ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIII, 4 (1973)

R. ZIELIŃSKI (Warszawa)

A SIMPLEX DESIGN FOR GRADIENT ESTIMATION IN QUADRATIC REGRESSION

1. The note deals with estimation of parameters $a_1, a_2, \ldots, a_k$ (the gradient at the origin) in quadratic regression

$$\mathrm{E}Y(x_1, x_2, \ldots, x_k) = \frac{\varrho}{\sqrt{k}} a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} x_i x_j,$$

ϱ being a given constant. It is well known that, for the estimation of $b_0, b_1, \ldots, b_k$ in the case of linear regression

$$\mathrm{E}Z(x_1, x_2, \ldots, x_k) = \frac{\varrho}{\sqrt{k}} b_0 + \sum_{i=1}^{k} b_i x_i,$$

the minimal design requires $k+1$ experiments; it is a *D-optimal, orthogonal* and *rotatable design* if the points $X_0, X_1, X_2, \ldots, X_k$ ($X_j = (x_{1j}, x_{2j}, \ldots, x_{kj})$) at which experiments are performed form a regular simplex in $\boldsymbol{E}^k$ (see, e.g., [1] and [2]). In the case of quadratic regression any such design yields biased estimators. We show that a random rotation of the simplex leads to unbiased estimators of $a_1, a_2, \ldots, a_k$.

2. Consider the family of independent random variables $Y(X)$, $X = (x_1, x_2, \ldots, x_k) \in \boldsymbol{E}^k$ with the mean

$$F(X) = \frac{\varrho}{\sqrt{k}} a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} x_i x_j$$

and a common variance σ^2. The points X are referred to as possible levels of feasible experiments and the random variable $Y(X)$ as the outcome of the experiment performed at X. A simplex design is a set of points $X_0, X_1, \ldots, X_k$ which are vertices of a regular simplex; let the simplex be centered at the origin and let $(\sum_{i=1}^{k} x_{ij}^2)^{1/2} = \varrho$ (the radius of the simplex). Denote by $D = (d_{ij})$ the design matrix: $d_{0j} = \varrho/\sqrt{k}$, $d_{ij} = x_{ij}$ for $i = 1, 2, \ldots, k$ and $j = 0, 1, 2, \ldots, k$.

In the case of linear regression

$$\mathrm{E}\,Z(x_1, x_2, \ldots, x_k) = \frac{\varrho}{\sqrt{k}} b_0 + \sum_{i=1}^{k} b_i x_i$$

the least squares estimator of $b = (b_0, b_1, \ldots, b_k)$ is $\hat{b} = (DD^T)^{-1} DY_D$, where $Y_D = \big(Y(X_0), Y(X_1), \ldots, Y(X_k)\big)$. The estimator $(DD^T)^{-1}DY_D$ when applied to $a = (a_0, a_1, \ldots, a_k)$ is a biased one.

Consider the following procedure. Let D^0 be the design matrix for a given simplex and X^0 a given vertex of the simplex. Let Ω be the sphere with radius ϱ centered at the origin. Sample a point $\omega \in \Omega$ according to the uniform distribution on Ω. Transform the simplex into a new one by the rotation which transfers X^0 into ω. Denote the new design matrix by $D(\omega)$. Let $\hat{a} = (\hat{a}_0, \hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k)$ be the estimator of $a = (a_0, a_1, a_2, \ldots, a_k)$ defined as follows:

$$\hat{a} = [D(\omega) D^T(\omega)]^{-1} D(\omega)\, Y_{D(\omega)}.$$

Then $\mathrm{E}\,\hat{a}_i = a_i$ for $i = 1, 2, \ldots, k$.

To prove the assertion note that

$$\begin{aligned}\mathrm{E}\,\hat{a} &= \mathrm{E}\big[\mathrm{E}\{(D(\omega)D^T(\omega))^{-1} D(\omega)\, Y_{D(\omega)}|\omega\}\big] \\ &= \mathrm{E}\big[(D(\omega)D^T(\omega))^{-1} D(\omega)\,\mathrm{E}\{Y_{D(\omega)}|\omega\}\big],\end{aligned}$$

where $\mathrm{E}\{Y_{D(\omega)}|\omega\}$ is a vector with i-th component equal to $F\big(X_i(\omega)\big)$, $X_i(\omega) = \big(x_{1i}(\omega), x_{2i}(\omega), \ldots, x_{ki}(\omega)\big)$ being the i-th experimental point after the above-mentioned rotation. The points $X_i(\omega)$, $i = 0, 1, 2, .., k$, form a simplex for any $\omega \in \Omega$. Hence $\big(D(\omega)D^T(\omega)\big)^{-1}$ is a diagonal matrix with diagonal elements equal to $k\varrho^{-2}/(k+1)$. For $\hat{a}_i$ we have

$$\mathrm{E}\,\hat{a}_i = \frac{k}{k+1} \varrho^{-2} \sum_{j=0}^{k} \mathrm{E}\big[x_{ij}(\omega) F\big(X_j(\omega)\big)\big].$$

In the paper [3] the following Theorem 2 is proved:

Let

$$\hat{g}_i = \frac{k}{k+1} \varrho^{-2} \sum_{j=0}^{k} x_{ij}(\omega) F\big(X_j(\omega)\big)$$

be an estimator of the gradient $g = (g_1, g_2, \ldots, g_k)$ of a function F. Then

$$\mathrm{E}\hat{g}_i = g_i + \frac{\varrho^2}{6(k+2)} \sum_{j=1}^{k} (f_{ijj} + f_{jij} + f_{jji}),$$

where f_{ijj}, f_{jij} and f_{jji} are the third partial derivates of the function F.

In our problem all third partial derivates of F are equal to zero, and hence the assertion follows.

References

[1] G. E. P. Box, *Multifactor designs of first order*, Biometrika 39 (1952), p. 49-57.

[2] T. I. Golikova and N. G. Mikeshina (Т. И. Голикова и Н. Г. Микешина), *Свойства D-оптимальных планов и методы их построения* in *Новые идеи в планировании эксперимента*, Москва 1969.

[3] R. Zieliński, *A randomized finite-differential estimator of the gradient*, Preprint no. 41, Institute of Mathematics, Polish Academy of Sciences, June 1972 (to appear in Algorytmy).

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA

Received on 13. 10. 1972

R. ZIELIŃSKI (Warszawa)

PLANY SYMPLEKSOWE DLA SZACOWANIA GRADIENTU REGRESJI DRUGIEGO STOPNIA

STRESZCZENIE

W pracy rozważa się zagadnienie szacowania parametrów $a_1, a_2, \ldots, a_k$ regresji drugiego stopnia

$$\mathrm{E}Y = \frac{\varrho}{\sqrt{k}} a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} x_i x_j,$$

gdzie ϱ jest pewną stałą. Jak wiadomo, dla oszacowania współczynników $b_0, b_1, \ldots, b_k$ regresji liniowej

$$\mathrm{E}Z = \frac{\varrho}{\sqrt{k}} b_0 + \sum_{i=1}^{k} b_i x_i,$$

minimalne plany wymagają wykonania $k+1$ eksperymentów. Gdy punkty $X_0, X_1, \ldots$ $\ldots, X_k$ są wierzchołkami regularnego sympleksu w E^k, plany takie są D-optymalne, ortogonalne i mają symetrię obrotową. W przypadku regresji kwadratowej, takie plany prowadzą do estymatorów obciążonych. W pracy pokazano, że przez losowy obrót planu sympleksowego uzyskuje się nieobciążone oszacowanie współczynników $a_1, a_2, \ldots, a_k$.