Ridge analysis through profile likelihoods

Valeria Sambucini and Ludovico Piccinato

Abstract A likelihood approach is considered as an alternative to the classical ridge analysis when the estimated stationary point of a quadratic response surface is a saddle point or it is not located inside the experimental region. We compute the path of maximum profile likelihood for the optimum point on the contours of hyperspheres centered at the design center. Plausibility regions around this path can provide information about the nature of the system inside the experimental region and about suitable regions where to conduct further experiments.

Key words: maximum likelihood path, plausibility regions, profile likelihood, response surface methodology, ridge analysis

1 Introduction

In Response Surface Methodology quadratic models are the most frequently used to approximate the relationship between a response variable and k continuous factors over a bounded region. The aim is to determine the settings for the independent variables that result in the optimum value of the response.

Without loss of generality, let us assume that we are interested in the maximum point of the true surface. From the fitted quadratic model, the stationary point of the surface is estimated. It can be a maximum, a minimum or a saddle point. Therefore, a *canonical analysis* is typically performed to determine the nature of the estimated stationary point, $\hat{\xi}$. If it turns out to be a maximum point and lies inside the experimental region, it is reasonable to suggest it as a candidate for

Valeria Sambucini
Sapienza Università di Roma, e-mail: valeria.sambucini@uniroma1.it

Ludovico Piccinato
Sapienza Università di Roma, e-mail: ludovico.piccinato@uniroma1.it

operating conditions and further inferential tools are used to assess the accuracy of the estimate. If, instead, $\hat{\xi}$ is maximum but located far away of the experimental region or if it is a saddle point, further experiments should be conducted. In this latter situation, a *ridge analysis*, firstly suggested by Hoerl [4] and further investigated by Draper [2], is typically implemented. This technique uses the method of Lagrange multipliers to maximize the estimated response on the contours of hyperspheres of increasing radius centered at the origin of the factor space (the design center), obtaining the path of highest response. For details about the mathematical development of this methodology readers are referred to [1]. Furthermore, Gilmour and Draper [3] provide confidence intervals for the maximum response at fixed radius R and obtain a confidence region around the path of highest response by varying R.

In this paper we use a likelihood approach to propose an alternative method to handle situations where the classical ridge analysis is typically implemented. More specifically, we shall use the profile likelihood for the optimum to determine the locus of points of maximum likelihood, under the constraint that each point lies on the contours of an hypersphere of a certain radius. The uncertainty around this locus of points can be assessed by computing plausibility regions. Different likelihood tools, such as for instance integrated likelihoods, could be similarly used. As for the ridge analysis, this procedure is performed with a double aim. First of all, it is of practical interest to determine the nature of the system inside the experimental region, when a saddle point is observed. Secondly, it can be used as a tool to indicate the direction in which further experimentation should be performed.

2 Maximum profile likelihood path and plausibility regions

Given n observations on the response, we consider a useful reparametrization of the standard quadratic model, where the stationary point appears explicitly (see [5]),

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

Here $\mathbf{y}$ is the vector of responses, $\alpha = (\alpha_0, \alpha_1, ..., \alpha_k, \alpha_{12}, ..., \alpha_{k-1,k})$ is a vector of $p' = 1 + k + \frac{k(k-1)}{2}$ coefficients and ε is the vector of random errors. The coordinates of the stationary point $\xi = (\xi_1, ..., \xi_k)$ are contained in the $n \times p'$ matrix $\mathbf{X}_\xi$ whose generic row, for $i = 1, ..., n$, is given by

$$[1, (x_{i1} - \xi_1)^2, ..., (x_{ik} - \xi_k)^2, (x_{i1} - \xi_1)(x_{i2} - \xi_2), ..., (x_{i,k-1} - \xi_{k-1})(x_{ik} - \xi_k)].$$

In the likelihood approach, a common way to eliminate the influence of the nuisance parameters is to use conditional maximization, obtaining the profile likelihood for the parameter of interest. In our case, the profile likelihood for ξ (see [6]) is

$$L_p(\xi) \propto \left(\mathbf{y} - \mathbf{X}_\xi \hat{\alpha}(\xi)\right)^T \left(\mathbf{y} - \mathbf{X}_\xi \hat{\alpha}(\xi)\right),$$

where $\hat{\alpha}(\xi) = (\mathbf{X}_\xi^T \mathbf{X}_\xi)^{-1} \mathbf{X}_\xi^T \mathrm{y}$ is the maximum likelihood estimator of α for a fixed ξ. The Hessian matrix of the estimated response surface written as function of ξ is $2\hat{\mathbf{A}}(\xi)$, where $\hat{\mathbf{A}}(\xi)$ is a $k \times k$ symmetric matrix with ith diagonal element equal to $\hat{\alpha}_i(\xi)$ and (ij)th off-diagonal element equal to $\frac{1}{2}\hat{\alpha}_{ij}(\xi)$. Therefore, by studying the sign of the eigenvalues of $\hat{\mathbf{A}}(\xi)$ it is possible to explore the nature of any fixed value of ξ. In particular the set $\Xi_{max} = \{\xi : \hat{\mathbf{A}}(\xi) \text{ is negative definite}\}$ contains all stationary points that, according to the data, turn out to be points of maxima.

As an alternative to ridge analysis, we propose to consider the path of maximum profile likelihood by identifying the points ξ with highest value of $L_p(\xi)$ on hyperspheres of increasing radius and centered at the origin $\xi = (0,...,0)$. Note that it is necessary to restrict the research to the points in the set Ξ_{max} in order to provide the path associated to stationary points that are maxima. Moreover, to assess the uncertainty around this path, we suggest to use plausibility regions that are given by all the parameter values for which the relative profile likelihood is greater than or equal to a fixed level q, with $0 \leq q \leq 1$. Specifically, in our case, the plausibility regions must be computed conditionally on hyperspheres. In practice, for a given radius R and a level q, we consider all the points $\xi \in \Xi_{max}$ that reside on an hypersphere of radius R and are such that $\dfrac{L_p(\xi)}{L_p(\hat{\xi}_R)} \geq q$, where $\hat{\xi}_R$ is the point on the path at radius R. Then, by considering different values of R, we get a region of most plausible values for the points along the path of maximum profile likelihood. This region can provide useful information regarding where future experiments should be performed. Notice that the region we get using this procedure is larger than the not-constrained profile likelihood region obtained for the same level q.

3 Numerical examples

In this Section we consider two different experiments which involve two factors. The first one is described by [1] (Exercise 7.11). A central composite design (CCD) with two centre runs is used to maximize the true response surface. The estimated stationary point, $\hat{\xi} = (1.138, 1.300)$, is a maximum point located outside the experimental region, that is the circle of radius $\sqrt{2}$ (see left panel of Figure 1). To provide a second example, we consider the first 14 runs of the experiment given in [3], which correspond to a CCD with four centre points. Therefore, also in this case the experimental region is the circle of radius $\sqrt{2}$, but the estimated stationary point is $\hat{\xi} = (0.278, 0.345)$ and turns out to be a saddle point (see right panel of Figure 1). Thus, in both cases additional explorations are recommended.

The proposed procedure is implemented using an algorithm, developed using the statistical software R, which creates grids of points on circles centered at the origin, to compute both the maximum profile likelihood path and its plausibility region of level $q = 0.15$. In both Figures, different gray levels denote the areas of the stationary point parameter space which correspond to possible maximum points (the set Ξ_{max}), saddle points and minimum points. Of course the path and the plausibility region are completely enclosed in Ξ_{max} and, in particular for the first experiment, $\hat{\xi}$ lies on the path. Moreover, as it is reasonable, the plausibility band becomes larger with the distance from the design.

It is possible to show that the likelihood path is pretty much the same as the path of highest response obtained through the classical ridge analysis. This does not hold, however, for the plausibility regions, since the family of such regions appears quite different from the family of the confidence regions (see [3]). This can be expected as a consequence of the different inferential techniques adopted. Note moreover that the likelihood technique allows a direct selection of the possible maximum points among the stationary points.

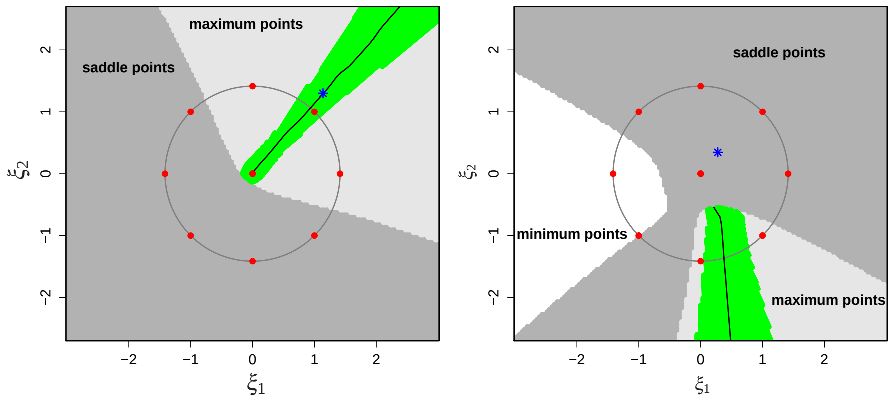

Fig. 1 Maximum profile likelihood path and plausibility region of level $q = 0.15$ for the two examples considered in Section 3. Symbols $\bullet$ denote the experimental runs and $*$ denotes the estimated stationary point.

References

1. Box, G.E.P., Draper, N.R.: Response Surfaces, Mixtures, and Ridge Analyses. 2nd Ed., Wiley, New York (2007)
2. Draper N.R.: Ridge analysis of response surfaces. Technometrics. **5**, 469–479 (1963)
3. Gilmour, S., Draper, N.R.: Confidence regions around the ridge of optimal response on fitted second-order response surfaces. Technometrics **45**, 333–339 (2003)
4. Hoerl, A.E.: Optimum solution of many variables equations. Chem. Eng. Prog. **55**, 69–78 (1959)
5. Sambucini, V.: A reference prior for the analysis of a response surface. J. Stat. Plan. Inference **137**, 1119–1128 (2007)
6. Sambucini, V., Piccinato, L.: Likelihood and Bayesian approaches to inference for the stationary point of a quadratic response surface. Can. J. Stat. **36**, 223–238 (2008)