

CHAPTER 8

Optimization in Statistics

Optimization is an essential feature in many problems in statistics. This is apparent in almost all fields of statistics. Here are few examples, some of which will be discussed in more detail in this chapter.

1. In the theory of estimation, an estimator of an unknown parameter is sought that satisfies a certain optimality criterion such as minimum variance, maximum likelihood, or minimum average risk (as in the case of a Bayes estimator). Some of these criteria were already discussed in Section 7.11. For example, in regression analysis, estimates of the parameters of a fitted model are obtained by minimizing a certain expression that measures the closeness of the fit of the model. One common example of such an expression is the sum of the squared residuals (these are deviations of the predicted response values, as specified by the model, from the corresponding observed response values). This particular expression is used in the method of ordinary least squares. A more general class of parameter estimators is the class of M -estimators. See Huber (1973, 1981). The name “ M -estimator” comes from “generalized maximum likelihood.” They are based on the idea of replacing the squared residuals by another symmetric function of the residuals that has a unique minimum at zero. For example, minimizing the sum of the absolute values of the residuals produces the so-called *least absolute values* (LAV) estimators.
2. Estimates of the variance components associated with random or mixed models are obtained by using several methods. In some of these methods, the estimates are given as solutions to certain optimization problems as in maximum likelihood (ML) estimation and minimum norm quadratic unbiased estimation (MINQUE). In the former method, the likelihood function is maximized under the assumption of normally distributed data [see Hartley and Rao (1967)]. A completely different approach is used in the latter method, which was proposed by Rao (1970, 1971). This method does not require the normality assumption.

For a review of methods of estimating variance components, see Khuri and Sahai (1985).

3. In statistical inference, tests are constructed so that they are optimal in a certain sense. For example, in the Neyman–Pearson lemma (see, for example, Roussas, 1973, Chapter 13), a test is obtained by minimizing the probability of Type II error while holding the probability of Type I error at a certain level.
4. In the field of response surface methodology, design settings are chosen to minimize the prediction variance inside a region of interest, or to minimize the bias that occurs from fitting the “wrong” model. Other optimality criteria can also be considered. For example, under the D -optimality criterion, the determinant of the variance–covariance matrix of the least-squares estimator of the vector of unknown parameters (of a fitted model) is minimized with respect to the design settings.
5. Another objective of response surface methodology is the determination of optimum operating conditions on the input variables that produce maximum, or minimum, response values inside a region of interest. For example, in a particular chemical reaction setting, it may be of interest to determine the reaction temperature and the reaction time that maximize the percentage yield of a product. Optimum seeking methods in response surface methodology will be discussed in detail in Section 8.3.
6. Several response variables may be observed in an experiment for each setting of a group of input variables. Such an experiment is called a multiresponse experiment. In this case, optimization involves a number of response functions and is therefore referred to as simultaneous (or multiresponse) optimization. For example, it may be of interest to maximize the yield of a certain chemical compound while reducing the production cost. Multiresponse optimization will be discussed in Section 8.7.
7. In multivariate analysis, a large number of measurements may be available as a result of some experiment. For convenience in the analysis and interpretation of such data, it would be desirable to work with fewer of the measurements, without loss of much information. This problem of data reduction is dealt with by choosing certain linear functions of the measurements in an optimal manner. Such linear functions are called *principal components*.

Optimization of a multivariable function was discussed in Chapter 7. However, there are situations in which the optimum cannot be obtained explicitly by simply following the methods described in Chapter 7. Instead, iterative procedures may be needed. In this chapter, we shall first discuss some commonly used iterative optimization methods. A number of these methods require the explicit evaluation of the partial derivatives of the function to be optimized (objective function). These

are referred to as the gradient methods. Three other optimization techniques that rely solely on the values of the objective function will also be discussed. They are called direct search methods.

8.1. THE GRADIENT METHODS

Let $f(\mathbf{x})$ be a real-valued function of k variables $x_1, x_2, \dots, x_k$, where $\mathbf{x} = (x_1, x_2, \dots, x_k)'$. The gradient methods are based on approximating $f(\mathbf{x})$ with a low-degree polynomial, usually of degree one or two, using Taylor's expansion. The first- and second-order partial derivatives of $f(\mathbf{x})$ are therefore assumed to exist at every point $\mathbf{x}$ in the domain of f . Without loss of generality, we shall consider that f is to be minimized.

8.1.1. The Method of Steepest Descent

This method is based on a first-order approximation of $f(\mathbf{x})$ with a polynomial of degree one using Taylor's theorem (see Section 7.5). Let $\mathbf{x}_0$ be an initial point in the domain of $f(\mathbf{x})$. Let $\mathbf{x}_0 + t\mathbf{h}_0$ be a neighboring point, where $t\mathbf{h}_0$ represents a small change in the direction of a unit vector $\mathbf{h}_0$ (that is, $t > 0$). The corresponding change in $f(\mathbf{x})$ is $f(\mathbf{x}_0 + t\mathbf{h}_0) - f(\mathbf{x}_0)$. A first-order approximation of this change is given by

$$f(\mathbf{x}_0 + t\mathbf{h}_0) - f(\mathbf{x}_0) \approx t\mathbf{h}_0'\nabla f(\mathbf{x}_0), \quad (8.1)$$

as can be seen from applying formula (7.27). If the objective is to minimize $f(\mathbf{x})$, then $\mathbf{h}_0$ must be chosen so as to obtain the largest value for $-t\mathbf{h}_0'\nabla f(\mathbf{x}_0)$. This is a constrained maximization problem, since $\mathbf{h}_0$ has unit length. For this purpose we use the method of Lagrange multipliers. Let F be the function

$$F = -t\mathbf{h}_0'\nabla f(\mathbf{x}_0) + \lambda(\mathbf{h}_0'\mathbf{h}_0 - 1).$$

By differentiating F with respect to the elements of $\mathbf{h}_0$ and equating the derivatives to zero we obtain

$$\mathbf{h}_0 = \frac{t}{2\lambda} \nabla f(\mathbf{x}_0). \quad (8.2)$$

Using the constraint $\mathbf{h}_0'\mathbf{h}_0 = 1$, we find that λ must satisfy the equation

$$\lambda^2 = \frac{t^2}{4} \|\nabla f(\mathbf{x}_0)\|_2^2, \quad (8.3)$$

where $\|\nabla f(\mathbf{x}_0)\|_2$ is the Euclidean norm of $\nabla f(\mathbf{x}_0)$. In order for $-t\mathbf{h}_0'\nabla f(\mathbf{x}_0)$

to have a maximum, λ must be negative. From formula (8.3) we then have

$$\lambda = -\frac{t}{2}\|\nabla f(\mathbf{x}_0)\|_2.$$

By substituting this expression in formula (8.2) we get

$$\mathbf{h}_0 = -\frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|_2}. \quad (8.4)$$

Thus for a given $t > 0$, we can achieve a maximum reduction in $f(\mathbf{x}_0)$ by moving from $\mathbf{x}_0$ in the direction specified by $\mathbf{h}_0$ in formula (8.4). The value of t is now determined by performing a linear search in the direction of $\mathbf{h}_0$. This is accomplished by increasing the value of t (starting from zero) until no further reduction in the values of f is obtained. Let such a value of t be denoted by t_0 . The corresponding value of $\mathbf{x}$ is given by

$$\mathbf{x}_1 = \mathbf{x}_0 - t_0 \frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|_2}.$$

Since the direction of $\mathbf{h}_0$ is in general not toward the location $\mathbf{x}^*$ of the true minimum of f , the above process must be performed iteratively. Thus if at stage i we have an approximation $\mathbf{x}_i$ for $\mathbf{x}^*$, then at stage $i + 1$ we have the approximation

$$\mathbf{x}_{i+1} = \mathbf{x}_i + t_i \mathbf{h}_i, \quad i = 0, 1, 2, \dots,$$

where

$$\mathbf{h}_i = -\frac{\nabla f(\mathbf{x}_i)}{\|\nabla f(\mathbf{x}_i)\|_2}, \quad i = 0, 1, 2, \dots,$$

and t_i is determined by a linear search in the direction of $\mathbf{h}_i$, that is, t_i is the value of t that minimizes $f(\mathbf{x}_i + t\mathbf{h}_i)$. Note that if it is desired to maximize f , then for each i (≥ 0) we need to move in the direction of $-\mathbf{h}_i$. In this case, the method is called the method of steepest ascent.

Convergence of the method of steepest descent can be very slow, since frequent changes of direction may be necessary. Another reason for slow convergence is that the direction of $\mathbf{h}_i$ at the i th iteration may be nearly perpendicular to the direction toward the minimum. Furthermore, the method becomes inefficient when the first-order approximation of f is no longer adequate. In this case, a second-order approximation should be attempted. This will be described in the next section.

8.1.2. The Newton–Raphson Method

Let $\mathbf{x}_0$ be an initial point in the domain of $f(\mathbf{x})$. By a Taylor's expansion of f in a neighborhood of $\mathbf{x}_0$ (see Theorem 7.5.1), it is possible to approximate $f(\mathbf{x})$ with the quadratic function $\phi(\mathbf{x})$ given by

$$\phi(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)' \nabla f(\mathbf{x}_0) + \frac{1}{2!} (\mathbf{x} - \mathbf{x}_0)' \mathbf{H}_f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0), \quad (8.5)$$

where $\mathbf{H}_f(\mathbf{x}_0)$ is the Hessian matrix of f evaluated at $\mathbf{x}_0$.

On the basis of formula (8.5) we can obtain a reasonable approximation to the minimum of $f(\mathbf{x})$ by using the minimum of $\phi(\mathbf{x})$. If $\phi(\mathbf{x})$ attains a local minimum at $\mathbf{x}_1$, then we must necessarily have $\nabla \phi(\mathbf{x}_1) = \mathbf{0}$ (see Section 7.7), that is,

$$\nabla f(\mathbf{x}_0) + \mathbf{H}_f(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0) = \mathbf{0}. \quad (8.6)$$

If $\mathbf{H}_f(\mathbf{x}_0)$ is nonsingular, then from equation (8.6) we obtain

$$\mathbf{x}_1 = \mathbf{x}_0 - \mathbf{H}_f^{-1}(\mathbf{x}_0) \nabla f(\mathbf{x}_0).$$

If we now approximate $f(\mathbf{x})$ with another quadratic function, by again applying Taylor's expansion in a neighborhood of $\mathbf{x}_1$, and then repeat the same process as before with $\mathbf{x}_1$ used instead of $\mathbf{x}_0$, we obtain the point

$$\mathbf{x}_2 = \mathbf{x}_1 - \mathbf{H}_f^{-1}(\mathbf{x}_1) \nabla f(\mathbf{x}_1).$$

Further repetitions of this process lead to a sequence of points, $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots$, such that

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{H}_f^{-1}(\mathbf{x}_i) \nabla f(\mathbf{x}_i), \quad i = 0, 1, 2, \dots \quad (8.7)$$

The Newton–Raphson method requires finding the inverse of the Hessian matrix $\mathbf{H}_f$ at each iteration. This can be computationally involved, especially if the number of variables, k , is large. Furthermore, the method may fail to converge if $\mathbf{H}_f(\mathbf{x}_i)$ is not positive definite. This can occur, for example, when $\mathbf{x}_i$ is far from the location $\mathbf{x}^*$ of the true minimum. If, however, the initial point $\mathbf{x}_0$ is close to $\mathbf{x}^*$, then convergence occurs at a rapid rate.

8.1.3. The Davidon–Fletcher–Powell Method

This method is basically similar to the one in Section 8.1.1 except that at the i th iteration we have

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \theta_i \mathbf{G}_i \nabla f(\mathbf{x}_i), \quad i = 0, 1, 2, \dots,$$

where $\mathbf{G}_i$ is a positive definite matrix that serves as the i th approximation to the inverse of the Hessian matrix $\mathbf{H}_f(\mathbf{x}_i)$, and θ_i is a scalar determined by a linear search from $\mathbf{x}_i$ in the direction of $-\mathbf{G}_i \nabla f(\mathbf{x}_i)$, similar to the one for the steepest descent method. The initial choice $\mathbf{G}_0$ of the matrix $\mathbf{G}$ can be any positive definite matrix, but is usually taken to be the identity matrix. At the $(i + 1)$ st iteration, $\mathbf{G}_i$ is updated by using the formula

$$\mathbf{G}_{i+1} = \mathbf{G}_i + \mathbf{L}_i + \mathbf{M}_i, \quad i = 0, 1, 2, \dots,$$

where

$$\mathbf{L}_i = - \frac{\mathbf{G}_i [\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)] [\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)]' \mathbf{G}_i}{[\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)]' \mathbf{G}_i [\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)]},$$

$$\mathbf{M}_i = - \frac{\theta_i [\mathbf{G}_i \nabla f(\mathbf{x}_i)] [\mathbf{G}_i \nabla f(\mathbf{x}_i)]'}{[\mathbf{G}_i \nabla f(\mathbf{x}_i)]' [\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)]}.$$

The justification for this method is given in Fletcher and Powell (1963). See also Bunday (1984, Section 4.3). Note that if $\mathbf{G}_i$ is initially chosen as the identity, then the first increment is in the steepest descent direction $-\nabla f(\mathbf{x}_0)$.

This is a powerful optimization method and is considered to be very efficient for most functions.

8.2. THE DIRECT SEARCH METHODS

The direct search methods do not require the evaluation of any partial derivatives of the objective function. For this reason they are suited for situations in which it is analytically difficult to provide expressions for the partial derivatives, such as the minimization of the maximum absolute deviation. Three such methods will be discussed here, namely, the Nelder–Mead simplex method, Price's controlled random search procedure, and generalized simulated annealing.

8.2.1. The Nelder–Mead Simplex Method

Let $f(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \dots, x_k)'$, be the function to be minimized. The simplex method is based on a comparison of the values of f at the $k + 1$ vertices of a general simplex followed by a move away from the vertex with the highest function value. By definition, a general simplex is a geometric figure formed by a set of $k + 1$ points called vertices in a k -dimensional space. Originally, the simplex method was proposed by Spendley, Hext, and Himsworth (1962), who considered a regular simplex, that is, a simplex with mutually equidistant points such as an equilateral triangle in a two-dimensional space ($k = 2$). Nelder and Mead (1965) modified this method by

allowing the simplex to be nonregular. This modified version of the simplex method will be described here.

The simplex method follows a sequential search procedure. As was mentioned earlier, it begins by evaluating f at the $k+1$ points that form a general simplex. Let these points be denoted by $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k+1}$. Let f_h and f_l denote, respectively, the largest and the smallest of the values $f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_{k+1})$. Let us also denote the points where f_h and f_l are attained by $\mathbf{x}_h$ and $\mathbf{x}_l$, respectively.

Obviously, if we are interested in minimizing f , then a move away from $\mathbf{x}_h$ will be in order. Let us therefore define $\mathbf{x}_c$ as the centroid of all the points with the exclusion of $\mathbf{x}_h$. Thus

$$\mathbf{x}_c = \frac{1}{k} \sum_{i \neq h} \mathbf{x}_i.$$

In order to move away from $\mathbf{x}_h$, we reflect $\mathbf{x}_h$ with respect to $\mathbf{x}_c$ to obtain the point $\mathbf{x}_h^*$. More specifically, the latter point is defined by the relation

$$\mathbf{x}_h^* - \mathbf{x}_c = r(\mathbf{x}_c - \mathbf{x}_h),$$

or equivalently,

$$\mathbf{x}_h^* = (1+r)\mathbf{x}_c - r\mathbf{x}_h,$$

where r is a positive constant called the reflection coefficient and is given by

$$r = \frac{\|\mathbf{x}_h^* - \mathbf{x}_c\|_2}{\|\mathbf{x}_c - \mathbf{x}_h\|_2}.$$

The points $\mathbf{x}_h$, $\mathbf{x}_c$, and $\mathbf{x}_h^*$ are depicted in Figure 8.1. Let us consider the

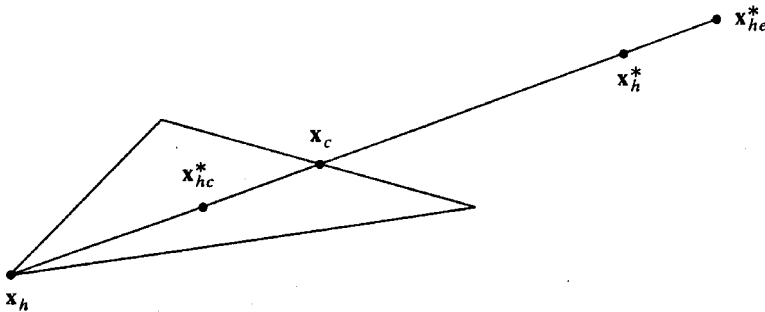


Figure 8.1. A two-dimensional simplex with the reflection ($\mathbf{x}_h^*$), expansion ($\mathbf{x}_{he}^*$), and contraction ($\mathbf{x}_{hc}^*$) points.

following cases:

- a. If $f_l < f(\mathbf{x}_h^*) < f_h$, replace $\mathbf{x}_h$ by $\mathbf{x}_h^*$ and start the process again with the new simplex (that is, evaluate f at the vertices of the simplex which has the same points as the original simplex, but with $\mathbf{x}_h^*$ substituted for $\mathbf{x}_h$).
- b. If $f(\mathbf{x}_h^*) < f_l$, then the move from $\mathbf{x}_c$ to $\mathbf{x}_h^*$ is in the right direction and should therefore be expanded. In this case, $\mathbf{x}_h^*$ is expanded to $\mathbf{x}_{he}^*$ defined by the relation

$$\mathbf{x}_{he}^* - \mathbf{x}_c = \gamma(\mathbf{x}_h^* - \mathbf{x}_c),$$

that is,

$$\mathbf{x}_{he}^* = \gamma \mathbf{x}_h^* + (1 - \gamma) \mathbf{x}_c,$$

where $\gamma (> 1)$ is an expansion coefficient given by

$$\gamma = \frac{\|\mathbf{x}_{he}^* - \mathbf{x}_c\|_2}{\|\mathbf{x}_h^* - \mathbf{x}_c\|_2}$$

(see Figure 8.1). This operation is called expansion. If $f(\mathbf{x}_{he}^*) < f_l$, replace $\mathbf{x}_h$ by $\mathbf{x}_{he}^*$ and restart the process. However, if $f(\mathbf{x}_{he}^*) > f_l$, then expansion is counterproductive. In this case, $\mathbf{x}_{he}^*$ is dropped, $\mathbf{x}_h$ is replaced by $\mathbf{x}_h^*$, and the process is restarted.

- c. If upon reflecting $\mathbf{x}_h$ to $\mathbf{x}_h^*$ we discover that $f(\mathbf{x}_h^*) > f(\mathbf{x}_i)$ for all $i \neq h$, then replacing $\mathbf{x}_h$ by $\mathbf{x}_h^*$ would leave $f(\mathbf{x}_h^*)$ as the maximum in the new simplex. In this case, a new $\mathbf{x}_h$ is defined to be either the old $\mathbf{x}_h$ or $\mathbf{x}_h^*$, whichever has the lower value. A point $\mathbf{x}_{hc}^*$ is then found such that

$$\mathbf{x}_{hc}^* - \mathbf{x}_c = \beta(\mathbf{x}_h - \mathbf{x}_c),$$

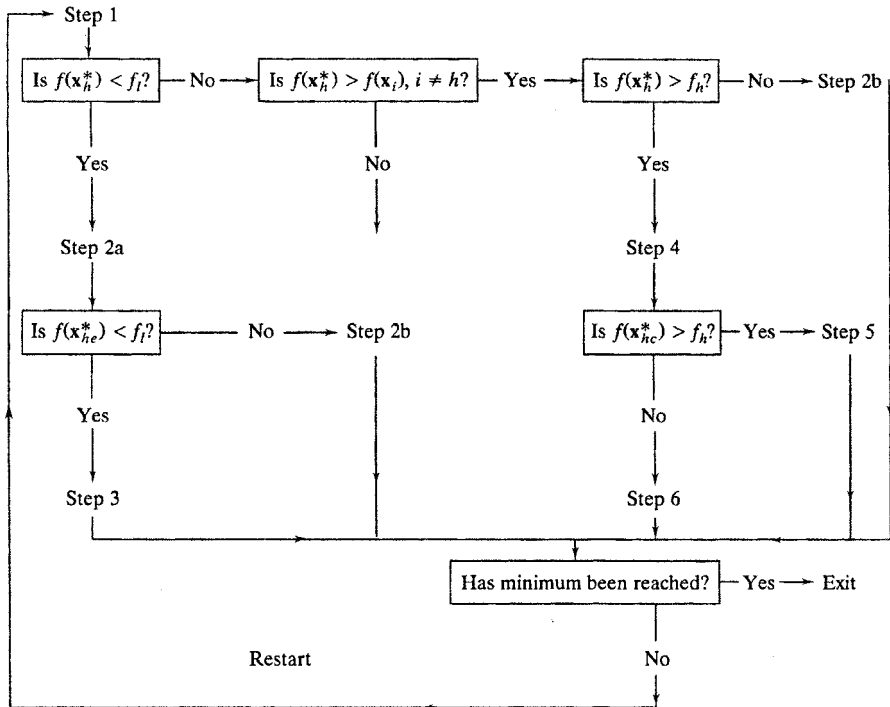
that is,

$$\mathbf{x}_{hc}^* = \beta \mathbf{x}_h + (1 - \beta) \mathbf{x}_c,$$

where β ($0 < \beta < 1$) is a contraction coefficient given by

$$\beta = \frac{\|\mathbf{x}_{hc}^* - \mathbf{x}_c\|_2}{\|\mathbf{x}_h - \mathbf{x}_c\|_2}.$$

Next, $\mathbf{x}_{hc}^*$ is substituted for $\mathbf{x}_h$ and the process is restarted unless $f(\mathbf{x}_{hc}^*) > \min[f_h, f(\mathbf{x}_h^*)]$, that is, the contracted point is worse than the better of f_h and $f(\mathbf{x}_h^*)$. When such a contraction fails, the size of the simplex is reduced by halving the distance of each point of the simplex from $\mathbf{x}_l$, where, if we recall, $\mathbf{x}_l$ is the point generating the lowest function value. Thus $\mathbf{x}_i$ is replaced by $\mathbf{x}_l + \frac{1}{2}(\mathbf{x}_i - \mathbf{x}_l)$, that is, by $\frac{1}{2}(\mathbf{x}_i + \mathbf{x}_l)$. The process is then restarted with the new reduced simplex.



- Step 1. Select initial points, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k+1}$, and calculate $f(\mathbf{x}_i)$, $i = 1, 2, \dots, k + 1$. Determine $\mathbf{x}_l$, $\mathbf{x}_h$ and calculate $\mathbf{x}_c = \sum_{i \neq h} \mathbf{x}_i / k$. Select $r > 0$, say $r = \frac{1}{2}, \frac{2}{3}$, or 1, find $\mathbf{x}_h^* = (1 + r)\mathbf{x}_c - r\mathbf{x}_h$, and calculate $f(\mathbf{x}_h^*)$.
- Step 2. (a) Calculate $\mathbf{x}_{hc}^* = \gamma \mathbf{x}_h^* + (1 - \gamma)\mathbf{x}_c$ by choosing $\gamma > 1$, say $\gamma = 1.5$, then calculate $f(\mathbf{x}_{hc}^*)$.
(b) Replace $\mathbf{x}_h$ with $\mathbf{x}_h^*$.
- Step 3. Replace $\mathbf{x}_h$ with $\mathbf{x}_{hc}^*$.
- Step 4. Calculate $\mathbf{x}_{hc}^* = \beta \mathbf{x}_h + (1 - \beta)\mathbf{x}_c$ by choosing $0 < \beta < 1$, say $\beta = 0.5$, then calculate $f(\mathbf{x}_{hc}^*)$.
- Step 5. Replace all the $\mathbf{x}_i$'s with $(\mathbf{x}_i + \mathbf{x}_l)/2$.
- Step 6. Replace $\mathbf{x}_h$ with $\mathbf{x}_{hc}^*$.

Figure 8.2. Flow diagram for the Nelder-Mead simplex method. *Source:* Nelder and Mead (1965). Reproduced with permission of Oxford University Press.

Thus at each stage in the minimization process, $\mathbf{x}_h$, the point at which f has the highest value, is replaced by a new point according to one of three operations, namely, reflection, contraction, and expansion. As an aid to illustrating this step-by-step procedure, a flow diagram is shown in Figure 8.2. This flow diagram is similar to one given by Nelder and Mead (1965, page 309). Figure 8.2 lists the explanations of steps 1 through 6.

The criterion used to stop the search procedure is based on the variation in the function values over the simplex. At each step, the

standard error of these values in the form

$$s = \left[\frac{\sum_{i=1}^{k+1} (f_i - \bar{f})^2}{k} \right]^{1/2}$$

is calculated and compared with some preselected value d , where $f_1, f_2, \dots, f_{k+1}$ denote the function values at the vertices of the simplex at hand and $\bar{f} = \sum_{i=1}^{k+1} f_i / (k+1)$. The search is halted when $s < d$. The reasoning behind this criterion is that when $s < d$, all function values are very close together. This hopefully indicates that the points of the simplex are near the minimum.

Bunday (1984) provided the listing of a computer program which can be used to implement the steps described in the flow diagram.

Olsson and Nelson (1975) demonstrated the usefulness of this method by using it to solve six minimization problems in statistics. The robustness of the method itself and its advantages relative to other minimization techniques were reported in Nelson (1973).

8.2.2. Price's Controlled Random Search Procedure

The controlled random search procedure was introduced by Price (1977). It is capable of finding the absolute (or global) minimum of a function within a constrained region R . It is therefore well suited for a multimodal function, that is, a function that has several local minima within the region R .

The essential features of Price's algorithm are outlined in the flow diagram of Figure 8.3. A predetermined number, N , of trial points are randomly chosen inside the region R . The value of N must be greater than k , the number of variables. The corresponding function values are obtained and stored in an array $\mathbf{A}$ along with the coordinates of the N chosen points. At each iteration, $k+1$ distinct points, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k+1}$, are chosen at random from the N points in storage. These $k+1$ points form a simplex in a k -dimensional space. The point $\mathbf{x}_{k+1}$ is arbitrarily taken as the pole (designated vertex) of the simplex, and the next trial point $\mathbf{x}_t$ is obtained as the image (reflection) point of the pole with respect to the centroid $\mathbf{x}_c$ of the remaining k points. Thus

$$\mathbf{x}_t = 2\mathbf{x}_c - \mathbf{x}_{k+1}.$$

The point $\mathbf{x}_t$ must satisfy the constraints of the region R . The value of the function f at $\mathbf{x}_t$ is then compared with $f_{\max}$, the largest function value in storage. Let $\mathbf{x}_{\max}$ denote the point at which $f_{\max}$ is achieved. If $f(\mathbf{x}_t) < f_{\max}$, then $\mathbf{x}_{\max}$ is replaced in the array $\mathbf{A}$ by $\mathbf{x}_t$. If $\mathbf{x}_t$ fails to satisfy the constraints of the region R , or if $f(\mathbf{x}_t) > f_{\max}$, then $\mathbf{x}_t$ is discarded and a new point is

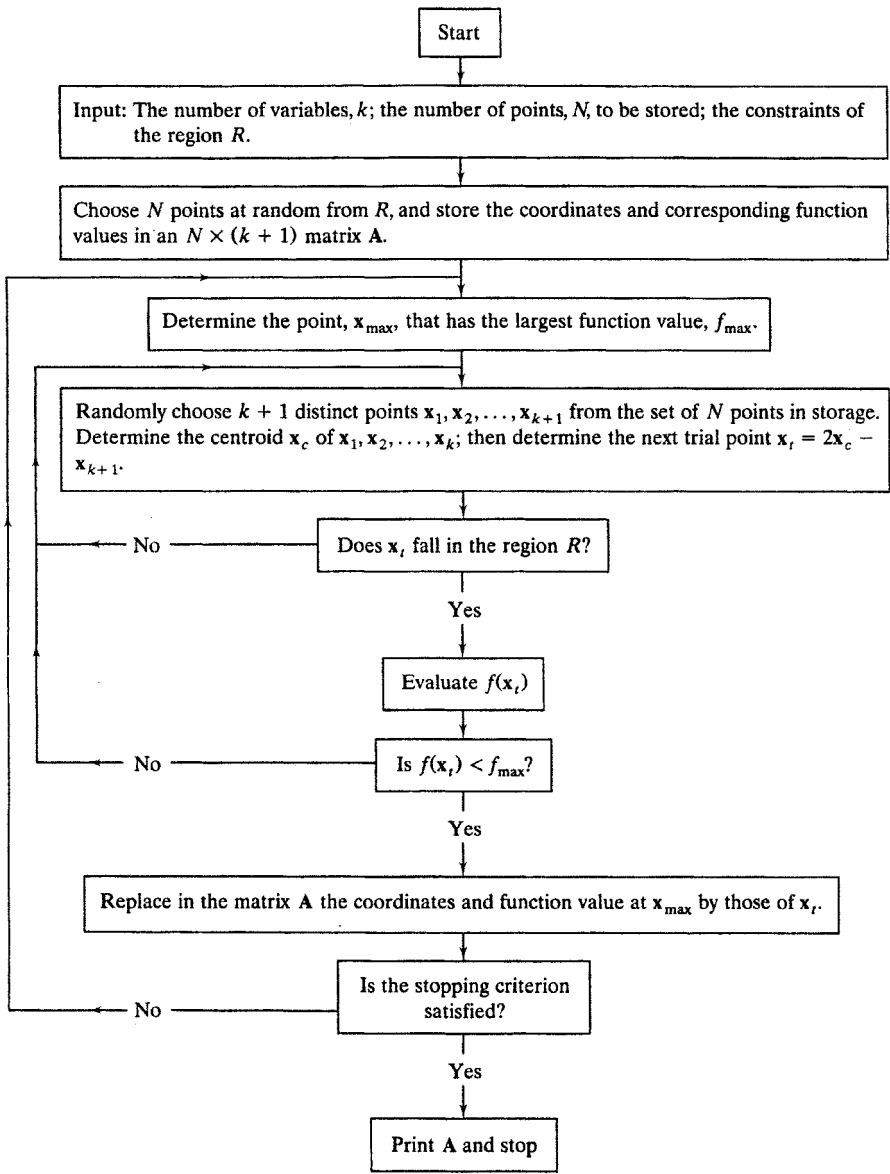


Figure 8.3. A flow diagram for Price's procedure. *Source:* Price (1977). Reproduced with permission of Oxford University Press.

chosen by following the same procedure as the one used to obtain $\mathbf{x}_i$. As the algorithm proceeds, the N points in storage tend to cluster around points at which the function values are lower than the current value of $f_{\max}$. Price did not specify a particular stopping rule. He left it to the user to do so. A possible stopping criterion is to terminate the search when the N points in storage cluster in a small region of the k -dimensional space, that is, when $f_{\max}$ and $f_{\min}$ are close together, where $f_{\min}$ is the smallest function value in storage. Another possibility is to stop after a specified number of function evaluations have been made. In any case, the rate of convergence of the procedure depends on the value of N , the complexity of the function f , the nature of the constraints, and the way in which the set of trial points is chosen.

Price's procedure is simple and does not necessarily require a large value of N . It is sufficient that N should increase linearly with k . Price chose, for example, the value $N = 50$ for $k = 2$. The value $N = 10k$ has proved useful for many functions. Furthermore, the region constraints can be quite complex. A FORTRAN program for the implementation of Price's algorithm was written by Conlon (1991).

8.2.3. The Generalized Simulated Annealing Method

This method derives its name from the annealing of metals, in which many final crystalline configurations (corresponding to different energy states) are possible, depending on the rate of cooling (see Kirkpatrick, Gelatt, and Vechhi, 1983). The method can be applied to find the absolute (or global) optimum of a multimodal function f within a constrained region R in a k -dimensional space.

Bohachevsky, Johnson, and Stein (1986) presented a generalization of the method of simulated annealing for function optimization. The following is a description of their algorithm for function minimization (a similar one can be used for function maximization): Let f_m be some tentative estimate of the minimum of f over the region R . The method proceeds according to the following steps (reproduced with permission of the American Statistical Association):

1. Select an initial point $\mathbf{x}_0$ in R . This point can be chosen at random or specified depending on available information.
2. Calculate $f_0 = f(\mathbf{x}_0)$. If $|f_0 - f_m| < \epsilon$, where ϵ is a specified small constant, then stop.
3. Choose a random direction of search by generating k independent standard normal variates $z_1, z_2, \dots, z_k$; then compute the elements of the random vector $\mathbf{u} = (u_1, u_2, \dots, u_k)'$, where

$$u_i = \frac{z_i}{(z_1^2 + z_2^2 + \dots + z_k^2)^{1/2}}, \quad i = 1, 2, \dots, k,$$

and k is the number of variables in the function.

4. Set $\mathbf{x}_1 = \mathbf{x}_0 + \Delta r \mathbf{u}$, where Δr is the size of a step to be taken in the direction of $\mathbf{u}$. The magnitude of Δr depends on the properties of the objective function and on the desired accuracy.
5. If $\mathbf{x}_1$ does not belong to R , return to step 3. Otherwise, compute $f_1 = f(\mathbf{x}_1)$ and $\Delta f = f_1 - f_0$.
6. if $f_1 \leq f_0$, set $\mathbf{x}_0 = \mathbf{x}_1$ and $f_0 = f_1$. If $|f_0 - f_m| < \epsilon$, stop. Otherwise, go to step 3.
7. If $f_1 > f_0$, set a probability value given by $p = \exp(-\beta f_0^g \Delta f)$, where β is a positive number such that $0.50 < \exp(-\beta \Delta f) < 0.90$, and g is an arbitrary negative number. Then, generate a random number v from the uniform distribution $U(0, 1)$. If $v \geq p$, go to step 3. Otherwise, if $v < p$, set $\mathbf{x}_0 = \mathbf{x}_1$, $f_0 = f_1$, and go to step 3.

From steps 6 and 7 we note that beneficial steps (that is, $f_1 \leq f_0$) are accepted unconditionally, but detrimental steps ($f_1 > f_0$) are accepted according to a probability value p described in step 7. If $v < p$, then the step leading to $\mathbf{x}_1$ is accepted; otherwise, it is rejected and a step in a new random direction is attempted. Thus the probability of accepting an increment of f depends on the size of the increment: the larger the increment, the smaller the probability of its acceptance.

Several possible values of the tentative estimate f_m can be attempted. For a given f_m we proceed with the search until $f - f_m$ becomes negative. Then, we decrease f_m , continue the search, and repeat the process when necessary. Bohachevsky, Johnson, and Stein gave an example in optimal design theory to illustrate the application of their algorithm.

Price's (1977) controlled random search algorithm produces results comparable to those of simulated annealing, but with fewer tuning parameters. It is also better suited for problems with constrained regions.

8.3. OPTIMIZATION TECHNIQUES IN RESPONSE SURFACE METHODOLOGY

Response surface methodology (RSM) is an area in the design and analysis of experiments. It consists of a collection of techniques that encompasses:

1. Conducting a series of experiments based on properly chosen settings of a set of input variables, denoted by $x_1, x_2, \dots, x_k$, that influence a response of interest y . The choice of these settings is governed by certain criteria whose purpose is to produce adequate and reliable information about the response. The collection of all such settings constitutes a matrix $\mathbf{D}$ of order $n \times k$, where n is the number of experimental runs. The matrix $\mathbf{D}$ is referred to as a response surface design.

2. Determining a mathematical model that best fits the data collected under the design chosen in (1). Regression techniques can be used to evaluate the adequacy of fit of the model and to conduct appropriate tests concerning the model's parameters.
3. Determining optimal operating conditions on the input variables that produce maximum (or minimum) response value within a region of interest R .

This last aspect of RSM can help the experimenter in determining the best combinations of the input variables that lead to desirable response values. For example, in drug manufacturing, two drugs are tested with regard to reducing blood pressure in humans. A series of clinical trials involving a certain number of high blood pressure patients is set up, and each patient is given some predetermined combination of the two drugs. After a period of time the patient's blood pressure is checked. This information can be used to find the specific combination of the drugs that results in the greatest reduction in the patient's blood pressure within some specified time interval.

In this section we shall describe two well-known optimum-seeking procedures in RSM. These include the method of steepest ascent (or descent) and ridge analysis.

8.3.1. The Method of Steepest Ascent

This is an adaptation of the method described in Section 8.1.1 to a response surface environment; here the objective is to increase the value of a certain response function.

The method of steepest ascent requires performing a sequence of sets of trials. Each set is obtained as a result of proceeding sequentially along a path of maximum increase in the values of a given response y , which can be observed in an experiment. This method was first introduced by Box and Wilson (1951) for the general area of RSM.

The procedure of steepest ascent depends on approximating a response surface with a hyperplane in some restricted region. The hyperplane is represented by a first-order model which can be fitted to a data set obtained as a result of running experimental trials using a first-order design such as a complete 2^k factorial design, where k is the number of input variables in the model. A fraction of this design can also be used if k is large [see, for example, Section 3.3.2 in Khuri and Cornell (1996)]. The fitted first-order model is then used to determine a path along which one may initially observe increasing response values. However, due to curvature in the response surface, the initial increase in the response will likely be followed by a leveling off, and then a decrease. At this stage, a new series of experiments is performed (using again a first-order design) and the resulting data are used to fit another first-order model. A new path is determined along which

increasing response values may be observed. This process continues until it becomes evident that little or no additional increase in the response can be gained.

Let us now consider more specific details of this sequential procedure. Let $y(\mathbf{x})$ be a response function that depends on k input variables, $x_1, x_2, \dots, x_k$, which form the elements of a vector $\mathbf{x}$. Suppose that in some restricted region $y(\mathbf{x})$ is adequately represented by a first-order model of the form

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \epsilon, \quad (8.8)$$

where $\beta_0, \beta_1, \dots, \beta_k$ are unknown parameters and ϵ is a random error. This model is fitted using data collected under a first-order design (for example, a 2^k factorial design or a fraction thereof). The data are utilized to calculate the least-squares estimates $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ of the model's parameters. These are elements of $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, where $\mathbf{X} = [\mathbf{1}_n : \mathbf{D}]$ with $\mathbf{1}_n$ being a vector of ones of order $n \times 1$, $\mathbf{D}$ is the design matrix of order $n \times k$, and $\mathbf{y}$ is the corresponding vector of response values. It is assumed that the random errors associated with the n response values are independently distributed with means equal to zero and a common variance σ^2 . The predicted response $\hat{y}(\mathbf{x})$ is then given by

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i. \quad (8.9)$$

The input variables are coded so that the design center coincides with the origin of the coordinates system.

The next step is to move a distance of r units away from the design center (or the origin) such that a maximum increase in $\hat{y}$ can be obtained. To determine the direction to be followed to achieve such an increase, we need to maximize $\hat{y}(\mathbf{x})$ subject to the constraint $\sum_{i=1}^k x_i^2 = r^2$ using the method of Lagrange multipliers. Consider therefore the function

$$Q(\mathbf{x}) = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i - \lambda \left(\sum_{i=1}^k x_i^2 - r^2 \right), \quad (8.10)$$

where λ is a Lagrange multiplier. Setting the partial derivatives of Q equal to zero produces the equations

$$x_i = \frac{1}{2\lambda} \hat{\beta}_i, \quad i = 1, 2, \dots, k.$$

For a maximum, λ must be positive. Using the equality constraint, we conclude that

$$\lambda = \frac{1}{2r} \left(\sum_{i=1}^k \hat{\beta}_i^2 \right)^{1/2}.$$

A local maximum is then achieved at the point whose coordinates are given by

$$x_i = \frac{r\hat{\beta}_i}{\left(\sum_{i=1}^k \hat{\beta}_i^2\right)^{1/2}}, \quad i = 1, 2, \dots, k,$$

which can be written as

$$x_i = re_i, \quad i = 1, 2, \dots, k, \quad (8.11)$$

where $e_i = \hat{\beta}_i(\sum_{i=1}^k \hat{\beta}_i^2)^{-1/2}$, $i = 1, 2, \dots, k$. Thus $\mathbf{e} = (e_1, e_2, \dots, e_k)'$ is a unit vector in the direction of $(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)'$. Equations (8.11) indicate that at a distance of r units away from the origin, a maximum increase in $\hat{y}$ occurs along a path in the direction of $\mathbf{e}$. Since this is the only local maximum on the hypersphere of radius r , it must be the absolute maximum.

If the actual response value (that is, the value of y) at the point $\mathbf{x} = r\mathbf{e}$ exceeds its value at the origin, then a move along the path determined by $\mathbf{e}$ is in order. A series of experiments is then conducted to obtain response values at several points along the path until no additional increase in the response is evident. At this stage, a new first-order model is fitted using data collected under a first-order design centered at a point in the vicinity of the point at which that first drop in the response was observed along the path. This model leads to a new direction similar to the one given by formula (8.11). As before, a series of experiments are conducted along the new path until no further increase in the value of y can be observed. The process of moving along different paths continues until it becomes evident that little or no additional increase in y can be gained. This usually occurs when the first-order model becomes inadequate as the method progresses, due to curvature in the response surface. It is therefore necessary to test each fitted first-order model for lack of fit at every stage of the process. This can be accomplished by taking repeated observations at the center of each first-order design and at possibly some other design points in order to obtain an independent estimate of the error variance that is needed for the lack of fit test [see, for example, Sections 2.6 and 3.4 in Khuri and Cornell (1996)]. If the lack of fit test is significant, indicating an inadequate model, then the process is stopped and a more elaborate experiment must be conducted to fit a higher-order model, as will be seen in the next section.

Examples that illustrate the application of the method of steepest ascent can be found in Box and Wilson (1951), Bayne and Rubin (1986, Section 5.2), Khuri and Cornell (1996, Chapter 5), and Myers and Khuri (1979). In the last reference, the authors present a stopping rule along a path that takes into account random error variation in the observed response. We recall that a search along a path is discontinued as soon as a drop in the response is first

observed. Since response values are subject to random error, the decision to stop can be premature due to a false drop in the observed response. The stopping rule by Myers and Khuri (1979) protects against taking too many observations along a path when in fact the true mean response (that is, the mean of y) is decreasing. It also protects against stopping prematurely when the true mean response is increasing.

It should be noted that the procedure of steepest ascent is not invariant with respect to the scales of the input variables $x_1, x_2, \dots, x_k$. This is evident from the fact that a path taken by the procedure is determined by the least-squares estimates $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ [see equations (8.11)], which depend on the scales of the x_i 's.

There are situations in which it is of interest to determine conditions that lead to a decrease in the response, instead of an increase. For example, in a chemical investigation it may be desired to decrease the level of impurity or the unit cost. In this case, a path of steepest descent will be needed. This can be accomplished by changing the sign of the response y , followed by an application of the method of steepest ascent. Thus any steepest descent problem can be handled by the method of steepest ascent.

8.3.2. The Method of Ridge Analysis

The method of steepest ascent is most often used as a maximum-region-seeking procedure. By this we mean that it is used as a preliminary tool to get quickly to the region where the maximum of the mean response is located. Since the first-order approximation of the mean response will eventually break down, a better estimate of the maximum can be obtained by fitting a second-order model in the region of the maximum. The method of ridge analysis, which was introduced by Hoerl (1959) and formalized by Draper (1963), is used for this purpose.

Let us suppose that inside a region of interest R , the true mean response is adequately represented by the second-order model

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^{k-1} \sum_{\substack{j=2 \\ i < j}}^k \beta_{ij} x_i x_j + \sum_{i=1}^k \beta_{ii} x_i^2 + \epsilon, \quad (8.12)$$

where the β 's are unknown parameters and ϵ is a random error with mean zero and variance σ^2 . Model (8.12) can be written as

$$y(\mathbf{x}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \epsilon, \quad (8.13)$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$ and $\mathbf{B}$ is a symmetric $k \times k$ matrix of the form

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \frac{1}{2}\beta_{13} & \cdots & \frac{1}{2}\beta_{1k} \\ & \beta_{22} & \frac{1}{2}\beta_{23} & \cdots & \frac{1}{2}\beta_{2k} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \frac{1}{2}\beta_{k-1,k} \\ \text{symmetric} & & & & \beta_{kk} \end{bmatrix}.$$

Least-squares estimates of the parameters in model (8.13) can be obtained by using data collected according to a second-order design. A description of potential second-order designs can be found in Khuri and Cornell (1996, Chapter 4).

Let $\hat{\beta}_0$, $\hat{\boldsymbol{\beta}}$, and $\hat{\mathbf{B}}$ denote the least-squares estimates of β_0 , $\boldsymbol{\beta}$, and $\mathbf{B}$, respectively. The predicted response $\hat{y}(\mathbf{x})$ inside the region R is then given by

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0 + \mathbf{x}'\hat{\boldsymbol{\beta}} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x}. \quad (8.14)$$

The input variables are coded so that the design center coincides with the origin of the coordinates system.

The method of ridge analysis is used to find the optimum (maximum or minimum) of $\hat{y}(\mathbf{x})$ on concentric hyperspheres of varying radii inside the region R . It is particularly useful in situations in which the unconstrained optimum of $\hat{y}(\mathbf{x})$ falls outside the region R , or if a saddle point occurs inside R .

Let us now proceed to optimize $\hat{y}(\mathbf{x})$ subject to the constraint

$$\sum_{i=1}^k x_i^2 = r^2, \quad (8.15)$$

where r is the radius of a hypersphere centered at the origin and is contained inside the region R . Using the method of Lagrange multipliers, let us consider the function

$$F = \hat{y}(\mathbf{x}) - \lambda \left(\sum_{i=1}^k x_i^2 - r^2 \right), \quad (8.16)$$

where λ is a Lagrange multiplier. Differentiating F with respect to x_i

($i = 1, 2, \dots, k$) and equating the partial derivatives to zero, we obtain

$$\begin{aligned}\frac{\partial F}{\partial x_1} &= 2(\hat{\beta}_{11} - \lambda)x_1 + \hat{\beta}_{12}x_2 + \dots + \hat{\beta}_{1k}x_k + \hat{\beta}_1 = 0, \\ \frac{\partial F}{\partial x_2} &= \hat{\beta}_{12}x_1 + 2(\hat{\beta}_{22} - \lambda)x_2 + \dots + \hat{\beta}_{2k}x_k + \hat{\beta}_2 = 0, \\ &\vdots \\ \frac{\partial F}{\partial x_k} &= \hat{\beta}_{1k}x_1 + \hat{\beta}_{2k}x_2 + \dots + 2(\hat{\beta}_{kk} - \lambda)x_k + \hat{\beta}_k = 0.\end{aligned}$$

These equations can be expressed as

$$(\hat{\mathbf{B}} - \lambda \mathbf{I}_k) \mathbf{x} = -\frac{1}{2} \hat{\boldsymbol{\beta}}. \quad (8.17)$$

Equations (8.15) and (8.17) need to be solved for $x_1, x_2, \dots, x_k$ and λ . This traditional approach, however, requires calculations that are somewhat involved. Draper (1963) proposed the following simpler, yet equivalent procedure:

- i. Regard r as a variable, but fix λ instead.
- ii. Insert the selected value of λ in equation (8.17) and solve for $\mathbf{x}$. The solution is used in steps iii and iv.
- iii. Compute $r = (\mathbf{x}'\mathbf{x})^{1/2}$.
- iv. Evaluate $\hat{y}(\mathbf{x})$.

Several values of λ can give rise to several stationary points which lie on the same hypersphere of radius r . This can be seen from the fact that if λ is chosen to be different from any eigenvalue of $\hat{\mathbf{B}}$, then equation (8.17) has a unique solution given by

$$\mathbf{x} = -\frac{1}{2}(\hat{\mathbf{B}} - \lambda \mathbf{I}_k)^{-1} \hat{\boldsymbol{\beta}}. \quad (8.18)$$

By substituting $\mathbf{x}$ in equation (8.15) we obtain

$$\hat{\boldsymbol{\beta}}'(\hat{\mathbf{B}} - \lambda \mathbf{I}_k)^{-2} \hat{\boldsymbol{\beta}} = 4r^2. \quad (8.19)$$

Hence, each value of r gives rise to at most $2k$ corresponding values of λ .

The choice of λ has an effect on the nature of the stationary point. Some values of λ produce points at each of which $\hat{y}$ has a maximum. Other values of λ cause $\hat{y}$ to have minimum values. More specifically, suppose that λ_1 and

λ_2 are two values substituted for λ in equation (8.18). Let $\mathbf{x}_1, \mathbf{x}_2$ and r_1, r_2 be the corresponding values of $\mathbf{x}$ and r , respectively. The following results, which were established in Draper (1963), can be helpful in selecting the value of λ that produces a particular type of stationary point:

RESULT 1. If $r_1 = r_2$ and $\lambda_1 > \lambda_2$, then $\hat{y}_1 > \hat{y}_2$, where $\hat{y}_1$ and $\hat{y}_2$ are the values of $\hat{y}(\mathbf{x})$ at $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively.

This result means that for two stationary points that have the same distance from the origin, $\hat{y}$ will be larger at the stationary point with the larger value of λ .

RESULT 2. Let $\mathbf{M}$ be the matrix of second-order partial derivatives of F in formula (8.16), that is,

$$\mathbf{M} = 2(\hat{\mathbf{B}} - \lambda \mathbf{I}_k). \quad (8.20)$$

If $r_1 = r_2$, and if $\mathbf{M}$ is positive definite for $\mathbf{x}_1$ and is indefinite (that is, neither positive definite nor negative definite) for $\mathbf{x}_2$, then $\hat{y}_1 < \hat{y}_2$.

RESULT 3. If λ_1 is larger than the largest eigenvalue of $\hat{\mathbf{B}}$, then the corresponding solution $\mathbf{x}_1$ in formula (8.18) is a point of absolute maximum for $\hat{y}$ on a hypersphere of radius $r_1 = (\mathbf{x}'_1 \mathbf{x}_1)^{1/2}$. If, on the other hand, λ_1 is smaller than the smallest eigenvalue of $\hat{\mathbf{B}}$, then $\mathbf{x}_1$ is a point of absolute minimum for $\hat{y}$ on the same hypersphere.

On the basis of Result 3 we can select several values of λ that exceed the largest eigenvalue of $\hat{\mathbf{B}}$. The resulting values of the k elements of $\mathbf{x}$ and $\hat{y}$ can be plotted against the corresponding values of r . This produces $k + 1$ plots called ridge plots (see Myers, 1976, Section 5.3). They are useful in that an experimenter can determine, for a particular r , the maximum of $\hat{y}$ within a region R and the operating conditions (that is, the elements of $\mathbf{x}$) that give rise to the maximum. Similar plots can be obtained for the minimum of $\hat{y}$ (here, values of λ that are smaller than the smallest eigenvalue of $\hat{\mathbf{B}}$ must be chosen). Obviously, the portions of the ridge plots that fall outside R should not be considered.

EXAMPLE 8.3.1. An experiment was conducted to investigate the effects of three fertilizer ingredients on the yield of snap beans under field conditions. The fertilizer ingredients and actual amounts applied were nitrogen (N), from 0.94 to 6.29 lb/plot; phosphoric acid (P_2O_5), from 0.59 to 2.97 lb/plot; and potash (K_2O), from 0.60 to 4.22 lb/plot. The response of interest is the average yield in pounds per plot of snap beans.

Five levels of each fertilizer were used. The levels are coded using the following linear transformations:

$$x_1 = \frac{X_1 - 3.62}{1.59}, \quad x_2 = \frac{X_2 - 1.78}{0.71}, \quad x_3 = \frac{X_3 - 2.42}{1.07}.$$

Here, X_1 , X_2 , and X_3 denote the actual levels of nitrogen, phosphoric acid, and potash, respectively, used in the experiment, and x_1, x_2, x_3 the corresponding coded values. In this particular coding scheme, 3.62, 1.78, and 2.42 are the averages of the experimental levels of X_1 , X_2 , and X_3 , respectively, that is, they represent the centers of the values of nitrogen, phosphoric acid, and potash, respectively. The denominators of x_1 , x_2 , and x_3 were chosen so that the second and fourth levels of each X_i correspond to the values -1 and 1 , respectively, for x_i ($i = 1, 2, 3$). One advantage of such a coding scheme is to make the levels of the three fertilizers scale free (this is necessary in general, since the input variables can have different units of measurement). The measured and coded levels for the three fertilizers are shown below:

Fertilizer	Levels of x_i ($i = 1, 2, 3$)				
	-1.682	-1.000	0.000	1.000	1.682
N	0.94	2.03	3.62	5.21	6.29
P ₂ O ₅	0.59	1.07	1.78	2.49	2.97
K ₂ O	0.60	1.35	2.42	3.49	4.22

Combinations of the levels of the three fertilizers were applied according to the experimental design shown in Table 8.1, in which the design settings are given in terms of the coded levels. Six center-point replications were run in order to obtain an estimate of the experimental error variance. This particular design is called a central composite design [for a description of this design and its properties, see Khuri and Cornell (1996, Section 4.5.3)], which has the rotatability property. By this we mean that the prediction variance, that is, $\text{Var}[\hat{y}(\mathbf{x})]$, is constant at all points that are equidistant from the design center [see Khuri and Cornell (1996, Section 2.8.3) for more detailed information concerning rotatability]. The corresponding response (yield) values are given in Table 8.1.

A second-order model of the form given by formula (8.12) was fitted to the data set in Table 8.1. Thus in terms of the coded variables we have the model

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^3 \beta_i x_i + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \sum_{i=1}^3 \beta_{ii} x_i^2 + \epsilon. \quad (8.21)$$

Table 8.1. The Coded and Actual Settings of the Three Fertilizers and the Corresponding Response Values

x_1	x_2	x_3	N	P_2O_5	K_2O	Yield y
-1	-1	-1	2.03	1.07	1.35	11.28
1	-1	-1	5.21	1.07	1.35	8.44
-1	1	-1	2.03	2.49	1.35	13.19
1	1	-1	5.21	2.49	1.35	7.71
-1	-1	1	2.03	1.07	3.49	8.94
1	-1	1	5.21	1.07	3.49	10.90
-1	1	1	2.03	2.49	3.49	11.85
1	1	1	5.21	2.49	3.49	11.03
-1.682	0	0	0.94	1.78	2.42	8.26
1.682	0	0	6.29	1.78	2.42	7.87
0	-1.682	0	3.62	0.59	2.42	12.08
0	1.682	0	3.62	2.97	2.42	11.06
0	0	-1.682	3.62	1.78	0.60	7.98
0	0	1.682	3.62	1.78	4.22	10.43
0	0	0	3.62	1.78	2.42	10.14
0	0	0	3.62	1.78	2.42	10.22
0	0	0	3.62	1.78	2.42	10.53
0	0	0	3.62	1.78	2.42	9.50
0	0	0	3.62	1.78	2.42	11.53
0	0	0	3.62	1.78	2.42	11.02

Source: A. I. Khuri and J. A. Cornell (1996). Reproduced with permission of Marcel Dekker, Inc.

The resulting prediction equation is given by

$$\begin{aligned}\hat{y}(\mathbf{x}) = & 10.462 - 0.574x_1 + 0.183x_2 + 0.456x_3 - 0.678x_1x_2 + 1.183x_1x_3 \\ & + 0.233x_2x_3 - 0.676x_1^2 + 0.563x_2^2 - 0.273x_3^2.\end{aligned}\quad (8.22)$$

Here, $\hat{y}(\mathbf{x})$ is the predicted yield at the point $\mathbf{x} = (x_1, x_2, x_3)'$. Equation (8.22) can be expressed in matrix form as in equation (8.14), where $\hat{\boldsymbol{\beta}} = (-0.574, 0.183, 0.456)'$ and $\hat{\mathbf{B}}$ is the matrix

$$\hat{\mathbf{B}} = \begin{bmatrix} -0.676 & -0.339 & 0.592 \\ -0.339 & 0.563 & 0.117 \\ 0.592 & 0.117 & -0.273 \end{bmatrix}.\quad (8.23)$$

The coordinates of the stationary point $\mathbf{x}_0$ of $\hat{y}(\mathbf{x})$ satisfy the equation

$$\begin{aligned}\frac{\partial \hat{y}}{\partial x_1} &= -0.574 + 2(-0.676x_1 - 0.339x_2 + 0.592x_3) = 0, \\ \frac{\partial \hat{y}}{\partial x_2} &= 0.183 + 2(-0.339x_1 + 0.563x_2 + 0.117x_3) = 0, \\ \frac{\partial \hat{y}}{\partial x_3} &= 0.456 + 2(0.592x_1 + 0.117x_2 - 0.273x_3) = 0,\end{aligned}$$

which can be expressed as

$$\hat{\boldsymbol{\beta}} + 2\hat{\mathbf{B}}\mathbf{x}_0 = \mathbf{0}.\tag{8.24}$$

Hence,

$$\mathbf{x}_0 = -\frac{1}{2}\hat{\mathbf{B}}^{-1}\hat{\boldsymbol{\beta}} = (-0.394, -0.364, -0.175)'.$$

The eigenvalues of $\hat{\mathbf{B}}$ are $\tau_1 = 0.6508$, $\tau_2 = 0.1298$, $\tau_3 = -1.1678$. The matrix $\hat{\mathbf{B}}$ is therefore neither positive definite nor negative definite, that is, $\mathbf{x}_0$ is a saddle point (see Corollary 7.7.1). This point falls inside the experimental region R , which, in the space of the coded variables x_1, x_2, x_3 , is a sphere centered at the origin of radius $\sqrt{3}$.

Let us now apply the method of ridge analysis to maximize $\hat{y}$ inside the region R . For this purpose we choose values of λ [the Lagrange multiplier in equation (8.16)] larger than $\tau_1 = 0.6508$, the largest eigenvalue of $\hat{\mathbf{B}}$. For each such value of λ , equation (8.17) has a solution for $\mathbf{x}$ that represents a point of absolute maximum of $\hat{y}(\mathbf{x})$ on a sphere of radius $r = (\mathbf{x}'\mathbf{x})^{1/2}$ inside R . The results are displayed in Table 8.2. We note that at the point $(-0.558, 1.640, 0.087)$, which is located near the periphery of the region R , the maximum value of $\hat{y}$ is 13.021. By expressing the coordinates of this point in terms of the actual values of the three fertilizers we obtain $X_1 = 2.733$ lb/plot, $X_2 = 2.944$ lb/plot, and $X_3 = 2.513$ lb/plot. We conclude that a combination of nitrogen, phosphoric acid, and potash fertilizers at the rates

Table 8.2. Ridge Analysis Values

λ	1.906	1.166	0.979	0.889	0.840	0.808	0.784	0.770	0.754	0.745	0.740
x_1	-0.106	-0.170	-0.221	-0.269	-0.316	-0.362	-0.408	-0.453	-0.499	-0.544	-0.558
x_2	0.102	0.269	0.438	0.605	0.771	0.935	1.099	1.263	1.426	1.589	1.640
x_3	0.081	0.110	0.118	0.120	0.117	0.113	0.108	0.102	0.096	0.089	0.087
r	0.168	0.337	0.505	0.673	0.841	1.009	1.177	1.346	1.514	1.682	1.734
$\hat{y}$	10.575	10.693	10.841	11.024	11.243	11.499	11.790	12.119	12.484	12.886	13.021

of 2.733, 2.944, and 2.513 lb/plot, respectively, results in an estimated maximum yield of snap beans of 13.021 lb/plot.

8.3.3. Modified Ridge Analysis

Optimization of $\hat{y}(\mathbf{x})$ on a hypersphere S by the method of ridge analysis is justified provided that the prediction variance on S is relatively small. Furthermore, it is desirable that this variance remain constant on S . If not, then it is possible to obtain poor estimates of the optimum response, especially when the dispersion in the prediction variances on S is large. Thus the reliability of ridge analysis as an optimum-seeking procedure depends very much on controlling the size and variability of the prediction variance. If the design is rotatable, then the prediction variance, $\text{Var}[\hat{y}(\mathbf{x})]$, is constant on S . It is then easy to attain small prediction variances by restricting the procedure to hyperspheres of small radii. However, if the design is not rotatable, then $\text{Var}[\hat{y}(\mathbf{x})]$ may vary widely on S , which, as was mentioned earlier, can adversely affect the quality of estimation of the optimum response. This suggests that the prediction variance should be given serious consideration in the strategy of ridge analysis if the design used is not rotatable.

Khuri and Myers (1979) proposed a certain modification to the method of ridge analysis: one that optimizes $\hat{y}(\mathbf{x})$ subject to a particular constraint on the prediction variance. The following is a description of their proposed modification:

Consider model (8.12), which can be written as

$$y(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\gamma} + \epsilon, \quad (8.25)$$

where

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= (1, x_1, x_2, \dots, x_k, x_1x_2, x_1x_3, \dots, x_{k-1}x_k, x_1^2, x_2^2, \dots, x_k^2)', \\ \boldsymbol{\gamma} &= (\beta_0, \beta_1, \beta_2, \dots, \beta_k, \beta_{12}, \beta_{13}, \dots, \beta_{k-1,k}, \beta_{11}, \beta_{22}, \dots, \beta_{kk})'. \end{aligned}$$

The predicted response is given by

$$\hat{y}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\hat{\boldsymbol{\gamma}}, \quad (8.26)$$

where $\hat{\boldsymbol{\gamma}}$ is the least-squares estimator of $\boldsymbol{\gamma}$, namely,

$$\hat{\boldsymbol{\gamma}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \quad (8.27)$$

where $\mathbf{X} = [\mathbf{f}(\mathbf{x}_1), \mathbf{f}(\mathbf{x}_2), \dots, \mathbf{f}(\mathbf{x}_n)]'$ with $\mathbf{x}_i$ being the vector of design settings at the i th experimental run ($i = 1, 2, \dots, n$, where n is the number of runs used in the experiment), and $\mathbf{y}$ is the corresponding vector of n observations. Since

$$\text{Var}(\hat{\boldsymbol{\gamma}}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2, \quad (8.28)$$

where σ^2 is the error variance, then from equation (8.26), the prediction variance is of the form

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1} \mathbf{f}(\mathbf{x}). \quad (8.29)$$

The number of unknown parameters in model (8.25) is $p = (k+1)(k+2)/2$, where k is the number of input variables. Let $\nu_1, \nu_2, \dots, \nu_p$ denote the eigenvalues of $\mathbf{X}'\mathbf{X}$. Then from equation (8.29) and Theorem 2.3.16 we have

$$\frac{\sigma^2 \mathbf{f}'(\mathbf{x}) \mathbf{f}(\mathbf{x})}{\nu_{\max}} \leq \text{Var}[\hat{y}(\mathbf{x})] \leq \frac{\sigma^2 \mathbf{f}'(\mathbf{x}) \mathbf{f}(\mathbf{x})}{\nu_{\min}},$$

where $\nu_{\min}$ and $\nu_{\max}$ are, respectively, the smallest and largest of the ν_i 's. This double inequality shows that the prediction variance can be inflated if $\mathbf{X}'\mathbf{X}$ has small eigenvalues. This occurs when the columns of $\mathbf{X}$ are multicollinear (see, for example, Myers, 1990, pages 125–126 and Chapter 8). Now, by the spectral decomposition theorem (Theorem 2.3.10), $\mathbf{X}'\mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$, where $\mathbf{V}$ is an orthogonal matrix of orthonormal eigenvectors of $\mathbf{X}'\mathbf{X}$ and $\mathbf{\Lambda} = \text{Diag}(\nu_1, \nu_2, \dots, \nu_p)$ is a diagonal matrix of eigenvalues of $\mathbf{X}'\mathbf{X}$. Equation (8.29) can then be written as

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \sum_{j=1}^p \frac{[\mathbf{f}'(\mathbf{x}) \mathbf{v}_j]^2}{\nu_j}, \quad (8.30)$$

where $\mathbf{v}_j$ is the j th column of $\mathbf{V}$ ($j = 1, 2, \dots, p$). If we denote the elements of $\mathbf{v}_j$ by $v_{0j}, v_{1j}, \dots, v_{kj}, v_{12j}, v_{13j}, \dots, v_{k-1kj}, v_{11j}, v_{22j}, \dots, v_{kkj}$, then $\mathbf{f}'(\mathbf{x}) \mathbf{v}_j$ can be expressed as

$$\mathbf{f}'(\mathbf{x}) \mathbf{v}_j = v_{0j} + \mathbf{x}' \boldsymbol{\tau}_j + \mathbf{x}' \mathbf{T}_j \mathbf{x}, \quad j = 1, 2, \dots, p, \quad (8.31)$$

where $\boldsymbol{\tau}_j = (v_{1j}, v_{2j}, \dots, v_{kj})'$ and

$$\mathbf{T}_j = \begin{bmatrix} v_{11j} & \frac{1}{2}v_{12j} & \frac{1}{2}v_{13j} & \cdots & \frac{1}{2}v_{1kj} \\ & v_{22j} & \frac{1}{2}v_{23j} & \cdots & \frac{1}{2}v_{2kj} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \frac{1}{2}v_{k-1kj} \\ \text{symmetric} & & & & v_{kkj} \end{bmatrix}, \quad j = 1, 2, \dots, p.$$

We note that the form of $\mathbf{f}'(\mathbf{x}) \mathbf{v}_j$, as given by formula (8.31), is identical to that of a second-order model. Formula (8.30) can then be written as

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \sum_{j=1}^p \frac{(v_{0j} + \mathbf{x}' \boldsymbol{\tau}_j + \mathbf{x}' \mathbf{T}_j \mathbf{x})^2}{\nu_j}. \quad (8.32)$$

As was noted earlier, small values of ν_j ($j = 1, 2, \dots, p$) cause $\hat{y}(\mathbf{x})$ to have large variances.

To reduce the size of the prediction variance within the region explored by ridge analysis, we can consider putting constraints on the portion of $\text{Var}[\hat{y}(\mathbf{x})]$ that corresponds to $\nu_{\min}$. It makes sense to optimize $\hat{y}(\mathbf{x})$ subject to the constraints

$$\mathbf{x}'\mathbf{x} = r^2, \quad (8.33)$$

$$|v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}| \leq q, \quad (8.34)$$

where v_{0m} , $\boldsymbol{\tau}_m$, and $\mathbf{T}_m$ are the values of v_0 , $\boldsymbol{\tau}$, and $\mathbf{T}$ that correspond to $\nu_{\min}$. Here, q is a positive constant chosen small enough to offset the small value of $\nu_{\min}$. Khuri and Myers (1979) suggested that q be equal to the largest value taken by $|v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}|$ at the n design points. The rationale behind this rule of thumb is that the prediction variance is smaller at the design points than at other points in the experimental region.

The modification suggested by Khuri and Myers (1979) amounts to adding the constraint (8.34) to the usual procedure of ridge analysis. In this way, some control can be maintained on the size of prediction variance during the optimization process. The mathematical algorithm needed for this constrained optimization is based on a technique introduced by Myers and Carter (1973) for a dual response system in which a *primary* second-order response function is optimized subject to the condition that a *constrained* second-order response function takes on some specified or desirable values. Here, the primary response is $\hat{y}(\mathbf{x})$ and the constrained response is $v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}$.

Myers and Carter's (1973) procedure is based on the method of Lagrange multipliers, which uses the function

$$L = \hat{\beta}_0 + \mathbf{x}'\hat{\boldsymbol{\beta}} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x} - \mu(v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x} - \omega) - \lambda(\mathbf{x}'\mathbf{x} - r^2),$$

where $\hat{\beta}_0$, $\hat{\boldsymbol{\beta}}$, and $\hat{\mathbf{B}}$ are the same as in model (8.14), μ and λ are Lagrange multipliers, ω is such that $|\omega| \leq q$ [see inequality (8.34)], and r is the radius of a hypersphere centered at the origin and contained inside a region of interest R . By differentiating L with respect to $x_1, x_2, \dots, x_k$ and equating the derivatives to zero, we obtain

$$(\hat{\mathbf{B}} - \mu\mathbf{T}_m - \lambda\mathbf{I}_k)\mathbf{x} = \frac{1}{2}(\mu\boldsymbol{\tau}_m - \hat{\boldsymbol{\beta}}). \quad (8.35)$$

As in the method of ridge analysis, to solve equation (8.35), values of μ and λ are chosen directly in such a way that the solution represents a point of maximum (or minimum) for $\hat{y}(\mathbf{x})$. Thus for a given value of μ , the matrix of second-order partial derivatives of L , namely $2(\hat{\mathbf{B}} - \mu\mathbf{T}_m - \lambda\mathbf{I}_k)$, is made negative definite [and hence a maximum of $\hat{y}(\mathbf{x})$ is achieved] by selecting λ

larger than the largest eigenvalue of $\hat{\mathbf{B}} - \mu \mathbf{T}_m$. Values of λ smaller than the smallest eigenvalue of $\hat{\mathbf{B}} - \mu \mathbf{T}_m$ should be considered in order for $\hat{y}(\mathbf{x})$ to attain a minimum. It follows that for such an assignment of values for μ and λ , the corresponding solution of equation (8.35) produces an optimum for $\hat{y}$ subject to a fixed $r = (\mathbf{x}'\mathbf{x})^{1/2}$ and a fixed value of $v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}$.

EXAMPLE 8.3.2. An attempt was made to design an experiment from which one could find conditions on concentration of three basic substances that maximize a certain mechanical modular property of a solid propellant. The initial intent was to construct and use a central composite design (see Khuri and Cornell, 1996, Section 4.5.3) for the three components in the system. However, certain experimental difficulties prohibited the use of the design as planned, and the design used led to problems with multicollinearity as far as the fitting of the second-order model is concerned. The design settings and corresponding response values are given in Table 8.3.

In this example, the smallest eigenvalue of $\mathbf{X}'\mathbf{X}$ is $\nu_{\min} = 0.0321$. Correspondingly, the values of v_{0m} , $\boldsymbol{\tau}_m$, and $\mathbf{T}_m$ in inequality (8.34) are $v_{0m} = -0.2935$, $\boldsymbol{\tau}_m = (0.0469, 0.4081, 0.4071)'$, and

$$\mathbf{T}_m = \begin{bmatrix} 0.1129 & 0.0095 & 0.2709 \\ 0.0095 & -0.1382 & -0.0148 \\ 0.2709 & -0.0148 & 0.6453 \end{bmatrix}.$$

As for q in inequality (8.34), values of $|v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}|$ were computed at each of the 15 design points in Table 8.3. The largest value was found to

Table 8.3. Design Settings and Response Values for Example 8.3.2

x_1	x_2	x_3	y
− 1.020	− 1.402	− 0.998	13.5977
0.900	0.478	− 0.818	12.7838
0.870	− 1.282	0.882	16.2780
− 0.950	0.458	0.972	14.1678
− 0.930	− 1.242	− 0.868	9.2461
0.750	0.498	− 0.618	17.0167
0.830	− 1.092	0.732	13.4253
− 0.950	0.378	0.832	16.0967
1.950	− 0.462	0.002	14.5438
− 2.150	− 0.402	− 0.038	20.9534
− 0.550	0.058	− 0.518	11.0411
− 0.450	1.378	0.182	21.2088
0.150	1.208	0.082	25.5514
0.100	1.768	− 0.008	33.3793
1.450	− 0.342	0.182	15.4341

Source: Khuri and Myers (1979). Reproduced with permission of the American Statistical Association.

Table 8.4. Results of Modified Ridge Analysis

r	0.848	1.162	1.530	1.623	1.795	1.850	1.904	1.935	2.000
$ \omega $	0.006	0.074	0.136	1.139	0.048	0.086	0.126	0.146	0.165
$\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$	1.170	1.635	2.922	3.147	1.305	2.330	3.510	4.177	5.336
x_1	0.410	0.563	0.773	0.785	0.405	0.601	0.750	0.820	0.965
x_2	0.737	1.015	1.320	1.422	1.752	1.751	1.750	1.752	1.752
x_3	0.097	0.063	0.019	0.011	0.000	0.000	0.012	0.015	-0.030
$\hat{y}(\mathbf{x})$	22.420	27.780	35.242	37.190	37.042	40.222	42.830	44.110	46.260

Source: Khuri and Myers (1979). Reproduced with permission of the American Statistical Association.

Table 8.5. Results of Standard Ridge Analysis

r	0.140	0.379	0.698	0.938	1.146	1.394	1.484	1.744	1.944	1.975	2.000
$ \omega $	0.241	0.124	0.104	0.337	0.587	0.942	1.085	1.553	1.958	2.025	2.080
$\omega^2/\nu_{\min}$	1.804	0.477	0.337	3.543	10.718	27.631	36.641	75.163	119.371	127.815	134.735
$\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$	2.592	1.554	2.104	6.138	14.305	32.787	42.475	83.38	129.834	138.668	145.907
x_1	0.037	0.152	0.352	0.515	0.660	0.835	0.899	1.085	1.227	1.249	1.265
x_2	0.103	0.255	0.422	0.531	0.618	0.716	0.749	0.845	0.916	0.927	0.936
x_3	0.087	0.235	0.431	0.577	0.705	0.858	0.912	1.074	1.197	1.217	1.232
$\hat{y}(\mathbf{x})$	12.796	16.021	21.365	26.229	31.086	37.640	40.197	48.332	55.147	56.272	57.176

Source: Khuri and Myers (1979). Reproduced with permission of the American Statistical Association.

be 0.087. Hence, the value of $|v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}|$ should not grow much larger than 0.09 in the experimental region. Furthermore, r in equation (8.33) must not exceed the value 2, since most of the design points are contained inside a sphere of radius 2. The results of maximizing $\hat{y}(\mathbf{x})$ subject to this dual constraint are given in Table 8.4. For the sake of comparison, the results of applying the standard procedure of ridge analysis (that is, without the additional constraint concerning $v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}$) are displayed in Table 8.5.

It is clear from Tables 8.4 and 8.5 that the extra constraint concerning $v_{0m} + \mathbf{x}'\boldsymbol{\tau}_m + \mathbf{x}'\mathbf{T}_m\mathbf{x}$ has profoundly improved the precision of $\hat{y}$ at the estimated maxima. At a specified radius, the value of $\hat{y}$ obtained under standard ridge analysis is higher than the one obtained under modified ridge analysis. However, the prediction variance values under the latter procedure are much smaller, as can be seen from comparing Tables 8.4 and 8.5. While the tradeoff that exists between a high response value and a small prediction variance is a bit difficult to cope with from a decision making standpoint, there is a clear superiority of the results displayed in Table 8.4. For example, one would hardly choose any operating conditions in Table 8.5 that indicate $\hat{y} \geq 50$, due to the accompanying large prediction variances. On the other hand, Table 8.4 reveals that at radius $r = 2.000$, $\hat{y} = 46.26$ with $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$

$= 5.336$, while a rival set of coordinates at $r = 1.744$ for standard ridge analysis gives $\hat{y} = 48.332$ with $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2 = 83.38$.

Row 3 of Table 8.5 gives values of $\omega^2/\nu_{\min}$, which should be compared with the corresponding values in row 4 of the same table. One can easily see that in this example, $\omega^2/\nu_{\min}$ accounts for a large portion of $\text{Var}[\hat{y}(\mathbf{x})]/\sigma^2$.

8.4. RESPONSE SURFACE DESIGNS

We recall from Section 8.3 that one of the objectives of response surface methodology is the selection of a response surface design according to a certain optimality criterion. The design selection entails the specification of the settings of a group of input variables that can be used as experimental runs in a given experiment.

The proper choice of a response surface design can have a profound effect on the success of a response surface exploration. To see this, let us suppose that the fitted model is linear of the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (8.36)$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations, $\mathbf{X}$ is an $n \times p$ known matrix that depends on the design settings, $\boldsymbol{\beta}$ is a vector of p unknown parameters, and $\boldsymbol{\epsilon}$ is a vector of random errors in the elements of $\mathbf{y}$. Typically, $\boldsymbol{\epsilon}$ is assumed to have the normal distribution $N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, where σ^2 is unknown. In this case, the vector $\boldsymbol{\beta}$ is estimated by the least-squares estimator $\hat{\boldsymbol{\beta}}$, which is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \quad (8.37)$$

If $x_1, x_2, \dots, x_k$ are the input variables for the model under consideration, then the predicted response at a point $\mathbf{x} = (x_1, x_2, \dots, x_k)'$ in a region of interest R is written as

$$\hat{y}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\hat{\boldsymbol{\beta}}, \quad (8.38)$$

where $\mathbf{f}(\mathbf{x})$ is a $p \times 1$ vector whose first element is equal to one and whose remaining $p - 1$ elements are functions of $x_1, x_2, \dots, x_k$. These functions are in the form of powers and cross products of powers of the x_i 's up to degree d . In this case, the model is said to be of order d . At the u th experimental run, $\mathbf{x}_u = (x_{u1}, x_{u2}, \dots, x_{uk})'$ and the corresponding response value is y_u ($u = 1, 2, \dots, n$). Then $n \times k$ matrix $\mathbf{D} = [\mathbf{x}_1: \mathbf{x}_2: \dots: \mathbf{x}_n]'$ is the design matrix. Thus by a choice of design we mean the specification of the elements of $\mathbf{D}$.

If model (8.36) is correct, then $\hat{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}$ and its variance-covariance matrix is given by

$$\text{Var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1} \sigma^2. \quad (8.39)$$

Hence, from formula (8.38), the prediction variance can be written as

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x}). \quad (8.40)$$

The design $\mathbf{D}$ is rotatable if $\text{Var}[\hat{y}(\mathbf{x})]$ remains constant at all points that are equidistant from the design center, as we may recall from Section 8.3. The input variables are coded so that the center of the design coincides with the origin of the coordinates system (see Khuri and Cornell, 1996, Section 2.8).

8.4.1. First-Order Designs

If model (8.36) is of the first order (that is, $d = 1$), then the matrix $\mathbf{X}$ is of the form $\mathbf{X} = [\mathbf{1}_n; \mathbf{D}]$, where $\mathbf{1}_n$ is a vector of ones of order $n \times 1$ [see model (8.8)]. The input variables can be coded in such a way that the sum of the elements in each column of $\mathbf{D}$ is equal to zero. Consequently, the prediction variance in formula (8.40) can be written as

$$\text{Var}[\hat{y}(\mathbf{x})] = \sigma^2 \left[\frac{1}{n} + \mathbf{x}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{x} \right]. \quad (8.41)$$

Formula (8.41) clearly shows the dependence of the prediction variance on the design matrix.

A reasonable criterion for the choice of $\mathbf{D}$ is the minimization of $\text{Var}[\hat{y}(\mathbf{x})]$, or equivalently, the minimization of $\mathbf{x}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{x}$ within the region R . To accomplish this we first note that for any $\mathbf{x}$ in the region R ,

$$\mathbf{x}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{x} \leq \|\mathbf{x}\|_2^2 \|(\mathbf{D}'\mathbf{D})^{-1}\|_2, \quad (8.42)$$

where $\|\mathbf{x}\|_2 = (\mathbf{x}'\mathbf{x})^{1/2}$ and $\|(\mathbf{D}'\mathbf{D})^{-1}\|_2 = [\sum_{i=1}^k \sum_{j=1}^k (d^{ij})^2]^{1/2}$ is the Euclidean norm of $(\mathbf{D}'\mathbf{D})^{-1}$ with d^{ij} being its (i, j) th element ($i, j = 1, 2, \dots, k$). Inequality (8.42) follows from applying Theorems 2.3.16 and 2.3.20. Thus by choosing the design $\mathbf{D}$ so that it minimizes $\|(\mathbf{D}'\mathbf{D})^{-1}\|_2$, the quantity $\mathbf{x}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{x}$, and hence the prediction variance, can be reduced throughout the region R .

Theorem 8.4.1. For a given number n of experimental runs, $\|(\mathbf{D}'\mathbf{D})^{-1}\|_2$ attains its minimum if the columns $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_k$ of $\mathbf{D}$ are such that $\mathbf{d}_i'\mathbf{d}_j = 0$, $i \neq j$, and $\mathbf{d}_i'\mathbf{d}_i$ is as large as possible inside the region R .

Proof. We have that $\mathbf{D} = [\mathbf{d}_1: \mathbf{d}_2: \dots: \mathbf{d}_k]$. The elements of $\mathbf{d}_i$ are the n design settings of the input variable x_i ($i = 1, 2, \dots, k$). Suppose that the region R places the following restrictions on these settings:

$$\mathbf{d}_i'\mathbf{d}_i \leq c_i^2, \quad i = 1, 2, \dots, k, \quad (8.43)$$

where c_i is some fixed constant. This means that the spread of the design in the direction of the i th coordinate axis is bounded by c_i^2 ($i = 1, 2, \dots, k$). $\square$

Now, if d_{ii} denotes the i th diagonal element of $\mathbf{D}'\mathbf{D}$, then $d_{ii} = \mathbf{d}'_i \mathbf{d}_i$ ($i = 1, 2, \dots, k$). Furthermore,

$$d^{ii} \geq \frac{1}{d_{ii}}, \quad i = 1, 2, \dots, k, \quad (8.44)$$

where d^{ii} is the i th diagonal element of $(\mathbf{D}'\mathbf{D})^{-1}$.

To prove inequality (8.44), let $\mathbf{D}_i$ be a matrix of order $n \times (k - 1)$ obtained from $\mathbf{D}$ by removing its i th column $\mathbf{d}_i$ ($i = 1, 2, \dots, k$). The cofactor of d_{ii} in $\mathbf{D}'\mathbf{D}$ is then $\det(\mathbf{D}'_i \mathbf{D}_i)$. Hence, from Section 2.3.3,

$$d^{ii} = \frac{\det(\mathbf{D}'_i \mathbf{D}_i)}{\det(\mathbf{D}'\mathbf{D})}, \quad i = 1, 2, \dots, k. \quad (8.45)$$

There exists an orthogonal matrix $\mathbf{E}_i$ of order $k \times k$ (whose determinant has an absolute value of one) such that the first column of $\mathbf{D}\mathbf{E}_i$ is $\mathbf{d}_i$ and the remaining columns are the same as those of $\mathbf{D}_i$, that is,

$$\mathbf{D}\mathbf{E}_i = [\mathbf{d}_i : \mathbf{D}_i], \quad i = 1, 2, \dots, k.$$

It follows that [see property 7 in Section 2.3.3]

$$\begin{aligned} \det(\mathbf{D}'\mathbf{D}) &= \det(\mathbf{E}'_i \mathbf{D}' \mathbf{D} \mathbf{E}_i) \\ &= \det(\mathbf{D}'_i \mathbf{D}_i) [\mathbf{d}'_i \mathbf{d}_i - \mathbf{d}'_i \mathbf{D}_i (\mathbf{D}'_i \mathbf{D}_i)^{-1} \mathbf{D}'_i \mathbf{d}_i]. \end{aligned}$$

Hence, from (8.45) we obtain

$$d^{ii} = \left[d_{ii} - \mathbf{d}'_i \mathbf{D}_i (\mathbf{D}'_i \mathbf{D}_i)^{-1} \mathbf{D}'_i \mathbf{d}_i \right]^{-1}, \quad i = 1, 2, \dots, k. \quad (8.46)$$

Inequality (8.44) now follows from formula (8.46), since $\mathbf{d}'_i \mathbf{D}_i (\mathbf{D}'_i \mathbf{D}_i)^{-1} \mathbf{D}'_i \mathbf{d}_i \geq 0$. We can therefore write

$$\|(\mathbf{D}'\mathbf{D})^{-1}\|_2 \geq \left[\sum_{i=1}^k (d^{ii})^2 \right]^{1/2} \geq \left[\sum_{i=1}^k \frac{1}{d_{ii}^2} \right]^{1/2}.$$

Using the restrictions (8.43) we then have

$$\|(\mathbf{D}'\mathbf{D})^{-1}\|_2 \geq \left[\sum_{i=1}^k \frac{1}{c_i^4} \right]^{1/2}.$$

Equality is achieved if the columns of $\mathbf{D}$ are orthogonal to one another and $d_{ii} = c_i^2$ ($i = 1, 2, \dots, k$). This follows from the fact that $d^{ii} = 1/d_{ii}$ if and only if $\mathbf{d}_i'\mathbf{D}_i = 0$ ($i = 1, 2, \dots, k$), as can be seen from formula (8.46).

Definition 8.4.1. A design for fitting a first-order model is said to be orthogonal if its columns are orthogonal to one another. $\square$

Corollary 8.4.1. For a given number n of experimental runs, $\text{Var}(\hat{\beta}_i)$ attains a minimum if and only if the design is orthogonal, where $\hat{\beta}_i$ is the least-squares estimator of β_i in model (8.8), $i = 1, 2, \dots, k$.

Proof. This follows directly from Theorem 8.4.1 and the fact that $\text{Var}(\hat{\beta}_i) = \sigma^2 d^{ii}$ ($i = 1, 2, \dots, k$), as can be seen from formula (8.39). $\square$

From Theorem 8.4.1 and Corollary 8.4.1 we conclude that an orthogonal design for fitting a first-order model has optimal variance properties. Another advantage of orthogonal first-order designs is that the effects of the k input variables in model (8.8), as measured by the values of the β_i 's ($i = 1, 2, \dots, k$), can be estimated independently. This is because the off-diagonal elements of the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ in formula (8.39) are zero. This means that the elements of $\hat{\boldsymbol{\beta}}$ are uncorrelated and hence statistically independent under the assumption of normality of the random error vector $\boldsymbol{\epsilon}$ in model (8.36).

Examples of first-order orthogonal designs are given in Khuri and Cornell (1996, Chapter 3). Prominent among these designs are the 2^k factorial design (each input variable has two levels, and the number of all possible combinations of these levels is 2^k) and the Plackett-Burman design, which was introduced in Plackett and Burman (1946). In the latter design, the number of design points is equal to $k + 1$, which must be a multiple of 4.

8.4.2. Second-Order Designs

These designs are used to fit second-order models of the form given by (8.12). Since the number of parameters in this model is $p = (k + 1)(k + 2)/2$, the number of experimental runs (or design points) in a second-order design must at least be equal to p . The most frequently used second-order designs include the 3^k design (each input variable has three levels, and the number of all possible combinations of these levels is 3^k), the central composite design (CCD), and the Box-Behnken design.

The CCD was introduced by Box and Wilson (1951). It is made up of a factorial portion consisting of a 2^k factorial design, an axial portion of k pairs of points with the i th pair consisting of two symmetric points on the i th coordinate axis ($i = 1, 2, \dots, k$) at a distance of α (> 0) from the design center (which coincides with the center of the coordinates system by the coding scheme), and n_0 (≥ 1) center-point runs. The values of α and n_0 can be chosen so that the CCD acquires certain desirable features (see, for example, Khuri and Cornell, 1996, Section 4.5.3). In particular, if $\alpha = F^{1/4}$, where F denotes the number of points in the factorial portion, then the CCD is rotatable. The choice of n_0 can affect the stability of the prediction variance.

The Box–Behnken design, introduced in Box and Behnken (1960), is a subset of a 3^k factorial design and, in general, requires many fewer points. It also compares favorably with the CCD. A thorough description of this design is given in Box and Draper (1987, Section 15.4).

Other examples of second-order designs are given in Khuri and Cornell (1996, Chapter 4).

8.4.3. Variance and Bias Design Criteria

We have seen that the minimization of the prediction variance represents an important criterion for the selection of a response surface design. This criterion, however, presumes that the fitted model is correct. There are many situations in which bias in the predicted response can occur due to fitting the wrong model. We refer to this as model bias.

Box and Draper (1959, 1963) presented convincing arguments in favor of recognizing bias as an important design criterion—in certain cases, even more important than the variance criterion.

Consider again model (8.36). The response value at a point $\mathbf{x} = (x_1, x_2, \dots, x_k)'$ in a region R is represented as

$$y(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \epsilon, \quad (8.47)$$

where $\mathbf{f}'(\mathbf{x})$ is the same as in model (8.38). While it is hoped that model (8.47) is correct, there is always a fear that the true model is different. Let us therefore suppose that in reality the true mean response at $\mathbf{x}$, denoted by $\eta(\mathbf{x})$, is given by

$$\eta(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \mathbf{g}'(\mathbf{x})\boldsymbol{\delta} \quad (8.48)$$

where the elements of $\mathbf{g}'(\mathbf{x})$ depend on $\mathbf{x}$ and consist of powers and cross products of powers of $x_1, x_2, \dots, x_k$ of degree $d' > d$, with d being the order of model (8.47), and $\boldsymbol{\delta}$ is a vector of q unknown parameters. For a given

design $\mathbf{D}$ of n experimental runs, we then have the model

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta},$$

where $\boldsymbol{\eta}$ is the vector of true means (or expected values) of the elements of $\mathbf{y}$ at the n design points, $\mathbf{X}$ is the same as in model (8.36), and $\mathbf{Z}$ is a matrix of order $n \times q$ whose u th row is equal to $\mathbf{g}'(\mathbf{x}_u)$. Here, $\mathbf{x}'_u$ denotes the u th row of $\mathbf{D}$ ($u = 1, 2, \dots, n$).

At each point $\mathbf{x}$ in R , the mean squared error (MSE) of $\hat{y}(\mathbf{x})$, where $\hat{y}(\mathbf{x})$ is the predicted response as given by formula (8.38), is defined as

$$\text{MSE}[\hat{y}(\mathbf{x})] = E[\hat{y}(\mathbf{x}) - \eta(\mathbf{x})]^2.$$

This can be expressed as

$$\text{MSE}[\hat{y}(\mathbf{x})] = \text{Var}[\hat{y}(\mathbf{x})] + \text{Bias}^2[\hat{y}(\mathbf{x})], \quad (8.49)$$

where $\text{Bias}[\hat{y}(\mathbf{x})] = E[\hat{y}(\mathbf{x})] - \eta(\mathbf{x})$. The fundamental philosophy of Box and Draper (1959, 1963) is centered around the consideration of the integrated mean squared error (IMSE) of $\hat{y}(\mathbf{x})$. This is denoted by J and is defined in terms of a k -tuple Riemann integral over the region R , namely,

$$J = \frac{n\Omega}{\sigma^2} \int_R \text{MSE}[\hat{y}(\mathbf{x})] d\mathbf{x}, \quad (8.50)$$

where $\Omega^{-1} = \int_R d\mathbf{x}$ and σ^2 is the error variance. The partitioning of $\text{MSE}[\hat{y}(\mathbf{x})]$ as in formula (8.49) enables us to separate J into two parts:

$$J = \frac{n\Omega}{\sigma^2} \int_R \text{Var}[\hat{y}(\mathbf{x})] d\mathbf{x} + \frac{n\Omega}{\sigma^2} \int_R \text{Bias}^2[\hat{y}(\mathbf{x})] d\mathbf{x} = V + B. \quad (8.51)$$

The quantities V and B are called the average variance and average squared bias of $\hat{y}(\mathbf{x})$, respectively. Both V and B depend on the design $\mathbf{D}$. Thus a reasonable choice of design is one that minimizes (1) V alone, (2) B alone, or (3) $J = V + B$.

Now, using formula (8.40), V can be written as

$$\begin{aligned} V &= n\Omega \int_R \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1} \mathbf{f}(\mathbf{x}) d\mathbf{x} \\ &= \text{tr} \left\{ n(\mathbf{X}'\mathbf{X})^{-1} \left[\Omega \int_R \mathbf{f}(\mathbf{x}) \mathbf{f}'(\mathbf{x}) d\mathbf{x} \right] \right\} \\ &= \text{tr} \left[n(\mathbf{X}'\mathbf{X})^{-1} \boldsymbol{\Gamma}_{11} \right], \end{aligned} \quad (8.52)$$

where

$$\Gamma_{11} = \Omega \int_R \mathbf{f}(\mathbf{x}) \mathbf{f}'(\mathbf{x}) d\mathbf{x}.$$

As for B , we note from formula (8.37) that

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\eta} \\ &= \boldsymbol{\beta} + \mathbf{A}\boldsymbol{\delta}, \end{aligned}$$

where $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Z}$. Thus from formula (8.38) we have

$$E[\hat{y}(\mathbf{x})] = \mathbf{f}'(\mathbf{x})(\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\delta}).$$

Using the expression for $\eta(\mathbf{x})$ in formula (8.48), B can be written as

$$\begin{aligned} B &= \frac{n\Omega}{\sigma^2} \int_R [\mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \mathbf{f}'(\mathbf{x})\mathbf{A}\boldsymbol{\delta} - \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} - \mathbf{g}'(\mathbf{x})\boldsymbol{\delta}]^2 d\mathbf{x} \\ &= \frac{n\Omega}{\sigma^2} \int_R [\mathbf{f}'(\mathbf{x})\mathbf{A}\boldsymbol{\delta} - \mathbf{g}'(\mathbf{x})\boldsymbol{\delta}]^2 d\mathbf{x} \\ &= \frac{n\Omega}{\sigma^2} \int_R \boldsymbol{\delta}' [\mathbf{A}'\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x})] [\mathbf{f}'(\mathbf{x})\mathbf{A} - \mathbf{g}'(\mathbf{x})] \boldsymbol{\delta} d\mathbf{x} \\ &= \frac{n\Omega}{\sigma^2} \int_R \boldsymbol{\delta}' [\mathbf{A}'\mathbf{f}(\mathbf{x})\mathbf{f}'(\mathbf{x})\mathbf{A} - \mathbf{g}(\mathbf{x})\mathbf{f}'(\mathbf{x})\mathbf{A} - \mathbf{A}'\mathbf{f}(\mathbf{x})\mathbf{g}'(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{g}'(\mathbf{x})] \boldsymbol{\delta} d\mathbf{x} \\ &= \frac{n}{\sigma^2} \boldsymbol{\delta}' \boldsymbol{\Delta} \boldsymbol{\delta}, \end{aligned} \tag{8.53}$$

where

$$\boldsymbol{\Delta} = \mathbf{A}'\Gamma_{11}\mathbf{A} - \Gamma'_{12}\mathbf{A} - \mathbf{A}'\Gamma_{12} + \Gamma_{22},$$

$$\Gamma_{12} = \Omega \int_R \mathbf{f}(\mathbf{x}) \mathbf{g}'(\mathbf{x}) d\mathbf{x},$$

$$\Gamma_{22} = \Omega \int_R \mathbf{g}(\mathbf{x}) \mathbf{g}'(\mathbf{x}) d\mathbf{x}.$$

The matrices Γ_{11} , Γ_{12} , and Γ_{22} are called region moments. By adding and subtracting the matrix $\Gamma'_{12}\Gamma_{11}^{-1}\Gamma_{12}$ from $\boldsymbol{\Delta}$ in formula (8.53), B can be expressed as

$$B = \frac{n}{\sigma^2} \boldsymbol{\delta}' [(\Gamma_{22} - \Gamma'_{12}\Gamma_{11}^{-1}\Gamma_{12}) + (\mathbf{A} - \Gamma_{11}^{-1}\Gamma_{12})'\Gamma_{11}(\mathbf{A} - \Gamma_{11}^{-1}\Gamma_{12})] \boldsymbol{\delta}. \tag{8.54}$$

We note that the design $\mathbf{D}$ affects only the second expression inside brackets on the right-hand side of formula (8.54). Thus to minimize B , the design $\mathbf{D}$ should be chosen such that

$$\mathbf{A} - \Gamma_{11}^{-1} \Gamma_{12} = \mathbf{0}. \quad (8.55)$$

Since $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Z}$, a sufficient (but not necessary) condition for the minimization of B is

$$\mathbf{M}_{11} = \Gamma_{11}, \quad \mathbf{M}_{12} = \Gamma_{12}, \quad (8.56)$$

where $\mathbf{M}_{11} = (1/n)\mathbf{X}'\mathbf{X}$, $\mathbf{M}_{12} = (1/n)\mathbf{X}'\mathbf{Z}$ are the so-called design moments. Thus a sufficient condition for the minimization of B is the equality of the design moments, $\mathbf{M}_{11}$ and $\mathbf{M}_{12}$, to the corresponding region moments, Γ_{11} and Γ_{12} .

The minimization of $J = V + B$ is not possible without the specification of δ/σ . Box and Draper (1959, 1963) showed that unless V is considerably larger than B , the optimal design that minimizes J has characteristics similar to those of a design that minimizes just B .

Examples of designs that minimize V alone or B alone can be found in Box and Draper (1987, Chapter 13), Khuri and Cornell (1996, Chapter 6), and Myers (1976, Chapter 9).

8.5. ALPHABETIC OPTIMALITY OF DESIGNS

Let us again consider model (8.47), which we now assume to be correct, that is, the true mean response, $\eta(\mathbf{x})$, is equal to $\mathbf{f}'(\mathbf{x})\boldsymbol{\beta}$. In this case, the matrix $\mathbf{X}'\mathbf{X}$ plays an important role in the determination of an optimal design, since the elements of $(\mathbf{X}'\mathbf{X})^{-1}$ are proportional to the variances and covariances of the least-squares estimators of the model's parameters [see formula (8.39)].

The mathematical theory of optimal designs, which was developed by Kiefer (1958, 1959, 1960, 1961, 1962a, b), is concerned with the choice of designs that minimize certain functions of the elements of $(\mathbf{X}'\mathbf{X})^{-1}$. The kernel of Kiefer's approach is based on the concept of design measure, which represents a generalization of the traditional design concept. So far, each of the designs that we have considered for fitting a response surface model has consisted of a set of n points in a k -dimensional space ($k \geq 1$). Suppose that $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are distinct points of an n -point design ($m \leq n$) with the l th ($l = 1, 2, \dots, m$) point being replicated n_l (≥ 1) times (that is, n_l repeated observations are taken at this point). The design can therefore be regarded as a collection of points in a region of interest R with the l th point being assigned the weight n_l/n ($l = 1, 2, \dots, m$), where $n = \sum_{l=1}^m n_l$. Kiefer generalized this setup using the so-called continuous design measure, which is

basically a probability measure $\xi(\mathbf{x})$ defined on R and satisfies the conditions

$$\xi(\mathbf{x}) \geq 0 \quad \text{for all } \mathbf{x} \in R \quad \text{and} \quad \int_R d\xi(\mathbf{x}) = 1.$$

In particular, the measure induced by a traditional design $\mathbf{D}$ with n points is called a discrete design measure and is denoted by ξ_n . It should be noted that while a discrete design measure is realizable in practice, the same is not true of a general continuous design measure. For this reason, the former design is called *exact* and the latter design is called *approximate*.

By definition, the moment matrix of a design measure ξ is a symmetric matrix of the form $\mathbf{M}(\xi) = [m_{ij}(\xi)]$, where

$$m_{ij}(\xi) = \int_R f_i(\mathbf{x}) f_j(\mathbf{x}) d\xi(\mathbf{x}). \quad (8.57)$$

Here, $f_i(\mathbf{x})$ is the i th element of $\mathbf{f}(\mathbf{x})$ in formula (8.47), $i = 1, 2, \dots, p$. For a discrete design measure ξ_n , the (i, j) th element of the moment matrix is

$$m_{ij}(\xi_n) = \frac{1}{n} \sum_{l=1}^m n_l f_i(\mathbf{x}_l) f_j(\mathbf{x}_l), \quad (8.58)$$

where m is the number of distinct design points and n_l is the number of replications at the l th point ($l = 1, 2, \dots, m$). In this special case, the matrix $\mathbf{M}(\xi)$ reduces to the usual moment matrix $(1/n)\mathbf{X}'\mathbf{X}$, where $\mathbf{X}$ is the same matrix as in formula (8.36).

For a general design measure ξ , the standardized prediction variance, denoted by $d(\mathbf{x}, \xi)$, is defined as

$$d(\mathbf{x}, \xi) = \mathbf{f}'(\mathbf{x}) [\mathbf{M}(\xi)]^{-1} \mathbf{f}(\mathbf{x}), \quad (8.59)$$

where $\mathbf{M}(\xi)$ is assumed to be nonsingular. In particular, for a discrete design measure ξ_n , the prediction variance in formula (8.40) is equal to $(\sigma^2/n)d(\mathbf{x}, \xi_n)$.

Let H denote the class of all design measures defined on the region R . A prominent design criterion that has received a great deal of attention is that of D -optimality, in which the determinant of $\mathbf{M}(\xi)$ is maximized. Thus a design measure ξ_d is D -optimal if

$$\det[\mathbf{M}(\xi_d)] = \sup_{\xi \in H} \det[\mathbf{M}(\xi)]. \quad (8.60)$$

The rationale behind this criterion has to do with the minimization of the generalized variance of the least-squares estimator $\hat{\boldsymbol{\beta}}$ of the parameter vector $\boldsymbol{\beta}$. By definition, the generalized variance of $\hat{\boldsymbol{\beta}}$ is the same as the determinant of the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$. This is based on the fact that

under the normality assumption, the content (volume) of a fixed-level confidence region on β is proportional to $[\det(\mathbf{X}'\mathbf{X})]^{-1/2}$. The review articles by St. John and Draper (1975), Ash and Hedayat (1978), and Atkinson (1982, 1988) contain many references on D -optimality.

Another design criterion that is closely related to D -optimality is G -optimality, which is concerned with the prediction variance. By definition, a design measure ξ_g is G -optimal if it minimizes over H the maximum standardized prediction variance over the region R , that is,

$$\sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi_g) = \inf_{\xi \in H} \left\{ \sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi) \right\}. \quad (8.61)$$

Kiefer and Wolfowitz (1960) showed that D -optimality and G -optimality, as defined by formulas (8.60) and (8.61), are equivalent. Furthermore, a design measure ξ^* is G -optimal (or D -optimal) if and only if

$$\sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi^*) = p, \quad (8.62)$$

where p is the number of parameters in the model. Formula (8.62) can be conveniently used to determine if a given design measure is D -optimal, since in general $\sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi) \geq p$ for any design measure $\xi \in H$. If equality can be achieved by a design measure, then it must be G -optimal, and hence D -optimal.

EXAMPLE 8.5.1. Consider fitting a second-order model in one input variable x over the region $R = [-1, 1]$. In this case, model (8.47) takes the form $y(x) = \beta_0 + \beta_1 x + \beta_{11} x^2 + \epsilon$, that is, $\mathbf{f}'(x) = (1, x, x^2)$. Suppose that the design measure used is defined as

$$\xi(x) = \begin{cases} \frac{1}{3}, & x = -1, 0, 1, \\ 0 & \text{otherwise.} \end{cases} \quad (8.63)$$

Thus ξ is a discrete design measure that assigns one-third of the experimental runs to each of the points -1 , 0 , and 1 . This design measure is D -optimal. To verify this claim, we first need to determine the values of the elements of the moment matrix $\mathbf{M}(\xi)$. Using formula (8.58) with $n_l/n = \frac{1}{3}$ for $l = 1, 2, 3$, we find that $m_{11} = 1$, $m_{12} = 0$, $m_{13} = \frac{2}{3}$, $m_{22} = \frac{2}{3}$, $m_{23} = 0$, and $m_{33} = \frac{2}{3}$. Hence,

$$\mathbf{M}(\xi) = \begin{bmatrix} 1 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} \end{bmatrix},$$

$$\mathbf{M}^{-1}(\xi) = \begin{bmatrix} 3 & 0 & -3 \\ 0 & \frac{3}{2} & 0 \\ -3 & 0 & \frac{9}{2} \end{bmatrix}.$$

By applying formula (8.59) we find that $d(x, \xi) = 3 - \frac{9}{2}x^2 + \frac{9}{2}x^4$, $-1 \leq x \leq 1$. We note that $d(x, \xi) \leq 3$ for all x in $[-1, 1]$ with $d(0, \xi) = 3$. Thus $\sup_{x \in R} d(x, \xi) = 3$. Since 3 is the number of parameters in the model, then by condition (8.62) we conclude that the design measure defined by formula (8.63) is D -optimal.

In addition to the D - and G -optimality criteria, other variance-related design criteria have also been investigated. These include A - and E -optimality. By definition, a design measure is A -optimal if it maximizes the trace of $\mathbf{M}(\xi)$. This is equivalent to minimizing the sum of the variances of the least-squares estimators of the fitted model's parameters. In E -optimality, the smallest eigenvalue of $\mathbf{M}(\xi)$ is maximized. The rationale behind this criterion is based on the fact that

$$d(\mathbf{x}, \xi) \leq \frac{\mathbf{f}'(\mathbf{x})\mathbf{f}(\mathbf{x})}{\lambda_{\min}},$$

as can be seen from formula (8.59), where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{M}(\xi)$. Hence, $d(\mathbf{x}, \xi)$ can be reduced by maximizing $\lambda_{\min}$.

The efficiency of a design measure $\zeta \in H$ with respect to a D -optimal design is defined as

$$D\text{-efficiency} = \left\{ \frac{\det[\mathbf{M}(\zeta)]}{\sup_{\xi \in H} \det[\mathbf{M}(\xi)]} \right\}^{1/p},$$

where p is the number of parameters in the model. Similarly, the G -efficiency of ζ is defined as

$$G\text{-efficiency} = \frac{p}{\sup_{\mathbf{x} \in R} d(\mathbf{x}, \zeta)}.$$

Both D - and G -efficiency values fall within the interval $[0, 1]$. The closer these values are to 1, the more efficient their corresponding designs are. Lucas (1976) compared several second-order designs (such as central composite and Box-Behnken designs) on the basis of their D - and G -efficiency values.

The equivalence theorem of Kiefer and Wolfowitz (1960) can be applied to construct a D -optimal design using a sequential procedure. This procedure is described in Wynn (1970, 1972) and goes as follows: Let $\mathbf{D}_{n_0}$ denote an initial response surface design with n_0 points, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_0}$, for which the matrix $\mathbf{X}'\mathbf{X}$ is nonsingular. A point $\mathbf{x}_{n_0+1}$ is found in the region R such that

$$d(\mathbf{x}_{n_0+1}, \xi_{n_0}) = \sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi_{n_0}),$$

where ξ_{n_0} is the discrete design measure that represents $\mathbf{D}_{n_0}$. By augmenting $\mathbf{D}_{n_0}$ with $\mathbf{x}_{n_0+1}$ we obtain the design $\mathbf{D}_{n_0+1}$. Then, another point $\mathbf{x}_{n_0+2}$ is chosen such that

$$d(\mathbf{x}_{n_0+2}, \xi_{n_0+1}) = \sup_{\mathbf{x} \in R} d(\mathbf{x}, \xi_{n_0+1}),$$

where ξ_{n_0+1} is the discrete design measure that represents $\mathbf{D}_{n_0+1}$. The point $\mathbf{x}_{n_0+2}$ is then added to $\mathbf{D}_{n_0+1}$ to obtain the design $\mathbf{D}_{n_0+2}$. By continuing this process we obtain a sequence of discrete design measures, namely, $\xi_{n_0}, \xi_{n_0+1}, \xi_{n_0+2}, \dots$. Wynn (1970) showed that this sequence converges to the D -optimal design ξ_d , that is,

$$\det[\mathbf{M}(\xi_{n_0+n})] \rightarrow \det[\mathbf{M}(\xi_d)]$$

as $n \rightarrow \infty$. An example is given in Wynn (1970, Section 5) to illustrate this sequential procedure.

The four design criteria, A -, D -, E -, and G -optimality, are referred to as alphabetic optimality. More detailed information about these criteria can be found in Atkinson (1982, 1988), Fedorov (1972), Pazman (1986), and Silvey (1980). Recall that to perform an actual experiment, one must use a discrete design. It is possible to find a discrete design measure ξ_n that approximates an optimal design measure. The approximation is good whenever n is large with respect to p (the number of parameters in the model).

Note that the equivalence theorem of Kiefer and Wolfowitz (1960) applies to general design measures and not necessarily to discrete design measures, that is, D - and G -optimality criteria are not equivalent for the class of discrete design measures. Optimal n -point discrete designs, however, can still be found on the basis of maximizing the determinant of $\mathbf{X}'\mathbf{X}$, for example. In this case, finding an optimal n -point design requires a search involving nk variables, where k is the number of input variables. Several algorithms have been introduced for this purpose. For example, the DETMAX algorithm by Mitchell (1974) is used to maximize $\det(\mathbf{X}'\mathbf{X})$. A review of algorithms for constructing optimal discrete designs can be found in Cook and Nachtsheim (1980) (see also Johnson and Nachtsheim, 1983).

One important criticism of the alphabetic optimality approach is that it is set within a rigid framework governed by a set of assumptions. For example, a specific model for the response function must be assumed as the “true” model. Optimal design measures can be quite sensitive to this assumption. Box (1982) presented a critique to this approach. He argued that in a response surface situation, it may not be realistic to assume that a model such as (8.47) represents the true response function exactly. Some protection against bias in the model should therefore be considered when choosing a response surface design. On the other hand, Kiefer (1975) criticized certain aspects of the preoccupation with bias, pointing out examples in which the variance criterion is compromised for the sake of the bias criterion. It follows

that design selection should be guided by more than one single criterion (see Kiefer, 1975, page 286; Box, 1982, Section 7). A reasonable approach is to select compromise designs that are sufficiently good (but not necessarily optimal) from the viewpoint of several criteria that are important to the user.

8.6. DESIGNS FOR NONLINEAR MODELS

The models we have considered so far in the area of response surface methodology were linear in the parameters; hence the term linear models. There are, however, many experimental situations in which linear models do not adequately represent the true mean response. For example, the growth of an organism is more appropriately depicted by a nonlinear model. By definition, a nonlinear model is one of the form

$$y(\mathbf{x}) = h(\mathbf{x}, \boldsymbol{\theta}) + \epsilon, \quad (8.64)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_k)'$ is a vector of k input variables, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$ is a vector of p unknown parameters, ϵ is a random error, and $h(\mathbf{x}, \boldsymbol{\theta})$ is a known function, nonlinear in at least one element of $\boldsymbol{\theta}$. An example of a nonlinear model is

$$h(x, \boldsymbol{\theta}) = \frac{\theta_1 x}{\theta_2 + x}.$$

Here, $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ and θ_2 is a nonlinear parameter. This particular model is known as the Michaelis–Menten model for enzyme kinetics. It relates the initial velocity of an enzymatic reaction to the substrate concentration x .

In contrast to linear models, nonlinear models have not received a great deal of attention in response surface methodology, especially in the design area. The main design criterion for nonlinear models is the D -optimality criterion, which actually applies to a linearized form of the nonlinear model. More specifically, this criterion depends on the assumption that in some neighborhood of a specified value $\boldsymbol{\theta}_0$ of $\boldsymbol{\theta}$, the function $h(\mathbf{x}, \boldsymbol{\theta})$ is approximately linear in $\boldsymbol{\theta}$. In this case, a first-order Taylor's expansion of $h(\mathbf{x}, \boldsymbol{\theta})$ yields the following approximation of $h(\mathbf{x}, \boldsymbol{\theta})$:

$$h(\mathbf{x}, \boldsymbol{\theta}) \approx h(\mathbf{x}, \boldsymbol{\theta}_0) + \sum_{i=1}^p (\theta_i - \theta_{i0}) \frac{\partial h(\mathbf{x}, \boldsymbol{\theta}_0)}{\partial \theta_i}.$$

Thus if $\boldsymbol{\theta}$ is close enough to $\boldsymbol{\theta}_0$, then we have approximately the linear model

$$z(\mathbf{x}) = \sum_{i=1}^p \psi_i \frac{\partial h(\mathbf{x}, \boldsymbol{\theta}_0)}{\partial \theta_i} + \epsilon, \quad (8.65)$$

where $z(\mathbf{x}) = y(\mathbf{x}) - h(\mathbf{x}, \boldsymbol{\theta}_0)$, and ψ_i is the i th element of $\boldsymbol{\psi} = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ ($i = 1, 2, \dots, p$).

For a given design consisting of n experimental runs, model (8.65) can be written in vector form as

$$\mathbf{z} = \mathbf{H}(\boldsymbol{\theta}_0)\boldsymbol{\psi} + \boldsymbol{\epsilon}, \quad (8.66)$$

where $\mathbf{H}(\boldsymbol{\theta}_0)$ is an $n \times p$ matrix whose (u, i) th element is $\partial h(\mathbf{x}_u, \boldsymbol{\theta}_0)/\partial \theta_i$ with $\mathbf{x}_u$ being the vector of design settings for the k input variables at the u th experimental run ($i = 1, 2, \dots, p$; $u = 1, 2, \dots, n$). Using the linearized form given by model (8.66), a design is chosen to maximize the determinant $\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)]$. This is known as the Box–Lucas criterion (see Box and Lucas, 1959).

It can be easily seen that a nonlinear design obtained on the basis of the Box–Lucas criterion depends on the value of $\boldsymbol{\theta}_0$. This is an undesirable characteristic of nonlinear models, since a design is supposed to be used for estimating the unknown parameter vector $\boldsymbol{\theta}$. By contrast, designs for linear models are not dependent on the fitted model's parameters. Several procedures have been proposed for dealing with the problem of design dependence on the parameters of a nonlinear model. These procedures are mentioned in the review article by Myers, Khuri, and Carter (1989). See also Khuri and Cornell (1996, Section 10.5).

EXAMPLE 8.6.1. Let us again consider the Michaelis–Menten model mentioned earlier. The partial derivatives of $h(x, \boldsymbol{\theta})$ with respect to θ_1 and θ_2 are

$$\begin{aligned} \frac{\partial h(x, \boldsymbol{\theta})}{\partial \theta_1} &= \frac{x}{\theta_2 + x}, \\ \frac{\partial h(x, \boldsymbol{\theta})}{\partial \theta_2} &= \frac{-\theta_1 x}{(\theta_2 + x)^2}. \end{aligned}$$

Suppose that it is desired to find a two-point design that consists of the settings x_1 and x_2 using the Box–Lucas criterion. In this case,

$$\begin{aligned} \mathbf{H}(\boldsymbol{\theta}_0) &= \begin{bmatrix} \frac{\partial h(x_1, \boldsymbol{\theta}_0)}{\partial \theta_1} & \frac{\partial h(x_1, \boldsymbol{\theta}_0)}{\partial \theta_2} \\ \frac{\partial h(x_2, \boldsymbol{\theta}_0)}{\partial \theta_1} & \frac{\partial h(x_2, \boldsymbol{\theta}_0)}{\partial \theta_2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{x_1}{\theta_{20} + x_1} & \frac{-\theta_{10} x_1}{(\theta_{20} + x_1)^2} \\ \frac{x_2}{\theta_{20} + x_2} & \frac{-\theta_{10} x_2}{(\theta_{20} + x_2)^2} \end{bmatrix}, \end{aligned}$$

where θ_{10} and θ_{20} are the elements of $\boldsymbol{\theta}_0$. In this example, $\mathbf{H}(\boldsymbol{\theta}_0)$ is a square matrix. Hence,

$$\begin{aligned}\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)] &= \{\det[\mathbf{H}(\boldsymbol{\theta}_0)]\}^2 \\ &= \frac{\theta_{10}^2 x_1^2 x_2^2 (x_2 - x_1)^2}{(\theta_{20} + x_1)^4 (\theta_{20} + x_2)^4}.\end{aligned}\quad (8.67)$$

To determine the maximum of this determinant, let us first equate its partial derivatives with respect to x_1 and x_2 to zero. It can be verified that the solution of the resulting equations (that is, the stationary point) falls outside the region of feasible values for x_1 and x_2 (both x_1 and x_2 must be nonnegative). Let us therefore restrict our search for the maximum within the region $R = \{(x_1, x_2) | 0 \leq x_1 \leq x_{\max}, 0 \leq x_2 \leq x_{\max}\}$, where $x_{\max}$ is the maximum allowable substrate concentration. Since the partial derivatives of the determinant in formula (8.67) do not vanish in R , then its maximum must be attained on the boundary of R . On $x_1 = 0$, or $x_2 = 0$, the value of the determinant is zero. If $x_1 = x_{\max}$, then

$$\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)] = \frac{\theta_{10}^2 x_{\max}^2 x_2^2 (x_2 - x_{\max})^2}{(\theta_{20} + x_{\max})^4 (\theta_{20} + x_2)^4}.$$

It can be verified that this function of x_2 has a maximum at the point $x_2 = \theta_{20} x_{\max} / (2\theta_{20} + x_{\max})$ with a value given by

$$\max_{x_1 = x_{\max}} \{\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)]\} = \frac{\theta_{10}^2 x_{\max}^6}{16\theta_{20}^2 (\theta_{20} + x_{\max})^6}.\quad (8.68)$$

Similarly, if $x_2 = x_{\max}$, then

$$\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)] = \frac{\theta_{10}^2 x_{\max}^2 x_1^2 (x_{\max} - x_1)^2}{(\theta_{20} + x_1)^4 (\theta_{20} + x_{\max})^4},$$

which attains the same maximum value as in formula (8.68) at the point $x_1 = \theta_{20} x_{\max} / (2\theta_{20} + x_{\max})$. We conclude that the maximum of $\det[\mathbf{H}'(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)]$ over the region R is achieved when $x_1 = x_{\max}$ and $x_2 = \theta_{20} x_{\max} / (2\theta_{20} + x_{\max})$, or when $x_1 = \theta_{20} x_{\max} / (2\theta_{20} + x_{\max})$ and $x_2 = x_{\max}$.

We can clearly see in this example the dependence of the design settings on θ_2 , but not on θ_1 . This is attributed to the fact that θ_1 appears linearly in the model, but θ_2 does not. In this case, the model is said to be partially nonlinear. Its D -optimal design depends only on those parameters that do not appear linearly. More details concerning partially nonlinear models can be found in Khuri and Cornell (1996, Section 10.5.3).

8.7. MULTIRESPONSE OPTIMIZATION

By definition, a multiresponse experiment is one in which a number of responses can be measured for each setting of a group of input variables. For example, in a skim milk extrusion process, the responses, y_1 = percent residual lactose and y_2 = percent ash, are known to depend on the input variables, x_1 = pH level, x_2 = temperature, x_3 = concentration, and x_4 = time (see Fichtali, Van De Voort, and Khuri, 1990).

As in single-response experiments, one of the objectives of a multiresponse experiment is the determination of conditions on the input variables that optimize the predicted responses. The definition of an optimum in a multiresponse situation, however, is more complex than in the single-response case. The reason for this is that when two or more response variables are considered simultaneously, the meaning of an optimum becomes unclear, since there is no unique way to order the values of a multiresponse function. To overcome this difficulty, Khuri and Conlon (1981) introduced a multiresponse optimization technique called the generalized distance approach. The following is an outline of this approach:

Let r be the number of responses, and n be the number of experimental runs for all the responses. Suppose that these responses can be represented by the linear models

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i, \quad i = 1, 2, \dots, r,$$

where $\mathbf{y}_i$ is a vector of observations on the i th response, $\mathbf{X}$ is a known matrix of order $n \times p$ and rank p , $\boldsymbol{\beta}_i$ is a vector of p unknown parameters, and $\boldsymbol{\epsilon}_i$ is a random error vector associated with the i th response ($i = 1, 2, \dots, r$). It is assumed that the rows of the error matrix $[\boldsymbol{\epsilon}_1: \boldsymbol{\epsilon}_2: \dots: \boldsymbol{\epsilon}_r]$ are statistically independent with each having a zero mean vector and a common variance-covariance matrix $\boldsymbol{\Sigma}$. Note that the matrix $\mathbf{X}$ is assumed to be the same for all the responses.

Let $x_1, x_2, \dots, x_k$ be input variables that influence the r responses. The predicted response value at a point $\mathbf{x} = (x_1, x_2, \dots, x_k)'$ in a region R for the i th response is given by $\hat{y}_i(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\hat{\boldsymbol{\beta}}_i$, where $\hat{\boldsymbol{\beta}}_i = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_i$ is the least-squares estimator of $\boldsymbol{\beta}_i$ ($i = 1, 2, \dots, r$). Here, $\mathbf{f}'(\mathbf{x})$ is of the same form as a row of $\mathbf{X}$, except that it is evaluated at the point $\mathbf{x}$. It follows that

$$\text{Var}[\hat{y}_i(\mathbf{x})] = \sigma_{ii}\mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x}), \quad i = 1, 2, \dots, r,$$

$$\text{Cov}[\hat{y}_i(\mathbf{x}), \hat{y}_j(\mathbf{x})] = \sigma_{ij}\mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x}), \quad i \neq j = 1, 2, \dots, r,$$

where σ_{ij} is the (i, j) th element of $\boldsymbol{\Sigma}$. The variance-covariance matrix of $\hat{\mathbf{y}}(\mathbf{x}) = [\hat{y}_1(\mathbf{x}), \hat{y}_2(\mathbf{x}), \dots, \hat{y}_r(\mathbf{x})]'$ is then of the form

$$\text{Var}[\hat{\mathbf{y}}(\mathbf{x})] = \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x})\boldsymbol{\Sigma}.$$

Since Σ is in general unknown, an unbiased estimator, $\hat{\Sigma}$, of Σ can be used instead, where

$$\hat{\Sigma} = \frac{1}{n-p} \mathbf{Y}' [\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \mathbf{Y},$$

and $\mathbf{Y} = [\mathbf{y}_1: \mathbf{y}_2: \cdots: \mathbf{y}_r]$. The matrix $\hat{\Sigma}$ is nonsingular provided that $\mathbf{Y}$ is of rank $r \leq n-p$. An estimate of $\text{Var}[\hat{\mathbf{y}}(\mathbf{x})]$ is then given by

$$\widehat{\text{Var}}[\hat{\mathbf{y}}(\mathbf{x})] = \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x})\hat{\Sigma}. \quad (8.69)$$

Let ϕ_i denote the optimum value of $\hat{y}_i(\mathbf{x})$ optimized individually over the region R ($i = 1, 2, \dots, r$). Let $\Phi = (\phi_1, \phi_2, \dots, \phi_r)'$. These individual optima do not in general occur at the same location in R . To achieve a *compromise* optimum, we need to find $\mathbf{x}$ that minimizes $\rho[\hat{\mathbf{y}}(\mathbf{x}), \Phi]$, where ρ is some metric that measures the distance of $\hat{\mathbf{y}}(\mathbf{x})$ from Φ . One possible choice for ρ is the metric

$$\rho[\hat{\mathbf{y}}(\mathbf{x}), \Phi] = \left[[\hat{\mathbf{y}}(\mathbf{x}) - \Phi]' \{ \widehat{\text{Var}}[\hat{\mathbf{y}}(\mathbf{x})] \}^{-1} [\hat{\mathbf{y}}(\mathbf{x}) - \Phi] \right]^{1/2},$$

which, by formula (8.69), can be written as

$$\rho[\hat{\mathbf{y}}(\mathbf{x}), \Phi] = \left[\frac{[\hat{\mathbf{y}}(\mathbf{x}) - \Phi]' \hat{\Sigma}^{-1} [\hat{\mathbf{y}}(\mathbf{x}) - \Phi]}{\mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\mathbf{x})} \right]^{1/2}. \quad (8.70)$$

We note that $\rho = 0$ if and only if $\hat{\mathbf{y}}(\mathbf{x}) = \Phi$, that is, when all the responses attain their individual optima at the same point; otherwise, $\rho > 0$. Such a point (if it exists) is called a point of *ideal* optimum. In general, an ideal optimum rarely exists.

In order to have conditions that are as close as possible to an ideal optimum, we need to minimize ρ over the region R . Let us suppose that the minimum occurs at the point $\mathbf{x}_0 \in R$. Then, at $\mathbf{x}_0$ the experimental conditions can be described as being near optimal for each of the r response functions. We therefore refer to $\mathbf{x}_0$ as a point of compromise optimum.

Note that the elements of Φ in formula (8.70) are random variables since they are the individual optima of $\hat{y}_1(\mathbf{x}), \hat{y}_2(\mathbf{x}), \dots, \hat{y}_r(\mathbf{x})$. If the variation associated with Φ is large, then the metric ρ may not accurately measure the deviation of $\hat{\mathbf{y}}(\mathbf{x})$ from the true ideal optimum. In this case, some account should be taken of the randomness of Φ in the development of the metric. To do so, let $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_r)'$, where ζ_i is the optimum value of the true mean of the i th response optimized individually over the region R ($i = 1, 2, \dots, r$). Let D_ζ be a confidence region for ζ . For a fixed $\mathbf{x} \in R$ and whenever $\zeta \in D_\zeta$, we obviously have

$$\rho[\hat{\mathbf{y}}(\mathbf{x}), \zeta] \leq \max_{\boldsymbol{\eta} \in D_\zeta} \rho[\hat{\mathbf{y}}(\mathbf{x}), \boldsymbol{\eta}]. \quad (8.71)$$

The right-hand side of this inequality serves as an upper bound on $\rho[\hat{\mathbf{y}}(\mathbf{x}), \boldsymbol{\zeta}]$, which represents the distance of $\hat{\mathbf{y}}(\mathbf{x})$ from the true ideal optimum. It follows that

$$\min_{\mathbf{x} \in R} \rho[\hat{\mathbf{y}}(\mathbf{x}), \boldsymbol{\zeta}] \leq \min_{\mathbf{x} \in R} \left\{ \max_{\boldsymbol{\eta} \in D_{\boldsymbol{\zeta}}} \rho[\hat{\mathbf{y}}(\mathbf{x}), \boldsymbol{\eta}] \right\}. \quad (8.72)$$

The right-hand side of this inequality provides a conservative measure of distance between the compromise and ideal optima.

The confidence region $D_{\boldsymbol{\zeta}}$ can be determined in a variety of ways. Khuri and Conlon (1981) considered a rectangular confidence region of the form

$$\gamma_{1i} \leq \zeta_i \leq \gamma_{2i}, \quad i = 1, 2, \dots, r,$$

where

$$\begin{aligned} \gamma_{1i} &= \phi_i - g_i(\boldsymbol{\xi}_i) MS_i^{1/2} t_{\alpha/2, n-p}, \\ \gamma_{2i} &= \phi_i + g_i(\boldsymbol{\xi}_i) MS_i^{1/2} t_{\alpha/2, n-p}, \end{aligned} \quad (8.73)$$

where MS_i is the error mean square for the i th response, $\boldsymbol{\xi}_i$ is the point at which $\hat{\mathbf{y}}_i(\mathbf{x})$ attains the individual optimum ϕ_i , $t_{\alpha/2, n-p}$ is the upper $(\alpha/2) \times 100$ th percentile of the t -distribution with $n-p$ degrees of freedom, and $g_i(\boldsymbol{\xi}_i)$ is given by

$$g_i(\boldsymbol{\xi}_i) = \left[\mathbf{f}'(\boldsymbol{\xi}_i)(\mathbf{X}'\mathbf{X})^{-1}\mathbf{f}(\boldsymbol{\xi}_i) \right]^{1/2}, \quad i = 1, 2, \dots, r.$$

Khuri and Conlon (1981) showed that such a rectangular confidence region has approximately a confidence coefficient of at least $1 - \alpha^*$, where $\alpha^* = 1 - (1 - \alpha)^r$.

It should be noted that the evaluation of the right-hand side of inequality (8.72) requires that $\rho[\hat{\mathbf{y}}(\mathbf{x}), \boldsymbol{\eta}]$ be maximized first with respect to $\boldsymbol{\eta}$ over $D_{\boldsymbol{\zeta}}$ for a given $\mathbf{x} \in R$. The maximum value thus obtained, being a function of $\mathbf{x}$, is then minimized over the region R . A computer program for the implementation of this min-max procedure is described in Conlon and Khuri (1992). A complete electronic copy of the code, along with examples, can be downloaded from the Internet at <ftp://ftp.stat.ufl.edu/pub/mr.tar.Z>.

Numerical examples that illustrate the application of the generalized distance approach for multiresponse optimization can be found in Khuri and Conlon (1981) and Khuri and Cornell (1996, Chapter 7).

8.8. MAXIMUM LIKELIHOOD ESTIMATION AND THE EM ALGORITHM

We recall from Section 7.11.2 that the maximum likelihood (ML) estimates of a set of parameters, $\theta_1, \theta_2, \dots, \theta_p$, for a given distribution maximize the

likelihood function of a sample, $X_1, X_2, \dots, X_{n_2}$ of size n from the distribution. The ML estimates of the θ_i 's denoted by $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$, can be found by solving the likelihood equations (the likelihood function must be differentiable and unimodal)

$$\frac{\partial \log [L(\mathbf{x}, \hat{\boldsymbol{\theta}})]}{\partial \theta_i} = 0, \quad i = 1, 2, \dots, p, \quad (8.74)$$

where $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p)'$, $\mathbf{x} = (x_1, x_2, \dots, x_n)'$, and $L(\mathbf{x}, \boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta})$ with $f(\mathbf{x}, \boldsymbol{\theta})$ being the density function (or probability mass function) of $\mathbf{X} = (X_1, X_2, \dots, X_n)'$. Note that $f(\mathbf{x}, \boldsymbol{\theta})$ can be written as $\prod_{i=1}^n g(x_i, \boldsymbol{\theta})$, where $g(x, \boldsymbol{\theta})$ is the density function (or probability mass function) associated with the distribution.

Equations (8.74) may not have a closed-form solution. For example, consider the so-called truncated Poisson distribution whose probability mass function is of the form (see Everitt, 1987, page 29)

$$g(x, \theta) = \frac{e^{-\theta} \theta^x}{(1 - e^{-\theta}) x!}, \quad x = 1, 2, \dots \quad (8.75)$$

In this case,

$$\begin{aligned} \log L(\mathbf{x}, \theta) &= \log \left[\prod_{i=1}^n g(x_i, \theta) \right] \\ &= -n\theta + (\log \theta) \sum_{i=1}^n x_i - \sum_{i=1}^n \log x_i! - n \log(1 - e^{-\theta}). \end{aligned}$$

Hence,

$$\frac{\partial L^*(\mathbf{x}, \theta)}{\partial \theta} = -n + \frac{1}{\theta} \sum_{i=1}^n x_i - \frac{ne^{-\theta}}{1 - e^{-\theta}}, \quad (8.76)$$

where $L^*(\mathbf{x}, \theta) = \log L(\mathbf{x}, \theta)$ is the log-likelihood function. The likelihood equation, which results from equating the right-hand side of formula (8.76) to zero, has no closed-form solution for θ .

In general, if equations (8.74) do not have a closed-form solution, then, as was seen in Section 8.1, iterative methods can be applied to maximize $L(\mathbf{x}, \boldsymbol{\theta})$ [or $L^*(\mathbf{x}, \boldsymbol{\theta})$]. Using, for example, the Newton–Raphson method (see Section 8.1.2), if $\hat{\boldsymbol{\theta}}_0$ is an initial estimate of $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}_i$ is the estimate at the i th iteration, then by applying formula (8.7) we have

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i - \mathbf{H}_{L^*}^{-1}(\mathbf{x}, \hat{\boldsymbol{\theta}}_i) \nabla L^*(\mathbf{x}, \hat{\boldsymbol{\theta}}_i), \quad i = 0, 1, 2, \dots,$$

where $\mathbf{H}_L^*(\mathbf{x}, \boldsymbol{\theta})$ and $\nabla L^*(\mathbf{x}, \boldsymbol{\theta})$ are, respectively, the Hessian matrix and gradient vector of the log-likelihood function. Several iterations can be made until a certain convergence criterion is satisfied. A modification of this procedure is the so-called Fisher's method of scoring, where $\mathbf{H}_L^*$ is replaced by its expected value, that is,

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i - \left\{ E[\mathbf{H}_L^*(\mathbf{x}, \hat{\boldsymbol{\theta}}_i)] \right\}^{-1} \nabla L^*(\mathbf{x}, \hat{\boldsymbol{\theta}}_i), \quad i = 0, 1, 2, \dots \quad (8.77)$$

Here, the expected value is taken with respect to the given distribution.

EXAMPLE 8.8.1. (Everitt, 1987, pages 30–31). Consider the truncated Poisson distribution described in formula (8.75). In this case, since we only have one parameter θ , the gradient takes the form $\nabla L^*(\mathbf{x}, \boldsymbol{\theta}) = \partial L^*(\mathbf{x}, \theta) / \partial \theta$, which is given by formula (8.76). Hence, the Hessian matrix is

$$\begin{aligned} \mathbf{H}_L^*(\mathbf{x}, \theta) &= \frac{\partial^2 L^*(\mathbf{x}, \theta)}{\partial \theta^2} \\ &= -\frac{1}{\theta^2} \sum_{i=1}^n x_i + \frac{ne^{-\theta}}{(1 - e^{-\theta})^2}. \end{aligned}$$

Furthermore, if X denotes the truncated Poisson random variable, then

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} \frac{xe^{-\theta}\theta^x}{(1 - e^{-\theta})x!} \\ &= \frac{\theta}{1 - e^{-\theta}}. \end{aligned}$$

Thus

$$\begin{aligned} E[\mathbf{H}_L^*(\mathbf{x}, \theta)] &= E\left[\frac{\partial^2 L^*(\mathbf{x}, \theta)}{\partial \theta^2} \right] \\ &= -\frac{1}{\theta^2} \frac{n\theta}{1 - e^{-\theta}} + \frac{ne^{-\theta}}{(1 - e^{-\theta})^2} \\ &= \frac{ne^{-\theta}(1 + \theta) - n}{\theta(1 - e^{-\theta})^2}. \end{aligned}$$

Suppose now we have the sample 1, 2, 3, 4, 5, 6 from this distribution. Let $\hat{\theta}_0 = 1.5118$ be an initial estimate of θ . Several iterations are made by applying formula (8.77), and the results are shown in Table 8.6. The final

Table 8.6. Fisher’s Method of Scoring for the Truncated Poisson Distribution

Iteration	$\frac{\partial L^*}{\partial \theta}$	$\frac{\partial^2 L^*}{\partial \theta^2}$	θ	L^*
1	−685.5137	−1176.7632	1.5118	−1545.5549
2	−62.0889	−1696.2834	0.9293	−1303.3340
3	−0.2822	−1750.5906	0.8927	−1302.1790
4	0.0012	−1750.8389	0.8925	−1302.1792

Source: Everitt (1987, page 31). Reproduced with permission of Chapman and Hall, London.

estimate of θ is 0.8925, which is considered to be the maximum likelihood estimate of θ for the given sample. The convergence criterion used here is $|\hat{\theta}_{i+1} - \hat{\theta}_i| < 0.001$.

8.8.1. The EM Algorithm

The EM algorithm is a general iterative procedure for maximum likelihood estimation in incomplete data problems. This encompasses situations involving missing data, or when the actual data are viewed as forming a subset of a larger system of quantities.

The term EM was introduced by Dempster, Laird, and Rubin (1977). The reason for this terminology is that each iteration in this algorithm consists of two steps called the expectation step (*E*-step) and the maximization step (*M*-step). In the *E*-step, the conditional expectations of the missing data are found given the observed data and the current estimates of the parameters. These expected values are then substituted for the missing data and used to complete the data. In the *M*-step, maximum likelihood estimation of the parameters is performed in the usual manner using the completed data. More generally, missing sufficient statistics can be estimated rather than the individual missing data. The estimated parameters are then used to reestimate the missing data (or missing sufficient statistics), which in turn lead to new parameter estimates. This defines an iterative procedure, which can be carried out until convergence is achieved.

More details concerning the theory of the EM algorithm can be found in Dempster, Laird, and Rubin (1977), and in Little and Rubin (1987, Chapter 7). The following two examples, given in the latter reference, illustrate the application of this algorithm:

EXAMPLE 8.8.2. (Little and Rubin, 1987, pages 130–131). Consider a sample of size n from a normal distribution with a mean μ and a variance σ^2 . Suppose that $x_1, x_2, \dots, x_m$ are observed data and that $X_{m+1}, X_{m+2}, \dots, X_n$ are missing data. Let $\mathbf{x}_{\text{obs}} = (x_1, x_2, \dots, x_m)'$. For $i = m + 1, m + 2, \dots, n$, the expected value of X_i given $\mathbf{x}_{\text{obs}}$ and $\boldsymbol{\theta} = (\mu, \sigma^2)'$ is μ . Now, from Example 7.11.3, the log-likelihood function for the complete data

set is

$$L^*(\mathbf{x}, \boldsymbol{\theta}) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2\mu \sum_{i=1}^n x_i + n\mu^2 \right), \quad (8.78)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)'$. We note that $\sum_{i=1}^n X_i^2$ and $\sum_{i=1}^n X_i$ are sufficient statistics. Therefore, to apply the E -step of the algorithm, we only have to find the conditional expectations of these statistics given $\mathbf{x}_{\text{obs}}$ and the current estimate of $\boldsymbol{\theta}$. We thus have

$$E\left(\sum_{i=1}^n X_i | \hat{\boldsymbol{\theta}}_j, \mathbf{x}_{\text{obs}}\right) = \sum_{i=1}^m x_i + (n-m)\hat{\mu}_j, \quad j = 0, 1, 2, \dots, \quad (8.79)$$

$$E\left(\sum_{i=1}^n X_i^2 | \hat{\boldsymbol{\theta}}_j, \mathbf{x}_{\text{obs}}\right) = \sum_{i=1}^m x_i^2 + (n-m)(\hat{\mu}_j^2 + \hat{\sigma}_j^2), \quad j = 0, 1, 2, \dots, \quad (8.80)$$

where $\hat{\boldsymbol{\theta}}_j = (\hat{\mu}_j, \hat{\sigma}_j^2)'$ is the estimate of $\boldsymbol{\theta}$ at the j th iteration with $\hat{\boldsymbol{\theta}}_0$ being an initial estimate.

From Section 7.11 we recall that the maximum likelihood estimates of μ and σ^2 based on the complete data set are $(1/n)\sum_{i=1}^n x_i$ and $(1/n)\sum_{i=1}^n x_i^2 - [(1/n)\sum_{i=1}^n x_i]^2$. Thus in the M -step, these same expressions are used, except that the current expectations of the sufficient statistics in formulas (8.79) and (8.80) are substituted for the missing data portion of the sufficient statistics. In other words, the estimates of μ and σ^2 at the $(j+1)$ th iteration are given by

$$\hat{\mu}_{j+1} = \frac{1}{n} \left[\sum_{i=1}^m x_i + (n-m)\hat{\mu}_j \right], \quad j = 0, 1, 2, \dots, \quad (8.81)$$

$$\hat{\sigma}_{j+1}^2 = \frac{1}{n} \left[\sum_{i=1}^m x_i^2 + (n-m)(\hat{\mu}_j^2 + \hat{\sigma}_j^2) \right] - \hat{\mu}_{j+1}^2, \quad j = 0, 1, 2, \dots \quad (8.82)$$

By setting $\hat{\mu}_j = \hat{\mu}_{j+1} = \hat{\mu}$ and $\hat{\sigma}_j = \hat{\sigma}_{j+1} = \hat{\sigma}$ in equations (8.81) and (8.82), we find that the iterations converge to

$$\begin{aligned} \hat{\mu} &= \frac{1}{m} \sum_{i=1}^m x_i, \\ \hat{\sigma}^2 &= \frac{1}{m} \sum_{i=1}^m x_i^2 - \hat{\mu}^2, \end{aligned}$$

which are the maximum likelihood estimates of μ and σ^2 from $\mathbf{x}_{\text{obs}}$.

The EM algorithm is unnecessary in this example, since the maximum likelihood estimates of μ and σ^2 can be obtained explicitly.

EXAMPLE 8.8.3. (Little and Rubin, 1987, pages 131–132). This example was originally given in Dempster, Laird, and Rubin (1977). It involves a multinomial $\mathbf{x} = (x_1, x_2, x_3, x_4)'$ with cell probabilities $(\frac{1}{2} - \frac{1}{2}\theta, \frac{1}{4}\theta, \frac{1}{4}\theta, \frac{1}{2})'$, where $0 \leq \theta \leq 1$. Suppose that the observed data consist of $\mathbf{x}_{\text{obs}} = (38, 34, 125)'$ such that $x_1 = 38, x_2 = 34, x_3 + x_4 = 125$. The likelihood function for the complete data is

$$L(\mathbf{x}, \theta) = \frac{(x_1 + x_2 + x_3 + x_4)!}{x_1!x_2!x_3!x_4!} \left(\frac{1}{2} - \frac{1}{2}\theta\right)^{x_1} \left(\frac{1}{4}\theta\right)^{x_2} \left(\frac{1}{4}\theta\right)^{x_3} \left(\frac{1}{2}\right)^{x_4}.$$

The log-likelihood function is of the form

$$\begin{aligned} L^*(\mathbf{x}, \theta) = \log \left[\frac{(x_1 + x_2 + x_3 + x_4)!}{x_1!x_2!x_3!x_4!} \right] &+ x_1 \log\left(\frac{1}{2} - \frac{1}{2}\theta\right) \\ &+ x_2 \log\left(\frac{1}{4}\theta\right) + x_3 \log\left(\frac{1}{4}\theta\right) + x_4 \log\left(\frac{1}{2}\right). \end{aligned}$$

By differentiating $L^*(\mathbf{x}, \theta)$ with respect to θ and equating the derivative to zero we obtain

$$-\frac{x_1}{1 - \theta} + \frac{x_2}{\theta} + \frac{x_3}{\theta} = 0.$$

Hence, the maximum likelihood estimate of θ for the complete data set is

$$\hat{\theta} = \frac{x_2 + x_3}{x_1 + x_2 + x_3}. \quad (8.83)$$

Let us now find the conditional expectations of X_1, X_2, X_3, X_4 given the observed data and the current estimate of θ :

$$E(X_1 | \hat{\theta}_i, \mathbf{x}_{\text{obs}}) = 38,$$

$$E(X_2 | \hat{\theta}_i, \mathbf{x}_{\text{obs}}) = 34,$$

$$E(X_3 | \hat{\theta}_i, \mathbf{x}_{\text{obs}}) = \frac{125 \left(\frac{1}{4}\hat{\theta}_i\right)}{\frac{1}{2} + \frac{1}{4}\hat{\theta}_i},$$

$$E(X_4 | \hat{\theta}_i, \mathbf{x}_{\text{obs}}) = \frac{125 \left(\frac{1}{2}\right)}{\frac{1}{2} + \frac{1}{4}\hat{\theta}_i}.$$

Table 8.7. The EM Algorithm for Example 8.8.3

Iteration	$\hat{\theta}$
0	0.500000000
1	0.608247423
2	0.624321051
3	0.626488879
4	0.626777323
5	0.626815632
6	0.626820719
7	0.626821395
8	0.626821484

Source: Little and Rubin (1987, page 132). Reproduced with permission of John Wiley & Sons, Inc.

Thus at the $(i + 1)$ st iteration we have

$$\hat{\theta}_{i+1} = \frac{34 + (125)\left(\frac{1}{4}\hat{\theta}_i\right)/\left(\frac{1}{2} + \frac{1}{4}\hat{\theta}_i\right)}{38 + 34 + (125)\left(\frac{1}{4}\hat{\theta}_i\right)/\left(\frac{1}{2} + \frac{1}{4}\hat{\theta}_i\right)}, \tag{8.84}$$

as can be seen from applying formula (8.83) using the conditional expectation of X_3 instead of x_3 . Formula (8.84) can be used iteratively to obtain the maximum likelihood estimate of θ on the basis of the observed data. Using an initial estimate $\hat{\theta}_0 = \frac{1}{2}$, the results of this iterative procedure are given in Table 8.7. Note that if we set $\hat{\theta}_{i+1} = \hat{\theta}_i = \hat{\theta}$ in formula (8.84) we obtain the quadratic equation,

$$197\hat{\theta}^2 - 15\hat{\theta} - 68 = 0$$

whose only positive root is $\hat{\theta} = 0.626821498$, which is very close to the value obtained in the last iteration in Table 8.7.

8.9. MINIMUM NORM QUADRATIC UNBAISED ESTIMATION OF VARIANCE COMPONENTS

Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \sum_{i=1}^c \mathbf{U}_i\boldsymbol{\beta}_i, \tag{8.85}$$

where $\mathbf{y}$ is a vector of n observations; $\boldsymbol{\alpha}$ is a vector of fixed effects; $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_c$ are vectors of random effects; $\mathbf{X}, \mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_c$ are known matrices of constants with $\boldsymbol{\beta}_c = \boldsymbol{\epsilon}$, the vector of random errors; and $\mathbf{U}_c = \mathbf{I}_n$. We assume that the $\boldsymbol{\beta}_i$'s are uncorrelated with zero mean vectors and variance-covariance matrices $\sigma_i^2\mathbf{I}_{m_i}$, where m_i is the number of columns of

U_i ($i = 1, 2, \dots, c$). The variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_c^2$ are referred to as variance components. Model (8.85) can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{U}\boldsymbol{\beta}, \quad (8.86)$$

where $\mathbf{U} = [\mathbf{U}_1: \mathbf{U}_2: \dots: \mathbf{U}_c]$, $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \dots, \boldsymbol{\beta}'_c)'$. From model (8.86) we have

$$\begin{aligned} E(\mathbf{y}) &= \mathbf{X}\boldsymbol{\alpha}, \\ \text{Var}(\mathbf{y}) &= \sum_{i=1}^c \sigma_i^2 \mathbf{V}_i, \end{aligned} \quad (8.87)$$

with $\mathbf{V}_i = \mathbf{U}_i \mathbf{U}'_i$.

Let us consider the estimation of a linear function of the variance components, namely, $\sum_{i=1}^c a_i \sigma_i^2$, where the a_i 's are known constants, by a quadratic estimator of the form $\mathbf{y}'\mathbf{A}\mathbf{y}$. Here, $\mathbf{A}$ is a symmetric matrix to be determined so that $\mathbf{y}'\mathbf{A}\mathbf{y}$ satisfies certain criteria, which are the following:

1. *Translation Invariance.* If instead of $\boldsymbol{\alpha}$ we consider $\boldsymbol{\gamma} = \boldsymbol{\alpha} - \boldsymbol{\alpha}_0$, then from model (8.86) we have

$$\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}_0 = \mathbf{X}\boldsymbol{\gamma} + \mathbf{U}\boldsymbol{\beta}.$$

In this case, $\sum_{i=1}^c a_i \sigma_i^2$ is estimated by $(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}_0)'\mathbf{A}(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}_0)$. The estimator $\mathbf{y}'\mathbf{A}\mathbf{y}$ is said to be translation invariant if

$$\mathbf{y}'\mathbf{A}\mathbf{y} = (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}_0)'\mathbf{A}(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}_0).$$

In order for this to be true we must have

$$\mathbf{A}\mathbf{X} = \mathbf{0}. \quad (8.88)$$

2. *Unbiasedness.* $E(\mathbf{y}'\mathbf{A}\mathbf{y}) = \sum_{i=1}^c a_i \sigma_i^2$. Using a result in Searle (1971, Theorem 1, page 55), the expected value of the quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ is given by

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) = \boldsymbol{\alpha}'\mathbf{X}'\mathbf{A}\mathbf{X}\boldsymbol{\alpha} + \text{tr}[\mathbf{A}\text{Var}(\mathbf{y})], \quad (8.89)$$

since $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\alpha}$. From formulas (8.87), (8.88), and (8.89) we then have

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) = \sum_{i=1}^c \sigma_i^2 \text{tr}(\mathbf{A}\mathbf{V}_i). \quad (8.90)$$

By comparison with $\sum_{i=1}^c a_i \sigma_i^2$, the condition for unbiasedness is

$$a_i = \text{tr}(\mathbf{A}\mathbf{V}_i), \quad i = 1, 2, \dots, c. \quad (8.91)$$

3. *Minimum Norm.* If $\beta_1, \beta_2, \dots, \beta_c$ in model (8.85) were observable, then a natural unbiased estimator of $\sum_{i=1}^c a_i \sigma_i^2$ would be $\sum_{i=1}^c a_i \beta_i' \beta_i / m_i$, since $E(\beta_i' \beta_i) = \text{tr}(\mathbf{I}_{m_i} \sigma_i^2) = m_i \sigma_i^2$, $i = 1, 2, \dots, c$. This estimator can be written as $\beta' \Delta \beta$, where Δ is the block-diagonal matrix

$$\Delta = \text{Diag} \left(\frac{a_1}{m_1} \mathbf{I}_{m_1}, \frac{a_2}{m_2} \mathbf{I}_{m_2}, \dots, \frac{a_c}{m_c} \mathbf{I}_{m_c} \right).$$

The difference between this estimator and $\mathbf{y}' \mathbf{A} \mathbf{y}$ is

$$\mathbf{y}' \mathbf{A} \mathbf{y} - \beta' \Delta \beta = \beta' (\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta) \beta,$$

since $\mathbf{A} \mathbf{X} = \mathbf{0}$. This difference can be made small by minimizing the Euclidean norm $\|\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta\|_2$.

The quadratic estimator $\mathbf{y}' \mathbf{A} \mathbf{y}$ is said to be a minimum norm quadratic unbiased estimator (MINQUE) of $\sum_{i=1}^c a_i \sigma_i^2$ if the matrix $\mathbf{A}$ is determined so that $\|\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta\|_2$ attains a minimum subject to the conditions given in formulas (8.88) and (8.91). Such an estimator was introduced by Rao (1971, 1972).

The minimization of $\|\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta\|_2$ is equivalent to that of $\text{tr}(\mathbf{A} \mathbf{V} \mathbf{A} \mathbf{V})$, where $\mathbf{V} = \sum_{i=1}^c \mathbf{V}_i$. The reason for this is the following:

$$\begin{aligned} \|\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta\|_2^2 &= \text{tr}[(\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta)(\mathbf{U}' \mathbf{A} \mathbf{U} - \Delta)] \\ &= \text{tr}(\mathbf{U}' \mathbf{A} \mathbf{U} \mathbf{U}' \mathbf{A} \mathbf{U}) - 2 \text{tr}(\mathbf{U}' \mathbf{A} \mathbf{U} \Delta) + \text{tr}(\Delta^2). \end{aligned} \quad (8.92)$$

Now,

$$\begin{aligned} \text{tr}(\mathbf{U}' \mathbf{A} \mathbf{U} \Delta) &= \text{tr}(\mathbf{A} \mathbf{U} \Delta \mathbf{U}') \\ &= \text{tr} \left(\mathbf{A} \sum_{i=1}^c \mathbf{U}_i \frac{a_i}{m_i} \mathbf{I}_{m_i} \mathbf{U}_i' \right) \\ &= \text{tr} \left(\sum_{i=1}^c \frac{a_i}{m_i} \mathbf{A} \mathbf{U}_i \mathbf{U}_i' \right) \\ &= \text{tr} \left(\sum_{i=1}^c \frac{a_i}{m_i} \mathbf{A} \mathbf{V}_i \right) \\ &= \sum_{i=1}^c \frac{a_i}{m_i} \text{tr}(\mathbf{A} \mathbf{V}_i) \\ &= \sum_{i=1}^c \frac{a_i^2}{m_i}, \quad \text{by (8.91)} \\ &= \text{tr}(\Delta^2). \end{aligned}$$

Formula (8.92) can then be written as

$$\begin{aligned}\|\mathbf{U}'\mathbf{A}\mathbf{U} - \mathbf{\Delta}\|_2^2 &= \text{tr}(\mathbf{U}'\mathbf{A}\mathbf{U}\mathbf{U}'\mathbf{A}\mathbf{U}) - \text{tr}(\mathbf{\Delta}^2) \\ &= \text{tr}(\mathbf{A}\mathbf{V}\mathbf{A}\mathbf{V}) - \text{tr}(\mathbf{\Delta}^2),\end{aligned}$$

since $\mathbf{V} = \sum_{i=1}^c \mathbf{V}_i = \sum_{i=1}^c \mathbf{U}_i \mathbf{U}_i' = \mathbf{U}\mathbf{U}'$. The trace of $\mathbf{\Delta}^2$ does not involve $\mathbf{A}$; hence the problem of MINQUE reduces to finding $\mathbf{A}$ that minimizes $\text{tr}(\mathbf{A}\mathbf{V}\mathbf{A}\mathbf{V})$ subject to conditions (8.88) and (8.91). Rao (1971) showed that the solution to this optimization problem is of the form

$$\mathbf{A} = \sum_{i=1}^c \lambda_i \mathbf{R}\mathbf{V}_i \mathbf{R}, \quad (8.93)$$

where

$$\mathbf{R} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1} \mathbf{X})^{-} \mathbf{X}'\mathbf{V}^{-1}$$

with $(\mathbf{X}'\mathbf{V}^{-1} \mathbf{X})^{-}$ being a generalized inverse of $\mathbf{X}'\mathbf{V}^{-1} \mathbf{X}$, and the λ_i 's are obtained from solving the equations

$$\sum_{i=1}^c \lambda_i \text{tr}(\mathbf{R}\mathbf{V}_i \mathbf{R}\mathbf{V}_j) = a_j, \quad j = 1, 2, \dots, c,$$

which can be expressed as

$$\boldsymbol{\lambda}' \mathbf{S} = \mathbf{a}', \quad (8.94)$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_c)'$, $\mathbf{S}$ is the $c \times c$ matrix (s_{ij}) with $s_{ij} = \text{tr}(\mathbf{R}\mathbf{V}_i \mathbf{R}\mathbf{V}_j)$, and $\mathbf{a} = (a_1, a_2, \dots, a_c)'$. The MINQUE of $\sum_{i=1}^c a_i \sigma_i^2$ can then be written as

$$\begin{aligned}\mathbf{y}' \left(\sum_{i=1}^c \lambda_i \mathbf{R}\mathbf{V}_i \mathbf{R} \right) \mathbf{y} &= \sum_{i=1}^c \lambda_i \mathbf{y}' \mathbf{R}\mathbf{V}_i \mathbf{R} \mathbf{y} \\ &= \boldsymbol{\lambda}' \mathbf{q},\end{aligned}$$

where $\mathbf{q} = (q_1, q_2, \dots, q_c)'$ with $q_i = \mathbf{y}' \mathbf{R}\mathbf{V}_i \mathbf{R} \mathbf{y}$ ($i = 1, 2, \dots, c$). But, from formula (8.94), $\boldsymbol{\lambda}' = \mathbf{a}' \mathbf{S}^{-}$, where $\mathbf{S}^{-}$ is a generalized inverse of $\mathbf{S}$. Hence, $\boldsymbol{\lambda}' \mathbf{q} = \mathbf{a}' \mathbf{S}^{-} \mathbf{q} = \mathbf{a}' \hat{\boldsymbol{\sigma}}$, where $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_c^2)'$ is a solution of the equation

$$\mathbf{S} \hat{\boldsymbol{\sigma}} = \mathbf{q}. \quad (8.95)$$

This equation has a unique solution if and only if the individual variance components are unbiasedly estimable (see Rao, 1972, page

114). Thus the MINQUEs of the σ_i^2 's are obtained from solving equation (8.95).

If the random effects in model (8.85) are assumed to be normally distributed, then the MINQUEs of the variance components reduce to the so-called minimum variance quadratic unbiased estimators (MIVQUEs). An example that shows how to compute these estimators in the case of a random one-way classification model is given in Swallow and Searle (1978). See also Milliken and Johnson (1984, Chapter 19).

8.10. SCHEFFÉ'S CONFIDENCE INTERVALS

Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (8.96)$$

where $\mathbf{y}$ is a vector of n observations, $\mathbf{X}$ is a known matrix of order $n \times p$ and $\text{rank } r (\leq p)$, $\boldsymbol{\beta}$ is a vector of unknown parameters, and $\boldsymbol{\epsilon}$ is a random error vector. It is assumed that $\boldsymbol{\epsilon}$ has the normal distribution with a mean $\mathbf{0}$ and a variance-covariance matrix $\sigma^2 \mathbf{I}_n$. Let $\psi = \mathbf{a}'\boldsymbol{\beta}$ be an estimable linear function of the elements of $\boldsymbol{\beta}$. By this we mean that there exists a linear function $\mathbf{t}'\mathbf{y}$ of $\mathbf{y}$ such that $E(\mathbf{t}'\mathbf{y}) = \psi$, where $\mathbf{t}$ is some constant vector. A necessary and sufficient condition for ψ to be estimable is that $\mathbf{a}'$ belongs to the row space of $\mathbf{X}$, that is, $\mathbf{a}'$ is a linear combination of the rows of $\mathbf{X}$ (see, for example, Searle, 1971, page 181). Since the rank of $\mathbf{X}$ is r , the row space of $\mathbf{X}$, denoted by $\rho(\mathbf{X})$, is an r -dimensional subspace of the p -dimensional Euclidean space R^p . Thus $\mathbf{a}'\boldsymbol{\beta}$ estimable if and only if $\mathbf{a}' \in \rho(\mathbf{X})$.

Suppose that $\mathbf{a}'$ is an arbitrary vector in a q -dimensional subspace $\mathcal{L}$ of $\rho(\mathbf{X})$, where $q \leq r$. Then $\mathbf{a}'\hat{\boldsymbol{\beta}} = \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}$ is the best linear unbiased estimator of $\mathbf{a}'\boldsymbol{\beta}$, and its variance is given by

$$\text{Var}(\mathbf{a}'\hat{\boldsymbol{\beta}}) = \sigma^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-} \mathbf{a},$$

where $(\mathbf{X}'\mathbf{X})^{-}$ is a generalized inverse of $\mathbf{X}'\mathbf{X}$ (see, for example, Searle, 1971, pages 181–182). Both $\mathbf{a}'\hat{\boldsymbol{\beta}}$ and $\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-} \mathbf{a}$ are invariant to the choice of $(\mathbf{X}'\mathbf{X})^{-}$, since $\mathbf{a}'\boldsymbol{\beta}$ is estimable (see, for example, Searle, 1971, page 181). In particular, if $r = p$, then $\mathbf{X}'\mathbf{X}$ is of full rank and $(\mathbf{X}'\mathbf{X})^{-} = (\mathbf{X}'\mathbf{X})^{-1}$.

Theorem 8.10.1. Simultaneous $(1 - \alpha)100\%$ confidence intervals on $\mathbf{a}'\boldsymbol{\beta}$ for all $\mathbf{a}' \in \mathcal{L}$, where $\mathcal{L}$ is a q -dimensional subspace of $\rho(\mathbf{X})$, are of the form

$$\mathbf{a}'\hat{\boldsymbol{\beta}} \mp (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-} \mathbf{a}]^{1/2}, \quad (8.97)$$

where $F_{\alpha, q, n-r}$ is the upper α 100th percentile of the F -distribution with q and $n-r$ degrees of freedom, and MS_E is the error mean square given by

$$MS_E = \frac{1}{n-r} \mathbf{y}' [\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \mathbf{y}. \quad (8.98)$$

In Theorem 8.10.1, the word “simultaneous” means that with probability $1 - \alpha$, the values of $\mathbf{a}'\boldsymbol{\beta}$ for all $\mathbf{a}' \in \mathcal{L}$ satisfy the double inequality

$$\begin{aligned} \mathbf{a}'\hat{\boldsymbol{\beta}} - (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}]^{1/2} \\ \leq \mathbf{a}'\boldsymbol{\beta} \leq \mathbf{a}'\hat{\boldsymbol{\beta}} + (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}]^{1/2}. \end{aligned} \quad (8.99)$$

A proof of this theorem is given in Scheffé (1959, Section 3.5). Another proof is presented here using the method of Lagrange multipliers. This proof is based on the following lemma:

Lemma 8.10.1. Let C be the set $\{\mathbf{x} \in R^q | \mathbf{x}'\mathbf{A}\mathbf{x} \leq 1\}$, where $\mathbf{A}$ is a positive definite matrix of order $q \times q$. Then $\mathbf{x} \in C$ if and only if $|\mathbf{l}'\mathbf{x}| \leq (\mathbf{l}'\mathbf{A}^{-1}\mathbf{l})^{1/2}$ for all $\mathbf{l} \in R^q$.

Proof. Suppose that $\mathbf{x} \in C$. Since $\mathbf{A}$ is positive definite, the boundary of C is an ellipsoid in a q -dimensional space. For any $\mathbf{l} \in R^q$, let $\mathbf{e}$ be a unit vector in its direction. The projection of $\mathbf{x}$ on an axis in the direction of $\mathbf{l}$ is given by $\mathbf{e}'\mathbf{x}$. Consider optimizing $\mathbf{e}'\mathbf{x}$ with respect to $\mathbf{x}$ over the set C . The minimum and maximum values of $\mathbf{e}'\mathbf{x}$ are obviously determined by the end points of the projection of C on the $\mathbf{l}$ -axis. This is equivalent to optimizing $\mathbf{e}'\mathbf{x}$ subject to the constraint $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$, since the projection of C on the $\mathbf{l}$ -axis is the same as the projection of its boundary, the ellipsoid $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$. This constrained optimization problem can be solved by using the method of Lagrange multipliers.

Let $G = \mathbf{e}'\mathbf{x} + \lambda(\mathbf{x}'\mathbf{A}\mathbf{x} - 1)$, where λ is a Lagrange multiplier. By differentiating G with respect to $x_1, x_2, \dots, x_q$, where x_i is the i th element of $\mathbf{x}$ ($i = 1, 2, \dots, q$), and equating the derivatives to zero, we obtain the equation $\mathbf{e} + 2\lambda\mathbf{A}\mathbf{x} = \mathbf{0}$, whose solution is $\mathbf{x} = -(1/2\lambda)\mathbf{A}^{-1}\mathbf{e}$. If we substitute this value of $\mathbf{x}$ into the equation $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$ and then solve for λ , we obtain the two solutions $\lambda_1 = -\frac{1}{2}(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^{1/2}$, $\lambda_2 = \frac{1}{2}(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^{1/2}$. But, $\mathbf{e}'\mathbf{x} = -2\lambda$, since $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$. It follows that the minimum and maximum values of $\mathbf{e}'\mathbf{x}$ under the constraint $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$ are $-(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^{1/2}$ and $(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^{1/2}$, respectively. Hence,

$$|\mathbf{e}'\mathbf{x}| \leq (\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^{1/2}. \quad (8.100)$$

Since $\mathbf{l} = \|\mathbf{l}\|_2 \mathbf{e}$, where $\|\mathbf{l}\|_2$ is the Euclidean norm of $\mathbf{l}$, multiplying the two

sides of inequality (8.100) by $\|\mathbf{l}\|_2$ yields

$$|\mathbf{l}'\mathbf{x}| \leq (\mathbf{l}'\mathbf{A}^{-1}\mathbf{l})^{1/2}. \quad (8.101)$$

Vice versa, if inequality (8.101) is true for all $\mathbf{l} \in R^q$, then by choosing $\mathbf{l}' = \mathbf{x}'\mathbf{A}$ we obtain

$$|\mathbf{x}'\mathbf{A}\mathbf{x}| \leq (\mathbf{x}'\mathbf{A}\mathbf{A}^{-1}\mathbf{A}\mathbf{x})^{1/2},$$

which is equivalent to $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 1$, that is, $\mathbf{x} \in C$. $\square$

Proof of Theorem 8.10.1. Let $\mathbf{L}$ be a $q \times p$ matrix of rank q whose rows form a basis for the q -dimensional subspace $\mathcal{L}$ of $\rho(\mathbf{X})$. Since $\mathbf{y}$ in model (8.96) is distributed as $N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n)$, $\mathbf{L}\hat{\boldsymbol{\beta}} = \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ is distributed as $N[\mathbf{L}\boldsymbol{\beta}, \sigma^2\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']$. Thus the random variable

$$F = \frac{[\mathbf{L}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]' [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1} [\mathbf{L}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]}{qMS_E}$$

has the F -distribution with q and $n - r$ degrees of freedom (see, for example, Searle, 1971, page 190). It follows that

$$P(F \leq F_{\alpha, q, n-r}) = 1 - \alpha. \quad (8.102)$$

By applying Lemma 8.10.1 to formula (8.102) with $\mathbf{x} = \mathbf{L}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ and $\mathbf{A} = [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1}/(qMS_E F_{\alpha, q, n-r})$ we obtain the equivalent probability statement

$$P\left\{|\mathbf{l}'\mathbf{L}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})| \leq (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{l}'\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}\mathbf{l}]^{1/2} \quad \forall \mathbf{l} \in R^q\right\} = 1 - \alpha.$$

Let $\mathbf{a}' = \mathbf{l}'\mathbf{L}$. We then have

$$P\left\{|\mathbf{a}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})| \leq (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}]^{1/2} \quad \forall \mathbf{a}' \in \mathcal{L}\right\} = 1 - \alpha.$$

We conclude that the values of $\mathbf{a}'\boldsymbol{\beta}$ satisfy the double inequality (8.99) for all $\mathbf{a}' \in \mathcal{L}$ with probability $1 - \alpha$. Simultaneous $(1 - \alpha)100\%$ confidence intervals on $\mathbf{a}'\boldsymbol{\beta}$ are therefore given by formula (8.97). We refer to these intervals as Scheffé's confidence intervals.

Theorem 8.10.1 can be used to obtain simultaneous confidence intervals on all contrasts among the elements of $\boldsymbol{\beta}$. By definition, the linear function $\mathbf{a}'\boldsymbol{\beta}$ is a contrast among the elements of $\boldsymbol{\beta}$ if $\sum_{i=1}^p a_i = 0$, where a_i is the i th element of $\mathbf{a}$ ($i = 1, 2, \dots, p$). If $\mathbf{a}'$ is in the row space of $\mathbf{X}$, then it must belong to a q -dimensional subspace of $\rho(\mathbf{X})$, where $q = r - 1$. Hence, simul-

taneous $(1 - \alpha)100\%$ confidence intervals on all such contrasts can be obtained from formula (8.97) by replacing q with $r - 1$. $\square$

8.10.1. The Relation of Scheffé's Confidence Intervals to the F -Test

There is a relationship between the confidence intervals (8.97) and the F -test used to test the hypothesis $H_0: \mathbf{L}\boldsymbol{\beta} = \mathbf{0}$ versus $H_a: \mathbf{L}\boldsymbol{\beta} \neq \mathbf{0}$, where $\mathbf{L}$ is the matrix whose rows form a basis for the q -dimensional subspace $\mathcal{L}$ of $\rho(\mathbf{X})$. The test statistic for testing H_0 is given by (see Searle, 1971, Section 5.5)

$$F = \frac{\hat{\boldsymbol{\beta}}' \mathbf{L}' [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{L}']^{-1} \mathbf{L} \hat{\boldsymbol{\beta}}}{q MS_E},$$

which under H_0 has the F -distribution with q and $n - r$ degrees of freedom. The hypothesis H_0 can be rejected at the α -level of significance if $F > F_{\alpha, q, n-r}$. In this case, by Lemma 8.10.1, there exists at least one $\mathbf{l} \in R^q$ such that

$$|\mathbf{l}' \mathbf{L} \hat{\boldsymbol{\beta}}| > (q MS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{l}' \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{L} \mathbf{l}]^{1/2}. \quad (8.103)$$

It follows that the F -test rejects H_0 if and only if there exists a linear combination $\mathbf{a}' \hat{\boldsymbol{\beta}}$, where $\mathbf{a}' = \mathbf{l}' \mathbf{L}$ for some $\mathbf{l} \in R^q$, for which the confidence interval in formula (8.97) does not contain the value zero. In this case, $\mathbf{a}' \hat{\boldsymbol{\beta}}$ is said to be significantly different from zero.

It is easy to see that inequality (8.103) holds for some $\mathbf{l} \in R^q$ if and only if

$$\sup_{\mathbf{l} \in R^q} \frac{|\mathbf{l}' \mathbf{L} \hat{\boldsymbol{\beta}}|^2}{\mathbf{l}' \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{L} \mathbf{l}} > q MS_E F_{\alpha, q, n-r},$$

or equivalently,

$$\sup_{\mathbf{l} \in R^q} \frac{\mathbf{l}' \mathbf{G}_1 \mathbf{l}}{\mathbf{l}' \mathbf{G}_2 \mathbf{l}} > q MS_E F_{\alpha, q, n-r}, \quad (8.104)$$

where

$$\mathbf{G}_1 = \mathbf{L} \hat{\boldsymbol{\beta}} \hat{\boldsymbol{\beta}}' \mathbf{L}', \quad (8.105)$$

$$\mathbf{G}_2 = \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{L}'. \quad (8.106)$$

However, by Theorem 2.3.17,

$$\begin{aligned} \sup_{\mathbf{l} \in R^q} \frac{\mathbf{l}' \mathbf{G}_1 \mathbf{l}}{\mathbf{l}' \mathbf{G}_2 \mathbf{l}} &= e_{\max}(\mathbf{G}_2^{-1} \mathbf{G}_1) \\ &= \hat{\boldsymbol{\beta}}' \mathbf{L}' [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{L}']^{-1} \mathbf{L} \hat{\boldsymbol{\beta}}, \end{aligned} \quad (8.107)$$

where $e_{\max}(\mathbf{G}_2^{-1}\mathbf{G}_1)$ is the largest eigenvalue of $\mathbf{G}_2^{-1}\mathbf{G}_1$. The second equality in (8.107) is true because the nonzero eigenvalues of $[\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}'\mathbf{L}'$ are the same as those of $\hat{\boldsymbol{\beta}}'\mathbf{L}'[\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}$ by Theorem 2.3.9. Note that the latter expression is the numerator sum of squares of the F -test statistic for H_0 .

The eigenvector of $\mathbf{G}_2^{-1}\mathbf{G}_1$ corresponding to $e_{\max}(\mathbf{G}_2^{-1}\mathbf{G}_1)$ is of special interest. Let $\mathbf{I}^*$ be such an eigenvector. Then

$$\frac{\mathbf{I}^{*'}\mathbf{G}_1\mathbf{I}^*}{\mathbf{I}^{*'}\mathbf{G}_2\mathbf{I}^*} = e_{\max}(\mathbf{G}_2^{-1}\mathbf{G}_1). \quad (8.108)$$

This follows from the fact that $\mathbf{I}^*$ satisfies the equation

$$(\mathbf{G}_1 - e_{\max}\mathbf{G}_2)\mathbf{I}^* = \mathbf{0},$$

where $e_{\max}$ is an abbreviation for $e_{\max}(\mathbf{G}_2^{-1}\mathbf{G}_1)$. It is easy to see that $\mathbf{I}^*$ can be chosen to be the vector $\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}$, since

$$\begin{aligned} \mathbf{G}_2^{-1}\mathbf{G}_1(\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}) &= \mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}'\mathbf{L}'(\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}) \\ &= (\hat{\boldsymbol{\beta}}'\mathbf{L}'\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}})\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}} \\ &= e_{\max}\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}. \end{aligned}$$

This shows that $\mathbf{G}_2^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}$ is an eigenvector of $\mathbf{G}_2^{-1}\mathbf{G}_1$ for the eigenvalue $e_{\max}$.

From inequality (8.104) and formula (8.108) we conclude that if the F -test rejects H_0 at the α -level, then

$$|\mathbf{I}^{*'}\mathbf{L}\hat{\boldsymbol{\beta}}| > (qMS_E F_{\alpha, q, n-r})^{1/2} [\mathbf{I}^{*'}\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'\mathbf{I}^*]^{1/2}.$$

This means that the linear combination $\mathbf{a}^{*'}\hat{\boldsymbol{\beta}}$, where $\mathbf{a}^{*'} = \mathbf{I}^{*'}\mathbf{L}$, is significantly different from zero. Let us express $\mathbf{a}^{*'}\hat{\boldsymbol{\beta}}$ as

$$\mathbf{I}^{*'}\mathbf{L}\hat{\boldsymbol{\beta}} = \sum_{i=1}^q l_i^* \hat{\gamma}_i, \quad (8.109)$$

where l_i^* and $\hat{\gamma}_i$ are the i th elements of $\mathbf{I}^*$ and $\hat{\boldsymbol{\gamma}} = \mathbf{L}\hat{\boldsymbol{\beta}}$, respectively ($i = 1, 2, \dots, q$). If we divide $\hat{\gamma}_i$ by its estimated standard error $\hat{\kappa}_i$ [which is equal to the square root of the i th diagonal element of the variance-covariance matrix of $\mathbf{L}\hat{\boldsymbol{\beta}}$, namely, $\sigma^2\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'$ with σ^2 replaced by the error mean square MS_E in formula (8.98)], then formula (8.109) can be written as

$$\mathbf{I}^{*'}\mathbf{L}\hat{\boldsymbol{\beta}} = \sum_{i=1}^q l_i^* \hat{\kappa}_i \hat{\tau}_i, \quad (8.110)$$

where $\hat{\tau}_i = \hat{\gamma}_i / \hat{\kappa}_i$, $i = 1, 2, \dots, q$. Consequently, large values of $|l_i^*| \hat{\kappa}_i$ identify those elements of $\hat{\boldsymbol{\gamma}}$ that are influential contributors to the significance of

the F -test concerning H_0 . Note that the elements of $\boldsymbol{\gamma} = \mathbf{L}\boldsymbol{\beta}$ form a set of linearly independent estimable linear functions of $\boldsymbol{\beta}$.

We conclude from the previous arguments that the eigenvector $\mathbf{I}^*$, which corresponds to the largest eigenvalue of $\mathbf{G}_2^{-1}\mathbf{G}_1$, can be conveniently used to identify an estimable linear function of $\boldsymbol{\beta}$ that is significantly different from zero whenever the F -test rejects H_0 .

It should be noted that if model (8.96) is a response surface model (in this case, the matrix $\mathbf{X}$ in the model is of full column rank, that is, $r = p$) whose input variables, $x_1, x_2, \dots, x_k$, have different units of measurement, then these variables must be made scale free. This is accomplished as follows: If x_{ui} denotes the u th measurement on x_i , then we may consider the transformation

$$z_{ui} = \frac{x_{ui} - \bar{x}_i}{s_i}, \quad i = 1, 2, \dots, k; u = 1, 2, \dots, n,$$

where $\bar{x}_i = (1/n)\sum_{u=1}^n x_{ui}$, $s_i = [\sum_{u=1}^n (x_{ui} - \bar{x}_i)^2]^{1/2}$, and n is the total number of observations. One advantage of this scaling convention, besides making the input variables scale free, is that it can greatly improve the conditioning of the matrix $\mathbf{X}$ with regard to multicollinearity (see, for example, Belsley, Kuh, and Welsch, 1980, pages 183–185).

EXAMPLE 8.10.1. Let us consider the one-way classification model

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n_i, \quad (8.111)$$

where μ and α_i are unknown parameters with the latter representing the effect of the i th level of a certain factor at m levels; n_i observations are obtained at the i th level. The ϵ_{ij} 's are random errors assumed to be independent and normally distributed with zero means and a common variance σ^2 .

Model (8.111) can be represented in vector form as model (8.96). Here, $\mathbf{y} = (y_{11}, y_{12}, \dots, y_{1n_1}, y_{21}, y_{22}, \dots, y_{2n_2}, \dots, y_{m1}, y_{m2}, \dots, y_{mn_m})'$, $\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, \dots, \alpha_m)'$, and $\mathbf{X}$ is of order $n \times (m+1)$ of the form $\mathbf{X} = [\mathbf{1}_n : \mathbf{T}]$, where $\mathbf{1}_n$ is a vector of ones of order $n \times 1$, $n = \sum_{i=1}^m n_i$, and $\mathbf{T} = \text{Diag}(\mathbf{1}_{n_1}, \mathbf{1}_{n_2}, \dots, \mathbf{1}_{n_m})$. The rank of $\mathbf{X}$ is $r = m$. For such a model, the hypothesis of interest is

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_m,$$

which can be expressed as $H_0: \mathbf{L}\boldsymbol{\beta} = \mathbf{0}$, where $\mathbf{L}$ is a matrix of order

$(m-1) \times (m+1)$ and rank $m-1$ of the form

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

This hypothesis states that the factor under consideration has no effect on the response. Note that each row of $\mathbf{L}$ is a linear combination of the rows of $\mathbf{X}$. For example, the i th row of $\mathbf{L}$ —whose elements are equal to zero except for the second and the $(i+2)$ th elements, which are equal to 1 and -1 , respectively—is the difference between rows 1 and $\nu_i + 1$ of $\mathbf{X}$, where $\nu_i = \sum_{j=1}^i n_j$, $i = 1, 2, \dots, m-1$. Thus the rows of $\mathbf{L}$ form a basis for a q -dimensional subspace $\mathcal{L}$ of $\rho(\mathbf{X})$, the row space of $\mathbf{X}$, where $q = m-1$.

Let $\mu_i = \mu + \alpha_i$. Then μ_i is the mean of the i th level of the factor ($i = 1, 2, \dots, m$). Consider the contrast $\psi = \sum_{i=1}^m c_i \mu_i$, that is, $\sum_{i=1}^m c_i = 0$. We can write $\psi = \mathbf{a}'\boldsymbol{\beta}$, where $\mathbf{a}' = (0, c_1, c_2, \dots, c_m)$ belongs to a q -dimensional subspace of R^{m+1} . This subspace is the same as $\mathcal{L}$, since each row of $\mathbf{L}$ is of the form $(0, c_1, c_2, \dots, c_m)$ with $\sum_{i=1}^m c_i = 0$. Vice versa, if $\psi = \mathbf{a}'\boldsymbol{\beta}$ is such that $\mathbf{a}' = (0, c_1, c_2, \dots, c_m)$ with $\sum_{i=1}^m c_i = 0$, then $\mathbf{a}'$ can be expressed as

$$\mathbf{a}' = (-c_2, -c_3, \dots, -c_m)\mathbf{L},$$

since $c_1 = -\sum_{i=2}^m c_i$. Hence, $\mathbf{a}' \in \mathcal{L}$. It follows that $\mathcal{L}$ is a subspace associated with all contrasts among the means $\mu_1, \mu_2, \dots, \mu_m$ of the m levels of the factor.

Simultaneous $(1-\alpha)100\%$ confidence intervals on all contrasts of the form $\psi = \sum_{i=1}^m c_i \mu_i$ can be obtained by applying formula (8.97). Here, $q = m-1$, $r = m$, and a generalized inverse of $\mathbf{X}'\mathbf{X}$ is of the form

$$(\mathbf{X}'\mathbf{X})^- = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \mathbf{D} \end{bmatrix},$$

where $\mathbf{D} = \text{Diag}(n_1^{-1}, n_2^{-1}, \dots, n_m^{-1})$ and $\mathbf{0}$ is a zero vector of order $m \times 1$. Hence,

$$\begin{aligned} \mathbf{a}'\hat{\boldsymbol{\beta}} &= (0, c_1, c_2, \dots, c_m)(\mathbf{X}'\mathbf{X})^- \mathbf{X}'\mathbf{y} \\ &= \sum_{i=1}^m c_i \bar{y}_i, \end{aligned}$$

where $\bar{y}_i = (1/n_i) \sum_{j=1}^{n_i} y_{ij}$, $i = 1, 2, \dots, m$. Furthermore,

$$\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a} = \sum_{i=1}^m \frac{c_i^2}{n_i}.$$

By making the substitution in formula (8.97) we obtain

$$\sum_{i=1}^m c_i \bar{y}_i \mp [(m-1) MS_E F_{\alpha, m-1, n-m}]^{1/2} \left(\sum_{i=1}^m \frac{c_i^2}{n_i} \right)^{1/2}. \quad (8.112)$$

Now, if the F -test rejects H_0 at the α -level, then there exists a contrast $\sum_{i=1}^m c_i^* \bar{y}_i = \mathbf{a}^{*'} \hat{\boldsymbol{\beta}}$, which is significantly different from zero, that is, the interval (8.112) for $c_i = c_i^*$ ($i = 1, 2, \dots, m$) does not contain the value zero. Here, $\mathbf{a}^{*'} = \mathbf{l}^{*'} \mathbf{L}$, where $\mathbf{l}^* = \mathbf{G}_2^{-1} \mathbf{L} \boldsymbol{\beta}$ is an eigenvector of $\mathbf{G}_2^{-1} \mathbf{G}_1$ corresponding to $e_{\max}(\mathbf{G}_2^{-1} \mathbf{G}_1)$. We have that

$$\mathbf{G}_2^{-1} = [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1} = \left(\frac{1}{n_1} \mathbf{J}_{m-1} + \boldsymbol{\Lambda} \right)^{-1},$$

where $\mathbf{J}_{m-1}$ is a matrix of ones of order $(m-1) \times (m-1)$, and $\boldsymbol{\Lambda} = \text{Diag}(n_2^{-1}, n_3^{-1}, \dots, n_m^{-1})$. By applying the Sherman–Morrison–Woodbury formula (see Exercise 2.15), we obtain

$$\begin{aligned} \mathbf{G}_2^{-1} &= n_1(n_1 \boldsymbol{\Lambda} + \mathbf{1}_{m-1} \mathbf{1}_{m-1}')^{-1} \\ &= \boldsymbol{\Lambda}^{-1} - \frac{\boldsymbol{\Lambda}^{-1} \mathbf{1}_{m-1} \mathbf{1}_{m-1}' \boldsymbol{\Lambda}^{-1}}{n_1 + \mathbf{1}_{m-1}' \boldsymbol{\Lambda}^{-1} \mathbf{1}_{m-1}} \\ &= \text{Diag}(n_2, n_2, \dots, n_m) - \frac{1}{n} \begin{bmatrix} n_2 \\ n_3 \\ \vdots \\ n_m \end{bmatrix} \begin{bmatrix} n_2 & n_3 & \dots & n_m \end{bmatrix}. \end{aligned}$$

Also,

$$\hat{\boldsymbol{\gamma}} = \mathbf{L} \hat{\boldsymbol{\beta}} = \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \begin{bmatrix} \bar{y}_1 - \bar{y}_2 \\ \bar{y}_1 - \bar{y}_3 \\ \vdots \\ \bar{y}_1 - \bar{y}_m \end{bmatrix}.$$

It can be verified that the i th element of $\mathbf{l}^* = \mathbf{G}_2^{-1} \mathbf{L} \hat{\boldsymbol{\beta}}$ is given by

$$l_i^* = n_{i+1}(\bar{y}_1 - \bar{y}_{i+1}) - \frac{n_{i+1}}{n} \sum_{j=2}^m n_j(\bar{y}_1 - \bar{y}_j), \quad i = 1, 2, \dots, m-1. \quad (8.113)$$

The estimated standard error, $\hat{\kappa}_i$, of the i th element of $\hat{\boldsymbol{\gamma}}$ ($i = 1, 2, \dots, m-1$) is the square root of the i th diagonal element of $MS_E \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}' = [(1/n_1)\mathbf{J}_{m-1} + \mathbf{\Lambda}]MS_E$, that is,

$$\hat{\kappa}_i = \left[\left(\frac{1}{n_1} + \frac{1}{n_{i+1}} \right) MS_E \right]^{1/2}, \quad i = 1, 2, \dots, m-1.$$

Thus by formula (8.110), large values of $|l_i^*| \hat{\kappa}_i$ identify those elements of $\hat{\boldsymbol{\gamma}} = \mathbf{L}\hat{\boldsymbol{\beta}}$ that are influential contributors to the significance of the F -test. In particular, if the data set used to analyze model (8.111) is balanced, that is, $n_i = n/m$ for $i = 1, 2, \dots, m$, then

$$|l_i^*| \hat{\kappa}_i = |\bar{y}_{i+1} - \bar{y}_{..}| \left(\frac{2n}{m} MS_E \right)^{1/2}, \quad i = 1, 2, \dots, m-1,$$

where $\bar{y}_{..} = (1/m)\sum_{i=1}^m \bar{y}_i$.

Alternatively, the contrast $\mathbf{a}^{*'}\hat{\boldsymbol{\beta}}$ can be expressed as

$$\begin{aligned} \mathbf{a}^{*'}\hat{\boldsymbol{\beta}} &= \mathbf{l}^{*'}\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \left(0, \sum_{i=1}^{m-1} l_i^*, -l_1^*, -l_2^*, \dots, -l_{m-1}^* \right) (0, \bar{y}_1, \bar{y}_2, \dots, \bar{y}_m)' \\ &= \sum_{i=1}^m c_i^* \bar{y}_i, \end{aligned}$$

where l_i^* is given in formula (8.113) and

$$c_i^* = \begin{cases} \sum_{j=1}^{m-1} l_j^*, & i = 1, \\ -l_{i-1}^*, & i = 2, 3, \dots, m. \end{cases} \quad (8.114)$$

Since the estimated standard error of $\bar{y}_i$ is $(MS_E/n_i)^{1/2}$, $i = 1, 2, \dots, m$, by dividing $\bar{y}_i$ by this value we obtain

$$\mathbf{a}^{*'}\hat{\boldsymbol{\beta}} = \sum_{i=1}^m \left(\frac{1}{n_i} MS_E \right)^{1/2} c_i^* w_i,$$

where $w_i = \bar{y}_i / (MS_E/n_i)^{1/2}$ is a scaled value of $\bar{y}_i$ ($i = 1, 2, \dots, m$). Hence, large values of $(MS_E/n_i)^{1/2} |c_i^*|$ identify those $\bar{y}_i$'s that contribute significantly to the rejection of H_0 . In particular, for a balanced data set,

$$l_i^* = \frac{n}{m} (\bar{y}_{..} - \bar{y}_{i+1}), \quad i = 1, 2, \dots, m-1.$$

Thus from formula (8.114) we get

$$\left(\frac{MS_E}{n_i} \right)^{1/2} |c_i^*| = \left(\frac{n}{m} MS_E \right)^{1/2} |\bar{y}_{i.} - \bar{y}_{..}|, \quad i = 1, 2, \dots, m.$$

We conclude that large values of $|\bar{y}_{i.} - \bar{y}_{..}|$ are responsible for the rejection of H_0 by the F -test. This is consistent with the fact that the numerator sum of squares of the F -test statistic for H_0 is proportional to $\sum_{i=1}^m (\bar{y}_{i.} - \bar{y}_{..})^2$ when the data set is balanced.

FURTHER READING AND ANNOTATED BIBLIOGRAPHY

- Adby, P. R., and M. A. H. Dempster (1974). *Introduction of Optimization Methods*. Chapman and Hall, London. (This book is an introduction to nonlinear methods of optimization. It covers basic optimization techniques such as steepest descent and the Newton–Raphson method.)
- Ash, A., and A. Hedayat (1978). “An introduction to design optimality with an overview of the literature.” *Comm. Statist. Theory Methods*, **7**, 1295–1325.
- Atkinson, A. C. (1982). “Developments in the design of experiments.” *Internat. Statist. Rev.*, **50**, 161–177.
- Atkinson, A. C. (1988). “Recent developments in the methods of optimum and related experimental designs.” *Internat. Statist. Rev.*, **56**, 99–115.
- Bates, D. M., and D. G. Watts (1988). *Nonlinear Regression Analysis and its Applications*. Wiley, New York. (Estimation of parameters in a nonlinear model is addressed in Chaps. 2 and 3. Design aspects for nonlinear models are briefly discussed in Section 3.14.)
- Bayne, C. K., and I. B. Rubin (1986). *Practical Experimental Designs and Optimization Methods for Chemists*. VCH Publishers, Deerfield Beach, Florida. (Steepest ascent and the simplex method are discussed in Chap. 5. A bibliography of optimization and response surface methods, as actually applied in 17 major fields of chemistry, is provided in Chap. 7.)
- Belsley, D. A., E. Kuh, and R. E. Welsch (1980). *Regression Diagnostics*. Wiley, New York. (Chap. 3 is devoted to the diagnosis of multicollinearity among the columns of the matrix in a regression model. Multicollinearity renders the model’s least-squares parameter estimates less precise and less useful than would otherwise be the case.)
- Biles, W. E., and J. J. Swain (1980). *Optimization and Industrial Experimentation*. Wiley-Interscience, New York. (Chaps. 4 and 5 discuss optimization techniques that are directly applicable in response surface methodology.)
- Bohachevsky, I. O., M. E. Johnson, and M. L. Stein (1986). “Generalized simulated annealing for function optimization.” *Technometrics*, **28**, 209–217.
- Box, G. E. P. (1982). “Choice of response surface design and alphabetic optimality.” *Utilitas Math.*, **21B**, 11–55.
- Box, G. E. P., and D. W. Behnken (1960). “Some new three level designs for the study of quantitative variables.” *Technometrics*, **2**, 455–475.

- Box, G. E. P., and N. R. Draper (1959). "A basis for the selection of a response surface design." *J. Amer. Statist. Assoc.*, **55**, 622–654.
- Box, G. E. P., and N. R. Draper (1963). "The choice of a second order rotatable design." *Biometrika*, **50**, 335–352.
- Box, G. E. P., and N. R. Draper (1965). "The Bayesian estimation of common parameters from several responses." *Biometrika*, **52**, 355–365.
- Box, G. E. P., and N. R. Draper (1987). *Empirical Model-Building and Response Surfaces*. Wiley, New York. (Chap. 9 introduces the exploration of maxima with second-order models; the alphabetic optimality approach is critically considered in Chap. 14. Many examples are given throughout the book.)
- Box, G. E. P., and H. L. Lucas (1959). "Design of experiments in nonlinear situations." *Biometrika*, **46**, 77–90.
- Box, G. E. P., and K. B. Wilson (1951). "On the experimental attainment of optimum conditions." *J. Roy. Statist. Soc. Ser. B*, **13**, 1–45.
- Bunday, B. D. (1984). *Basic Optimization Methods*. Edward Arnold Ltd., Victoria, Australia. (Chaps. 3 and 4 discuss basic optimization techniques such as the Nelder–Mead simplex method and the Davidon–Fletcher–Powell method.)
- Conlon, M. (1991). "The controlled random search procedure for function optimization." Personal communication. (This is a FORTRAN file for implementing Price's controlled random search procedure.)
- Conlon, M., and A. I. Khuri (1992). "Multiple response optimization." Technical Report, Department of Statistics, University of Florida, Gainesville, Florida.
- Cook, R. D., and C. J. Nachtsheim (1980). "A comparison of algorithms for constructing exact D -optimal designs." *Technometrics*, **22**, 315–324.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm." *J. Roy. Statist. Soc. Ser. B*, **39**, 1–38.
- Draper, N. R. (1963). "Ridge analysis of response surfaces." *Technometrics*, **5**, 469–479.
- Everitt, B. S. (1987). *Introduction to Optimization Methods and Their Application in Statistics*. Chapman and Hall, London. (This book gives a brief introduction to optimization methods and their use in several areas of statistics. These include maximum likelihood estimation, nonlinear regression estimation, and applied multivariate analysis.)
- Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York. (This book is a translation of a monograph in Russian. It presents the mathematical apparatus of experimental design for a regression model.)
- Fichtali, J., F. R. Van De Voort, and A. I. Khuri (1990). "Multiresponse optimization of acid casein production." *J. Food Process Eng.*, **12**, 247–258.
- Fletcher, R. (1987). *Practical Methods of Optimization*, 2nd ed. Wiley, New York. (This book gives a detailed study of several unconstrained and constrained optimization techniques.)
- Fletcher, R., and M. J. D. Powell (1963). "A rapidly convergent descent method for minimization." *Comput. J.*, **6**, 163–168.
- Hartley, H. O., and J. N. K. Rao (1967). "Maximum likelihood estimation for the mixed analysis of variance model." *Biometrika*, **54**, 93–108.

- Hoerl, A. E. (1959). "Optimum solution of many variables equations." *Chem. Eng. Prog.*, **55**, 69–78.
- Huber, P. J. (1973). "Robust regression: Asymptotics, conjectures and Monte Carlo." *Ann. Statist.*, **1**, 799–821.
- Huber, P. J. (1981). *Robust Statistics*. Wiley, New York. (This book gives a solid foundation in robustness in statistics. Chap. 3 introduces and discusses M -estimation; Chap. 7 addresses M -estimation for a regression model.)
- Johnson, M. E., and C. J. Nachtsheim (1983). "Some guidelines for constructing exact D -optimal designs on convex design spaces." *Technometrics*, **25**, 271–277.
- Jones, E. R., and T. J. Mitchell (1978). "Design criteria for detecting model inadequacy." *Biometrika*, **65**, 541–551.
- Karson, M. J., A. R. Manson, and R. J. Hader (1969). "Minimum bias estimation and experimental design for response surfaces." *Technometrics*, **11**, 461–475.
- Khuri, A. I., and M. Conlon (1981). "Simultaneous optimization of multiple responses represented by polynomial regression functions." *Technometric*, **23**, 363–375.
- Khuri, A. I., and J. A. Cornell (1996). *Response Surfaces*, 2nd ed. Marcel Dekker, New York. (Optimization techniques in response surface methodology are discussed in Chap. 5.)
- Khuri, A. I., and R. H. Myers (1979). "Modified ridge analysis." *Technometrics*, **21**, 467–473.
- Khuri, A. I., and H. Sahai (1985). "Variance components analysis: A selective literature survey." *Internat. Statist. Rev.*, **53**, 279–300.
- Kiefer, J. (1958). "On the nonrandomized optimality and the randomized nonoptimality of symmetrical designs." *Ann. Math. Statist.*, **29**, 675–699.
- Kiefer, J. (1959). "Optimum experimental designs" (with discussion). *J. Roy. Statist. Soc. Ser. B*, **21**, 272–319.
- Kiefer, J. (1960). "Optimum experimental designs V , with applications to systematic and rotatable designs." In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1. University of California Press, Berkeley, pp. 381–405.
- Kiefer, J. (1961). "Optimum designs in regression problems II." *Ann. Math. Statist.*, **32**, 298–325.
- Kiefer, J. (1962a). "Two more criteria equivalent to D -optimality of designs." *Ann. Math. Statist.*, **33**, 792–796.
- Kiefer, J. (1962b). "An extremum result." *Canad. J. Math.*, **14**, 597–601.
- Kiefer, J. (1975). "Optimal design: Variation in structure and performance under change of criterion." *Biometrika*, **62**, 277–288.
- Kiefer, J., and J. Wolfowitz (1960). "The equivalence of two extremum problems." *Canad. J. Math.*, **12**, 363–366.
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vechhi (1983). "Optimization by simulated annealing." *Science*, **220**, 671–680.
- Little, R. J. A., and D. B. Rubin (1987). *Statistical Analysis with Missing Data*. Wiley, New York. (The theory of the EM algorithm is introduced in Chap. 7. The book presents a systematic approach to the analysis of data with missing values, where inferences are based on likelihoods derived from formal statistical models for the data.)

- Lucas, J. M. (1976). "Which response surface design is best." *Technometrics*, **18**, 411–417.
- Miller, R. G., Jr. (1981). *Simultaneous Statistical Inference*, 2nd ed. Springer-Verlag, New York. (Scheffé's simultaneous confidence intervals are derived in Chap. 2.)
- Milliken, G. A., and D. E. Johnson (1984). *Analysis of Messy Data*. Lifetime Learning Publications, Belmont, California. (This book presents several techniques and methods for analyzing unbalanced data.)
- Mitchell, T. J. (1974). "An algorithm for the construction of *D*-optimal experimental designs." *Technometrics*, **16**, 203–210.
- Myers, R. H. (1976). *Response Surface Methodology*. Author, Blacksburg, Virginia. (Chap. 5 discusses the determination of optimum operating conditions in response surface methodology; designs for fitting first-order and second-order models are discussed in Chaps. 6 and 7, respectively; Chap. 9 presents the *J*-criterion for choosing a response surface design.)
- Myers, R. H. (1990). *Classical and Modern Regression with Applications*, 2nd ed., PWS-Kent, Boston. (Chap. 3 discusses the effects and hazards of multicollinearity in a regression model. Methods for detecting and combating multicollinearity are given in Chap. 8.)
- Myers, R. H., and W. H. Carter, Jr. (1973). "Response surface techniques for dual response systems." *Technometrics*, **15**, 301–317.
- Myers, R. H., and A. I. Khuri (1979). "A new procedure for steepest ascent." *Comm. Statist. Theory Methods*, **8**, 1359–1376.
- Myers, R. H., A. I. Khuri, and W. H. Carter, Jr. (1989). "Response surface methodology: 1966–1988." *Technometrics*, **31**, 137–157.
- Nelder, J. A., and R. Mead (1965). "A simplex method for function minimization." *Comput. J.*, **7**, 308–313.
- Nelson, L. S. (1973). "A sequential simplex procedure for non-linear least-squares estimation and other function minimization problems." In *27th Annual Technical Conference Transaction*, American Society for Quality Control, pp. 107–117.
- Olsson, D. M., and L. S. Nelson (1975). "The Nelder–Mead simplex procedure for function minimization." *Technometrics*, **17**, 45–51.
- Pazman, A. (1986). *Foundations of Optimum Experimental Design*. D. Reidel, Dordrecht, Holland.
- Plackett, R. L., and J. P. Burman (1946). "The design of optimum multifactorial experiments." *Biometrika*, **33**, 305–325.
- Price, W. L. (1977). "A controlled random search procedure for global optimization." *Comput. J.*, **20**, 367–370.
- Rao, C. R. (1970). "Estimation of heteroscedastic variances in linear models." *J. Amer. Statist. Assoc.*, **65**, 161–172.
- Rao, C. R. (1971). "Estimation of variance and covariance components—MINQUE theory." *J. Multivariate Anal.*, **1**, 257–275.
- Rao, C. R. (1972). "Estimation of variance and covariance components in linear models." *J. Amer. Statist. Assoc.*, **67**, 112–115.
- Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley, Reading, Massachusetts.

- Rustagi, J. S., ed. (1979). *Optimizing Methods in Statistics*. Academic Press, New York.
- Scheffé, H. (1959). *The Analysis of Variance*. Wiley, New York. (This classic book presents the basic theory of analysis of variance, mainly in the balanced case.)
- Searle, S. R. (1971). *Linear Models*. Wiley, New York. (This book describes general procedures of estimation and hypothesis testing for linear models. Estimable linear functions for models that are not of full rank are discussed in Chap. 5.)
- Seber, G. A. F. (1984). *Multivariate Observations*, Wiley, New York. (This book gives a comprehensive survey of the subject of multivariate analysis and provides many useful references.)
- Silvey, S. D. (1980). *Optimal Designs*. Chapman and Hall, London.
- Spendley, W., G. R. Hext, and F. R. Himsworth (1962). "Sequential application of simplex designs in optimization and evolutionary operation." *Technometrics*, **4**, 441–461.
- St. John, R. C., and N. R. Draper (1975). "*D*-Optimality for regression designs: A review." *Technometrics*, **17**, 15–23.
- Swallow, W. H., and S. R. Searle (1978). "Minimum variance quadratic unbiased estimation (MIVQUE) of variance components." *Technometrics*, **20**, 265–272.
- Watson, G. S. (1964). "A note on maximum likelihood." *Sankhyā Ser. A*, **26**, 303–304.
- Wynn, H. P. (1970). "The sequential generation of *D*-optimum experimental designs." *Ann. Math. Statist.*, **41**, 1655–1664.
- Wynn, H. P. (1972). "Results in the theory and construction of *D*-optimum experimental designs." *J. Roy. Statist. Soc. Ser. B*, **34**, 133–147.
- Zanakis, S. H., and J. S. Rustagi, eds. (1982). *Optimization in Statistics*. North-Holland, Amsterdam, Holland. (This is Volume 19 in *Studies in the Management Sciences*. It contains 21 articles that address applications of optimization in three areas of statistics, namely, regression and correlation; multivariate data analysis and design of experiments; and statistical estimation, reliability, and quality control.)

EXERCISES

8.1. Consider the function

$$f(x_1, x_2) = 8x_1^2 - 4x_1x_2 + 5x_2^2.$$

Minimize $f(x_1, x_2)$ using the method of steepest descent with $\mathbf{x}_0 = (5, 2)'$ as an initial point.

8.2. Conduct a simulated steepest ascent exercise as follows: Use the function

$$\eta(x_1, x_2) = 47.9 + 3x_1 - x_2 + 4x_1^2 + 4x_1x_2 + 3x_2^2$$

as the true mean response, which depends on two input variables x_1 and x_2 . Generate response values by using the model

$$y(\mathbf{x}) = \eta(\mathbf{x}) + \epsilon,$$

where ϵ has the normal distribution with mean 0 and variance 2.25, and $\mathbf{x} = (x_1, x_2)'$. Fit a first-order model in x_1 and x_2 in a neighborhood of the origin using a 2^2 factorial design along with the corresponding simulated response values. Make sure that replications are taken at the origin in order to test for lack of fit of the fitted model. Determine the path of steepest ascent, then proceed along it using simulated response values. Conduct additional experiments as described in Section 8.3.1.

8.3. Two types of fertilizers were applied to experimental plots to assess their effects on the yield of a certain variety of potato. The design settings used in the experiment along with the corresponding yield values are given in the following table:

Original Settings		Coded Settings		Yield y
Fertilizer 1	Fertilizer 2	x_1	x_2	(lb/plot)
50.0	15.0	-1	-1	24.30
120.0	15.0	1	-1	35.82
50.0	25.0	-1	1	40.50
120.0	25.0	1	1	50.94
35.5	20.0	$-2^{1/2}$	0	30.60
134.5	20.0	$2^{1/2}$	0	42.90
85.0	12.9	0	$-2^{1/2}$	22.50
85.0	27.1	0	$2^{1/2}$	50.40
85.0	20.0	0	0	45.69

(a) Fit a second-order model in the coded variables

$$x_1 = \frac{F_1 - 85}{35}, \quad x_2 = \frac{F_2 - 20}{5}$$

to the yield data, where F_1 and F_2 are the original settings of fertilizers 1 and 2, respectively, used in the experiment.

(b) Apply the method of ridge analysis to determine the settings of the two fertilizers that are needed to maximize the predicted yield (in the space of the coded input variables, the region R is the interior and boundary of a circle centered at the origin with a radius equal to $2^{1/2}$).

8.4. Suppose that λ_1 and λ_2 are two values of the Lagrange multiplier λ used in the method of ridge analysis. Let $\hat{y}_1$ and $\hat{y}_2$ be the corresponding values of $\hat{y}$ on the two spheres $\mathbf{x}'\mathbf{x} = r_1^2$ and $\mathbf{x}'\mathbf{x} = r_2^2$, respectively. Show that if $r_1 = r_2$ and $\lambda_1 > \lambda_2$, then $\hat{y}_1 > \hat{y}_2$.

- 8.5.** Consider again Exercise 8.4. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be the stationary points corresponding to λ_1 and λ_2 , respectively. Consider also the matrix

$$\mathbf{M}(\mathbf{x}_i) = 2(\hat{\mathbf{B}} - \lambda_i \mathbf{I}), \quad i = 1, 2.$$

Show that if $r_1 = r_2$, $\mathbf{M}(\mathbf{x}_1)$ is positive definite, and $\mathbf{M}(\mathbf{x}_2)$ is indefinite, then $\hat{y}_1 < \hat{y}_2$.

- 8.6.** Consider once more the method of ridge analysis. Let $\mathbf{x}$ be a stationary point that corresponds to the radius r .

(a) Show that

$$r^3 \frac{\partial^2 r}{\partial \lambda^2} = 2r^2 \sum_{i=1}^k \left(\frac{\partial x_i}{\partial \lambda} \right)^2 + \left[r^2 \sum_{i=1}^k \left(\frac{\partial x_i}{\partial \lambda} \right)^2 - \left(\sum_{i=1}^k x_i \frac{\partial x_i}{\partial \lambda} \right)^2 \right],$$

where x_i is the i th element of $\mathbf{x}$ ($i = 1, 2, \dots, k$).

(b) Make use of part (a) to show that

$$\frac{\partial^2 r}{\partial \lambda^2} > 0 \quad \text{if } r \neq 0.$$

- 8.7.** Suppose that the “true” mean response $\eta(\mathbf{x})$ is represented by a model of order d_2 in k input variables $x_1, x_2, \dots, x_k$ of the form

$$\eta(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \mathbf{g}'(\mathbf{x})\boldsymbol{\delta},$$

where $\mathbf{x} = (x_1, x_2, \dots, x_k)'$. The fitted model is of order d_1 ($< d_2$) of the form

$$\hat{y}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\hat{\boldsymbol{\lambda}},$$

where $\hat{\boldsymbol{\lambda}}$ is an estimator of $\boldsymbol{\beta}$, not necessarily obtained by the method of least squares. Let $\boldsymbol{\gamma} = E(\hat{\boldsymbol{\lambda}})$.

- (a) Give an expression for B , the average squared bias of $\hat{y}(\mathbf{x})$, in terms of $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and $\boldsymbol{\gamma}$.
- (b) Show that B achieves its minimum value if and only if $\boldsymbol{\gamma}$ is of the form $\boldsymbol{\gamma} = \mathbf{C}\boldsymbol{\tau}$, where $\boldsymbol{\tau} = (\boldsymbol{\beta}', \boldsymbol{\delta}')'$ and $\mathbf{C} = [\mathbf{I}; \boldsymbol{\Gamma}_{11}^{-1} \boldsymbol{\Gamma}_{12}]$. The matrices $\boldsymbol{\Gamma}_{11}$ and $\boldsymbol{\Gamma}_{12}$ are the region moments used in formula (8.54).
- (c) Deduce from part (b) that B achieves its minimum value if and only if $\mathbf{C}\boldsymbol{\tau}$ is an estimable linear function (see Searle, 1971, Section 5.4).

- (d) Use part (c) to show that B achieves its minimum value if and only if there exists a matrix $\mathbf{L}$ such that $\mathbf{C} = \mathbf{L}[\mathbf{X}: \mathbf{Z}]$, where $\mathbf{X}$ and $\mathbf{Z}$ are matrices consisting of the values taken by $\mathbf{f}'(\mathbf{x})$ and $\mathbf{g}'(\mathbf{x})$, respectively, at n experimental runs.
- (e) Deduce from part (d) that B achieves its minimum for any design for which the row space of $[\mathbf{X}: \mathbf{Z}]$ contains the rows of $\mathbf{C}$.
- (f) Show that if $\hat{\boldsymbol{\lambda}}$ is the least-squares estimator of $\boldsymbol{\beta}$, that is, $\hat{\boldsymbol{\lambda}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, where $\mathbf{y}$ is the vector of response values at the n experimental runs, then the design property stated in part (e) holds for any design that satisfies the conditions described in equations (8.56).

[Note: This problem is based on an article by Karson, Manson, and Hader (1969), who introduced the so-called minimum bias estimation to minimize the average squared bias B .]

- 8.8. Consider again Exercise 8.7. Suppose that $\mathbf{f}'(\mathbf{x})\boldsymbol{\beta} = \beta_0 + \sum_{i=1}^3 \beta_i x_i$ is a first-order model in three input variables fitted to a data set obtained by using the design

$$\mathbf{D} = \begin{bmatrix} -g & -g & -g \\ g & g & -g \\ g & -g & g \\ -g & g & g \end{bmatrix},$$

where g is a scale factor. The region of interest is a sphere of radius 1. Suppose that the “true” model is of the form

$$\eta(\mathbf{x}) = \beta_0 + \sum_{i=1}^3 \beta_i x_i + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3.$$

- (a) Can g be chosen so that $\mathbf{D}$ satisfies the conditions described in equations (8.56)?
- (b) Can g be chosen so that $\mathbf{D}$ satisfies the minimum bias property described in part (e) of Exercise 8.7?

- 8.9. Consider the function

$$h(\boldsymbol{\delta}, \mathbf{D}) = \boldsymbol{\delta}' \boldsymbol{\Delta} \boldsymbol{\delta},$$

where $\boldsymbol{\delta}$ is a vector of unknown parameters as in model (8.48), $\boldsymbol{\Delta}$ is the matrix in formula (8.53), namely $\boldsymbol{\Delta} = \mathbf{A}'\boldsymbol{\Gamma}_{11}\mathbf{A} - \boldsymbol{\Gamma}'_{12}\mathbf{A} - \mathbf{A}'\boldsymbol{\Gamma}_{12} + \boldsymbol{\Gamma}_{22}$, and $\mathbf{D}$ is the design matrix.

- (a) Show that for a given $\mathbf{D}$, the maximum of $h(\boldsymbol{\delta}, \mathbf{D})$ over the region $\psi = \{\boldsymbol{\delta} | \boldsymbol{\delta}'\boldsymbol{\delta} \leq r^2\}$ is equal to $r^2 e_{\max}(\boldsymbol{\Delta})$, where $e_{\max}(\boldsymbol{\Delta})$ is the largest eigenvalue of $\boldsymbol{\Delta}$.
- (b) Deduce from part (a) a design criterion for choosing $\mathbf{D}$.

8.10. Consider fitting the model

$$y(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \epsilon,$$

where ϵ is a random error with a zero mean and a variance σ^2 . Suppose that the “true” mean response is given by

$$\eta(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\boldsymbol{\beta} + \mathbf{g}'(\mathbf{x})\boldsymbol{\delta}.$$

Let $\mathbf{X}$ and $\mathbf{Z}$ be the same matrices defined in part (d) of Exercise 8.7. Consider the function $\lambda(\boldsymbol{\delta}, \mathbf{D}) = \boldsymbol{\delta}'\mathbf{S}\boldsymbol{\delta}$, where

$$\mathbf{S} = \mathbf{Z}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Z},$$

and $\mathbf{D}$ is the design matrix. The quantity $\lambda(\boldsymbol{\delta}, \mathbf{D})/\sigma^2$ is the noncentrality parameter associated with the lack of fit F -test for the fitted model (see Khuri and Cornell, 1996, Section 2.6). Large values of λ/σ^2 increase the power of the lack of fit test. By formula (8.54), the minimum value of B is given by

$$B_{\min} = \frac{n}{\sigma^2} \boldsymbol{\delta}'\mathbf{T}\boldsymbol{\delta},$$

where $\mathbf{T} = \boldsymbol{\Gamma}_{22} - \boldsymbol{\Gamma}_{12}'\boldsymbol{\Gamma}_{11}^{-1}\boldsymbol{\Gamma}_{12}$. The fitted model is considered to be inadequate if there exists some constant $\kappa > 0$ such that $\boldsymbol{\delta}'\mathbf{T}\boldsymbol{\delta} \geq \kappa$. Show that

$$\inf_{\boldsymbol{\delta} \in \Phi} \boldsymbol{\delta}'\mathbf{S}\boldsymbol{\delta} = \kappa e_{\min}(\mathbf{T}^{-1}\mathbf{S}),$$

where $e_{\min}(\mathbf{T}^{-1}\mathbf{S})$ is the smallest eigenvalue of $\mathbf{T}^{-1}\mathbf{S}$ and Φ is the region $\{\boldsymbol{\delta} | \boldsymbol{\delta}'\mathbf{T}\boldsymbol{\delta} \geq \kappa\}$.

[Note: On the basis of this problem, we can define a new design criterion, that which maximizes $e_{\min}(\mathbf{T}^{-1}\mathbf{S})$ with respect to $\mathbf{D}$. A design chosen according to this criterion is called Λ_1 -optimal (see Jones and Mitchell, 1978).]

8.11. A second-order model of the form

$$y(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon$$

is fitted using a rotatable central composite design $\mathbf{D}$, which consists of a factorial 2^2 portion, an axial portion with an axial parameter $\alpha = 2^{1/2}$, and n_0 center-point replications. The settings of the 2^2 factorial portion are ± 1 . The region of interest R consists of the interior and boundary of a circle of radius $2^{1/2}$ centered at the origin.

- (a) Express V , the average variance of the predicted response given by formula (8.52), as a function of n_0 .
- (b) Can n_0 be chosen so that it minimizes V ?

8.12. Suppose that we have r response functions represented by the models

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i, \quad i = 1, 2, \dots, r,$$

where $\mathbf{X}$ is a known matrix of order $n \times p$ and rank p . The random error vectors have the same variance-covariance structure as in Section 8.7. Let $\mathbf{F} = E(\mathbf{Y}) = \mathbf{X}\mathbf{B}$, where $\mathbf{Y} = [\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_r]$ and $\mathbf{B} = [\boldsymbol{\beta}_1; \boldsymbol{\beta}_2; \dots; \boldsymbol{\beta}_r]$.

Show that the determinant of $(\mathbf{Y} - \mathbf{F})'(\mathbf{Y} - \mathbf{F})$ attains a minimum value when $\mathbf{B} = \hat{\mathbf{B}}$, where $\hat{\mathbf{B}}$ is obtained by replacing each $\boldsymbol{\beta}_i$ in $\mathbf{B}$ with $\hat{\boldsymbol{\beta}}_i = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_i$ ($i = 1, 2, \dots, r$).

[Note: The minimization of the determinant of $(\mathbf{Y} - \mathbf{F})'(\mathbf{Y} - \mathbf{F})$ with respect to $\mathbf{B}$ represents a general multiresponse estimation criterion known as the Box-Draper determinant criterion (see Box and Draper, 1965).]

8.13. Let $\mathbf{A}$ be a $p \times p$ matrix with nonnegative eigenvalues. Show that

$$\det(\mathbf{A}) \leq \exp[\text{tr}(\mathbf{A} - \mathbf{I}_p)].$$

[Note: This inequality is proved in an article by Watson (1964). It is based on the simple inequality $a \leq \exp(a - 1)$, which can be easily proved for any real number a .]

8.14. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ be a sample of n independently distributed random vectors from a p -variate normal distribution $N(\boldsymbol{\mu}, \mathbf{V})$. The corresponding likelihood function is

$$L = \frac{1}{(2\pi)^{np/2} [\det(\mathbf{V})]^{n/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right].$$

It is known that the maximum likelihood estimate of $\boldsymbol{\mu}$ is $\bar{\mathbf{x}}$, where $\bar{\mathbf{x}} = (1/n) \sum_{i=1}^n \mathbf{x}_i$ (see, for example, Seber, 1984, pages 59–61). Let $\mathbf{S}$ be the matrix

$$\mathbf{S} = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'.$$

Show that $\mathbf{S}$ is the maximum likelihood estimate of $\mathbf{V}$ by proving that

$$\begin{aligned} & \frac{1}{[\det(\mathbf{V})]^{n/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{V}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right] \\ & \leq \frac{1}{[\det(\mathbf{S})]^{n/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right], \end{aligned}$$

or equivalently,

$$[\det(\mathbf{S}\mathbf{V}^{-1})]^{n/2} \exp \left[-\frac{n}{2} \text{tr}(\mathbf{S}\mathbf{V}^{-1}) \right] \leq \exp \left[-\frac{n}{2} \text{tr}(\mathbf{I}_p) \right].$$

[Hint: Use the inequality given in Exercise 8.13.]

8.15. Consider the random one-way classification model

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, 2, \dots, a; j = 1, 2, \dots, n_i,$$

where the α_i 's and ϵ_{ij} 's are independently distributed as $N(0, \sigma_\alpha^2)$ and $N(0, \sigma_\epsilon^2)$. Determine the matrix $\mathbf{S}$ and the vector $\mathbf{q}$ in equation (8.95) that can be used to obtain the MINQUEs of σ_α^2 and σ_ϵ^2 .

8.16. Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{X}$ is a known matrix of order $n \times p$ and rank p , and $\boldsymbol{\epsilon}$ is normally distributed with a zero mean vector and a variance-covariance matrix $\sigma^2 \mathbf{I}_n$. Let $\hat{y}(\mathbf{x})$ denote the predicted response at a point $\mathbf{x}$ in a region of interest R .

Use Scheffé's confidence intervals given by formula (8.97) to obtain simultaneous confidence intervals on the mean response values at the points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ ($m \leq p$) in R . What is the joint confidence coefficient for these intervals?

8.17. Consider the fixed-effects two-way classification model

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \\ i &= 1, 2, \dots, a; j = 1, 2, \dots, b; k = 1, 2, \dots, m, \end{aligned}$$

where α_i and β_j are unknown parameters, $(\alpha\beta)_{ij}$ is the interaction effect, and ϵ_{ijk} is a random error that has the normal distribution with a zero mean and a variance σ^2 .

- (a) Use Scheffé's confidence intervals to obtain simultaneous confidence intervals on all contrasts among the μ_i 's, where $\mu_i = E(\bar{y}_{i..})$ and $\bar{y}_{i..} = (1/bm) \sum_{j=1}^b \sum_{k=1}^m y_{ijk}$.
- (b) Identify those $\bar{y}_{i..}$'s that are influential contributors to the significance of the F -test concerning the hypothesis $H_0: \mu_1 = \mu_2 = \dots = \mu_a$.