Júlia Volaufová; Peter Volauf

Estimation of a quadratic function of the parameter of the mean in a linear model

Aplikace matematiky, Vol. 34 (1989), No. 2, 155–160

Persistent URL: http://dml.cz/dmlcz/104343

34 (1989) APLIKACE MATEMATIKY No. 2, 155—160

ESTIMATION OF A QUADRATIC FUNCTION OF THE PARAMETER OF THE MEAN IN A LINEAR MODEL

JÚLIA VOLAUFOVÁ, PETER VOLAUF

(Received November 16, 1987)

Summary. The paper deals with an optimal estimation of the quadratic function $\boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta}$, where $\boldsymbol{\beta} \in \mathscr{R}^k$, $\mathbf{D}$ is a known $k \times k$ matrix, in the model $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$. The distribution of $\mathbf{Y}$ is assumed to be symmetric and to have a finite fourth moment. An explicit form of the best unbiased estimator is given for a special case of the matrix $\mathbf{X}$.

Keywords: Linear model, best unbiased quadratic estimator.

AMS Classification: 62F10, 62J05.

INTRODUCTION

Consider the following situation as an illustrating example. Let $y_1, y_2, \ldots, y^n$ denote n measurements of a length β of a side of a square $\mathbf{Q}$. The question is how to estimate the area $P(=\beta^2)$ of the square $\mathbf{Q}$ under the assumption that Y_i, $i = 1, 2, \ldots, n$ are i.i.d. The variance σ^2 of Y_i may or not be known. The answer to this simple question is not obvious. When the normality condition of Y_i is met and σ^2 is unknown, one can investigate e.g.

$$\hat{P}_1 = (\bar{Y})^2 - \frac{1}{n(n-1)} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

and

$$\hat{P}_2 = \bar{Y}^2 - \frac{1}{n(n+1)} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \, .$$

While $\hat{P}_1$ is an unbiased estimator, $\hat{P}_2$ is not, but has a smaller mean square error. Of course an "optimal" estimator depends on the criterion chosen. Our results concern the class of unbiased estimators. Another case is briefly mentioned in the concluding remark.

The above problem can be investigated as an estimation of the quadratic function of the mean parameter in a linear model. Under the normality assumption when σ^2 is a known scalar the best unbiased estimator of the polynomial function of the mean parameter is derived in [2].

In this paper we try to analyse some special situations in a linear model (with respect to the matrix $\mathbf{X}$) and to derive optimal estimators under rather general assumptions concerning the distribution of Y_i.

PRELIMINARIES

The linear model is commonly known in the form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ is an n-dimensional vector of observations, $\mathbf{X}$ is an $\boldsymbol{n} \times \boldsymbol{k}$ matrix with rank $r(\mathbf{X}) = k$ whose elements are not random and known and $\boldsymbol{\beta}$ is the k-dimensional unknown vector of parameters. The vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ is the $\boldsymbol{n}$-dimensional random vector of errors. We suppose that its components are i.i.d. and

$$\mathsf{E}(\varepsilon_i) = O\,, \quad \mathsf{E}(\varepsilon_i^2) = \sigma^2\,, \quad \mathsf{E}(\varepsilon_i^3) = O \quad \text{and} \quad \mathsf{E}(\varepsilon_i^4) = \sigma^4\gamma + 3\sigma^4 < \infty\,;$$

it means that just the symmetric distributions with finite fourth moments are considered.

We are interested in an estimate of the quadratic function

$$f(\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta} = \operatorname{tr} \mathbf{D}\boldsymbol{\beta}\boldsymbol{\beta}' \tag{2}$$

of the unknown parameter $\boldsymbol{\beta}$, where $\mathbf{D}$ is a known symmetric $k \times k$ matrix.

BEST UNBIASED QUADRATIC ESTIMATOR

a) σ^2 unknown

In this case we restrict ourselves to quadratic estimators of $f(\boldsymbol{\beta})$ in the form $\mathbf{Y}'\mathbf{A}\mathbf{Y}$, where $\mathbf{A}$ is a symmetric $n \times n$ matrix. The condition of unbiasedness means

$$\mathsf{E}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \sigma^2 \operatorname{tr} \mathbf{A} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{A}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta} \quad \forall \boldsymbol{\beta} \in \mathscr{R}^k \quad \forall \sigma \in \mathscr{R}^1$$

which yields

$$\begin{aligned} \operatorname{tr} \mathbf{A} &= \mathbf{O}\,, \\ \mathbf{X}'\mathbf{A}\mathbf{X} &= \mathbf{D}\,. \end{aligned} \tag{3}$$

Let us denote

$$\mathscr{A}_0^f = \{\mathbf{A} : \mathbf{A} = \mathbf{A}',\ \mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{D},\ \operatorname{tr} \mathbf{A} = \mathbf{O}\}\,.$$

Theorem 1. *The γ-locally best unbiased quadratic estimator of $f(\boldsymbol{\beta})$ of the form* (2) *in the model* (1) *is given by*

$$\hat{f}(\boldsymbol{\beta}) = \mathbf{Y}'\mathbf{A}_0\mathbf{Y}$$

where $\mathbf{A}_0 \in \mathscr{A}_0^f$ and $\mathbf{A}_0$ is the solution of the matrix equation

$$(4)\qquad 2\sigma^2 \mathbf{A} + \sigma^2\gamma \operatorname{Diag} \mathbf{A} + 2(\mathbf{A}\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}' + \mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}) - \sigma^2\gamma \mathbf{P}_{\mathbf{X}} \cdot \operatorname{Diag} \mathbf{A} \cdot \mathbf{P}_{\mathbf{X}} +$$

$$+ \frac{\sigma^2\gamma}{n-k} \operatorname{tr}(\operatorname{Diag} \mathbf{A} \cdot \mathbf{P}_{\mathbf{X}})\, \mathbf{M}_{\mathbf{X}} =$$

$$= 2\sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{D}(\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{X}' + 2(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{D}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}' +$$

$$+ \mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{D}(\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{X}') - \frac{2\sigma^2}{n-k} \operatorname{tr}(\mathbf{D}(\mathbf{X}'\mathbf{X})^{-1})\, \mathbf{M}_{\mathbf{X}},$$

where Diag $\mathbf{A}$ *is the diagonal matrix formed by the diagonal elements of the matrix* $\mathbf{A}$; $\mathbf{P}_{\mathbf{X}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{X}'$ *is the orthogonal projection matrix into the space spanned by the columns of the matrix* $\mathbf{X}$, *and* $\mathbf{M}_{\mathbf{X}} = \mathbf{I} - \mathbf{P}_{\mathbf{X}}$.

Proof. The variance of the estimator $\mathbf{Y}'\mathbf{A}\mathbf{Y}$, $\mathbf{A} \in \mathscr{A}_0^f$ is given by $\operatorname{var}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = 4\sigma^2\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}^2\mathbf{X}\boldsymbol{\beta} + 2\sigma^4 \operatorname{tr} \mathbf{A}^2 + \sigma^4\gamma \sum_{i=1}^{n} a_{ii}^2$, where a_{ii} is the i-th diagonal element of the matrix $\mathbf{A}$. Finding a γ-locally best estimator in the class $\{\mathbf{Y}'\mathbf{A}\mathbf{Y}\colon \mathbf{A} \in \mathscr{A}_0^f\}$ is equivalent to the minimization of the variance $\operatorname{var}(\mathbf{Y}'\mathbf{A}\mathbf{Y})$ under the constraints (3). It can be carried out by standard techniques of matrix derivatives, which yields the equation (4).

Remark 1. The matrix $\mathbf{A}_0$ in Theorem 1 is not given in an explicit form. Let us return to the example considered at the beginning. In this case the model has the form

$$(5)\qquad \mathbf{Y} = \mathbf{1}\beta + \varepsilon$$

where $\mathbf{1} = (1, 1, \ldots, 1)'$, β is one dimensional and $f(\beta) = \beta^2$ is the area of the square. Now the matrix equation (4) has a much simpler form and the solution of (4) is the matrix

$$(6)\qquad \mathbf{A}_0 = \frac{1}{n(n-1)}(\mathbf{1}\,\mathbf{1}' - \mathbf{I})\,.$$

The estimator

$$\mathbf{Y}'\mathbf{A}_0\mathbf{Y} = (\bar{Y})^2 - \frac{1}{n(n-1)} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

is a γ-uniformly best estimator of $P = \beta^2$ in the class $\{\mathbf{Y}'\mathbf{A}\mathbf{Y}\colon \mathbf{A} \in \mathscr{A}_0^f\}$ and its variance satisfies

$$\operatorname{var}(\mathbf{Y}'\mathbf{A}_0\mathbf{Y}) = \frac{2\sigma^4}{n(n-1)} + \frac{4\sigma^2\beta^2}{n}\,.$$

We point out that this estimator coincides with the best unbiased estimator of β^2 under the normality assumption and its variance does not depend on the parameter γ.

b) σ^2 known

As we have mentioned at the beginning of this paper, in the case when $\mathbf{Y}$ is normally distributed and the parameter σ^2 is known the following results are stated in [2]. The best unbiased estimator for $f(\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta}$ is given in the form

(7) $$\hat{f}(\boldsymbol{\beta}) = \hat{\boldsymbol{\beta}}'\mathbf{D}\hat{\boldsymbol{\beta}} - \sigma^2 \operatorname{tr} \mathbf{D}(\mathbf{X}'\mathbf{X})^{-1}, \quad \text{where} \quad \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

We note that the statistic $\hat{f}(\boldsymbol{\beta})$ includes an absolute term, i.e. it is of the form $\mathbf{Y}'\mathbf{A}\mathbf{Y} + c$, $c \in \mathscr{R}^1$. Under general conditions on the distribution of $\mathbf{Y}$ we shall concentrate our attention on the class of estimators

$$\mathscr{G}_1^f = \{\mathbf{Y}'\mathbf{A}\mathbf{Y} + \mathbf{b}'\mathbf{Y} + c \colon \mathbf{A} = \mathbf{A}',\ \mathbf{b} \in \mathscr{R}^n,\ c \in \mathscr{R}^1\}.$$

The following theorem is stated for the case $\mathbf{X} = (1, 1, \ldots, 1)'$.

Theorem 2. *The γ-locally best unbiased estimator for the function $f(\beta) = \beta^2$ in the model (5) has the form*

$$\hat{f}(\beta) = (\overline{Y})^2 - \frac{\gamma}{n(2n + (n-1)\gamma)} \sum_{i=1}^{n} (Y_i - \overline{Y})^2 - \frac{2\sigma^2}{2n + (n-1)\gamma},$$

where γ is the excess of the distribution and $\overline{Y}$ is the sample mean. The variance of $\hat{f}(\beta)$ is

$$\operatorname{var}(\hat{f}(\beta)) = 2\sigma^4 \frac{2\gamma(2n-1) + 4n + \gamma^2(n-1)}{n(2n + (n-1)\gamma)^2} + \frac{4\sigma^2\beta^2}{n}.$$

Proof. The condition of unbiasedness for the estimator $\boldsymbol{T} \in \mathscr{G}_1^f$ means

$$\mathsf{E}(\mathbf{Y}'\mathbf{A}\mathbf{Y} + \mathbf{b}'\mathbf{Y} + c) = \beta^2 \mathbf{1}'\mathbf{A}\mathbf{1} + \sigma^2 \operatorname{tr} \mathbf{A} + \beta \mathbf{b}'\mathbf{1} + c = \beta^2$$
$$\forall \beta \in \mathscr{R}^k \quad \forall \sigma^2 \in \mathscr{R}^+,$$

which yields

$$\mathbf{1}'\mathbf{A}\mathbf{1} = 1,$$
$$\mathbf{b}'\mathbf{1} = \mathbf{0},$$
$$c = -\sigma^2 \operatorname{tr} \mathbf{A}.$$

The variance of $\boldsymbol{T}$ is given by

(8) $$\operatorname{var}(\mathbf{Y}'\mathbf{A}\mathbf{Y} + \mathbf{b}'\mathbf{Y} + c) = 4\sigma^2\beta^2 \mathbf{1}'\mathbf{A}^2\mathbf{1} + 2\sigma^4 \operatorname{tr} \mathbf{A}^2 + \sigma^4\gamma \sum_{i=1}^{n} a_{ii}^2 + 4\sigma^2 \mathbf{b}'\mathbf{A}\mathbf{1}\beta + \sigma^2 \mathbf{b}'\mathbf{b}.$$

(The way how the calculate the variance of a quadratic form of $\mathbf{Y}$ can be found in [1].) Making use of the method of Lagrangian multipliers for minimizing the variance under the conditions of unbiasedness we get the following system of equations

$$(-2\beta^2)(\mathbf{A}\mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}'\mathbf{A}) + 2\sigma^2\mathbf{A} + \sigma^2\gamma \operatorname{Diag} \mathbf{A} - \frac{1}{n^2}(-4\beta^2 n + 2\sigma^2 + \sigma^2\gamma \operatorname{tr} \mathbf{A})\mathbf{1}\mathbf{1}' = \mathbf{0},$$

$$b = \mathbf{0}\,,$$
$$c = -\sigma^2 \operatorname{tr} \mathbf{A}\,.$$

The solution $\mathbf{A}_0$, c_0 can be derived by some tedious technique, which yields

$$\text{(9)} \qquad \mathbf{A}_0 = \frac{2+\gamma}{(n(2n+(n-l)\gamma)}\,\mathbf{11}' - \frac{\gamma}{n(2n+(n-l)\gamma)}\,\mathbf{I}\,,$$
$$c_0 = -\frac{2\sigma^2}{2n+(n-l)\gamma}\,.$$

The assertion of the previous Theorem 2 combined with the results stated in [2] enables us to generalize the problem for the simple variance-components model. Let the model be

$$\text{(10)} \qquad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon\,, \quad \mathsf{E}(\varepsilon) = \mathbf{0}\,, \quad \mathsf{E}(\varepsilon\varepsilon') = \vartheta_1\mathbf{V}_1 + \vartheta_2\mathbf{V}_2 = \mathbf{V}(\vartheta)\,,$$

where $\mathbf{Y}$ is normally distributed, $\mathbf{V}_i$, $i = 1, 2$ are symmetric matrices such that $\mathbf{V}(\vartheta)$ is p.d. $\forall\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)' \in \mathscr{R}^2$. Let $\boldsymbol{\beta}$, $\boldsymbol{\vartheta}$ be unknown vector parameters but let the ratio $\varrho = \vartheta_2/\vartheta_1$ be known. Denote $\mathbf{V}_\varrho = \mathbf{V}_1 + \varrho\mathbf{V}_2$. The model (10) can be expressed in a simpler form

$$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \vartheta_1 V_\varrho)\,.$$

The function of interest is $\hat{f}(\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta}$. If we denote

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}_\varrho^{-1}\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{V}_\varrho^{-1}\mathbf{Y} \quad \text{and} \quad \hat{\vartheta}_1 = \frac{1}{n-k}\,\mathbf{Y}'(\mathbf{V}_\varrho^{-1} - \mathbf{V}_\varrho^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_\varrho^{-1}\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{V}_\varrho^{-1})\,\mathbf{Y}$$

then the unbiased estimator of $f(\boldsymbol{\beta})$ is given by

$$\hat{f}(\boldsymbol{\beta}) = \hat{\boldsymbol{\beta}}'\mathbf{D}\hat{\boldsymbol{\beta}} - \hat{\vartheta}_1 \operatorname{tr}(\mathbf{D}(\mathbf{X}'\mathbf{V}_\varrho^{-1}\mathbf{X})^{-1})\,.$$

CONCLUDING REMARK

Instead of the criterion of minimal variance in the class of unbiased quadratic estimators one could choose the minimal mean square error (MSE) criterion or the optimal estimator. Consider the model (5) and let $\mathbf{Y} \sim \mathbf{N}_n(\mathbf{1}\beta, \sigma^2\mathbf{I})$, β, σ^2 both unknown. The MSE of the quadratic estimator from the class $\{\mathbf{Y}'\mathbf{A}\mathbf{Y}\colon \mathbf{A} = \mathbf{A}'\}$ is

$$\operatorname{MSE}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \mathsf{E}(\mathbf{Y}'\mathbf{A}\mathbf{Y} - \beta^2)^2\,, \quad \text{which is equivalent to}$$
$$\operatorname{MSE}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = (\operatorname{vec}\mathbf{A})'\,\mathsf{E}(\mathbf{Y}^{2\otimes}\mathbf{Y}'^{2\otimes})\operatorname{vec}\mathbf{A} - 2\beta^2\mathsf{E}(\mathbf{Y}'^{2\otimes})\operatorname{vec}\mathbf{A} + \beta^4\,,$$

where vec $\mathbf{A}$ is the vector formed by the columns of the matrix $\mathbf{A}$ written one below the other and $\mathbf{Y}^{2\otimes} = \mathbf{Y} \otimes \mathbf{Y}$ where the symbol $\otimes$ denotes the Kronecker multiplication of matrices or vectors. Minimizing the MSE $(\mathbf{Y}'\mathbf{A}\mathbf{Y})$ we conclude that the following equation holds:

$$\mathsf{E}(\mathbf{Y}^{2\otimes}\mathbf{Y}'^{2\otimes}) \operatorname{vec} \mathbf{A} = \beta^2 \mathsf{E}(\mathbf{Y}^{2\otimes}), \quad \text{and this yields}$$
$$\operatorname{vec} \mathbf{A}_0 = \beta^2[\mathsf{E}(\mathbf{Y}^{2\otimes}\mathbf{Y}'^{2\otimes})]^- \mathsf{E}(\mathbf{Y}^{2\otimes}).$$

To express a generalized inverse $[\mathsf{E}(\mathbf{Y}^{2\otimes}\mathbf{Y}'^{2\otimes})]^-$ is a rather complicated task and it seems that the solution vec $\mathbf{A}_0$ does depend on the unknown parameters β, σ. So the expression for the existing minimal MSE estimator in this simplest situation does not exist in a proper form.

References

[1] *J. Kleffe:* Simultaneous Estimation of Expectation and Covariance Matrix in Linear Models. Math. Operationsforsch. Statist., Ser. Statistics, Vol. 9 (1978) No. 3, 443—478.

[2] *J. Volaufová:* Estimation of Polynomials in the Regression Model. Aplikace matematiky, Vol. 27 (1982), No. 3, 223—231.

Súhrn

ODHAD KVADRATICKEJ FUNKCIE PARAMETRA STREDNEJ HODNOTY

JÚLIA VOLAUFOVÁ, PETER VOLAUF

Práca je venovaná optimálnym odhadom funkcie $\boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta}$, $\boldsymbol{\beta} \in \mathscr{R}^k$ v modeli $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, za predpokladu, že rozdelenie vektora $\mathbf{Y}$ je symetrické a má konečný štvrtý moment. Explicitne je odvodený tvar najlepšieho nevychýleného odhadu pre špeciálny prípad matice $\mathbf{X}$.

Резюме

ОЦЕНКА КВАДРАТИЧЕСКОЙ ФУНКЦИИ ПАРАМЕТРА СРЕДНЕГО ЗНАЧЕНИЯ

JÚLIA VOLAUFOVÁ, PETER VOLAUF

Работа занимается проблемой оптимальной оценки функций $\boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta}$, $\boldsymbol{\beta} \in \mathscr{R}^k$ в рамках модели $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ при условии, что распределение вектора $\mathbf{Y}$ симметрично с конечным четвертым моментом. Форма наилучшей несмещенной оценки выведена для матриц $\mathbf{X}$ специального вида.

Authors' addresses: RNDr. *Júlia Volaufová,* CSc., Ústav merania a meracej techniky CEFV SAV, Dúbravská 9, 842 19 Bratislava, RNDr. *Peter Volauf,* CSc., Katedra matematiky EF SVŠT, Mlynská dolina, 812 19 Bratislava.