\boldsymbol{d}}(\boldsymbol{x}) \leq k \, F(\alpha; k, n-p), \tag{2}$$

where $\hat{\mathbf{V}}_{\hat{\boldsymbol{d}}}$ is the estimate of the variance-covariance matrix of $\hat{\boldsymbol{d}}(\boldsymbol{x})$, n is the sample size, $p = 1 + 2k + \frac{k(k-1)}{2}$ is the number of regression coefficients and $F(\alpha; k, n-p)$ is the $1-\alpha$ quantile of the F distribution with k and $n-p$ degrees of freedom. Hereafter we refer to this confidence region as the "BH region". For more recent developments of procedures to compute constrained confidence regions for the optimal point, readers are referred to Stablein *et al.* (1983) and Peterson *et al.* (2002) among others.

Assuming that a unique stationary point $\boldsymbol{\xi}$ for the true quadratic surface exists, model (1) can be rewritten using a different parametrization, that is

$$y = \alpha_0 + (\boldsymbol{x} - \boldsymbol{\xi})^T \mathbf{A} (\boldsymbol{x} - \boldsymbol{\xi}) + \varepsilon, \tag{3}$$

where $\boldsymbol{x}$ and ε have the same meaning as in model (1), α_0 is the response value at $\boldsymbol{\xi}$ and

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \frac{1}{2}\alpha_{12} & \cdots & \frac{1}{2}\alpha_{1k} \\ \frac{1}{2}\alpha_{12} & \alpha_{22} & \cdots & \frac{1}{2}\alpha_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\alpha_{1k} & \frac{1}{2}\alpha_{2k} & \cdots & \alpha_{kk} \end{bmatrix}.$$

The one-to-one correspondence between the parameters of the two models allows to derive the MLEs of model (3) from the MLEs of model (1), using the relations $\hat{\alpha}_0 = \hat{\beta}_0 - \frac{1}{4}\hat{\boldsymbol{\beta}}^T\hat{\mathbf{B}}^{-1}\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\xi}} = -\frac{1}{2}\hat{\mathbf{B}}^{-1}\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{A}} = \hat{\mathbf{B}}$. Of course, if we employ model (3), the nature of $\hat{\boldsymbol{\xi}}$ is determined by studying the signs of the eigenvalues of $\hat{\mathbf{A}}$. The usefulness of this reparametrization lies in the fact that $\boldsymbol{\xi}$ appears explicitly in the model as the parameter of interest, while the other parameters, $\boldsymbol{\alpha} = (\alpha_0, \alpha_{11}, ..., \alpha_{kk}, \alpha_{12}, ..., \alpha_{k-1,k})^T$ and σ^2, can be treated as nuisance parameters. Sambucini (2007) and Sambucini and Piccinato (2008) introduce and employ this model to provide likelihood and Bayesian procedures to make inference on the location of the true optimum point, using both informative and non-informative prior distributions.

In this paper we construct unconstrained confidence regions for the stationary point by exploiting the asymptotic normal approximation to the sampling distribution of $\hat{\boldsymbol{\xi}}$. This distribution can be derived from model (3), as shown in Section 2. The results of a simulation study for $k = 2$ are also included to assess the accuracy of such normal approximation to $\hat{\boldsymbol{\xi}}$. Section 3 illustrates how to obtain approximated confidence regions for $\boldsymbol{\xi}$ based on the asymptotic distribution given in the previous section. In Sections 4 and 5 some real data examples with two and three input variables are used to compute and compare the proposed confidence regions with those proposed by Box and Hunter. In Section 6 we present simulation results about coverage rates and sizes of the confidence regions. Finally, a discussion is given in Section 7, with some concluding remarks.

2 Asymptotic distribution of the MLE of the stationary point

Let us assume that $n > k$ response values, say $y_1, y_2, ..., y_n$, are observed on various combinations of the input variables. Then, we may write model (3) in terms of the observations as

$$y_i = \alpha_0 + \sum_{j=1}^{k} \alpha_{jj}(x_{ij} - \xi_j)^2 + \sum_{j=1}^{k-1}\sum_{h=j+1}^{k} \alpha_{jh}(x_{ij} - \xi_j)(x_{ih} - \xi_h) + \varepsilon_i, \quad (i = 1, ..., n), \tag{4}$$

where x_{ij} denote the i-th level of regressor X_j and $\varepsilon_i \sim N(0, \sigma^2)$. It is also assumed that the errors ε_i are independent for each observation. We will denote by $\boldsymbol{\psi} = (\boldsymbol{\xi}, \boldsymbol{\alpha}, \sigma^2)$ the whole parameter vector, that has dimension p+1, where $p = 1 + 2k + \frac{k(k-1)}{2}$. Indeed, the parameter of interest $\boldsymbol{\xi}$ is a vector of k elements and the nuisance parameter $\boldsymbol{\alpha} = (\alpha_0, \alpha_{11}, ..., \alpha_{kk}, \alpha_{12}, ..., \alpha_{k-1,k})^T$ is a vector of dimension p', with $p' = 1 + k + \frac{k(k-1)}{2}$.

Let us denote by $\hat{\boldsymbol{\psi}}$ the maximum likelihood estimator of $\boldsymbol{\psi}$. The sampling distribution of $\hat{\boldsymbol{\psi}}$ is asymptotically approximated by a multivariate normal distribution with mean $\boldsymbol{\psi}$ and variance

and covariance matrix $\mathbf{V}_{\boldsymbol{\psi}}(\boldsymbol{\psi}) = \mathbf{H}_{\boldsymbol{\psi}}^{-1}(\boldsymbol{\psi})$, where $\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$ is the Fisher information matrix of the parameter vector $\boldsymbol{\psi}$. In Appendix A it is shown that

$$\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi}) = \mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\xi}, \boldsymbol{\alpha}, \sigma^2) = \frac{1}{\sigma^2} \left[\begin{array}{c|c|c} \mathbf{M}\mathbf{M}^T & \mathbf{M}\mathbf{X}_{\boldsymbol{\xi}} & \mathbf{0}_k \\ \hline (\mathbf{M}\mathbf{X}_{\boldsymbol{\xi}})^T & \mathbf{X}_{\boldsymbol{\xi}}^T\mathbf{X}_{\boldsymbol{\xi}} & \mathbf{0}_{p'} \\ \hline \mathbf{0}_k^T & \mathbf{0}_{p'}^T & \frac{n}{2\sigma^2} \end{array}\right], \tag{5}$$

where $\mathbf{0}_k$ is the zero vector of dimension k,

$$\mathbf{X}_{\boldsymbol{\xi}} = \left[\begin{array}{ccccccc} 1 & (x_{11}-\xi_1)^2 & \cdots & (x_{1k}-\xi_k)^2 & (x_{11}-\xi_1)(x_{12}-\xi_2) & \cdots & (x_{1,k-1}-\xi_{k-1})(x_{1k}-\xi_k) \\ 1 & (x_{21}-\xi_1)^2 & \cdots & (x_{2k}-\xi_k)^2 & (x_{21}-\xi_1)(x_{22}-\xi_2) & \cdots & (x_{2,k-1}-\xi_{k-1})(x_{2k}-\xi_k) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 1 & (x_{n1}-\xi_1)^2 & \cdots & (x_{nk}-\xi_k)^2 & (x_{n1}-\xi_1)(x_{n2}-\xi_2) & \cdots & (x_{n,k-1}-\xi_{k-1})(x_{nk}-\xi_k) \end{array}\right]$$

and $\mathbf{M}$ is the $k \times n$ matrix, whose generic column is the $k-$dimensional vector $2\mathbf{A}(\mathbf{x}^i - \boldsymbol{\xi})$, where $\boldsymbol{x}^i = (x_{i1}, x_{i2}, ..., x_{ik})^T$ is the combination of factors levels determined by the experimenter to take the experimental run y_i, $\forall\, i = 1, ..., n$.

We are in particular interested in the $k \times k$ upper left sub-matrix of $\mathbf{V}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$, which is given by (see Graybill, 1983)

$$\mathbf{V}_{\boldsymbol{\xi}}(\boldsymbol{\psi}) = \sigma^2 \left[\mathbf{M}\mathbf{M}^T - \mathbf{M}\mathbf{X}_{\boldsymbol{\xi}}(\mathbf{X}_{\boldsymbol{\xi}}^T\mathbf{X}_{\boldsymbol{\xi}})^{-1}\mathbf{X}_{\boldsymbol{\xi}}^T\mathbf{M}^T\right]^{-1}$$

and represents the asymptotic approximation of the variance covariance matrix of the parameter of interest $\boldsymbol{\xi}$. Hence, from the properties of the multivariate normal distribution, we have that

$$\hat{\boldsymbol{\xi}} \approx \mathbf{MN}_k(\boldsymbol{\xi}, \mathbf{V}_{\boldsymbol{\xi}}(\boldsymbol{\psi})), \tag{6}$$

where $\mathbf{MN}_k$ denotes the k-variate normal distribution.

2.1 A simulation study to evaluate the accuracy of the normal approximation

We perform a simulation study to examine the accuracy of normal approximation (6) when $k = 2$, by checking if the empirical sampling distribution of $\boldsymbol{\xi}$, obtained through simulation, is close to it.

Let us assume that we know the true parameters of model (4) and consider a rotatable central composite design (CCD) with four centre points. Therefore, at each step of the simulation study, we need to draw $n = 12$ experimental runs. More specifically, we obtain 10000 samples of $n = 12$ response values by perturbing each true response datum with random errors generated from a normal distribution with mean 0 and variance $\sigma^2 = 1$. We use each of these samples to fit a quadratic surface model and to compute the MLE of the stationary point, obtaining a sample of 10000 estimates. Then, given a fixed probability level of interest, say p, let us denote by E_p the ellipse of constant probability density based on the bivariate normal approximation and such that the probability of being inside the ellipse equals p. The accuracy of the normal approximation is evaluated by computing the relative error

$$RE = \frac{|p - \hat{p}|}{p}, \tag{7}$$

where $\hat{p}$ is the proportion of simulated estimates of the stationary point which are inside E_p, for different values of p.

The results of the simulations are showed in Table 1. The stationary point of the true response functions involved in the study is located at the design centre ($\boldsymbol{\xi} = (0, 0)$), inside the experimental region ($\boldsymbol{\xi} = (0.5, 0.5)$) and on the boundary of the experimental region ($\boldsymbol{\xi} = (1, 1)$). The surfaces with $\boldsymbol{\alpha} = (100, -8, -9, 6)$ and $\boldsymbol{\alpha} = (100, -2, -2, 0)$ have a maximum at $\boldsymbol{\xi}$, but the the first ones have a more elongated shape. When instead $\boldsymbol{\alpha} = (100, -3, 2, 4)$ the true stationary point turn out to be a saddle point. Let us notice that the relative errors in Table 1 are reasonably low. It is important to take into account that the sample size used in the study is quite small ($n = 12$). For instance, when $\boldsymbol{\xi} = (0, 0)$ and $\boldsymbol{\alpha} = (100, -2, -2, 0)$, if we replicate the generation of experimental runs at each factors combination (so that $n = 24$), the estimated relative errors are 0.0233, 0.0270, 0.0280, 0.0280 and 0.0214 for p equals to 0.30, 0.50, 0.70, 0.90 and 0.95, respectively.

Table 1: Relative errors (7), for different values of $\boldsymbol{\xi}$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_{11}, \alpha_{22}, \alpha_{12})$ and p.

$\boldsymbol{\alpha} = (100, -8, -9, 6)$, Eigenvalues($\mathbf{A}$) = $(-11.54,\ -5.46)$

$\boldsymbol{\xi}$	$p = 0.30$	$p = 0.50$	$p = 0.70$	$p = 0.90$	$p = 0.95$
$(0, 0)$	0.0070	0.0134	0.0073	0.0049	0.0046
$(0.5, 0.5)$	0.0060	0.0024	0.0031	0.0010	0.0040
$(1, 1)$	0.0013	0.0044	0.0033	0.0007	0.0076

$\boldsymbol{\alpha} = (100, -2, -2, 0)$, Eigenvalues($\mathbf{A}$) = $(-2,\ -2)$

$\boldsymbol{\xi}$	$p = 0.30$	$p = 0.50$	$p = 0.70$	$p = 0.90$	$p = 0.95$
$(0, 0)$	0.0270	0.0288	0.0427	0.0483	0.0436
$(0.5, 0.5)$	0.0097	0.0046	0.0219	0.0562	0.0598
$(1, 1)$	0.0120	0.0042	0.0219	0.0612	0.0669

$\boldsymbol{\alpha} = (100, -3, 2, 4)$, Eigenvalues($\mathbf{A}$) = $(2.70,\ -3.70)$

$\boldsymbol{\xi}$	$p = 0.30$	$p = 0.50$	$p = 0.70$	$p = 0.90$	$p = 0.95$
$(0, 0)$	0.0037	0.0044	0.0029	0.0113	0.0112
$(0.5, 0.5)$	0.0060	0.0026	0.0004	0.0147	0.0168
$(1, 1)$	0.0077	0.0074	0.0047	0.0182	0.0189

3 Approximated confidence regions for the stationary point

The asymptotic distribution (6) can be used to derive approximated confidence regions for the stationary point. From the properties of the multivariate normal distribution, such a region with $100(1 - \alpha)\%$ confidence level is given by of all the values $\boldsymbol{\xi}$ that satisfy

$$(\boldsymbol{\xi} - \hat{\boldsymbol{\xi}})^T \left[\mathbf{V}_{\boldsymbol{\xi}}(\hat{\boldsymbol{\psi}})\right]^{-1} (\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}) \leq k F(\alpha; k, n - p), \tag{8}$$

where $\mathbf{V}_{\boldsymbol{\xi}}(\hat{\boldsymbol{\psi}})$ is the plug-in estimator of the asymptotic variance and covariance matrix of $\hat{\boldsymbol{\xi}}$ and $F(\alpha; k, n - p)$ is the $1 - \alpha$ quantile of the F distribution with k and $n - p$ degrees of freedom.

Hereafter we refer to this asymptotic confidence region as the "AC region". Of course the case $k = 2$ is the simpler one to get a plot of the AC region: it is possible to display the edge of the region by drawing all points such that the inequality in (8) is an equality. When $k = 3$ it is still possible to plot the edge of the region in a $3-$dimensional space. However, in general for $k > 2$, it is convenient to show the point-by-point projections of the $k-$dimensional region into $2-$dimensional planes. The same graphical procedure can be used to represent the BH confidence regions.

In Sections 4 and 5 we provide some applications with data taken from the literature and conducted using the statistical software R (R Development Core Team, 2009). In particular, in Appendix B we illustrate a very user friendly software tool, which creates a graphical user interface (GUI) to easily plot the AC regions when $k = 2$.

4 Two factor experiments

In this Section we use the data of three examples taken from the statistical literature to illustrate the behavior of the AC regions for the stationary point, when two input variables are considered. The BH regions are also displayed for a comparison. In all the three examples the experimental design is a rotatable CCD with 4 centre points, which is the most widely used for fitting second order response surfaces. Moreover the focus of the experiment is on the maximum point of the surface. The data are provided in Table 2.

Table 2: Data of two factor experiments in Section 4

Data of Example 1		
x_1	x_2	**y**
-1	-1	43.33
1	-1	49.97
-1	1	45.94
1	1	43.38
0	0	49.73
0	0	50.07
$-\sqrt{2}$	0	42.79
$\sqrt{2}$	0	47.17
0	$-\sqrt{2}$	47.07
0	$\sqrt{2}$	45.06
0	0	49.43
0	0	50.22

Data of Example 2		
x_1	x_2	**y**
-1	-1	43
1	-1	78
-1	1	69
1	1	73
0	0	76
0	0	79
$-\sqrt{2}$	0	48
$\sqrt{2}$	0	78
0	$-\sqrt{2}$	65
0	$\sqrt{2}$	74
0	0	83
0	0	81

Data of Example 3		
x_1	x_2	**y**
-1	-1	33
1	-1	27
-1	1	28
1	1	52
0	0	45
0	0	41
$-\sqrt{2}$	0	26
$\sqrt{2}$	0	33
0	$-\sqrt{2}$	25
0	$\sqrt{2}$	37
0	0	42
0	0	46

4.1 Example 1

In a recent paper, Li *et al* (2007) used the desirability function approach to find the combinations of glucose concentration (coded factor X_1) and ammonium sulfate (coded factor X_2), which determine the best compromise between maximizing glutamine production and minimizing

glutamate accumulation. In this Section, we consider only one of the two response variables, i.e. the glutamine production (variable Y), to provide an application of the proposed confidence regions based on two factors.

Fitting the second order polynomial model, the F statistic to test the contribution of the quadratic terms to the model is highly significant (p-value< 0.0001) and the multiple R^2 is 0.987. The estimated stationary point is $\hat{\boldsymbol{\xi}} = (0.532, -0.557)$ and turns out to be a maximum point since the eigenvalues of matrix $\hat{\mathbf{B}}$ are $(\hat{\lambda}_1, \hat{\lambda}_2) = (-0.96, -3.32)$ (both negatives). Moreover we resort to the *double linear regression* (DLR) *method*, proposed by Bisgaard and Ankenman (1996), to compute approximated confidence intervals for the eigenvalues of $\mathbf{B}$. The 95% intervals we obtain, that are $(-1.37, -0.54)$ and $(-3.73, -2.90)$ for λ_1 and λ_2 respectively, do not contain zero and there is no reason to suspect the presence of a ridge in the surface. Let us also notice that the estimated stationary point lies inside the experimental region, which is the circle of radius $\sqrt{2}$. This is therefore a typical situation where it is convenient to construct confidence regions for the location of the stationary point. Figure 1(b) shows the 90% and 95% AC regions for $\boldsymbol{\xi}$. The BH regions of the same confidence levels are also displayed in Figure 1(a) for a comparison. The two procedures lead to very similar regions: those obtained using the AC procedure are slightly smaller and entirely contained inside the experimental region. In general, when there is a good fit of the model to the data and the fitted quadratic surface is not "flat" around $\hat{\boldsymbol{\xi}}$ (depending on the eigenvalues of matrix $\hat{\mathbf{B}}$) the BH and AC regions are quite similar in shape and size.

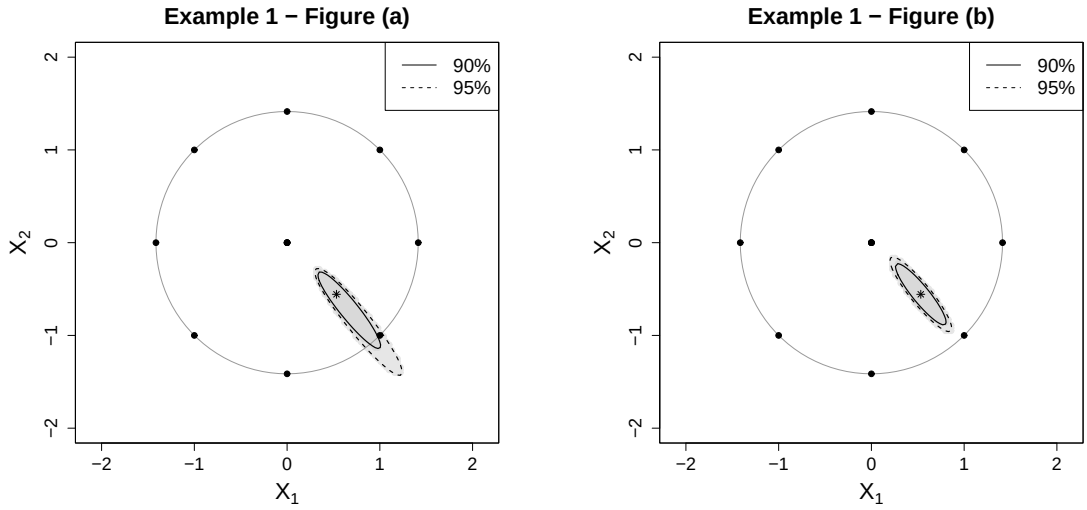


Figure 1: (a) BH confidence regions for $\boldsymbol{\xi}$. (b) AC regions for $\boldsymbol{\xi}$. Symbols $\bullet$ denote the experimental runs and $*$ denotes the estimated stationary point.

4.2 Example 2

The data of a chemical process experiment presented in Section 2.8 of Myers *et al.* (2009) are used to provide a second example with two factors. The input variables involved are the reaction

temperature and the reactant concentration (X_1 and X_2 as coded factors, respectively). The response variable to be maximized is the percent conversion of the chemical process of interest (variable Y). The fitted quadratic model has a very good fit (multiple $R^2 = 0.979$). The F-test for the contribution of the second order terms is significant (p-value< 0.001), meaning that it is necessary to add the quadratic term to the first order model. The stationary point of the fitted surface is $\hat{\boldsymbol{\xi}} = (0.626, -0.061)$, which is within the experimental region. The eigenvalues of $\hat{\mathbf{B}}$ are -2.674 and -11.078, indicating that $\hat{\boldsymbol{\xi}}$ is associated with maximum predicted response. By using the DLR method, we obtain the 95% approximated confidence intervals $(-5.084, -0.264)$ and $(-13.488, -8.668)$ for the eigenvalues of $\mathbf{B}$. Therefore, both eigenvalues are statistically significantly less than zero.

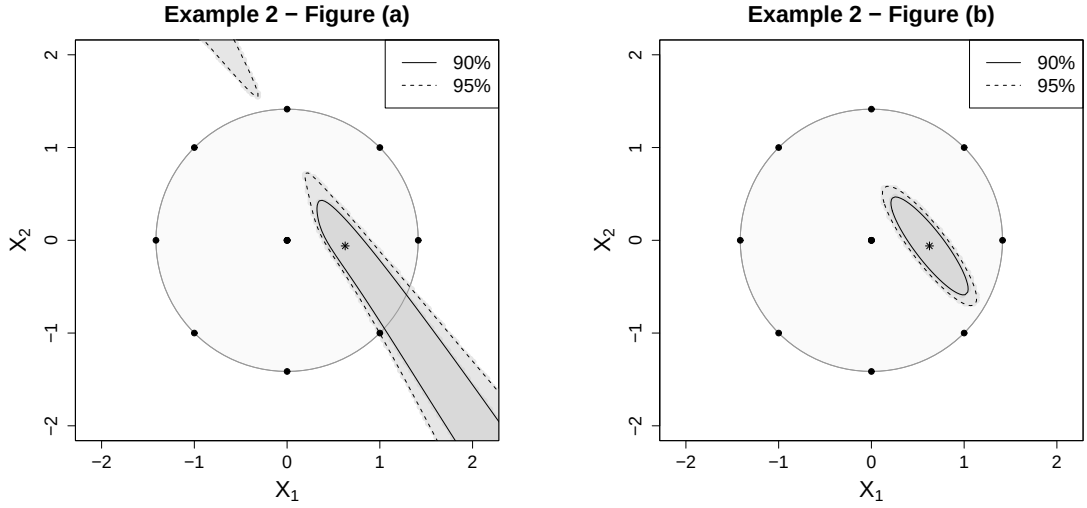


Figure 2: (a) BH confidence regions for $\boldsymbol{\xi}$. (b) AC regions for $\boldsymbol{\xi}$. Symbols $\bullet$ denote the experimental runs and $*$ denotes the estimated stationary point.

In Figure 2(a) we can see that the 95% BH confidence region for the stationary point splits into two disjoint and unbounded regions. The AC region of the same confidence level is instead bounded and entirely embodied in the experimental region (see Figure 2(b)). Actually these latter regions are always bounded and centered around the estimated stationary point, because of their construction procedure. Let us notice that BH regions like those just obtained are not unusual: these sets could consist of disconnected regions since some points could be associated with maximizing points, while other points could be associated with saddle points (see Theorem 2.1 of Peterson *et al.*, 2002). Confidence regions for $\boldsymbol{\xi}$ which are disjoint and open may indicate that we have not sufficient information to conclude that the stationary point exists. However, we agree with Lin and Peterson (2006, p. 76), who regard this behavior of the BH regions as looking rather odd, in cases (such as the current one) where there is a statistical evidence that matrix $\mathbf{B}$ is negative definite and, therefore, that the response is a concave function with a unique stationary point.

4.3 Example 3

The data from Box and Draper (2007, p. 252) are used as a third example with two factors. Fitting model (1), we find that the quadratic terms contribute significantly to the model (p-value$<$ 0.01) and the multiple R^2 is 0.918. The eigenvalues of matrix $\hat{\mathbf{B}}$ are -1.669 and -9.206, so that the estimated stationary point, which is at $\hat{\boldsymbol{\xi}} = (1.138, 1.300)$, is a maximum located in proximity but outside of the experimental region. The DLR method provides the following 95% approximated confidence intervals for the eigenvalues of $\hat{\mathbf{B}}$: $(-5.009, 1.671)$ and $(-12.547, -5.866)$. Actually, even the 75% approximated interval for the first eigenvalue contains the zero value.

In such a situation, where, although $\hat{\boldsymbol{\xi}}$ is a maximum, the statistical evidence suggests that matrix $\mathbf{B}$ is not necessarily negative definite, we find that the 90% and 95% AC regions are obviously bounded, but quite large, indicating flexibility in using operating conditions not exactly equal to $\hat{\boldsymbol{\xi}}$ (see Figure 3(b)). The 90% and 95% BH regions consist in two disconnected and open regions (see Figure 3(a)).

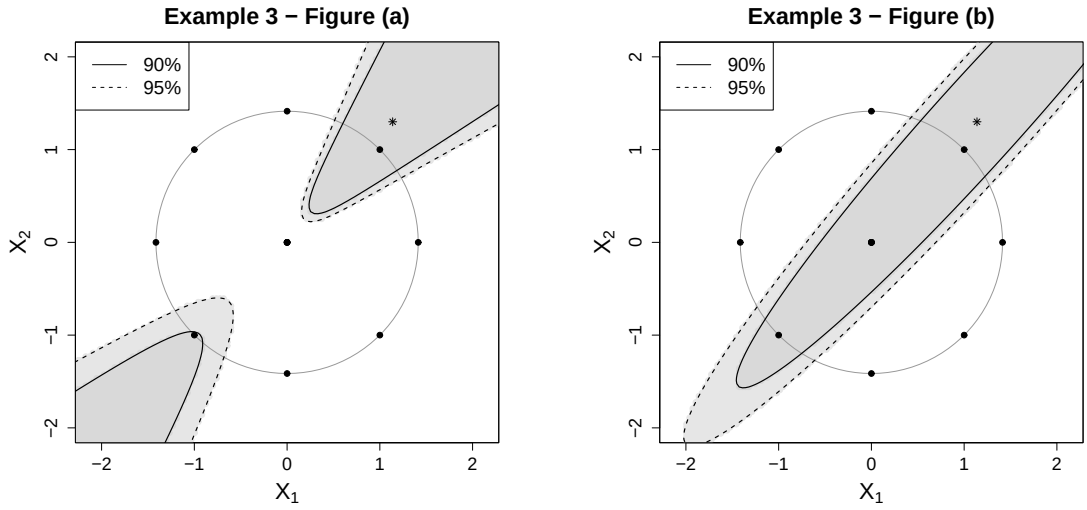


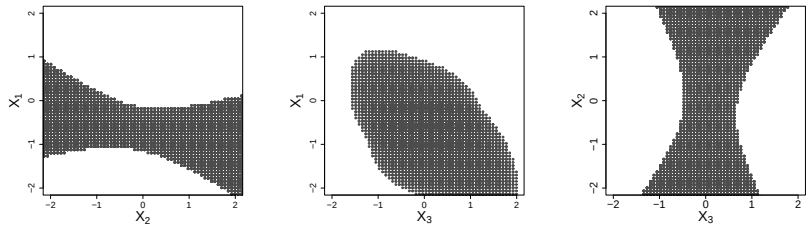
Figure 3: (a) BH confidence regions for $\boldsymbol{\xi}$. (b) AC regions for $\boldsymbol{\xi}$. Symbols $\bullet$ denote the experimental runs and $*$ denotes the estimated stationary point.

5 A three factor experiment

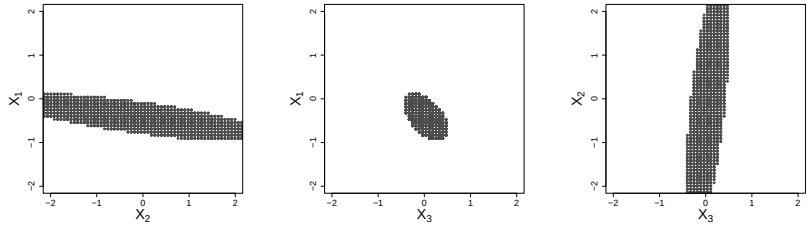
Let us consider the experiment about plant nutrients and optimum rates of fertilizer application provided by Hader *et al.* (1957). The study investigates how copper (X_1 as coded factor), molybdenum (X_2 as coded factor) and iron (X_3 as coded factor) affect lettuce growth (variable Y), measured as grams dry weight after six weeks in solutions of fertilizers containing different levels of the input variables. The goal is therefore to find the combination of factors levels that maximizes Y. Note that the experimental design is not exactly rotatable, since a coded value for X_3 has been changed because of a minor calculation error (see Table E11.22b, p.386, of Box

and Draper (2007), where the reader is referred to, in order to get the data). Actually four sets of experimental runs are conducted, which correspond to as many combinations of nitrogen and iron. In particular, in this section we consider the observed response values which correspond to the nitrogen source $NH_4^+ + NO_3^-$ and the iron source Fe^{3+} .

From the fitted second order surface, we obtain that the F-test for the contribution of the quadratic terms is significant (p-value< 0.001) and $R^2 = 0.903$. The estimated stationary point is at $\hat{\boldsymbol{\xi}} = (-0.416, -0.172, 0.036)$ and it represents a point of maximum, all the eigenvalues of $\hat{\mathbf{B}}$ being negative. However, not all these eigenvalues are statistically significantly less than zero, since one of the 95% approximated confidence intervals obtained through the DLR method is $(-2.577, 1.385)$. Figure 4 shows the point-by-point projections of the 95% BH and AC regions for $\boldsymbol{\xi}$. As expected the AC regions are smaller than the BH ones. Anyway, it is evident that both procedures indicate great flexibility in choosing the level of factor X_2 which maximizes the response variable. It is worth pointing out that if we look at the contributions of the single second order terms to the full quadratic model, we find that the regression coefficient β_{22} is not significantly different from zero (p-value $\cong 0.43$). The other two quadratic terms, instead, contribute significantly to the model (both p-values are less than 0.001).



(a) BH regions



(b) AC regions

Figure 4: Point-by-point projections of 95% confidence regions for $\boldsymbol{\xi}$

If we fit the full quadratic model using only variables X_1 and X_3, there is a strong statistical evidence that both the single second order terms contribute significantly to the model (p-values< 0.0001). The estimated stationary point is $(\hat{\xi}_1, \hat{\xi}_3) = (-0.446, 0.044)$ and turns out to be a maximum. In Figure 5(a) we can see that the 95% BH confidence region is considerably tighter than the point-by-point projection of the three-dimensional 95% BH region on the $X_1 - X_3$ plane showed in Figure 4. In this case the stationary point is estimated very well and the two procedures provide confidence regions quite similar.

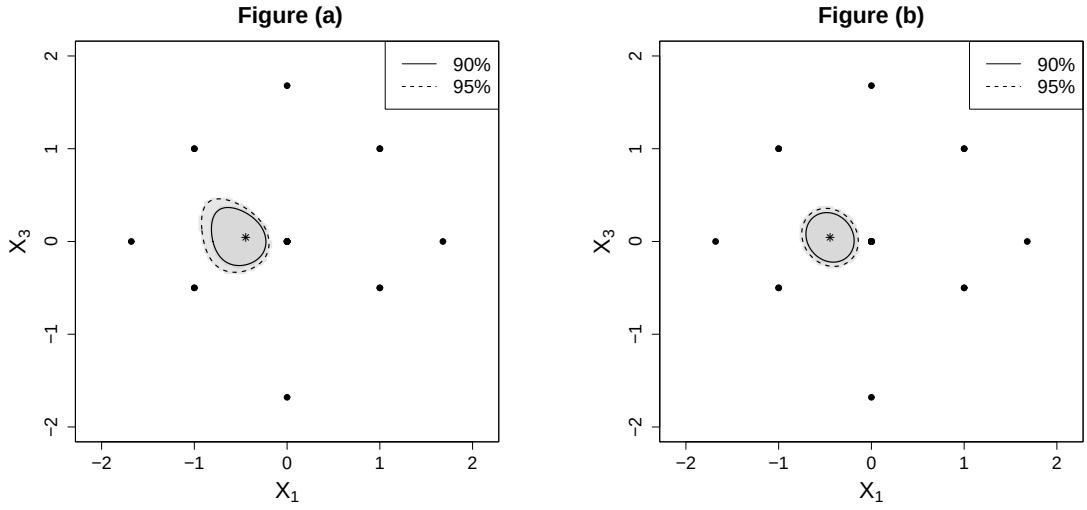


Figure 5: Confidence regions for (ξ_1, ξ_3) after removing variable X_2: (a) BH regions, (b) AC regions. Symbols $\bullet$ denote the experimental runs and $*$ denotes the estimated stationary point.

6 Coverage rates

A simulation study is undertaken to have an empirical check on the coverage probabilities of the proposed confidence regions. Specifically, given a true quadratic response function with two factors, we generate 1000 samples of size 12 from a rotatable CCD with four centre points. The simulated errors are drawn from a normal distribution with zero mean and variance one. These samples are used to compute the coverage rates of the 95% AC regions. Moreover, in order to check if high coverage rates are related to excessively wide confidence regions, we compute the proportion of simulated confidence regions, among those containing the true stationary point, that are completely embodied in the experimental region (that is the circle of radius $\sqrt{2}$, denoted by $\mathcal{R}_1$), in the circle of radius 3 (denoted by $\mathcal{R}_2$) or in the circle of radius 5 (denoted by $\mathcal{R}_3$). The same computations are performed for the 95% BH regions.

Table 3 shows the results of the simulations for different values of $\boldsymbol{\xi}$ and $\boldsymbol{\alpha}$. In particular, the values for $\boldsymbol{\alpha}$ have been chosen in order to take into account different shapes and curvatures, ensuring that the true stationary point is a point of maximum. The true values of $\boldsymbol{\xi}$ involved lie at the design centre ($\boldsymbol{\xi} = (0, 0)$), inside the experimental region ($\boldsymbol{\xi} = (0.5, 0.5)$) and on the boundary of the experimental region ($\boldsymbol{\xi} = (1, 1)$). In most cases, both the procedures lead to empirical coverage probabilities pretty close to the nominal value 0.95. Actually those of the AC regions seem more changeable and tend to decrease as (i) the true stationary point moves away from the design centre and (ii) the true response function is more "flat" around $\boldsymbol{\xi}$. However, let us notice that, when (i) or (ii) occur, the percentage of simulated fitted surface where the the estimated stationary point lies outside the experimental region or where it doesn't turn out to be a maximum (although $\boldsymbol{\xi}$ is a maximum) tend to increase. When such situations appear in real practical experiments, it is not of interest to construct confidence regions for the stationary point, but it is appropriate to carry out further experimental runs.

As regards the width of the confidence regions, in all the cases considered in Table 3, the proportion of the AC regions that contain the true stationary point and are at the same time close inside the experimental region (or $\mathcal{R}_2$, or $\mathcal{R}_3$) is considerably higher than that obtained for the BH regions. Consider for instance the simulated results when $\boldsymbol{\alpha} = (100, -2, -2, 0)$ and $\boldsymbol{\xi} = (0.5, 0.5)$. The coverage rates are both very close to 0.95. However the 58% of the AC regions containing $\boldsymbol{\xi}$ are inside the experimental region, while for the BH regions the percentage is only about 21%. Of course, when a confidence region contains the true parameter, a smaller size is preferred.

Table 3: Coverage rates for 95% BH and AC regions and proportion of regions containing the true stationary point and contained in $\mathcal{R}_1$, $\mathcal{R}_2$ or $\mathcal{R}_3$, for different values of $\boldsymbol{\xi}$ and $\boldsymbol{\alpha}$.

		$\boldsymbol{\alpha} = (100, -2, -2, 0)$ Eigenvalues $\mathbf{A} = (-2,\ -2)$		$\boldsymbol{\alpha} = (100, -8, -9, 6)$ Eigenvalues $\mathbf{A} = (-11.54,\ -5.46)$		$\boldsymbol{\alpha} = (100, -2, -6, 4.5)$ Eigenvalues $\mathbf{A} = (-7.01,\ -0.99)$	
$\boldsymbol{\xi}$		BH	AC	BH	AC	BH	AC
$(0, 0)$	**Coverage Rate**	**0.951**	**0.963**	**0.951**	**0.953**	**0.951**	**0.970**
	% regions inside $\mathcal{R}_1$	0.695	0.998	1	1	0.218	0.834
	% regions inside $\mathcal{R}_2$	0.758	1	1	1	0.268	0.951
	% regions inside $\mathcal{R}_3$	0.775	1	1	1	0.292	0.973
$(0.5, 0.5)$	**Coverage Rate**	**0.956**	**0.951**	**0.956**	**0.952**	**0.956**	**0.941**
	% regions inside $\mathcal{R}_1$	0.218	0.580	0.988	0.999	0.062	0.344
	% regions inside $\mathcal{R}_2$	0.581	0.978	1	1	0.157	0.745
	% regions inside $\mathcal{R}_3$	0.691	0.998	1	1	0.230	0.877
$(1, 1)$	**Coverage Rate**	**0.952**	**0.951**	**0.952**	**0.951**	**0.952**	**0.927**
	% regions inside $\mathcal{R}_1$	0	0	0	0	0	0.001
	% regions inside $\mathcal{R}_2$	0.288	0.703	0.999	1	0.071	0.417
	% regions inside $\mathcal{R}_3$	0.520	0.962	1	1	0.145	0.700

Note: $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ denote respectively the experimental region (i.e. the circle of radius $\sqrt{2}$), the circle of radius 3 and the circle of radius 5.

7 Discussion

In response surface methodology confidence regions on the stationary point are useful to determine how precise our point estimate $\hat{\boldsymbol{\xi}}$ is and how much flexibility is available in choosing optimum conditions. In this paper we exploit a reparametrization of the standard full quadratic model to derive the asymptotic normal distribution of $\hat{\boldsymbol{\xi}}$ and then we use this distribution to construct approximated confidence regions for $\boldsymbol{\xi}$.

Some examples, where two or three factors are involved and the interest is to locate the

maximum point, are showed in Sections 4 and 5. Unlike the standard regions for $\boldsymbol{\xi}$ obtained through the BH procedure, the confidence regions we propose are always bounded because of their construction and have in general smaller areas. A simulation study about coverage rates presented in Section 6 clearly shows how BH and AC regions, with very similar observed coverage probabilities, considerably differ in terms of size. Assuming that the interest is focused on the maximum point of the true surface, let us remind that it is reasonable to construct a confidence region for $\boldsymbol{\xi}$ when the estimated stationary point is actually a maximum and is located inside the experimental region. In these cases confidence intervals for the eigenvalues of matrix $\mathbf{B}$, which can be obtained using different methods, are also computed to assess if the response is a statistically concave quadratic function or not. When these intervals indicate that there is a strong statistical evidence that matrix $\mathbf{B}$ is negative definite, we suggest to use the AC regions rather than the BH ones: in many cases the two procedures give very similar regions (see the numerical example in Subsection 4.1), but it is also possible that the BH method produces two separate and unbounded regions (see the numerical example in Subsection 4.2).

ACKNOWLEDGEMENT

The author is extremely grateful to Ludovico Piccinato for his interest and advice on this paper.

APPENDIX

A Derivation of the Fisher information matrix $\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$

The likelihood function for the whole parameter vector $\boldsymbol{\psi} = (\boldsymbol{\xi}, \boldsymbol{\alpha}, \sigma^2)$ is given by

$$\begin{aligned} L &= L(\boldsymbol{\xi}, \boldsymbol{\alpha}, \sigma^2 | \boldsymbol{y}, \boldsymbol{x}^1, ..., \boldsymbol{x}^n) = (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \\ &\quad \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \big[y_i - \alpha_0 - \sum_{j=1}^{k} \alpha_{jj}(x_{ij} - \xi_j)^2 - \sum_{j=1}^{k-1} \sum_{h=j+1}^{k} \alpha_{jh}(x_{ij} - \xi_j)(x_{ih} - \xi_h) \big]^2 \Big\}, \end{aligned}$$

where $\boldsymbol{y} = (y_1, y_2, ..., y_n)$ and $\boldsymbol{x}^i = (x_{i1}, x_{i2}, ..., x_{ik})^T$. With a slight abuse of notation, let us denote the generic element of the $(p+1) \times (p+1)$ Fisher information matrix by

$$(\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi}))_{i,j} = -\mathbb{E}\Big[\frac{\partial^2}{\partial\psi_i \psi_j} \log L \Big], \quad i, j = 1, ..., p+1.$$

In order to obtain $\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$, it is convenient to partition it into the following blocks

$$\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi}) = \left[\begin{array}{c|c|c} \mathbf{H}_{\xi\xi}(\boldsymbol{\psi}) & \mathbf{H}_{\xi\alpha}(\boldsymbol{\psi}) & \mathbf{H}_{\xi\sigma^2}(\boldsymbol{\psi}) \\ \hline \mathbf{H}^T_{\xi\alpha}(\boldsymbol{\psi}) & \mathbf{H}_{\alpha\alpha}(\boldsymbol{\psi}) & \mathbf{H}_{\alpha\sigma^2}(\boldsymbol{\psi}) \\ \hline \mathbf{H}^T_{\xi\sigma^2}(\boldsymbol{\psi}) & \mathbf{H}^T_{\alpha\sigma^2}(\boldsymbol{\psi}) & \mathbf{H}_{\sigma^2\sigma^2}(\boldsymbol{\psi}) \end{array} \right].$$

After some simple calculations, we get that

$$-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u^2}\log L\Big] = \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big\{\big[2\alpha_{uu}(x_{iu}-\xi_u)+\sum_{l=1}^{u-1}\alpha_{lu}(x_{il}-\xi_l)+\sum_{l=u+1}^{k}\alpha_{ul}(x_{il}-\xi_l)\big]^2\Big\},$$

for $u = 1, ..., k$. We also obtain that

$$\begin{aligned}-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\xi_s}\log L\Big] &= \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big\{\big[2\alpha_{ss}(x_{is}-\xi_s)+\sum_{l=1}^{s-1}\alpha_{ls}(x_{il}-\xi_l)+\sum_{l=s+1}^{k}\alpha_{sl}(x_{il}-\xi_l)\big]\cdot\\ &\quad\big[2\alpha_{uu}(x_{iu}-\xi_u)+\sum_{l=1}^{u-1}\alpha_{lu}(x_{il}-\xi_l)+\sum_{l=u+1}^{k}\alpha_{ul}(x_{il}-\xi_l)\big]\Big\},\end{aligned}$$

for $u = 1, ..., k$ and $s = 1, ..., k$, with $s \neq u$. Thus in matrix notation we have that the $k \times k$ upper left submatrix of $\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$ is

$$\mathbf{H}_{\xi\xi}(\boldsymbol{\psi}) = \frac{1}{\sigma^2}\mathbf{M}\mathbf{M}^T.$$

In order to compute the $k \times p'$ block matrix $\mathbf{H}_{\xi\alpha}(\boldsymbol{\psi})$, we made the following computations

$$-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\alpha_0}\log L\Big] = \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big\{\big[2\alpha_{uu}(x_{iu}-\xi_u)+\sum_{l=1}^{u-1}\alpha_{lu}(x_{il}-\xi_l)^2+\sum_{l=u+1}^{k}\alpha_{ul}(x_{il}-\xi_l)\big]\Big\},$$

$$-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\alpha_{uu}}\log L\Big] =$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big\{\big[2\alpha_{uu}(x_{iu}-\xi_u)+\sum_{l=1}^{u-1}\alpha_{lu}(x_{il}-\xi_l)^2+\sum_{l=u+1}^{k}\alpha_{ul}(x_{il}-\xi_l)\big](x_{iu}-\xi_u)^2\Big\},$$

$$-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\alpha_{su}}\log L\Big] = -\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\alpha_{us}}\log L\Big] =$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big\{\big[2\alpha_{uu}(x_{iu}-\xi_u)+\sum_{l=1}^{u-1}\alpha_{lu}(x_{il}-\xi_l)^2+\sum_{l=u+1}^{k}\alpha_{ul}(x_{il}-\xi_l)\big](x_{is}-\xi_s)(x_{iu}-\xi_u)\Big\},$$

for $u = 1, ..., k$ and $s = 1, ..., k$, with $s \neq u$. Therefore in matrix notation we get

$$\mathbf{H}_{\xi\alpha}(\boldsymbol{\psi}) = \frac{1}{\sigma^2}\mathbf{M}\mathbf{X}_{\boldsymbol{\xi}}.$$

Moreover $-\mathbb{E}\Big[\frac{\partial^2}{\partial\xi_u\partial\sigma^2}\log L\Big] = 0$ and hence $\mathbf{H}_{\xi\sigma^2}(\boldsymbol{\psi}) = \mathbf{0}_k$.

Finally, let us notice that the model (4) may be conveniently written in matrix notation as

$$\boldsymbol{y} = \mathbf{X}_{\boldsymbol{\xi}}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \text{with } \boldsymbol{\varepsilon} \sim \mathbf{MN}_n(\mathbf{0}_n, \sigma^2\mathbf{I}_n).$$

For a given value of $\boldsymbol{\xi}$, this model has the structure of a normal linear model with design matrix $\mathbf{X}_{\boldsymbol{\xi}}$ and hence we can refer to the Fisher information matrix of such a model, obtaining

$$\left[\begin{array}{c|c} \mathbf{H}_{\alpha\alpha}(\boldsymbol{\psi}) & \mathbf{H}_{\alpha\sigma^2}(\boldsymbol{\psi}) \\ \hline \mathbf{H}^T_{\alpha\sigma^2}(\boldsymbol{\psi}) & \mathbf{H}_{\sigma^2\sigma^2}(\boldsymbol{\psi}) \end{array}\right] = \frac{1}{\sigma^2}\left[\begin{array}{c|c} \mathbf{X}^T_{\boldsymbol{\xi}}\mathbf{X}_{\boldsymbol{\xi}} & \mathbf{0}_{p'} \\ \hline \mathbf{0}^T_{p'} & \frac{n}{2\sigma^2} \end{array}\right].$$

We have therefore obtained that Fisher information matrix $\mathbf{H}_{\boldsymbol{\psi}}(\boldsymbol{\psi})$ is given by (5).

B A tool for computing the proposed confidence regions

We provide a very user friendly software tool for the display of the AC regions for the stationary point when $k = 2$. The program, coded in R-programming language, uses the package *gWidgets* (Verzani, 2007), which offers a relatively simple way of writing GUIs. This package is therefore supposed to be correctly installed within R. The file containing the R code is called "ConfRegStatPoint.R" and is available upon request.

Before using the program, the user must create a white-space separated text file, named "data.txt" and without column names, which contains the experimental data. This file must have n rows and 3 columns: each row represents an experimental run and contains the values x_{i1}, x_{i2} and y_i (for $i = 1, ..., n$), that is the two factors levels in coded units and the corresponding observed response value.

When running R, it is necessary to change the default starting directory to the one containing the file "data.txt". It can be easily done by choosing "Change directory" under the File menu and selecting the directory of interest. Then, from the same menu, select "Source R code" and open the file "ConfRegStatPoint.R". An information dialog will pop-up, providing a message about the location and the nature of the estimated stationary point, as shown in Figure 6.

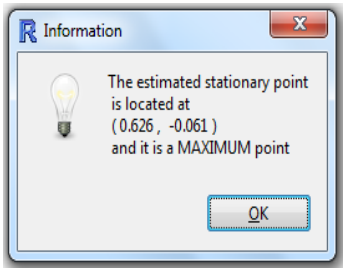


Figure 6: Information dialog about the location and the nature of $\hat{\boldsymbol{\xi}}$

After clicking on the button "ok", a window, such as that shown in the left panel of Figure 7, will open and will allow the user to specify the confidence level of the region and some plot-style options. The confidence level can be selected in a sequence of values from 0.5 to 0.99 and increment 0.01. It is possible to pick the lower and upper limits of the axes for both input variables: a fixed list of values is available. Moreover the user can decide to display in the graph

- the experimental runs, denoted by the symbol •, by selecting the item TRUE from the "Experimental Points" option;

- the stationary point of the fitted surface, denoted by the symbol $*$, by choosing the item TRUE from the "Estimated Stationary Point" option.

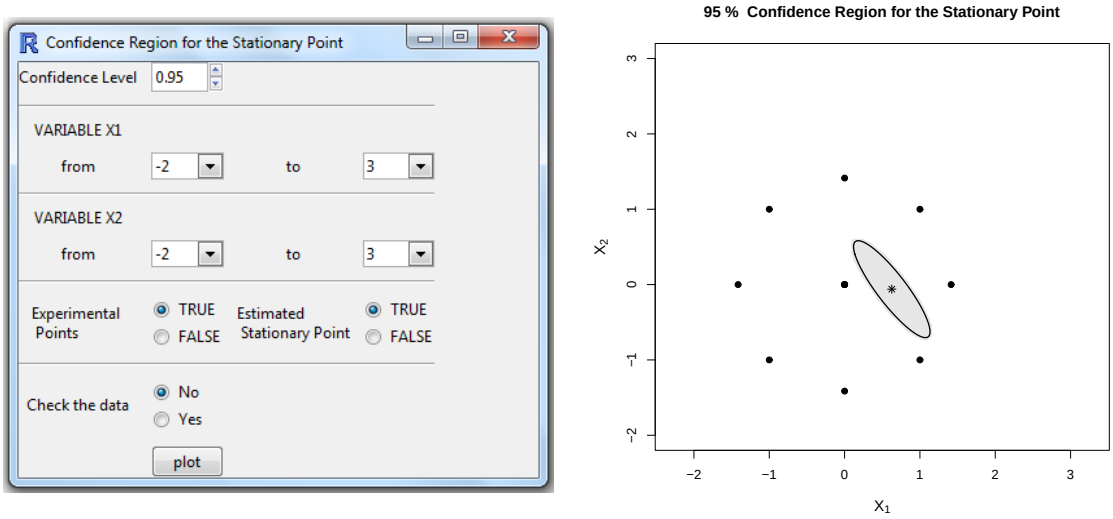


Figure 7: AC region for the experiment in Subsection 4.2, obtained using the R Gui described in Appendix B.

Finally, if one choose "Yes" from the "Check the data" option, a window showing the data appears, letting the user control if the data have been correctly inserted and imported in R. By clicking on the button "plot", the graph of the AC region is obtained. An example is provided in Figure 7, where in the right panel the confidence region for the experiment in Subsection 4.2 is plotted, using the options selected in the GUI represented in the left panel.

References

Bisgaard, S. and Ankenman, B. (1996): Standard errors for the eigenvalues in second-order response surface models, *Technometrics*, **38**, No 3, 238-246.

Box, G.E.P. and Draper, N.R. (2007). *Response Surfaces, Mixtures, and Ridge Analyses*, 2nd Ed., Wiley, New York.

Box, G.E.P. and Hunter, J.S. (1954). A confidence region for the solution of a set of simultaneous equations with an application to experimental design. *Biometrika*, **41**, 190-199.

Box, G.E.P. and Wilson, K.B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society*, **B**, **13**, 1-45.

Carter, W.H., Chinchilli, V.M., Myers, R.H. and Campbell, E.D. (1986): Confidence intervals and an improved ridge analysis of response surfaces, *Technometrics*, **28**, 339-346.

Carter, W.H., Chinchilli, V.M. and Campbell, E.D. (1990): A large-sample confidence region useful in characterizing the stationary point of a quadratic response surface, *Technometrics*, **32**, 425-435.

Graybill, F.A. (1983). *Matrices with Applications in Statistics*, 2nd Ed., Wadsworth, Belmont, USA.

Hader, R.J., Harward, M.E., Mason, D.D. and Moore D.P. (1957). An Investigation of Some of the Relationships between Copper, Iron, and Molybdenum in the Growth and Nutrition of Lettuce: I. Experimental Design and Statistical Methods for Characterizing the Response Surface. *Soil Science Society of America Journal* **21**, 5964.

Khuri, A.I. and Cornell, J.A. (1996). *Response Surfaces: Design and Analyses*, 2nd Ed., rev. and expanded, Marcel Dekker, New York.

Li, J., Ma, C., Ma, Y., Li, Y., Zhou, W. and Xu, P. (2007). Medium optimization by combination of response surface methodology and desirability function: an application in glutamine production. *Appl. Microbiol. Biotechnol.*, **74**, 563-571.

Lin, D.K.J. and Peterson, J.J. (2006). Statistical inference for response surface optima. In *Response surface methodology and related topics*, Edited by A.I. Khuri, Chapter 4, World Scientific, Singapore.

Myers, R.H., Montgomery, D.C. and Anderson-Cook, C.M (2009). *Response Surface Methodology: Process and Produce Optimization Using Designed Experiments*, 3rd Ed., Wiley, New York.

Peterson, J.J. (1993). A general approach to ridge analysis with confidence intervals. *Technometrics*, **35**, 204-214.

Peterson, J.J., Cahya, S. and del Castillo, E. (2002). A general approach to confidence regions for optimal factor levels of response surfaces. *Biometrics*, **58**, 422-431.

R Development Core Team (2009). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Sambucini, V. (2007). A reference prior for the analysis of a response surface. *Journal of Statistical Planning and Inference*, **137**, 1119-1128.

Sambucini, V. and Piccinato, L. (2008). Likelihood and Bayesian approaches to inference for the stationary point of a quadratic response surface. *Canadian Journal of Statistics*, **36**, 223-238.

Stablein, D.M., Carter, W.H. and Wampler, G.L. (1983). Confidence regions for constrained optimain response surface experiments. *Biometrics*, **39**, 759-763.

Verzani, J. (2007). An introduction to gWidgets. *R News*, **7**(3), 26-33.