Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Optimality VIS-À-VIS robustness in mixture models with heteroscedastic error

Madhura Mandal & Ganesh Dutta

To cite this article: Madhura Mandal & Ganesh Dutta (11 Oct 2023): Optimality VIS-À-VIS robustness in mixture models with heteroscedastic error, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2023.2265588

To link to this article: https://doi.org/10.1080/03610918.2023.2265588

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2023.2265588

Check for updates

Optimality VIS-À-VIS robustness in mixture models with heteroscedastic error

Madhura Mandal[a] and Ganesh Dutta[b]

[a]Department of Statistics, Vivekananda Mahavidyalaya, Burdwan-713103, West Bengal, India; [b]Department of Statistics, Basanti Devi College, 147B, Rash Behari Avenue, Kolkata-700029, West Bengal, India

ABSTRACT

Mixture models and designs are used in situations where the response depends on the proportions of the factors (components). Optimum designs were derived for mixture models with fixed regression parameters under homoscedastic error variance by several authors. In this paper, an attempt has been made to find D- and A-optimum designs for the estimation of model parameters with heteroscedastic error variance. It is assumed that the error variance has constant value for all points equidistant from the center of the design. Equivalence theorem plays an important role in this investigation. Robustness of the standard optimum designs in the homoscedastic case under changes in the error variances has also been studied.

ARTICLE HISTORY
Received 19 May 2021
Accepted 6 September 2023

AMS (2000) SUBJECT CLASSIFICATION: 62K25

KEYWORDS
Mixture models; Heteroscedastic error variance; Optimum designs; Robustness

1. Introduction

In a mixture experiment, the response depends on the proportions $x_1, \ldots, x_q$ of a number of components varying in the factor space

$$\mathcal{X} : \left\{ x_i \geq 0, \;\; i = 1, 2, \ldots, q; \;\; \sum x_i = 1 \right\}. \tag{1.1}$$

Schefé (1958, 1963) introduced canonical models of different degrees to represent the response function η_x. He also introduced Simplex Lattice Designs and Simplex Centroid Designs in such situations. Optimality of mixture designs for the estimation of parameters of the response function was considered by several authors, see e.g. Kiefer (1961), Galil and Kiefer (1977), Draper and Pukelsheim (1999). Optimum design for the estimation of some non-linear function of the parameters of a mixture model has also been considered (see e.g. Pal and Mandal (2006)). Optimality aspects of mixture designs can be found in the recent monograph by Sinha et al. (2014).

Mixture models can be applied in experiments of different fields such as fruit punch with watermelon, pineapple and orange; chick feeding with protein, fat and carbohydrate; concrete batches where hardness is measured from a cement consisting of three primary raw materials : clay, limestone and fly ash ; surface resistivity of paper coatings of different blends of chemicals etc. (see Cornell (1977)). In most of these investigations, it was assumed that the error variance is homoscedastic. However, in many practical situations the error variance, as in response surface model, may be heteroscedastic and may well be a function of the mixture components. Optimality in the context of response surface model under such circumstances has been considered by several authors (see e.g. Atkinson and Cook (1995), Dette and Wong (1996), Rodríguez and Ortiz (2005)). Analysis of experiments with heteroscedastic

CONTACT Ganesh Dutta duttaganesh78@gmail.com Department of Statistics, Basanti Devi College, 147B, Rash Behari Avenue, Kolkata-700029, West Bengal, India.

error variance in a mixture model has been considered in Cornell (2002). Cornell (2002) considered three distinct variances corresponding to three sets of design points with different radii and compared the unweighted and weighted estimates of the variance components using data from an experiment consisting of ground beef and peanut meal patties to illustrate the comparison. For a more detailed discussion see Cornell (2002). However the author did not consider the optimality aspect of the design. Yan, Zhang, and Peng (2017) have also considered heteroscedastic error variance of exponential nature for additive mixture model and studied D- and A-optimality for the estimation of the model parameters. It seems Yan, Zhang, and Peng (2017) first considered the problem of determining optimum designs for mixture models with heteroscedastic error variance. They worked with three types of additive mixture models in canonical form of degree one (cf. Becker (1968, 1978)). Again, they used simple exponential function as error function. They established that under certain conditions, the direct sum of D- and A-optimal designs for homogeneous models in sub-mixture system is also D- and A-optimum for the additive mixture model. In this paper, we have considered the classical mixture model in canonical form of degree 1 and 2 introduced by Schefé (1968) and D- and A-optimal designs have been investigated. Moreover, we have considered a particular form of error function which is constant for all points equidistant from the center of the design.

In this paper, assuming an error function which takes constant values at points equidistant from the centroid of the simplex, optimum designs are derived using D- and A-optimality criteria. It will be seen that, for the error functions assumed, the problem is invariant with respect to the components. Because of the invariance, for the two optimality criteria considered, we restrict to the class of invariant designs. Restricting to the invariant subclass $\mathcal{D}_0$ of the (q, m) simplex designs, optimum weights at the support points are determined using D- and A- optimality criteria. A (q, m) simplex design for q components consists of points defined by the following coordinate settings: the proportions assumed by each components takes the $m+1$ equally spaced values from 0 to 1 i.e. $x_i = 0, \frac{1}{m}, \frac{2}{m}, \ldots, 1$ for $i = 1, 2, \ldots, q$ satisfying $\sum_{i=1}^{q} x_i = 1$. Optimality of such designs in the entire class is then examined using equivalence theorem. Moreover, the performance of optimum designs under homoscedastic error variance has been examined when the true error variance is heteroscedastic. Only the first and second order models are considered.

The paper is organized as follows. In Sec. 2, the problem will be formulated and D-and A-optimum designs will be derived in a subclass $\mathcal{D}_0$ of the (q, m) simplex designs. Optimality of the derived designs will be examined in the whole class in Sec. 3. In Sec. 4, robustness of the optimum designs with homoscedastic error variance will be investigated, when the true error variance is heteroscedastic. The concluding remarks are presented in Sec. 5.

2. Optimum designs in $\mathcal{D}_0$

In this section, we first derive the expression of the criterion function for D-and A-optimality criteria under heteroscedastic error variance. We restrict to first and second degree mixture models in canonical form as introduced by Schefé (1958):

$$E(y|\boldsymbol{x}) = \sum \beta_i x_i, \tag{2.1}$$

$$E(y|\boldsymbol{x}) = \sum \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j. \tag{2.2}$$

Instead of assuming an uncorrelated homoscedastic error variance, the following uncorrelated but heteroscedastic error variance is assumed which is constant at all points equidistant from the centroid of the simplex:

$$V(y|\boldsymbol{x}) = \frac{1}{\lambda(\boldsymbol{x})} = \exp\left((\boldsymbol{x} - \boldsymbol{x}_0)'(\boldsymbol{x} - \boldsymbol{x}_0)\right) \tag{2.3}$$

$$\boldsymbol{x} \in \mathcal{X} = \left\{x_i \geq 0,\ \ i = 1, 2, ..., q;\ \ \sum x_i = 1\right\},\ \ \boldsymbol{x}_0 = (1/q, 1/q, ..., 1/q)'. \tag{2.4}$$

We first restrict our considerations to the sub-class $\mathcal{D}_0$ of (q, m) simplex lattice designs and derive D- and A-optimum designs in this class. Optimality of such designs in the whole class is then examined using equivalence theorem.

2.1. First degree model

For the first degree model (2.1), consider a design with support points only at the vertices of the simplex. Since for the variance function (2.3), all the support points have the same variance, the information matrix remains same as in the homoscedastic case except for a scalar multiple. Hence the optimum design in the homoscedastic case remains optimum for the error variance assumed here for all the optimality criteria and in particular with respect to D- and A-optimality criteria.

2.2. Second degree model

Consider a second degree model in canonical form with q components given by (2.2). Because of the constraint $\sum x_i = 1$, model (2.2) can also be expressed as

$$\eta_{\boldsymbol{x}} = \sum \theta_{ii} x_i \left(x_i - \frac{1}{2}\right) + \sum_{i<j} \theta_{ij} x_i x_j = f'(\boldsymbol{x})\boldsymbol{\theta} \tag{2.5}$$

where $\boldsymbol{\theta} = (\theta_{11}, \theta_{22}, ..., \theta_{qq}, \theta_{12}, \theta_{13}, ..., \theta_{q-1,q})'$ and the parameters in (2.2) and (2.5) are related by

$$\boldsymbol{\beta} = \left(\beta_1, \beta_2, ..., \beta_q, \beta_{12}, \beta_{13}, ..., \beta_{q-1,q}\right)' = \boldsymbol{L\theta} \tag{2.6}$$

$\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are parameter vectors corresponding to the models (2.2) and (2.5) respectively; $\mathbf{L}$ is given by

$$\boldsymbol{L} = \begin{pmatrix} \frac{1}{2}\boldsymbol{I}_q & \boldsymbol{0} \\ -\boldsymbol{P} & \boldsymbol{I}_{C(q,2)} \end{pmatrix} \tag{2.7}$$

where $C(a, b)$ stands for the usual binomial coefficient involving positive integers, $a \geq b > 0$. Here $\mathbf{P}$ is a $C(q, 2) \times q$ matrix whose (i, j)-th row $(i < j)$ has element 1 in the ith and jth place and zero elsewhere i.e.

$$\boldsymbol{P}^{C(q,2)\times q} = \begin{matrix} \textit{Row no.} \\ (1,2) \\ (1,3) \\ (1,4) \\ \vdots \\ (q-1,q) \end{matrix} \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix}.$$

Consider a design ξ which puts equal mass $\frac{\alpha}{q}$ at each of the vertex points and a mass $\frac{1-\alpha}{C(q,2)}$ at each of the midpoints of the edges of the simplex. Let us denote such a subclass of designs by $\mathcal{D}_0$. Now the error function assumes two distinct values namely σ_1^2 at the vertex points and σ_2^2 at the midpoints of the edges with

$$\sigma_i^2 = \frac{1}{\lambda(\boldsymbol{x}_i)} = \exp\left\{(\boldsymbol{x}_i - \boldsymbol{x}_0)'(\boldsymbol{x}_i - \boldsymbol{x}_0)\right\};\ \ i = 1, 2 \tag{2.8}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ correspond to the vertex points and midpoints of the edges respectively. The moment matrix of such weighted $(q,2)$ simplex design for model (2.5), with heteroscedastic error variances (2.8), is given by

$$\begin{aligned} \boldsymbol{M} &= \text{Diag}\left(\frac{1}{4\sigma_1^2}\mathbf{1}'_q, \frac{1}{16\sigma_2^2}\mathbf{1}'_{C(q,2)}\right)\text{Diag}\left(\frac{\alpha}{q}\mathbf{1}'_q, \frac{1-\alpha}{C(q,2)}\mathbf{1}'_{C(q,2)}\right) \\ &= \text{Diag}\left(\frac{\alpha}{4q\sigma_1^2}\mathbf{1}'_q, \frac{1-\alpha}{16C(q,2)\sigma_2^2}\mathbf{1}'_{C(q,2)}\right) \end{aligned} \tag{2.9}$$

For the D-optimum design, we have to maximize the determinant of (2.9). Now

$$\text{Det}(\boldsymbol{M}) = \left(\frac{\alpha}{4q\sigma_1^2}\right)^q \times \left(\frac{1-\alpha}{16C(q,2)\sigma_2^2}\right)^{C(q,2)} \tag{2.10}$$

and (2.10) is maximized at

$$\alpha = \alpha_{0D} = \frac{q}{C(q+1,2)} = \frac{2}{q+1} \tag{2.11}$$

and the maximum value is given by

$$\text{Det}(\boldsymbol{M}) = \frac{1}{2^{q(2q+1)}(q(q+1))^{\frac{q(q+1)}{2}}\sigma_1^{2q}\sigma_2^{q(q-1)}} \tag{2.12}$$

Here we see that all the support points have equal weights, as is known otherwise also since the number of distinct support points is same as the number of parameters in the model. Moreover, it is independent of the form of the error function. Now it is only to check whether such a design ξ_D^* in $\mathcal{D}_0$ is D-optimum in the entire class or not for both models (2.2) and (2.5). This will be investigated in Sec. 3.

For the A-optimality criterion also, we first find A-optimum design ξ_A^* in $\mathcal{D}_0$ minimizing tr.$(\mathbf{M}^{-1})$, where $\mathbf{M}$ is given by (2.9). This will provide optimum design in $\mathcal{D}_0$ for the estimation of the parameters of the model (2.5). Then the status of this ξ_A^* will be examined in the entire class in Sec. 3. Now, it is easy to see that, tr.$(\mathbf{M}^{-1})$ is minimized at

$$\alpha = \alpha_{0A} = \frac{\sigma_1}{\sigma_1 + (q-1)\sigma_2} \tag{2.13}$$

and the minimum value is given by

$$\text{tr.}(\boldsymbol{M}^{-1}) = 4q^2(\sigma_1 + (q-1)\sigma_2)^2. \tag{2.14}$$

It is clearly seen that the optimum weights are not same at all the support points of the $(q,2)$ simplex design as in D-optimum design. This is true in the case of homoscedastic error variance also. Again, in the case of D-optimality criterion, optimum weights are independent of the error variance structure while, for the A-optimality criterion, it depends on the error variance. Such a design is A-optimum in $\mathcal{D}_0$ for the estimation of the parameters of the model (2.5). To find A-optimum design for the estimation of the parameters of the original model (2.2), we have to find a design minimizing tr.$(D(\widehat{\boldsymbol{\beta}}))$. Now because of the relation (2.6), we have

$$\text{tr.}(D(\widehat{\boldsymbol{\beta}})) = \text{tr.}(\boldsymbol{L}D(\widehat{\boldsymbol{\theta}})\boldsymbol{L}') = \text{tr.}(\boldsymbol{M}^{-1}\boldsymbol{L}'\boldsymbol{L}). \tag{2.15}$$

where $\mathbf{L}$ is given in (2.7). Now

$$\boldsymbol{L}'\boldsymbol{L} = \begin{pmatrix} \frac{4q-7}{4}\boldsymbol{I}_q + \boldsymbol{J}_q & -\boldsymbol{P}' \\ -\boldsymbol{P} & \boldsymbol{I}_{C(q,2)} \end{pmatrix}; \ \boldsymbol{J}_q = \mathbf{1}_q\mathbf{1}'_q \tag{2.16}$$

so that

$$\text{tr.}(D(\widehat{\boldsymbol{\beta}})) = \text{tr.}(\boldsymbol{M}^{-1}\boldsymbol{L}'\boldsymbol{L}) = \frac{(4q-3)q^2\sigma_1^2}{\alpha} + \frac{4q^2(q-1)^2\sigma_2^2}{1-\alpha} \geq q^2\left(\sigma_1\sqrt{4q-3} + 2(q-1)\sigma_2\right)^2 \quad (2.17)$$

Equality holds in (2.17) at

$$\alpha = \alpha_{0A}^* = \frac{\sigma_1(4q-3)^{1/2}}{\sigma_1(4q-3)^{1/2} + 2\sigma_2(q-1)}. \quad (2.18)$$

Thus a weighted (q,2) simplex design with α_{0A}^* given by (2.18) is A-optimum in $\mathcal{D}_0$ for the estimation of parameters of the model (2.2). In the next section, optimality of the design ξ_A^{**} will be examined in the entire class of designs using equivalence theorem.

3. Verification of optimality

In Sec. 2, for the second degree model, we have derived optimum designs using D- and A-optimality criteria in the subclass $\mathcal{D}_0$ of designs with support points only at the vertices and the mid-points of the edges of the simplex. In this section, we will examine the status of these designs in the entire class. Equivalence theorem plays a key role in this investigation. Because of the complexity of the problem due to heteroscedasticity, it is not possible to establish the optimality of the designs algebraically. Instead, we will examine the status of the optimum designs ξ_D^*, ξ_A^* and ξ_A^{**} obtained in Sec. 2 in the entire class numerically.

Kiefer and Wolfowitz (1960) first established the equivalence between D- and G-optimality criteria. (For different optimality criteria see Silvey (1980), Pukelsheim (1993)). Fedorov (1971) extended it to linear optimality criteria. Finally, Kiefer (1974) generalized it to any concave function of the moment/information matrix.

Suppose that $\Phi(\mathbf{M})$ is a real valued function defined on the moment matrix $\mathbf{M}(\xi)$ for the model $\eta(\mathbf{x}) = f(\mathbf{x})'\boldsymbol{\theta}$. Then the equivalence theorem can be stated as follows:

Theorem 3.1. If $\Phi(\mathbf{M})$ is a concave function of $\mathbf{M}(\xi)$, then ξ^* is Φ-optimal if and only if $F(\mathbf{M}(\xi^*), f(\mathbf{x})f(\mathbf{x})') \leq 0$ for all $\mathbf{x} \in X$.

If ξ^* is discrete with finite support then $F(\mathbf{M}(\xi^*), f(\mathbf{x})f(\mathbf{x})') = 0$, at all the support points.

Here $F(\mathbf{M}_1, \mathbf{M}_2)$ stands for the Fr é chet derivative of $\mathbf{M}_1$ in the direction $\mathbf{M}_2$.

For details see Silvey (1980). In particular, for the D- and A-optimality criteria, the equivalence theorem can be stated as follows (see Fedorov (1972)):

The equivalence theorem for D-optimality criterion:

Theorem 3.2. The following assertions:

i. ξ^* minimizes $\text{Det}(\mathbf{M}^{-1}(\xi^*))$
ii. ξ^* minimizes $\max_{\mathbf{x}} \lambda(\mathbf{x})f'(\mathbf{x})\mathbf{M}^{-1}(\xi^*)f(\mathbf{x})$
iii. $\max_{\mathbf{x}} \lambda(\mathbf{x})f'(\mathbf{x})\mathbf{M}^{-1}(\xi^*)f(\mathbf{x}) = p$

are equivalent, where p is the number of parameters in the model. Any linear combination of designs satisfying (*i*)-(*iii*) also satisfies (*i*)-(*iii*).

The equivalence theorem for A-optimality criterion:

Theorem 3.3. The following assertions:

i. ξ^* minimizes $\text{tr.}(\mathbf{M}^{-1}(\xi^*))$
ii. ξ^* minimizes $\max_{\mathbf{x}} \lambda(\mathbf{x})f'(\mathbf{x})\mathbf{M}^{-2}(\xi^*)f(\mathbf{x})$

iii. $\max_{\mathbf{x}} \lambda(\mathbf{x}) f'(\mathbf{x}) \mathbf{M}^{-2}(\xi_A^*) f(\mathbf{x}) = \text{tr.}(\mathbf{M}^{-1})$

are equivalent. Any linear combination of designs satisfying (i)-(iii) also satisfies (i)-(iii).

Using the equivalence theorem, Kiefer (1961) established the D-optimality of the $(q,2)$ simplex design. Afterwards, several authors used this theorem to establish the optimality/non-optimality of designs in different settings. Pal and Mandal (2007) used equivalence theorem for the problem of estimation of some specific non-linear functions in mixture models.

3.1. D-optimum designs

In Sec. 2, it is shown that the design ξ_D^* which puts equal weight α/q at each of the vertex points and a weight $(1-\alpha)/C(q,2)$ at each of the midpoints of the edges with α given by (2.11) is D-optimum in $\mathcal{D}_0$. Now the error function (2.3) assumes two distinct values namely σ_1^2 at the vertex points and σ_2^2 at the midpoints of the edges.

After a little algebra, we have

$$d(\boldsymbol{x}, \xi_D^*) = f'(\boldsymbol{x}) \boldsymbol{M}^{-1}(\xi_D^*) f(\boldsymbol{x}) = p\left[4\sigma_1^2 \left\{\sum_i x_i^2 \left(x_i - \frac{1}{2}\right)^2\right\} + 16\sigma_2^2 \left\{\sum_{i<j} x_i^2 x_j^2\right\}\right]. \tag{3.1}$$

Thus, ξ_D^* is D-optimum, iff $\lambda(\mathbf{x}) d(\mathbf{x}, \xi_D^*) \leq p$ or equivalently

$$\lambda(\boldsymbol{x})\left[4\sigma_1^2 \left\{\sum_i x_i^2 \left(x_i - \frac{1}{2}\right)^2\right\} + 16\sigma_2^2 \left\{\sum_{i<j} x_i^2 x_j^2\right\}\right] \leq 1 \tag{3.2}$$

for all $\mathbf{x} \in \mathcal{X}$. It is clearly seen that equality holds at the support points of the design ξ_D^*. Thus to establish the optimality of ξ_D^* in the entire class, it is enough to show that (3.2) holds for any arbitrary point of $\mathcal{X}$, other than the support points. However, since it is difficult to establish algebraically, we have examined it computationally for different values of q.

For $q=2$, because of the constraint $x_1 + x_2 = 1$, the condition (3.2) can be expressed in terms of a single variable x_1:

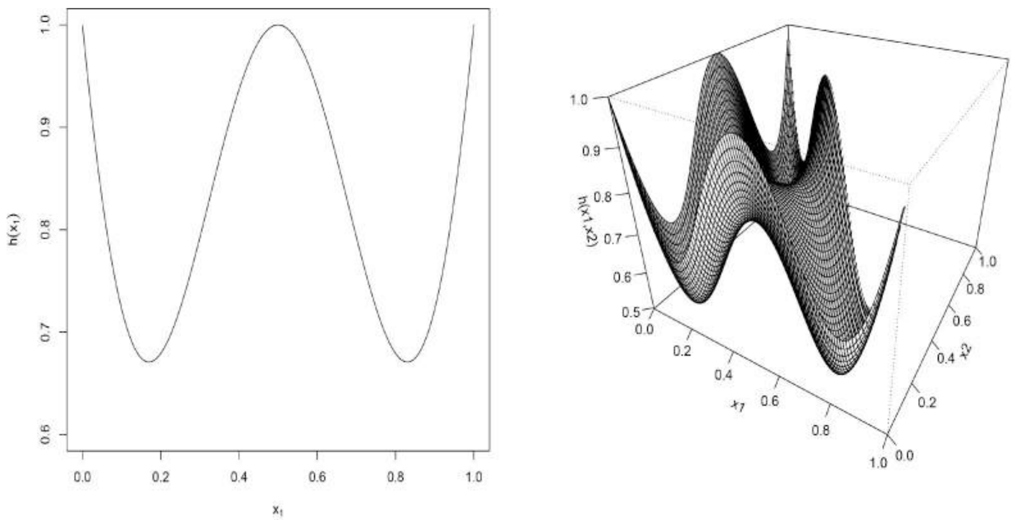
$$\begin{aligned} &4\left\{x_1^2\left(x_1 - \frac{1}{2}\right)^2 + (1-x_1)^2\left(x_1 - \frac{1}{2}\right)^2\right\} \exp\left\{1 - (x_1^2 + (1-x_1)^2)\right\} \\ &\quad + 16\left\{x_1^2(1-x_1)^2\right\} \exp\left\{\frac{1}{2} - (x_1^2 + (1-x_1)^2)\right\} \leq 1 \end{aligned} \tag{3.3}$$

From Figure 1a, it is clearly seen that the left hand side of (3.3) attains its maximum value 1 at the three points $x_1 = 0$, 1/2, 1 which validates (1,0), (1/2, 1/2), and (0,1) as the support points of the design and ξ_D^* is indeed D-optimum in the entire class of competing designs in $\mathcal{X}$.

Similarly, for $q = 3$, because of the constraint $\sum x_i = 1$, the left hand side of (3.2) can be expressed as a function of two variables x_1 and x_2:

$$\begin{aligned} &4\left\{x_1^2\left(x_1 - \frac{1}{2}\right)^2 + x_2^2\left(x_2 - \frac{1}{2}\right)^2 + (1-x_1-x_2)^2\left(x_1 + x_2 - \frac{1}{2}\right)^2\right\} \exp\left\{1 - (x_1^2 + x_2^2 + (1-x_1-x_2)^2)\right\} \\ &\quad + 16\left\{x_1^2 x_2^2 + (1-x_1-x_2)^2(x_1^2 + x_2^2)\right\} \exp\left\{\frac{1}{2} - (x_1^2 + x_2^2 + (1-x_1-x_2)^2)\right\} \leq 1 \end{aligned} \tag{3.4}$$

Here again, equality holds in (3.4) at the support points of ξ_D^*, which can be seen from Figure 1b also. Moreover, at all other points besides the support points, value of the left hand side of (3.4) is strictly less than 1 which guarantees the optimality of ξ_D^* in the entire class.



(a) Showing the graph of $h(x_1) = \frac{\lambda(\mathbf{x})d(\mathbf{x},\xi_D^*)}{3}$ for $x_1 \in [0,1]$

(b) Showing the graph of $h(x_1,x_2) = \frac{\lambda(\mathbf{x})d(\mathbf{x},\xi_D^*)}{6}$ for $x_1, x_2 \in [0,1]$

Figure 1. (a) Showing the graph of $h(x_1) = \frac{\lambda(\mathbf{x})d(\mathbf{x},\xi_D^*)}{3}$ for $x_1 \in [0,1]$ (b) Showing the graph of $h(x_1,x_2) = \frac{\lambda(\mathbf{x})d(\mathbf{x},\xi_D^*)}{6}$ for $x_1, x_2 \in [0,1]$.

Similarly, for $q = 4$, because of the constraint $\sum x_i = 1$, the left hand side of (3.2) can be expressed as a function of three variables x_1, x_2 and x_3:

$$4\left\{x_1^2\left(x_1-\frac{1}{2}\right)^2 + x_2^2\left(x_2-\frac{1}{2}\right)^2 + x_3^2\left(x_3-\frac{1}{2}\right)^2 + (1-x_1-x_2-x_3)^2\left(x_1+x_2+x_3-\frac{1}{2}\right)^2\right\}$$
$$\exp\left\{1-(x_1^2+x_2^2+x_3^2+(1-x_1-x_2-x_3)^2)\right\} + 16\left\{x_1^2(x_2^2+x_3^2)+x_2^2x_3^2+(1-x_1-x_2-x_3)^2(x_1^2+x_2^2+x_3^2)\right\}$$
$$\exp\left\{\frac{1}{2}-(x_1^2+x_2^2+x_3^2+(1-x_1-x_2-x_3)^2)\right\} \leq 1 \tag{3.5}$$

From (3.5), it is clear that $\lambda(\mathbf{x})d(\mathbf{x},\xi_D^*)$ is maximum at the support points of ξ_D^*, the (q,2) simplex design, with the maximum value p, the number of parameters and it is strictly less than the upper bound at all other points (see Appendix Table A1). Hence we conclude that the design ξ_D^* is D-optimum in the whole class for $q = 4$ also.

Remark 3.1. The design which is D-optimum for the estimation of the parameters of the model (2.5) is also D-optimum for the estimation of the parameters of the model (2.2). This is due to invariance of the D-optimality criterion with respect to the nonsingular transformation to the parameter vector.

Remark 3.2. Since $\lambda(\mathbf{x})\sigma_i^2 = \exp\{-(\mathbf{x}-\mathbf{x}_0)'(\mathbf{x}-\mathbf{x}_0)\}\ \exp\{(\mathbf{x}_i-\mathbf{x}_0)'(\mathbf{x}_i-\mathbf{x}_0)\} = \exp(-\mathbf{x}'\mathbf{x}+2\mathbf{x}'\mathbf{x}_0 -\mathbf{x}_0'\mathbf{x}_0)\ \exp(\mathbf{x}_i'\mathbf{x}_i-2\mathbf{x}_i'\mathbf{x}_0+\mathbf{x}_0'\mathbf{x}_0) = \exp\left(-\mathbf{x}'\mathbf{x}-\frac{1}{q}\right)\ \exp\left(\mathbf{x}_i'\mathbf{x}_i+\frac{1}{q}\right) = \exp(-\mathbf{x}'\mathbf{x})\ \exp(\mathbf{x}_i'\mathbf{x}_i);\ \ i = 1,2,$ instead of (2.1), one assumes the error function

$$V(y|\boldsymbol{x}) = \frac{1}{\lambda(\boldsymbol{x})} = \exp(\boldsymbol{x}'\boldsymbol{x}), \tag{3.6}$$

the necessary and sufficient condition of the equivalence theorem remains same for any arbitrary q. Hence the derived designs are D-optimum in the entire class for the error function (3.6) also.

Remark 3.3. In general, it is difficult to establish the optimality of the derived designs in the entire class using equivalence theorem algebraically. Here we have checked the optimality of the derived designs in the whole class numerically for $q=2$, 3 and 4. It may be conjectured that the design ξ_D^* is D-optimum in the entire class for all q.

3.2. A-optimum design

Here again, as in the case of D-optimality criterion, first A-optimum design is derived in the class of (q,2) simplex design under heteroscedastic error variance (3.1). Next, optimality of the derived design is investigated in the entire class *via* equivalence theorem. The information matrix $\mathbf{M}(\xi_A^*)$ of the design ξ_A^* for the model (2.5) is given in (2.9).

In Sec. 2, we have observed that ξ_A^* is A-optimum in $\mathcal{D}_0$ when $\alpha = \alpha_{0A}$, given by (2.13). To check optimality of the derived design in the entire class we have to verify it through equivalence theorem which, after a brief algebra, reduces to

$$\lambda(\boldsymbol{x})\left[\left(\frac{4q\sigma_1^2}{\alpha}\right)^2\left\{\sum_i x_i^2\left(x_i-\frac{1}{2}\right)^2\right\}+\left(\frac{16C(q,2)\sigma_2^2}{1-\alpha}\right)^2\left\{\sum_{i<j}x_i^2x_j^2\right\}\right]\le 4q^2(\sigma_1+(q-1)\sigma_2)^2. \quad (3.7)$$

For $\alpha=\alpha_{0A}$, (3.7) simplifies to

$$\lambda(\boldsymbol{x})\left[4\sigma_1^2\left\{\sum_i x_i^2\left(x_i-\frac{1}{2}\right)^2\right\}+16\sigma_2^2\left\{\sum_{i<j}x_i^2x_j^2\right\}\right]\le 1. \quad (3.8)$$

Since (3.8) is same as (3.2), we conclude that ξ_A^* is A-optimum in the entire class for $q=2,3,4$.

Remark 3.4. As in Remark 3.3, it may be conjectured that the design ξ_A^* is A-optimum in the entire class for $q>4$.

Remark 3.5. As Remark 3.1, it is not difficult to show that instead of (2.3), one assumes the error function (3.6), the necessary and sufficient condition of the equivalence theorem remains same for any arbitrary q. Hence the derived designs are A-optimum in the entire class for both the error functions (2.3) and (3.6).

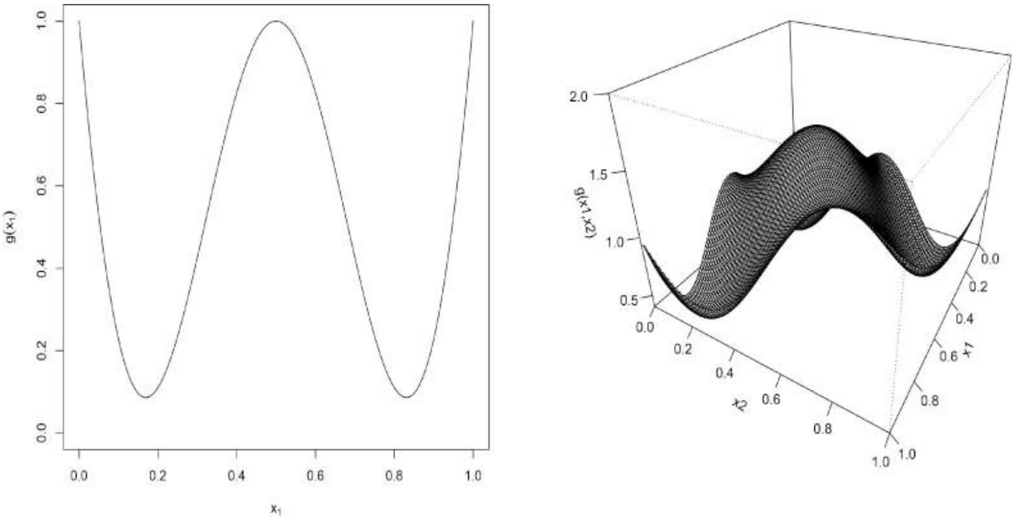
Remark 3.6. Unlike D-optimality, the design which is A-optimum for the estimation of the parameters of the model (2.5) is not A-optimum for the estimation of the parameters of the model (2.2). It has been seen in Sec. 2 that the optimum weights at the support points differ in the two situations (cf. equations (2.13) and (2.18)). It is to be noted that the condition mentioned in the equivalence theorem correspond to the parameters of the model (2.5). If one is interested in establishing the A-optimality for the estimation of the parameters of the model (2.2), the condition (3.7) is to be modified accordingly and it takes the form

$$\lambda(\boldsymbol{x})f'(\boldsymbol{x})\boldsymbol{M}^{-1}(\xi_A^{**})\boldsymbol{L}'\boldsymbol{L}\boldsymbol{M}^{-1}(\xi_A^{**})f(\boldsymbol{x})\le \text{tr.}(\boldsymbol{L}\boldsymbol{M}^{-1}(\xi_A^{**})\boldsymbol{L}')=\text{tr.}(\boldsymbol{M}^{-1}(\xi_A^{**})\boldsymbol{L}'\boldsymbol{L}) \quad (3.9)$$

which can equivalently be expressed as

$$\begin{aligned}\lambda(\boldsymbol{x})\Bigg[&\left(\frac{4q\sigma_1^2}{\alpha}\right)^2 f_1'(\boldsymbol{x})\left\{\frac{4q-7}{4}\boldsymbol{I}_q+\boldsymbol{J}_q\right\}f_1(\boldsymbol{x})+\left\{\frac{8q(q-1)\sigma_2^2}{1-\alpha}\right\}^2 f_2'(\boldsymbol{x})f_2(\boldsymbol{x})\\&-\frac{64q^2(q-1)}{\alpha(1-\alpha)}\sigma_1^2\sigma_2^2 f_1'(\boldsymbol{x})\boldsymbol{P}'f_2(\boldsymbol{x})\Bigg]\le \text{tr.}\left(\boldsymbol{M}^{-1}(\xi_A^{**})\boldsymbol{L}'\boldsymbol{L}\right),\end{aligned} \quad (3.10)$$

where $f(\mathbf{x})=(f_1'(\mathbf{x}),f_2'(\mathbf{x}))'$.



(a) Showing the graph of $g(x_1)$ for $x_1 \in [0,1]$

(b) Showing the graph of $g(x_1, x_2)$ for $x_1, x_2 \in [0,1]$

Figure 2. (a) Showing the graph of $g(x_1)$ for $x_1 \in [0,1]$ (b) Showing the graph of $g(x_1, x_2)$ for $x_1, x_2 \in [0,1]$.

In Sec. 2, we have observed that ξ_A^{**} is A-optimum in $\mathcal{D}_0$ when $\alpha = \alpha_{0A}^*$, given by (2.18). To check optimality of the derived design in the entire class we have to verify it through equivalence theorem which, after a brief algebra, reduces to

$$4\lambda(\boldsymbol{x})\Bigg[\sigma_1^2\sum_{i=1}^{q}x_i^2\left(x_i-\frac{1}{2}\right)^2+\frac{8\sigma_1^2}{4q-3}\sum_{i<j}^{q}x_ix_j\left(x_i-\frac{1}{2}\right)\left(x_j-\frac{1}{2}\right)+4\sigma_2^2\sum_{i<j}x_i^2x_j^2 \\ -\frac{8\sigma_1\sigma_2}{(4q-3)^{1/2}}\sum_{i=1}^{q}x_i^2\left(x_i-\frac{1}{2}\right)(1-x_i)\Bigg]\leq 1 \tag{3.11}$$

We denote left hand side of (3.11) by $g(x_1, x_2, ..., x_{q-1})$.

Here again, for $q = 2$, equality holds in (3.11) at the support points of ξ_A^{**}, which can be seen from Figure 2a. Moreover, at all other points besides the support points, value of the left hand side of (3.11) is strictly less than 1 which guarantees the optimality of ξ_A^{**} in the entire class. However for $q = 3$, we have seen from Figure 2b, value of the left hand side of (3.11) is greater than 1 at some other points besides the support points of ξ_A^{**}. Hence for $q = 3$, ξ_A^{**} is not A-optimum in the entire class which is also true for the homoscedastic case (See Galil and Kiefer (1977)). Also we have numerically checked that for $q = 4$, design ξ_A^{**} is not A-optimum in the entire class (see Appendix Table A1).

4. Robustness

In this section, we shall study the robustness of the usual D- and A-optimum designs under homoscedastic error variance, when the true error variance is heteroscedastic. For the D-optimality criterion, we have observed that the weights at the support points are independent of the form of error function so that the D-optimum designs are same for both the homoscedastic and heteroscedastic error variances. However for the A-optimality criterion the optimum weights are functions of the error function. To study the robustness of the optimum design under homoscedasticity when the error variance is really heteroscedastic one may use, as a measure of robustness, the ratio of the criterion functions of the two designs namely

$$R_{\xi_0|\xi_A^*} = \frac{\text{tr.}(\boldsymbol{M}^{-1}(\xi_A^*))}{\text{tr.}(\boldsymbol{M}^{-1}(\xi_0))}. \tag{4.1}$$

which gives an indication of the performance of the design ξ_0, the A-optimum design under homoscedasticity, against the design ξ_A^* for variation in error function.

Under homoscedasticity (i.e. $\sigma_1^2 = \sigma_2^2 = 1$), $\text{tr.}(\mathbf{M}^{-1}(\xi)) = \frac{4q^2}{\alpha} + \frac{4q^2(q-1)^2}{1-\alpha}$ and $\text{tr.}(\mathbf{M}^{-1}(\xi))$ is minimum when $\alpha = \frac{1}{q}$. Hence $\text{tr.}(\mathbf{M}^{-1}(\xi_0)) = 4q^3(\sigma_1^2 + (q-1)\sigma_2^2)$. Now from (4.1),

$$R_{\xi_0|\xi_A^*} = \frac{4q^2(\sigma_1 + (q-1)\sigma_2)^2}{4q^3(\sigma_1^2 + (q-1)\sigma_2^2)} = \frac{(\sigma_1 + (q-1)\sigma_2)^2}{q(\sigma_1^2 + (q-1)\sigma_2^2)}. \tag{4.2}$$

However we have seen in the previous section that A-optimum designs under heteroscedastic error variance do not exist for all q. Hence we compute (4.2) for different value of q and represent these in Table 1.

It is clearly seen that $R_{\xi_0|\xi_A^*} < 1$ for all q as expected but very close to 1.

Remark 4.1. Sensitivity of the criterion function (4.1) can be increased by slightly modifying the error function (error) to

$$V(y|\boldsymbol{x}) = \frac{1}{\lambda(\boldsymbol{x})} = \exp\left(\delta(\boldsymbol{x} - \boldsymbol{x}_0)'(\boldsymbol{x} - \boldsymbol{x}_0)\right) \tag{4.3}$$

where δ (> 0) is a constant. Then the robustness of ξ_0 can be examined for variation in δ. From Figure 3, it is clear that the value of the robustness criterion (4.1) is decreasing in δ for $q = 2$. Hence though the A- optimum designs under the homoscedastic set-up is robust with the error function assumed but is very sensitive with changes in the value of δ. Since it involves lot of numerical computations, we have not pursued it further. One may also be interested in

Table 1. The values of $R_{\xi_0|\xi_A^*}$ for different values of $q(\neq 3)$.

q	2	4	5	6	7	8	9	10
$R_{\xi_0|\xi_A^*}$	0.9848	0.9697	0.9689	0.9695	0.9706	0.9720	0.9733	0.9746

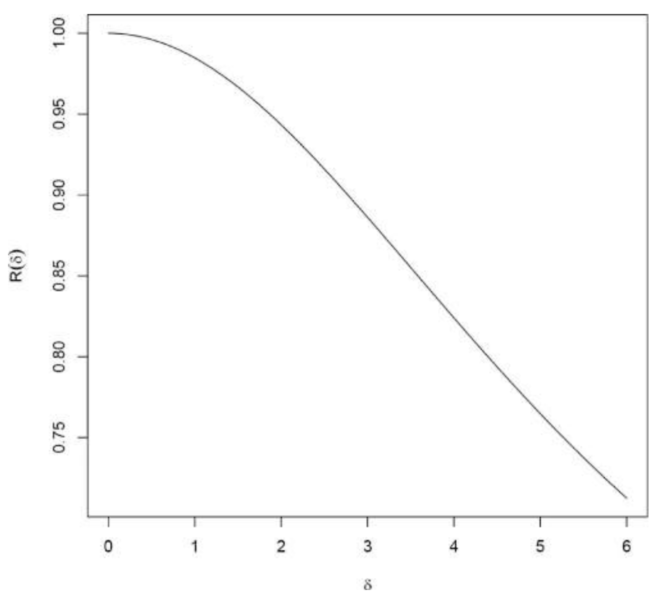


Figure 3. Showing the graph of $R_{\xi_0|\xi_A^*} = R(\delta)$ for $q = 2$.

estimating δ in addition to the estimation of the regression coefficients optimally (cf. Atkinson and Cook(1995)). It is intended to communicate it in a subsequent paper.

5. Concluding remarks

Optimum designs have been derived for the first and second degree models in a mixture experiment with D- and A-optimality criteria when the error variance is heteroscedastic. Since it is difficult to find optimum design for arbitrary form of error variance, we have assumed a specific form namely the exponential form of the error variance. It is seen that the support points are same as that in the homoscedastic case. Though the weights at the support points are same as in the homoscedastic case for D-optimality criterion but they differ for the A-optimality criterion. It may also be observed that the D-optimal design is independent of the form of the error function. It is seen that the weights are sensitive to changes in the multiplying factor in the exponent of the error function in terms of robustness of the design relative to the homoscedastic case. The problem for other choices of error functions remains open. One can pursue by extensive numerical computations to find optimum design in such cases.

Acknowledgement

The authors acknowledge the anonymous referees for their suggestions and valuable comments which improved the presentation of the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Atkinson, A. C., and R. D. Cook. 1995. D-optimum designs for heteroscedastic linear models. *Journal of the American Statistical Association* 90 (429):204–12. doi: 10.2307/2291144.

Becker, N. G. 1968. Models for the response of a mixture. *Journal of the Royal Statistical Society, Series B* 30 (2): 349–58. doi: 10.1111/j.2517-6161.1968.tb00735.x.

Becker, N. 1978. Models and designs for experiments with mixtures. *Australian Journal of Statistics* 20 (3):195–208. doi: 10.1111/j.1467-842X.1978.tb01102.x.

Cornell, J. A. 1977. Weighted versus unweighted estimates using Scheffes mixture model for symmetrical error variances patterns. *Technometrics* 19:237–47.

Cornell, J. A. 2002. *Experiments with mixtures*. 3rd ed. New York: Wiley.

Draper, N. R., and F. Pukelsheim. 1999. Kiefer ordering of simplex designs for first and second degree mixture models. *Journal of Statistical Planning and Inference* 79 (2):325–48. doi: 10.1016/S0378-3758(98)00263-8.

Dette, H., and W. K. Wong. 1996. Bayesian optimal designs for models with partially specified heteroscedastic structure. *Annals of Statistics* 24:2108–27.

Fedorov, V. V. 1971. Design of experiments for linear optimality criteria. *Theory of Probability & Its Applications* 16 (1):189–95. doi: 10.1137/1116019.

Fedorov, V. V. 1972. *Theory of optimal experiments*. New york: Academic.

Galil, Z., and J. Kiefer. 1977. Comparison of simplex designs for quadratic mixture models. *Technometrics* 19 (4): 445–53. doi: 10.1080/00401706.1977.10489584.

Kiefer, J. 1961. Optimum designs in regression problems II. *The Annals of Mathematical Statistics* 32 (1):298–325. doi: 10.1214/aoms/1177705160.

Kiefer, J. 1974. General equivalence theory for optimum designs (approximate theory). *Annals of Statistics* 22:849–79.

Kiefer, J., and J. Wolfowitz. 1960. The equivalence of two extremum problems. *Canadian Journal of Mathematics* 12:363–6. doi: 10.4153/CJM-1960-030-4.

Pal, M., and N. K. Mandal. 2006. Optimum designs for optimum mixtures. *Statistics & Probability Letters* 76 (13): 1369–79. doi: 10.1016/j.spl.2006.02.007.

Pal, M., and N. K. Mandal. 2007. Optimum mixture design via equivalence theorem. *Journal of Combinatorics, Information and System Sciences* 32 (2):107–26.
Pukelsheim, F. 1993. *Optimal design of experiments*. New York: Wiley.
Rodrı́guez, C., and I. Ortiz. 2005. D-optimum designs in multi-factor models with heteroscedastic errors. *Journal of Statistical Planning and Inference* 128 (2):623–31. doi: 10.1016/j.jspi.2003.12.013.
Schefé, H. 1958. Experiments with mixtures. *Journal of the Royal Statistical Society, Series B* 20:344–60.
Schefé, H. 1963. Simplex - centroid design for experiments with mixtures. *Journal of the Royal Statistical Society, Series B* 25:235–63.
Silvey, S. D. 1980. *Optimal design*. London: Chapman and Hall.
Sinha, B. K., N. K. Mandal, M. Pal, and P. Das. 2014. *Optimal mixture experiments*, Vol. 1028. Springer, India: Springer.
Yan, F., C. Zhang, and H. Peng. 2017. Optimal designs for additive mixture model with heteroscedastic errors. *Communications in Statistics - Theory and Methods* 46 (13):6401–11. doi: 10.1080/03610926.2015.1129414.

Appendix

Table A1. Table showing the values of $\frac{d(\mathbf{x},\xi^*)}{10}$ and $g(x_1,x_2,x_3)$ for $q=4$.

(a)

x_1	x_2	x_3	x_4	$\frac{d(\mathbf{x},\xi^*)}{10}$
0	0	0	1	1
0	0	0.25	0.75	0.7237469
0	0	0.5	0.5	1
0	0	0.75	0.25	0.7237469
0	0	1	0	1
0	0.25	0	0.75	0.7237469
0	0.25	0.25	0.5	0.6957787
0	0.25	0.5	0.25	0.6957787
0	0.25	0.75	0	0.7237469
0	0.5	0	0.5	1
0	0.5	0.25	0.25	0.6957787
0	0.5	0.5	0	1
0	0.75	0	0.25	0.7237469
0	0.75	0.25	0	0.7237469
0	1	0	0	1
0.25	0	0	0.75	0.7237469
0.25	0	0.25	0.5	0.6957787
0.25	0	0.5	0.25	0.6957787
0.25	0	0.75	0	0.7237469
0.25	0.25	0	0.5	0.6957787
0.25	0.25	0.25	0.25	0.6138220
0.25	0.25	0.5	0	0.6957787
0.25	0.5	0	0.25	0.6957787
0.25	0.5	0.25	0	0.6957787
0.25	0.75	0	0	0.7237469
0.5	0	0	0.5	1
0.5	0	0.25	0.25	0.6957787
0.5	0	0.5	0	1
0.5	0.25	0	0.25	0.6957787
0.5	0.25	0.25	0	0.6957787
0.5	0.5	0	0	1
0.75	0	0	0.25	0.7237469
0.75	0	0.25	0	0.7237469
0.75	0.25	0	0	0.7237469
1	0	0	0	1

(b)				
x_1	x_2	x_3	x_4	$g(x_1, x_2, x_3)$
0	0	0	1	1
0	0	0.25	0.75	0.4460669
0	0	0.5	0.5	1
0	0	0.75	0.25	0.4460669
0	0	1	0	1
0	0.25	0	0.75	0.4460669
0	0.25	0.25	0.5	1.0163991
0	0.25	0.5	0.25	1.0163991
0	0.25	0.75	0	0.4460669
0	0.5	0	0.5	1
0	0.5	0.25	0.25	1.0163991
0	0.5	0.5	0	1
0	0.75	0	0.25	0.4460669
0	0.75	0.25	0	0.4460669
0	1	0	0	1
0.25	0	0	0.75	0.4460669
0.25	0	0.25	0.50	1.0163991
0.25	0	0.50	0.25	1.0163991
0.25	0	0.75	0	0.4460669
0.25	0.25	0	0.5	1.0163991
0.25	0.25	0.25	0.25	1.4218662
0.25	0.25	0.5	0	1.0163991
0.25	0.5	0	0.25	1.0163991
0.25	0.5	0.25	0	1.0163991
0.25	0.75	0	0	0.4460669
0.5	0	0	0.5	1
0.5	0	0.25	0.25	1.0163991
0.5	0	0.5	0	1
0.5	0.25	0	0.25	1.0163991
0.5	0.25	0.25	0	1.0163991
0.5	0.5	0	0	1
0.75	0	0	0.25	0.4460669
0.75	0	0.25	0	0.4460669
0.75	0.25	0	0	0.4460669
1	0	0	0	1