▩ 연구논문

A Graphical Method for Evaluating Mixture Designs with respect to the Slope

Dae-Heung Jang · Ho-Jun Na
Dept. of Applied Mathematics, National Fisheries University of Pusan

Abstract

Good estimation of the slopes of the mixture response function may be important as well as estimation of mean mixture response. It is possible to evaluate and compare several mixture designs with respect to the slope. A graphical method is proposed that allows us to evaluate a given design's support for the fitted model in terms of slope variance. We can plot variances of slopes along Cox direction according to existence of restriction of simplex region when comparing several different mixture designs.

1. Introduction

In mixture experiments, the measured response is assumed to depend only on the relative proportions of the components present in the mixture. For mixture experiments, if we let x_i represent the proportion of the ith component in the mixture where the number of components is q, then

$$0 \le x_i \le 1, \qquad i = 1, 2, \cdots, q,$$

and

$$\sum_{i=1}^{q} x_i = 1.$$

The experimental region is a regular $(q-1)$-dimensional simplex. When additional constraints are imposed on the proportions in the form of lower and upper bounds

$$0 < L_i \leq x_i \leq U_i < 1, \qquad i = 1, 2, \cdots, q,$$

the experimental region becomes a subregion of the simplex.

The general form of the second-degree Scheff polynomial in q components is

$$\eta = \sum_{i=1}^{q} \beta_i x_i + \sum \sum_{i<j}^{q} \beta_{ij} x_i x_j \tag{1}$$

The parameter β_i is defined as the heights of the surface above the simplex at $x_i = 1$, $x_j = 0$, $j \neq i$, whereas β_{ij} is a measure of the departure of the surface from the plane along the edge $x_i + x_j = 1$, respectively.

In recent years, much emphasis has been placed by practitioners not on finding optimum condition but on finding region where there is demonstrated improvement in response over that achieved by current operating conditions. In many applications of response surface methodology, good estimation of the derivatives of the response functions may be as important or perhaps more important than estimation of mean response. Certainly, the computation of a stationary point in a second-order analysis or the use of gradient techniques-for example, steepest ascent or ridge analysis-depends heavenly on the partial derivatives of the estimated response functions with respect to the design variables.

The main stream in statistical work on the design of experiments has been on the comparison of treatments, that is on the estimation of treatment contrasts. Estimation of differences in response at different points in the factor space will often be of great importance. Herzberg(1967), Box and Draper(1980), Huda and Mukerjee(1984), Koske(1989,) and Park(1990) have considered problems associated with estimation of differences in response. If differences at points close together are involved, estimation of the local slopes (the rate of change) of the response surface is of interest. Atkinson(1970), Murty and Studden(1972), Ott and Mendenhall(1972), Myers and Lahoda(1975), Hader and Park(1978), Mukerjee and Huda(1985), and Park(1987) have considered problems associated with estimation of derivatives of the response function. Also in mixture experiments, estimation of the local slopes of the response surface and designs comparison with respect to the slope are important.

The alphabetic criteria(for example, A-, D-, E-, G- and V-optimality) have been proposed for choosing a design. Although these single-valued criteria provide algorithms contained in software packages (for example, the ACED package (Welch,1982,1985)) with useful basis for generating experimental designs, the resulting designs are optimal only in the strict sense of the particular criterion chosen. There is little reason to believe that these designs are necessarily best in terms of overall prediction properties. Kiefer(1975) argues that when selecting a design we should consider how well the design performs over every part of the region of interest. Along this same line, Giovanniti-Jensen and Myers(1989) proposed a variance-based graphical approach for standard response surface designs that considers plots of the maximum, the minimum, and the spherical average of the prediction variance over spheres through a region of interest. Jang and Park(1993) proposed a slope-variance-based graphical approach for standard response surface designs.

Generally, the mixture design employed for fitting a model over a constrained region consists of a subset of the extreme vertices, some of the edge midpoints, some face centroids (when $q \geq 4$). The extreme vertices, edge midpoints, face centroids and so on are called candidate points. The exact subset of points chosen from the list of candidate points depends on the form of the proposed model to be fitted as well as the shape of the constrained region. For first- and second-order models, alphabetic criteria tend to select boundary designs where all the design points are positioned on the boundaries of the region of interest. Often, these criteria fail to convey the true nature of the design's support of the fitted model in terms of the variance of the prediction equation over the region of the interest. Thus, Vining, Cornell, and Myers(1993) proposed the method which can evaluate and compare mixture designs with respect to prediction variance under unconstrained region.

When measuring the slope effect of component i under constrained region, Cox direction is used. The reason for measuring the rate of change of the surface along Cox direction is to learn more about the mixture systems' surface characteristics in terms of the shape of the surface, the proximity of the maximum or minimum on the surface and so on. Hence, it is necessary to obtain the estimates of the slopes and their variances along Cox direction which may be of great help in an attempt to get the characteristics of the mixture response surface.

The purpose of this paper is to propose a measure for evaluating mixture designs with respect to the slope. In this paper, we discuss a measure of the slope variance for mixture experiment and propose a graphical method for evaluating mixture designs with respect to the slope variance.

2. A Graphical Method to Evaluate Mixture Designs with respect to the Slope

Let us assume that mixture experiment model is represented by

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon},$$

where $\underline{y}$ is the $n\times 1$ vector of responses, X is the $n\times p(\leq n)$ matrix of the component proportions and cross-products between the proportions depending on the model, $\underline{\beta}$ is the $p\times 1$ vector of unknown coefficients, and $\underline{\varepsilon}$ is the $n\times 1$ vector of random errors.

When the mixture component proportions are restricted by lower and upper bounds, these restrictions make the reference mixture whose coordinates are the averages of the coordinates of the extreme vertices of constrained simplex, namely, the centroid of constrained simplex. When measuring the slope effect of component i and a reference mixture other than the centroid of the simplex is to be used, Cox direction is generally appropriate. Cox direction is an imaginary line projected from the reference mixture which is usually the centroid of the constrained region to the vertex $x_1 = 1$(See Cornell(1990).). We can measure the slope of surface along Cox direction with the help of the Scheff polynomials.

Vining, Cornell, and Myers(1993) proposed a graphical approach for evaluating mixture designs with respect to the response surface. This method evaluate and compare mixture designs using prediction variances along Cox direction under constrained region. Similarly, it is possible to evaluate and compare second mixture designs with respect to slope variance.

Under the constrained region, let us denote the proportions of the q components at the reference mixture by $\underline{c} = (c_1, c_2, \cdots, c_q)$, where $c_1 + c_2 + \cdots + c_q = 1$. When the proportion c_i of component i is changed by an amount Δ_i in Cox direction, so that the new proportion becomes

$$x_i = c_i + \Delta_i \tag{2}$$

then the proportions of the remaining $q-1$ components resulting from the change from c_i in the ith component, is

$$\begin{aligned} x_j &= c_j - \frac{\Delta_i c_j}{1-c_i} \\ &= c_j \times \frac{1-c_i-\Delta_i}{1-c_i}, \qquad j=1, 2, \cdots, q, \ j\neq i \end{aligned} \tag{3}$$

Note that the ratio of the proportions for components j and k, where x_j and x_k are defined by (3), is

$$\frac{x_j}{x_k} = \frac{c_j}{c_k}$$

which is the same value as the ratio of components j and k at the reference mixture $\underline{c}$. From (2),

$$\Delta_i = x_i - c_i$$

and

$$x_j = c_j \frac{1-c_i-\Delta_i}{1-c_i} \tag{4}$$
$$= c_j \frac{1-x_i}{1-c_i}.$$

Substituting the expression (4) for x_j into (1), the expected response at x_i on the Cox direction is

$$\eta = \beta_i x_i + \sum_{j\neq i}^{q} \beta_j \frac{c_j}{1-c_i}(1-x_i)$$
$$+ \sum_{l=1}^{i-1} \beta_{li} \frac{c_l}{1-c_i}(1-x_i)x_i + \sum_{j=i+1}^{q} \beta_{ij} x_i \frac{c_j}{1-c_i}(1-x_i)$$
$$+ \sum\sum_{\substack{j<i \\ j,\,k\neq i}}^{q} \beta_{jk} \frac{c_j c_k}{(1-c_i)^2}(1-x_i)^2.$$

The slope of η with respect to component i, evaluated at x_i, is

$$D_{x_i}\eta = \frac{\partial\eta}{\partial x_i} = \alpha_{0i} + \alpha_{1i} x_i,$$

where

$$\alpha_{0i} = \beta_i - \sum_{j\neq i}^{q} \beta_j \frac{c_j}{1-c_i} + \sum_{l=1}^{i-1} \beta_{li} \frac{c_l}{1-c_i} + \sum_{j=i+1}^{q} \beta_{ij} \frac{c_j}{1-c_i} \tag{5}$$
$$-2\sum\sum_{\substack{j<k \\ j,\,k\neq i}}^{q} \beta_{jk} \frac{c_j c_k}{(1-c_i)^2},$$

$$\alpha_{0i} = 2\Big[\sum_{\substack{j<k \\ j,\,k\neq i}}^{q} \sum \beta_{jk} \frac{c_j c_k}{(1-c_i)^2} - \sum_{l=1}^{i-1} \beta_{li} \frac{c_l}{1-c_i} - \sum_{j=i+1}^{q} \beta_{ij} \frac{c_j}{1-c_i} \Big]. \tag{6}$$

An estimate for the value of the slope at x_i of the second-degree surface is obtained by substituting the least squares estimate b_i and b_{ij} for β_i and β_{ij}, respectively. The estimate of $D_{x_i}\eta$ is

$$\hat{D}_{x_i}\eta = a_{0i} + a_{1i} x_i,$$

where a_{0i} and a_{1i} are estimators of α_{0i} and α_{1i}, respectively. A more general formula for the estimate of slope is written by

$$\hat{D}_{x_i}\eta = \underline{s}_i' \, \underline{b}, \qquad i = 1, 2, \cdots, q,$$

where $\underline{b}$ is the least squares estimator of $\underline{\beta}$. For $q=4$,

$$\begin{aligned} \underline{s}_1' = [\, & 1,\ -\frac{c_2}{1-c_1},\ -\frac{c_3}{1-c_1},\ -\frac{c_4}{1-c_1},\ \frac{c_2}{1-c_1}(1-2x_1), \\ & \frac{c_3}{1-c_1}(1-2x_1),\ \frac{c_4}{1-c_1}(1-2x_1),\ -2\frac{c_2 c_3}{(1-c_1)^2}(1-x_1), \\ & -2\frac{c_2 c_4}{(1-c_1)^2}(1-x_1),\ -2\frac{c_3 c_4}{(1-c_1)^2}(1-x_1)\,], \end{aligned}$$

$$\begin{aligned} \underline{s}_2' = [\, & -\frac{c_1}{1-c_2},\ 1,\ -\frac{c_3}{1-c_2},\ -\frac{c_4}{1-c_2},\ \frac{c_1}{1-c_2}(1-2x_2), \\ & -2\frac{c_1 c_3}{(1-c_2)^2}(1-x_2),\ -2\frac{c_1 c_4}{(1-c_2)^2}(1-x_2),\ \frac{c_3}{1-c_2}(1-2x_2), \\ & \frac{c_4}{1-c_2}(1-2x_2),\ -2\frac{c_3 c_4}{(1-c_2)^2}(1-x_2)\,], \end{aligned}$$

$$\begin{aligned} \underline{s}_3' = [\, & -\frac{c_1}{1-c_3},\ -\frac{c_2}{1-c_3},\ 1,\ -\frac{c_4}{1-c_3},\ -2\frac{c_1 c_2}{(1-c_3)^2}(1-x_3), \\ & \frac{c_1}{1-c_3}(1-2x_3),\ -2\frac{c_1 c_4}{(1-c_3)^2}(1-x_3),\ \frac{c_2}{1-c_3}(1-2x_3), \\ & -2\frac{c_2 c_4}{(1-c_3)^2}(1-x_3),\ \frac{c_4}{(1-c_3)}(1-2x_3)\,], \end{aligned}$$

$$\underline{s}_4' = [\ -\frac{c_1}{1-c_4},\ -\frac{c_2}{1-c_4},\ -\frac{c_3}{1-c_4},\ 1,\ -2\frac{c_1c_2}{(1-c_4)^2}(1-x_4),$$

$$-2\frac{c_1c_3}{(1-c_4)^2}(1-x_4),\ \frac{c_1}{1-c_4}(1-2x_4),\ -2\frac{c_2c_3}{(1-c_4)^2}(1-x_4),$$

$$\frac{c_2}{1-c_4}(1-2x_4),\ \frac{c_3}{1-c_4}(1-2x_4)\ ].$$

The variance of the estimate of $\hat{D}_{x_i}\eta = \underline{s}_i'\underline{b}$ is

$$Var[\hat{D}_{x_i}\eta] = \underline{s}_i'(X'X)^{-1}\underline{s}_i\sigma^2.$$

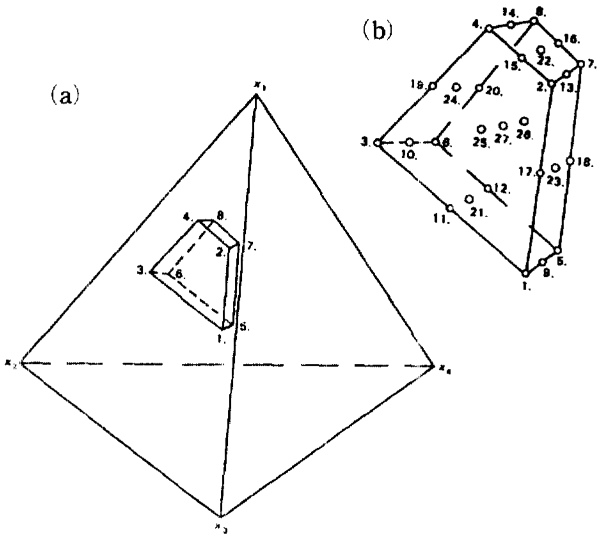
Let

$$SV(x_i) = \frac{Var[\hat{D}_{x_i}\eta]}{\sigma^2} = \underline{s}_i'(X'X)^{-1}\underline{s}_i.$$

A plot of $SV(x_i)$ along x_i's Cox direction, slope variance trace, is used to give comprehensive picture of the behavior of the variance of the slope estimate under constrained region and to evaluate and compare mixture designs with respect to slope variance under constrained region. As an alternative to a single-valued criterion, this graphical procedure is used to examine the relative strengths and weaknesses of mixture designs. Areas in the region of interest where slope estimation under constrained region is relative good and relative poor can be discussed with respect to every mixture design, and designs which have good overall performance with respect to the slope under constrained region are selected as the optimal designs. Therefore, as design selection criteria, we can consider the slope variance traces as well as the prediction variance traces.

3. Numerical Example

McLean and Anderson(1966) introduced the extreme vertices design and illustrated its use with their well-known flare data. The purpose of the experiment was to find the combination of the proportions of magnesium (x_1), sodium nitrate (x_2), strontium nitrate (x_3) and binder (x_4) for producing flares with maximum illumination. Engineering experience had indicated that the following constraints should be placed on each component $0.4 \le x_1 \le 0.6$, $0.1 \le x_2$, $x_3 \le 0.5$, $0.03 \le x_4 \le 0.08$. ⟨Figure 1⟩ shows the region of interest and candidate points.



〈 Figure 1 〉 (a) Region of intrerest and the eight extreme vertices
(b) candidate points for the McLean and Anderson flare experiment

〈Table 1〉 lists selected 15-point A-, D-, G-, V-optimal, and McLean-Anderson design that were generated from 27 candidate points. (See Vining, Cornell, and Myers(1993).) It is interesting that the McLean-Anderson design is far from optimal in term of any of the alphabetic design criteria since it does not contain any of the edge midpoints.

〈 Table 1 〉 Mixture experimental designs

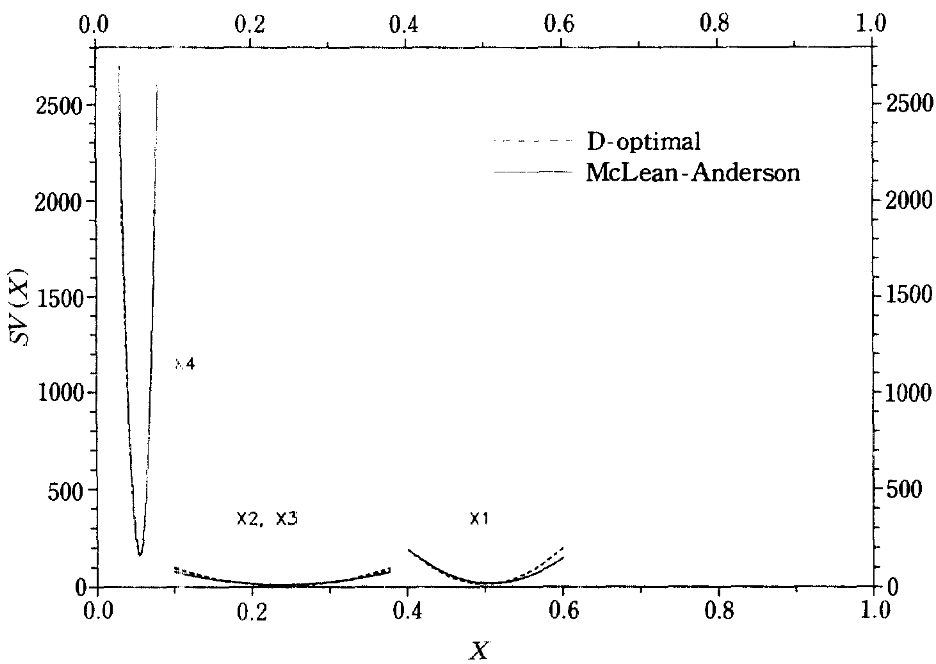
	M-A*	A	D	G	V
Extreme vertices, points 1-8	All	All	All	All	All
Midedge, points 9-20		9	9	9	10
		11	11	11	11
		13	13	14	14
		18	17	19	19
			18		
Face centroids,points 21-26	21	21	21	21	21
	22	24	24	23	23
	23	25		26	26
	24				
	25				
	26				
Overall centroid, point 27	27				

* M-A : McLean-Anderson design, A : A-optimal design, D : D-optimal design,
G : G-optimal design, V : V-optimal design

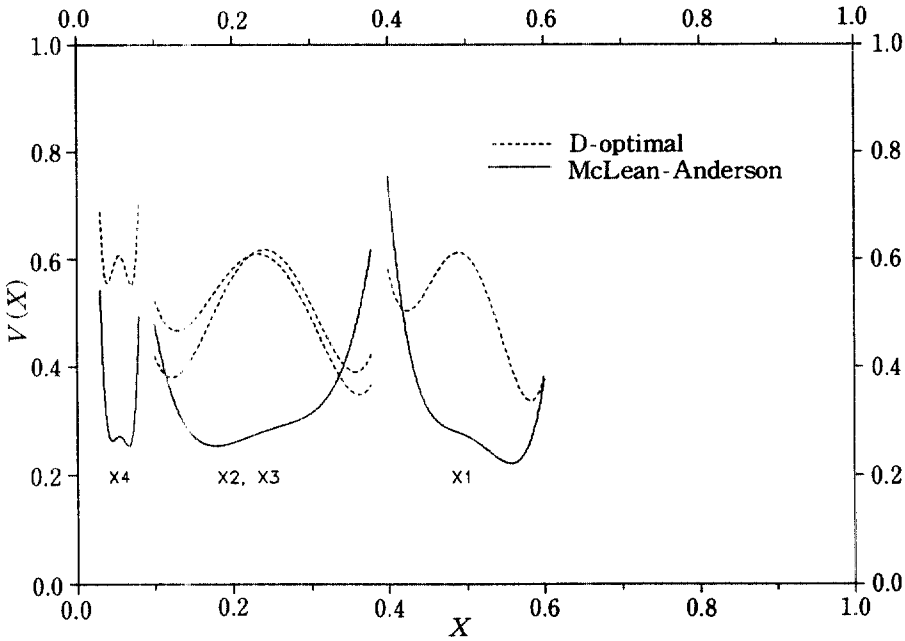
⟨Figure 2⟩ shows comparison of the slope variance traces for McLean-Anderson design and D-optimal design. For each design, we can know the values of $SV(x_i)$ along x_i's Cox direction. The values of $SV(x_2)$ along x_2's Cox direction are equal to the values of $SV(x_3)$ along x_3's Cox direction and the values of $SV(x_4)$ is very larger than the other values of $SV(x_i)$ in McLean-Anderson design and D-optimal design. The slope variance traces for McLean-Anderson design are similar to the slope variance traces for D-optimal design. The slope variance traces for McLean-Anderson design are similar to those for the alphabetic criteria designs because the slope variance traces for A-, G- and V-optimal designs are very similar to those for D-optimal design.

⟨Figure 3⟩ displays the plot of the prediction variance traces for the D-optimal design and the McLean-Anderson design. The original McLean-Anderson design is clearly superior to the D-optimal design and therefore to all the optimal designs because the prediction variance traces for A-, G- and V-optimal designs are very similarly to those for D-optimal design.

Consequently, McLean-Anderson design is superior to computer-generated designs with respect to the prediction variance and McLean-Anderson design is on a level with computer-generated designs with respect to the slope variance.



⟨ Figure 2 ⟩ Comparison of the slope variance traces for the McLean-Anderson design and D-optimal design



〈 Figure 3 〉 Comparison of the prediction variance traces for the D-optimal and McLean-Anderson designs

4. Conclusions

In the general mixture problem, the measured response is assumed to depend only on the relative proportions of the ingredients present in the mixture. In mixture experiments, estimation of the local slopes of the response surface and designs comparison with respect to the slope may be important. Single-valued criteria such as A-, D-, G- and V-optimality are used often in constructing and evaluating experimental designs. Unfortunately, these single-valued criteria often fail to convey the true nature of the design's support of the fitted model in terms of the variance of the prediction equation over the region of the interest.

In this paper, a measure for evaluating mixture designs with respect to the slope has been proposed and a graphical method have been proposed that allows us to evaluate a given design's support for the fitted model in terms of slope variance. That is, a plot of $SV(x_i)$ along x_i's Cox direction, slope variance trace, has used to give comprehensive picture of the behavior of the variance of the slope estimate and to evaluate and compare mixture designs with respect to slope variance.

The advantages of this method can be stated as follows :

(1) Since single-valued criteria do not always reveal the true picture of how a particular design will support the fitted model in terms of its prediction variance, our method can be used as an additional tool for these single-valued criteria.

(2) Graphical procedures such as plots of the slope variance trace along Cox direction across the experimental region are an important tool for evaluating mixture design.

(3) As design selection criteria, we can consider the slope variance traces as well as the prediction variance traces.

References

[1] Atkinson, A. C. (1970), "The Design of Experiments to Estimate the Slope of a Response Surface," *Biometrica*, 57, pp. 319－328.

[2] Box, G. E. P. and Draper, N. R. (1980), "The Variance Function of the Difference Between Two Estimated Response," *Journal of the Royal Statistical Society*, B 42, pp. 79－82.

[3] Cornell, J. A. (1990), *Experiments with Mixtures*, John Wily, New York.

[4] Giovanniti-Jensen, A. and Myers, R. H. (1989), "Graphical Assessment of the Prediction Capability of Response Surface Designs," *Technometrics*, 31, pp. 159－172.

[5] Hader, R. J. Park, S. H. (1978), "Slope-Rotatable Central Composite Designs," *Technometrics*, 20, pp. 413－417.

[6] Herzberg, A. M. (1967), "The Behaviour of the Variance Function of the Difference Between Two Estimated Responses," *Biometrica*, 71, pp. 173－178.

[7] Huda, S. and Mukerjee, R. (1984), "Minimizing the Maximum Variance of the Difference Between Two Estimated Responses," *Biometrica*, 71, pp. 381－385.

[8] Jang, D. H. and Park, S. H. (1993), "A Measure and a Graphical Method for Evaluating Slope Rotatability in Response Surface Designs," *Communications in Statistics－Theory and Methods*, 22(7), pp. 1849－1863.

[9] Kiefer, J. C. (1975), "Optimal Design: Variation in Structure and Performance under Change of Criterion," Biometrica, 62, pp. 277－288.

[10] Koske, J. K. A. (1989), "The Variance Function of the Difference Between Two Estimated Fourth Order Response Surfaces," *Journal of Statistical Planning and Inference*, 23, pp. 263－266.

[11] Mclean, R. A. and Anderson, V. L. (1966), "Extreme Vertices Design of Mixture Experiments," *Technometrics*, 8, pp. 447 – 454.

[12] Mukerjee, R. and Huda, S. J. (1985), "Minimax Second- and Third-Order Designs to Estimate the Slope of a Response Surface," *Biometrica*, 72, pp. 173 – 178.

[13] Murty, V. N. and Studden, W. J. (1972), "Optimal Designs for Estimating the Slope of a Response Surface," *Journal of the American Statistical Association*, 67, pp. 869 – 873.

[14] Myers, R. H. and Lahoda, S. J. (1975), "A Generalization of the Response Surface Mean Square Error Criterion with a Specific Application to the Slope," *Technometrics*, 17, pp. 481 – 486.

[15] Ott, L. and Mendenhall, W. (1972), "Designs for Estimating the Slope of a Second Order Linear Model," *Technometrics*, 14, pp. 341 – 353.

[16] Park, J. Y. (1990), "Designs for Estimating the Difference Between Two Responses," *Communications in Statistics – Theory and Methods*, 19(12), pp. 4773 – 4787.

[17] Park, S. H. (1987), "A Class of Multifactor Designs for Estimating the Slope of Response Surfaces," *Technometrics*, 29, pp. 449 – 453.

[18] Vining, G. G., Cornell, I. A. and Myers, R. H. (1993), "A Graphical Approach for Evaluating Mixture Designs," *Applied Statistics*, 42, pp. 127 – 138.

[19] Welch, W. J. (1982), "Branch-and-Bound Search for Experimental Designs Based on D-optimality and Other Criteria," *Technometrics*, 24, pp. 41 – 48.

[20] ________ (1985), ACED : "Algorithms for the Construction of Experimental Designs," *American Statistician*, 39, p. 146.