

## Chapter 2

# OPTIMIZATION OF FIRST ORDER MODELS

*One should not multiply explanations and causes unless it is strictly necessary*

—“William of Bakersville” in Umberto Eco’s *In the Name of the Rose*<sup>1</sup>

In Response Surface Methods, the optimal region to run a process is usually determined after a sequence of experiments is conducted and a series of empirical models are obtained. As mentioned in Chapter 1, in a new or poorly understood process it is likely that a first order model will fit well. The Box-Wilson methodology suggests the use of a steepest ascent technique coupled with lack of fit and curvature tests to move the process from a region of little curvature to one where curvature – and the presence of a stationary point – exists. In this chapter we discuss, at an elementary level, steepest ascent/descent methods for optimizing a process described by a first order model. We will assume readers are familiar with the linear regression material reviewed in Appendix A. More advanced techniques related to exploring a new region that incorporate the statistical inference issues into the optimization methods are discussed in Chapter 6.

<sup>1</sup>Paraphrasing a non-fictional Franciscan monk of the epoch, William of Ockham (1285–1347): *Essentia non sunt multiplicanda praeter necessitatem* (Entities should not be multiplied beyond necessity), a statement that became to be known as “Ockham’s razor”.

## 2.1 New Region Exploration

A first order model will serve as a good local approximation in a small region close to the initial operating conditions and far from where the process exhibits curvature. A fitted first order polynomial model in  $k$  factors has the form

$$\hat{y} = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j x_j.$$

Experimental designs used for fitting this type of models are discussed in more detail in Chapter 3. Usually, a two-level factorial experiment is conducted with repeated runs at the current operating conditions which serve as the origin of coordinates in coded factors.

The idea behind this region exploration “Phase” of RSM is to keep experimenting along the direction of steepest ascent (or descent, as required) until there is no further improvement in the response. At that point, a new fractional factorial experiment with center runs is conducted to determine a new search direction. This process is repeated until at some point significant curvature in  $\hat{y}$  or lack of fit in the 1st order model will be detected. In the new region, the operating conditions  $x_1, x_2, \dots, x_k$  are close to where a stationary point (and hopefully, a local maximum or minimum, as required) of  $y$  occurs. When significant curvature or lack of linear fit is detected, the experimenter should proceed to fit and optimize a higher order model, in what Box and Wilson called “canonical analysis”. Figure 2.1 illustrates a sequence of line searches when seeking a region where curvature exists in a problem with 2 factors (i.e.,  $k = 2$ ). In practical problems, more than 3 iterations of steepest ascent/descent are rare.

There are two main decisions we must make in the region exploration phase of RSM:

- 1 determine the search direction;
- 2 determine the length of the step to move from the current operating conditions.

These two decisions are linked, because a search in a direction where there is strong evidence the process will improve may proceed at larger steps than a search in a less reliable direction. In Chapter 6 we present some approaches on how to deal with these two aspects jointly.

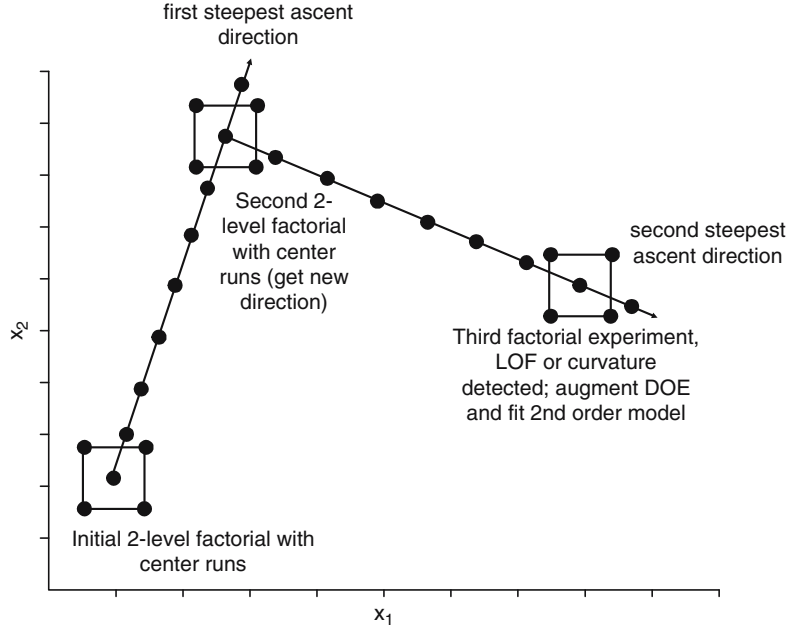


Figure 2.1. A sequence of line searches for a 2 factor optimization problem

## 2.2 Steepest Ascent/Descent Procedure

Suppose a first order model has been fitted and the model does not exhibit lack of fit due to second order terms. Then the direction of *maximum improvement* is given by:

- 1  $\nabla \hat{y}$ , if the objective is to *maximize*  $y$  (this gives the steepest ascent direction);
- 2  $-\nabla \hat{y}$ , if the objective is to *minimize*  $y$  (this gives the steepest descent direction).

The direction of the gradient is given by the values of the parameter estimates (excluding the intercept), that is,  $\hat{\beta}' = \mathbf{b}' = (b_1, b_2, \dots, b_k)$ . Since the parameter estimates depend on the scaling convention for the factors, the steepest ascent (descent) direction is also scale-dependent. That is, two experimenters using different scaling conventions will follow different paths

for process improvement. We note, however, that the region of the search, as given by the signs of the parameter estimates, does not change with scale<sup>2</sup>.

Let us consider without loss of generality a maximization problem. The coordinates of the factor settings on the direction of steepest ascent separated a distance  $\rho$  from the origin are obtained from solving:

$$\begin{aligned} \max \quad & b_1x_1 + b_2x_2 + \cdots + b_kx_k = \mathbf{b}'\mathbf{x} \\ \text{subject to:} \quad & \sum_{i=1}^k x_i^2 \leq \rho^2 \quad \text{or} \quad \mathbf{x}'\mathbf{x} \leq \rho^2. \end{aligned}$$

Unconstrained optimization of the first order model will evidently provide an unbounded solution, hence the constraint is added. The value of  $\rho$  is the step size, which is user-defined. A value of  $\rho = 1$  will give a point close to the edge of the experimental region where the model was fit (in coded factors). To solve the maximization problem, form the Lagrangian (see Appendix C for general results on optimality conditions):

$$\text{maximize} \quad L = \mathbf{b}'\mathbf{x} - \lambda(\mathbf{x}'\mathbf{x} - \rho^2)$$

where  $\lambda$  is a Lagrange multiplier. Compute the partials and equate them to zero

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{b} - 2\lambda\mathbf{x} = \mathbf{0}$$

$$\frac{\partial L}{\partial \lambda} = -(\mathbf{x}'\mathbf{x} - \rho^2) = 0.$$

These two equations have two unknowns (the vector  $\mathbf{x}$  and the scalar  $\lambda$ ) and thus can be solved yielding the desired solution:

$$\mathbf{x}^* = \rho \frac{\mathbf{b}}{\|\mathbf{b}\|}$$

or, in non-vector notation:

$$x_i^* = \rho \frac{b_i}{\sqrt{\sum_{j=1}^k b_j^2}} \quad i = 1, 2, \dots, k. \quad (2.1)$$

From this equation, we can see that any multiple  $\rho$  of the direction of the gradient (given by the unit vector  $\mathbf{b}/\|\mathbf{b}\|$ ) will lead to points on the steepest

<sup>2</sup>For a recent proposal on a scale-independent direction finding method, see Section 6.3.

ascent direction. For steepest descent, we use instead  $-b_i$  in the numerator of the equation above.

An equivalent and simple approach to finding the points on the steepest ascent path is as follows [110]:

- 1 Choose a step size in one of the controllable factors, say  $\Delta x_j$ . This may be the factor we are most comfortable varying or simply the one with the largest  $|b_j|$ .
- 2 Take a proportional step in all other factors, i.e.,

$$\frac{\Delta x_i}{b_i} = \frac{\Delta x_j}{b_j}, \quad i = 1, 2, \dots, k; \quad i \neq j$$

from which

$$\Delta x_i = \frac{b_i}{b_j} \Delta x_j, \quad i = 1, 2, \dots, k; \quad i \neq j. \quad (2.2)$$

These two approaches are equivalent since the second approach will result in a point on the path of steepest ascent located at a distance  $\rho = \sqrt{\sum_{i=1}^k \Delta x_i^2}$  (coded units) from the origin.

Starting at some given operating conditions (the origin in coded units), the recommended practice is to run a 2-level factorial experiment with center points replicated to allow to test for lack of fit and for curvature (see Appendix A). If there is evidence of lack of fit (LOF) or of curvature, we should add points to the 2-level factorial so that a second order model is estimable. With the second order model, we can then estimate the location of a local optima more precisely, as explained later in Chapter 4. We now illustrate the steepest ascent procedure.

**Example. Optimization of a Chemical Process.** It has been concluded (perhaps after a factor screening experiment) that the yield ( $y$ , in %) of a chemical process is mainly affected by the temperature ( $\xi_1$ , in  $^{\circ}\text{C}$ ) and by the reaction time ( $\xi_2$ , in minutes). Due to safety reasons, the region of operation is limited to

$$50 \leq \xi_1 \leq 250$$

$$150 \leq \xi_2 \leq 500$$

The process is currently run at a temperature of  $200^{\circ}\text{C}$  and a reaction time of 200 minutes. A process engineer decides to run a  $2^2$  full factorial experiment

with factor levels at

<i>factor</i>	<i>low</i>	<i>center</i>	<i>high</i>
$X_1$	170	200	230
$X_2$	150	200	250

Five repeated runs at the center were conducted to assess lack of fit (see Appendix A for details on lack of fit tests). The orthogonally coded factors are

$$x_1 = \frac{\xi_1 - 200}{30} \quad \text{and} \quad x_2 = \frac{\xi_2 - 200}{50}.$$

The experimental results are shown in Table 2.1. The corresponding ANOVA table for a first order polynomial model is given in Table 2.2.

Neither the single degree of freedom test of curvature nor the lack of fit test indicate a problem with the model. Furthermore, there is evidence that the first

Table 2.1. First 2-level experimental design, chemical experiment example

$x_1$	$x_2$	$\xi_1$	$\xi_2$	$Y(=yield)$
−1	−1	170	150	32.79
1	−1	230	150	24.07
−1	1	170	250	48.94
1	1	230	250	52.49
0	0	200	200	38.89
0	0	200	200	48.29
0	0	200	200	29.68
0	0	200	200	46.50
0	0	200	200	44.15

Table 2.2. ANOVA table for first DOE, chemical example

Source	SS	d.o.f.	MS	$F_0$	P-value
Model	503.3035	2	251.6517	4.810	0.0684
Curvature	8.1536	1	8.1536	0.1558	0.7093
Residual (Error)	261.5935	5	52.3187		
Lack of Fit	37.6382	1	37.6382	0.6722	0.4583
Pure Error	223.9553	4	55.9888		
Total	773.0506	8			

order model is significant. The fitted equation in coded factors, using ordinary least squares (OLS), is

$$\hat{Y} = 39.57 - 1.2925x_1 + 11.14x_2$$

Diagnostic checks (see Appendix A) show conformance to the regression assumptions, although the  $R^2$  value (0.6580) is not very high.

To maximize  $\hat{y}$ , we use the direction of steepest ascent. Suppose we select  $\rho = 1$ , since a point on the steepest ascent direction distanced one unit (in the coded units) from the origin is desired. Then from equation (2.1), the coordinates of the factor levels for the next run are given by:

$$x_1^* = \frac{\rho b_1}{\sqrt{\sum_{j=1}^2 b_j^2}} = \frac{(1)(-1.2925)}{\sqrt{(-1.2925)^2 + (11.14)^2}} = -0.1152$$

and

$$x_2^* = \frac{\rho b_2}{\sqrt{\sum_{j=1}^2 b_j^2}} = \frac{(1)(11.14)}{\sqrt{(-1.2925)^2 + (11.14)^2}} = 0.9933.$$

This means that to improve the process, for every  $(-0.1152)(30) = -3.456$  °C that temperature is varied (decreased), the reaction time should be varied by  $(0.9933)(50) = 49.66$  minutes.

Alternatively, we could have used instead the procedure that lead to expression (2.2):

- 1 Suppose we select  $\Delta\xi_2 = 50$  minutes. This can be based on process engineering considerations. It could have been felt that  $\Delta\xi_2 = 50$  does not move the process too far away from the current region of experimentation. This makes sense since the  $R^2$  value of 0.6580 for the fitted model is quite low, providing a steepest ascent direction not very reliable.
- 2  $\Delta x_2 = \frac{50}{50} = 1.0$ .
- 3  $\Delta x_1 = \frac{-1.2925}{11.14} = -0.1160$ .
- 4  $\Delta\xi_2 = (-1.160)(30) = -3.48$  °C.

Thus the step size in original units is  $\Delta\xi' = (-3.48$  °C, 50 minutes). ■

To conduct experiments along the direction of maximum improvement, we just continue selecting operating conditions using the same step size as selected before, adding the step  $\Delta\xi$  to the last point on the steepest ascent/descent direction  $\xi_i$ .

- 1 Given current operating conditions  $\xi'_0 = (\xi_1, \xi_2, \dots, \xi_k)$  and a step size  $\Delta\xi' = (\Delta\xi_1, \Delta\xi_2, \dots, \Delta\xi_k)$ , perform experiments at factor levels  $\xi_0 + \Delta\xi, \xi_0 + 2\Delta\xi, \xi_0 + 3\Delta\xi, \dots$  as long as improvement in the response  $y$  (decrease or increase, as desired) is observed.
- 2 Once a point has been reached where there is no further improvement, a new first order experiment (e.g. a 2-level factorial) should be performed with repeated center runs to assess lack of fit and curvature. If there is no significant evidence of lack of fit, the new first order model will provide a new search direction, and another iteration is performed. Otherwise, if either there is evidence of lack of fit or of (pure quadratic) curvature, the experimental design is augmented and a second order model should be fitted. That is, the experimenter should proceed with the next “Phase” of RSM.

**Example. Experimenting along the direction of maximum improvement.**

Let us apply the steepest ascent procedure to the Chemical Experiment analyzed earlier. Recall that a first order model was significant, and did not show lack of fit or evidence of curvature. We proceed to move the process in the direction of steepest ascent, by using the step size computed earlier.

Step 1:

Given  $\xi_0 = (200^\circ\text{C}, 200 \text{ minutes})$  and  $\Delta\xi = (-3.48^\circ\text{C}, 50 \text{ minutes})$ , perform experiments as shown on Table 2.3 (the step size in temperature was rounded to  $-3.5^\circ\text{C}$  for practical reasons). Since the goal was to maximize  $y$ , the point of maximum observed response is  $\xi_1 = 189.5^\circ\text{C}$ ,  $\xi_2 = 350 \text{ minutes}$ . Notice that the search was stopped after 2 consecutive drops in response. This was done to reassure us that the mean response was *actually decreasing* and to avoid

Table 2.3. Illustration of a steepest ascent search

	$\xi_1$	$\xi_2$	$x_1$	$x_2$	$y$ (=yield)
$\xi_0$	200	200	0	0	
$\xi_0 + \Delta\xi$	196.5	250	-0.1160	1	56.2
$\xi_0 + 2\Delta\xi$	193.0	300	-0.2333	2	71.49
$\xi_0 + 3\Delta\xi$	189.5	350	-0.3500	3	75.63
$\xi_0 + 4\Delta\xi$	186.0	400	-0.4666	4	72.31
$\xi_0 + 5\Delta\xi$	182.5	450	-0.5826	5	72.10

stopping the search too early due to noise. This issue calls for a formal stopping rule for steepest ascent, a topic we discuss in detail in Chapter 6.

Step 2:

A new  $2^2$  factorial experiment is performed with  $\xi' = (189.5, 350)$  as the origin. Using the same scaling factors as before, the new scaled controllable factors are:

$$x_1 = \frac{\xi_1 - 189.5}{30} \quad \text{and} \quad x_2 = \frac{\xi_2 - 350}{50}$$

Five center runs (at  $\xi_1 = 189.5, \xi_2 = 350$ ) were repeated to assess lack of fit and curvature. The experimental results are shown on Table 2.4. The corresponding ANOVA table for a linear model is shown on Table 2.5.

From the table, the linear effects model is significant and there is no evidence of lack of fit. However, there is a significant curvature effect, which implies

Table 2.4. Second 2-level factorial run, chemical experiment example

$x_1$	$x_2$	$\xi_1$	$\xi_2$	$y$ (=yield)
-1	-1	159.5	300	64.33
1	-1	219.5	300	51.78
-1	1	158.5	400	77.30
1	1	219.5	400	45.37
0	0	189.5	350	62.08
0	0	189.5	350	79.36
0	0	189.5	350	75.29
0	0	189.5	350	73.81
0	0	189.5	350	69.45

Table 2.5. ANOVA table for the second DOE, chemical example

Source	SS	d.o.f.	MS	$F_0$	P-value
Model	505.300	2	252.650	4.731	0.0703
Curvature	336.309	1	336.309	6.297	0.0539
Residual	267.036	5	53.407		
Lack of fit	93.857	1	93.857	2.168	0.2149
Pure Error	173.179	4	43.295		
Total	1108.646	8			

that we have moved the process operating conditions to a region where there is curvature, so we should proceed to the next phase in RSM, namely, fitting and optimizing a second order model. ■

The discussants of the Box and Wilson paper [29] pointed out that no clear stopping criterion existed in the methodology, to which the authors agreed. Although the discussants were referring to the final phase of RSM, the same can be said about a steepest ascent search. Using a simple “first drop” stopping rule (i.e., stop if the response drops for the first time) may miss the true maximum in the steepest ascent direction, as the observed drop may be due to noise. Box and Wilson noted this, mentioning that the observed response values could be used to test for the difference of the mean response, and warned that the steepest ascent procedure is useful under low noise conditions. Formal tests of hypothesis that consider “noise” in the ascent procedure were developed by later authors, and we present them in Chapter 6. A Bayesian stopping rule, much in the sense of the discussion of the Box-Wilson paper, has been developed by Gilmour and Mead [61]. We discuss it in detail in Chapter 12.

Considerable work has been conducted since the Box and Wilson paper [29] on using search directions other than the steepest ascent direction, and on computing a confidence “cone” around a given direction. If the cone is too wide, this is evidence to proceed cautiously along such direction, so the cone provides some guidance about how large one should choose the step size along the steepest ascent direction other than selecting either  $\rho$  or  $\Delta x_j$  arbitrarily. Again, we refer to Chapter 6 for details. Directions other than steepest ascent, which is well-known in the non-linear programming literature to have “zig-zagging” problems, have been proposed as well. We review one such proposal in Chapter 6.

A final comment on the steepest ascent procedure as proposed in Box and Wilson refers to the assumed model. If a fractional factorial experiment at two levels is conducted, why neglecting two factor interactions? The argument in favor of this model is that in the initial stage of experimentation interactions will be dominated by main effects. This might be true for pure quadratic terms, but it is not clear why it should be true for two factor interactions. The argument reminds us a point in the controversy around the Taguchi methods (see Chapter 9). Taguchi considers *a priori* that some interactions are insignificant, and this has been criticized – fairly in our view – to be unjustified in general.

Likewise, in steepest ascent it seems that the first order model is assumed to fit the optimization technique (steepest ascent) because otherwise one would have to rely on relatively more complex nonlinear programming techniques. A two level factorial of enough resolution may be adequate to estimate two-factor interactions which can be utilized to optimize a process more efficiently, even at an earlier stage. The success of steepest ascent in practice, as evidenced by the many papers that have reported its application, is due to the fact that the technique will achieve some improvement – maybe considerable improvement – with respect to an earlier, non-optimized state of the process. That this improvement is as large as it could have been obtained or that it was found at the minimum cost (i.e., with a minimum number of experiments) has not been shown, and this point has not been emphasized much in the literature. Thus, there is room for improvement in this topic. We further discuss steepest ascent and related methods at a somewhat more advanced point of view in Chapter 6.

### 2.3 Problems

- 1 Consider the first order model  $\hat{y} = 100 + 5x_1 + 8x_2 - 3x_3$ . This model was fit using an unreplicated  $2^3$  design in the coded variables  $-1 < x_i < 1, i = 1, 2, 3$ . The model fit was adequate and  $S_b^2 = 2.54$ . The region of exploration on the natural variables was

$\xi_1$  = temperature (100, 110 degrees C)

$\xi_2$  = time (1, 2 hrs.)

$\xi_3$  = pressure (50, 75 psi)

- a) Using  $x_2$  as the variable to define a step size along the steepest ascent direction, choose a step size large enough to take you to the boundary of the experimental region in that particular direction. Find and show the coordinates of this point on the path of steepest ascent in the coded variables  $x_i$ .
- b) Find and show the coordinates of this point on the path of steepest ascent from part a) using the natural variables.
- c) Find a unit vector that defines the path of steepest ascent.
- d) What step size multiplier for the unit vector in c) above would give the same point on the path of steepest ascent you found in parts a) and b)?

- e) Find the fraction of directions excluded by the 95% confidence cone of steepest ascent.
- 2 Consider the first order model  $\hat{y} = 14.4 - 5.09x_1 - 13.2x_2 - 4x_3$ , where  $y$  denotes the life of a tool, in minutes, as a function of 3 process variables. This model was fit using an unreplicated  $2^3$  design in the coded variables  $-1 < x_i < 1, i = 1, 2, 3$ . The model fit was adequate with  $S_b^2 = 1.74$ . The region of exploration on the natural variables was
- $\xi_1$  = speed of cut (650 to 800 units of surface per minute, sfm)  
 $\xi_2$  = cut feed (0.01 to 0.03 inches per revolution, ipr)  
 $\xi_3$  = depth of cut (0.05 to 0.20 in.)
- and the current operating conditions are the center of this region.
- a) For steepest ascent purposes, the Quality Engineer chooses to decrease the coded variable  $x_2$  one unit from the origin. Find the coordinates of the resulting point on the path of steepest ascent in all other coded variables  $x_i$ .
- b) Find and show the coordinates of this point on the path of steepest ascent from part a) using the natural variables.
- c) Find a unit vector that defines the path of steepest ascent.
- d) What step size multiplier for the unit vector in c) above would give the same point on the path of steepest ascent you found in parts a) and b)?
- e) Find the fraction of directions excluded by the 95% confidence cone of steepest ascent.
- 3 Consider the first order model:  $\hat{y} = 52.1 - 3.1x_1 + 6.4x_2 - 1.25x_3$ , where the variance of the parameter estimates,  $s_b^2$ , equals 0.4 computed based on 5 replications of the center point of the design. Does the point  $\mathbf{x}' = (-0.9, 1.0, -0.3)$  generate a direction vector inside the 95% confidence cone of steepest ascent?
- 4 Consider the tool-life experimental data shown in Table 2.6 [44]. Two factors were varied in a replicated  $2^2$  with three center runs. The tool life of the tool was considered by the first occurrence of 0.015 inch uniform flank wear, 0.004 inch crater depth, 0.030 inch localized wear, or catastrophic failure.

Table 2.6. Data for problems 4 and 5

Speed (sfm)	Feed (ipr)	Tool life (minutes)	
		Grade A	Grade B
500	0.015	67.0	84.4
500	0.015	101.9	91.2
500	0.015	63.6	66.7
800	0.015	23.5	16.0
800	0.015	17.6	15.2
800	0.015	21.3	17.6
500	0.027	17.9	24.6
500	0.027	25.3	15.3
500	0.027	25.4	30.4
800	0.027	0.4	1.1
800	0.027	0.6	0.5
800	0.027	0.5	0.9
650	0.021	21.4	11.8
650	0.021	19.2	8.9
650	0.021	22.6	10.6

a) Fit the best polynomial model you can to the “grade A” tool data. (Hint: transformations of the response may be necessary).

b) Find the direction of steepest ascent, and determine a point that would take the process to a distance approximately equal to the boundary of the current experimental region.

5 Consider the “grade B” tool life data in Table 2.6.

a) Fit the best polynomial model you can.

b) Determine a point on the direction of steepest ascent that would take the process to a distance approximately equal to the boundary of the current experimental region.

For the following problems, readers may wish to consult Appendix A.

6 Consider the following one factor experiment:

$\xi_1$	$y$ (observed)
650	7
800	18
650	6
800	11
725	10

a) Using regression analysis, find the least squares estimates and the estimated variance of the parameters estimates in the model:

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

where  $x_1$  is a coded factor corresponding to  $\xi_1$ .

b) Find Cook's  $D_i$  diagnostic statistics based on the studentized residuals ( $r_i$ ). Which point(s) seem to be outliers and/or influential?

c) Test the significance of the regression model. What are the null and the alternative hypotheses?

d) Test for lack of fit of the linear model. What are the null hypothesis and the alternative hypotheses?

7 Assume a 1st order model is to be fit in  $k$  factors. From  $\partial R(\beta)/\partial \beta_j = 0$ ,  $j = 0, 1, 2, \dots, k$ , find the set of normal equations. Do not use any matrix notation.

8 Consider the following “unit length” coding convention:

$$x_{iu} = \frac{\xi_{iu} - \bar{\xi}_i}{s_i}, \quad i = 1, \dots, k; \quad u = 1, 2, \dots, n$$

where

$$\bar{\xi}_i = \frac{\sum_{u=1}^n \xi_{iu}}{n}, \quad i = 1, \dots, k$$

and

$$s_i = \sqrt{\sum_{u=1}^n (\xi_{iu} - \bar{\xi}_i)^2}, \quad i = 1, \dots, k.$$

Show that the  $\mathbf{X}'\mathbf{X}$  matrix will contain the correlation matrix of the controllable factors. Assume a first order model for simplicity.

- 9 It is well-known in regression analysis that

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{Y} = (\mathbf{I} - \mathbf{H})\boldsymbol{\varepsilon}$$

- a) Interpret this expression using the concept of “projections” described in Appendix A. Draw a picture of the corresponding vectors to explain.
- b) Does this expression imply that  $\mathbf{Y} = \boldsymbol{\varepsilon}$ ? Explain.
- 10 Assume we fit a linear model with an intercept. Argue why  $\mathbf{1}'\mathbf{H} = \mathbf{1}'$  where  $\mathbf{1}$  is an  $n \times 1$  vector of ones. What does this imply for the average prediction,  $\bar{\hat{Y}}$  and the average response  $\bar{Y}$ ? Is  $\mathbf{1}'\mathbf{H} = \mathbf{1}'$  if the model has no intercept? Why?
- 11 Consider the basic ANOVA for a linear regression model. Show that the following expressions for the “corrected sums of squares” are true. Do *not* use the relation  $SS_{total} = SS_{reg} + SS_{error}$ :
- a)  $SS_{total} = (\mathbf{Y} - \bar{Y}\mathbf{1})'(\mathbf{Y} - \bar{Y}\mathbf{1}) = \mathbf{Y}'\mathbf{Y} - n\bar{Y}^2$ ;
- b)  $SS_{regression} = (\hat{\mathbf{Y}} - \bar{Y}\mathbf{1})'(\hat{\mathbf{Y}} - \bar{Y}\mathbf{1}) = \mathbf{Y}'\mathbf{X}\hat{\boldsymbol{\beta}} - n\bar{Y}^2$ . (Hint: use the result in Problem 10).
- 12 Show that  $\bar{h} = p/n$ . (Hint: use the fact that  $tr(\mathbf{AB}) = tr(\mathbf{BA})$ , where “tr” means the trace; see Appendix C.). Relate this fact to

$$\sum_{i=1}^n \frac{\text{Var } \hat{y}(\mathbf{x}_i)}{\sigma^2} = p.$$

What is this last expression saying in favor of simpler models (i.e., models with fewer parameters)?

- 13 Show that a random variable  $W \sim F_{1,n-p}$  is equivalent to  $W \sim t_{n-p}^2$ .



<http://www.springer.com/978-0-387-71434-9>

Process Optimization

A Statistical Approach

Del Castillo, E.

2007, XVIII, 462 p. 76 illus., Hardcover

ISBN: 978-0-387-71434-9