Estimation under restrictions with biased initial estimators

Natalja Lepik[1]

[1] University of Tartu, Estonia
e-mail: natalja.lepik@ut.ee

Abstract

Estimation under restrictions is considered with both unbiased and biased initial estimators. The properties of the restriction estimator are studied.

1 Introduction

The users of official statistics often require that estimates satisfy some certain restrictions. For example in the domain's case this requirement is that the estimators of the domain totals sum up to the population total or to its estimate. Another example is that quarterly estimates have to sum up to the yearly total. It is natural that such relationships are hold for the true population parameters, so they can be considered and used as a kind of the auxiliary information. Involving this information into the estimation process can improve the estimates.

One solution to the described situation is the general restriction estimator (GR) proposed by Knottnerus (2003) that is based on the unbiased initial estimators. The advantages of this GR estimator is the variance minimizing property among other linear estimators satisfying the same restrictions and using the same initial estimators in its construction. But it is well known that there are very many good estimators that are unbiased only asymptotically. In this paper we consider the biased initial estimators, and the new restriction estimator will be constructed.

2 General restriction estimator. Definition and properties

Let $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)'$ be the parameter vector under study that satisfy linear restrictions:

$$\mathbf{R}\boldsymbol{\theta} = \mathbf{c}, \tag{1}$$

where $\mathbf{R}$ is an $r \times k$ matrix of rank r and $\mathbf{c}$ is the r-dimensional vector of constants.

If for example we require that two population totals, say $t_1 = \sum_{i=1}^{N} y_{1i}$ and $t_2 = \sum_{i=1}^{N} y_{2i}$, need to sum up to the population total $t_3 = \sum_{i=1}^{N} y_{3i}$, where N is the size of the finite population and y_1, y_2, y_3 are some study variables, then

$$\mathbf{R} = (1, 1, -1),\ \boldsymbol{\theta} = (t_1, t_2, t_3)' \text{ and } \mathbf{c} = 0.$$

Theorem (Knottnerus, 2003, p. 328-329) Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, ..., \hat{\theta}_k)'$ be the vector of the unbiased estimators of the parameter vector $\boldsymbol{\theta}$ with a covariance matrix $\mathbf{V}$, such that $\mathbf{RVR}'$ can be inverted. Then the general restriction estimator $\hat{\boldsymbol{\theta}}_{GR}$ that satisfies restrictions (1), and the covariance matrix $\mathbf{V}_{GR}$ of the general restriction estimator are:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{GR} &= \hat{\boldsymbol{\theta}} + \mathbf{K}(\mathbf{c} - \mathbf{R}\hat{\boldsymbol{\theta}}), && (2) \\ \mathbf{V}_{GR} &= \mathbb{C}ov(\hat{\boldsymbol{\theta}}_{GR}) = (\mathbb{I} - \mathbf{KR})\mathbf{V}, && (3) \end{aligned}$$

where $\mathbb{I}$ is the $k \times k$ identity matrix and

$$\mathbf{K} = \mathbf{VR}'(\mathbf{RVR}')^{-1}. \quad (4)$$

As it was mentioned before, Knottnerus shows that the GR-estimator is a linear minimum variance estimator of $\boldsymbol{\theta}$, given $\hat{\boldsymbol{\theta}}$, and given the information that $\mathbf{R}\boldsymbol{\theta} = \mathbf{c}$. In Sõstra (2007, p. 45) it is also shown that GR-estimator is more effective than the initial estimator $\hat{\boldsymbol{\theta}}$, $\mathbf{V}_{GR} < \mathbf{V}$ in the sense of the Löwner ordering.
The case of the domain totals, when the domain totals need to sum up to the population total, is very thoroughly studied in Sõstra (2007).

3 Estimation with restrictions for biased initial estimators

Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, ..., \hat{\theta}_k)'$ be the vector of initial estimators so that,

$$\mathbb{E}\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \mathbf{b}, \quad (5)$$

where $\mathbf{b} = (b_1, ..., b_k)'$ is a bias-vector or simply bias. Denote the mean square error of $\hat{\boldsymbol{\theta}}$ by $\mathbf{M}$, so $\mathbf{M} = \mathbb{M}SE(\hat{\boldsymbol{\theta}})$. Then $\mathbf{M} = \mathbf{V} + \mathbf{bb}'$, where $\mathbf{V}$ is the true covariance matrix of $\hat{\boldsymbol{\theta}}$.

Let us consider the transformed parameter vector[1],

$$\boldsymbol{\vartheta} = \mathbf{M}^{-1/2}\boldsymbol{\theta}. \tag{6}$$

The corresponding estimator is

$$\hat{\boldsymbol{\vartheta}} = \mathbf{M}^{-1/2}\hat{\boldsymbol{\theta}} \tag{7}$$

with $\mathbb{M}SE(\hat{\boldsymbol{\vartheta}}) = \mathbb{I}$, where $\mathbb{I} : k \times k$ is the identity matrix.

Assume also that $\boldsymbol{\theta}$ satisfies restrictions,

$$\mathbf{R}\boldsymbol{\theta} = \mathbf{0},$$

where $\mathbf{R}$ is the $m \times k$ matrix of rank m. So, $\boldsymbol{\theta}$ is in a linear space specified by $\mathcal{N}_R = \{\boldsymbol{\theta}; \mathbf{R}\boldsymbol{\theta} = \mathbf{0}\}$.

In general, restrictions can be defined as $\mathbf{R}\boldsymbol{\theta} = \mathbf{c}$, where $\mathbf{c}$ can be different from $\mathbf{0}$. But in this case it's possible to choose a fixed $\boldsymbol{\theta}_0$ such that $\mathbf{R}\boldsymbol{\theta}_0 = \mathbf{c}$ and then

$$\mathbf{R}\boldsymbol{\theta} - \mathbf{c} = \mathbf{R}\boldsymbol{\theta} - \mathbf{R}\boldsymbol{\theta}_0 = \mathbf{R}(\boldsymbol{\theta} - \boldsymbol{\theta}_0),$$

so we can consider the parameter $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ instead of $\boldsymbol{\theta}$ for which $\mathbf{R}\tilde{\boldsymbol{\theta}} = \mathbf{0}$. Thus, it is no restriction to put $\mathbf{c} = \mathbf{0}$.

For the parameter $\boldsymbol{\vartheta}$ the corresponding restriction matrix is

$$\mathbf{L} = \mathbf{R} \cdot \mathbf{M}^{1/2}. \tag{8}$$

Due to (6) and (8) we have

$$\mathcal{N}_R = \{\boldsymbol{\theta} : \mathbf{R}\boldsymbol{\theta} = \mathbf{0}\} = \{\boldsymbol{\vartheta} : \mathbf{L}\boldsymbol{\vartheta} = \mathbf{0}\} = \mathcal{N}_L. \tag{9}$$

Let us project $\hat{\boldsymbol{\vartheta}}$ orthogonally on the restriction space $\mathcal{N}_L$:

$$\hat{\boldsymbol{\vartheta}}_L = \mathbf{P}_L\hat{\boldsymbol{\vartheta}}, \tag{10}$$

where

$$\mathbf{P}_L = \mathbb{I} - \mathbf{L}'(\mathbf{L}\mathbf{L}')^{-1}\mathbf{L} = \mathbb{I} - \mathbf{M}^{1/2}\mathbf{R}'(\mathbf{R}\mathbf{M}\mathbf{R}')^{-1}\mathbf{R}\mathbf{M}^{1/2}. \tag{11}$$

[1] For a positive-definite matrix $\mathbf{A}$ exists precisely one positive definite matrix $\mathbf{B}$ with $\mathbf{B}^2 = \mathbf{A}$, so we define $\mathbf{A}^{1/2} = \mathbf{B}$.

The estimator $\hat{\boldsymbol{\vartheta}}_L$ satisfies the restrictions $\mathbf{L}\hat{\boldsymbol{\vartheta}}_L = \mathbf{0}$. On the other hand, $\mathbf{L}\hat{\boldsymbol{\vartheta}}_L = \mathbf{R}\mathbf{M}^{1/2}\hat{\boldsymbol{\vartheta}}_L$, from which follows that the estimator

$$\hat{\boldsymbol{\theta}}_R = \boldsymbol{M}^{1/2}\hat{\boldsymbol{\vartheta}}_L \tag{12}$$

satisfies the restrictions $\mathbf{R}\hat{\boldsymbol{\theta}}_R = \mathbf{0}$. Do to (7) and (10), $\hat{\boldsymbol{\theta}}_R$ can be expressed through $\hat{\boldsymbol{\theta}}$:

$$\hat{\boldsymbol{\theta}}_R = \mathbf{M}^{1/2}\mathbf{P}_L\hat{\boldsymbol{\vartheta}} = \mathbf{M}^{1/2}\mathbf{P}_L\mathbf{M}^{-1/2}\hat{\boldsymbol{\theta}}. \tag{13}$$

The expectation of $\hat{\boldsymbol{\theta}}_R$ is then

$$\mathbb{E}(\hat{\boldsymbol{\theta}}_R) = \mathbf{M}^{1/2}\mathbf{P}_L\mathbf{M}^{-1/2}E(\hat{\boldsymbol{\theta}}). \tag{14}$$

Using (5) and (6) we get

$$\mathbb{E}(\hat{\boldsymbol{\theta}}_R) = \mathbf{M}^{1/2}\mathbf{P}_L\boldsymbol{\vartheta} + \mathbf{M}^{1/2}\mathbf{P}_L\mathbf{M}^{-1/2}\mathbf{b}.$$

Obviously for $\boldsymbol{\vartheta} \in \mathcal{N}_L$ we have $\mathbf{P}_L\boldsymbol{\vartheta} = \boldsymbol{\vartheta}$. From (6) it follows that $\mathbf{M}^{1/2}\boldsymbol{\vartheta} = \boldsymbol{\theta}$. Finally for the expectation of $\hat{\boldsymbol{\theta}}_R$ we have

$$\mathbb{E}(\hat{\boldsymbol{\theta}}_R) = \boldsymbol{\theta} + \mathbf{b}_R, \tag{15}$$

where $\mathbf{b}_R = \mathbf{M}^{1/2}\mathbf{P}_L\mathbf{M}^{-1/2}\mathbf{b}$ is the vector of biases of the restriction estimator $\hat{\boldsymbol{\theta}}_R$. From (13) and the property of MSE[2] we get for the mean square error matrix of the restriction estimator $\hat{\boldsymbol{\theta}}_R$:

$$\mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) = \mathbf{M}^{1/2}\mathbf{P}_L \cdot \mathbb{M}SE(\hat{\boldsymbol{\vartheta}}_L) \cdot \mathbf{P}_L\mathbf{M}^{1/2}.$$

Taking into consideration that $\mathbb{M}SE(\hat{\boldsymbol{\vartheta}}_L) = \mathbb{I}$ and that $\mathbf{P}_L^2 = \mathbf{P}_L$ we get

$$\mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) = \mathbf{M}^{1/2}\mathbf{P}_L\mathbf{M}^{1/2}. \tag{16}$$

From (11) we can rewrite the last expression as

$$\mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) = \left[\mathbb{I} - \mathbf{M}\mathbf{R}'(\mathbf{R}\mathbf{M}\mathbf{R}')^{-1}\mathbf{R}\right]\mathbf{M}. \tag{17}$$

Denote $\hat{\boldsymbol{e}} = \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_R$. Due to (7) and (13), we have

$$\hat{\boldsymbol{e}} = \mathbf{M}^{1/2}\hat{\boldsymbol{\vartheta}} - \mathbf{M}^{1/2}\mathbf{P}_L\hat{\boldsymbol{\vartheta}} = \mathbf{M}^{1/2}(\mathbb{I} - \mathbf{P}_L)\hat{\boldsymbol{\vartheta}}. \tag{18}$$

[2] It can be easily shown that $\mathbb{M}SE(\mathbf{A}\hat{\boldsymbol{\theta}}) = \mathbf{A}\mathbb{M}SE(\hat{\boldsymbol{\theta}})\mathbf{A}'$, where $\mathbf{A}$ is the constant matrix of the appropriate size.

Let us observe the MSE matrix between the error term $\hat{\boldsymbol{e}}$ and the restriction estimator $\hat{\boldsymbol{\theta}}_R$. First note that the parameter estimated by $\hat{\boldsymbol{e}}$ is a zero-vector since both $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}_R$ estimate the same parameter $\boldsymbol{\theta}$. From definition of the MSE matrix we have:

$$\mathbb{M}SE(\hat{\boldsymbol{\theta}}_R, \hat{\boldsymbol{e}}) = E\left[(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta})\hat{\boldsymbol{e}}'\right]. \tag{19}$$

By using MSE properties[3] and relationships (13), (18) the last equation can be simplified:

$$\begin{aligned}\mathbb{M}SE(\hat{\boldsymbol{\theta}}_R, \hat{\boldsymbol{e}}) &= \mathbb{M}SE\left[\mathbf{M}^{1/2}\mathbf{P}_L\hat{\boldsymbol{\vartheta}},\ \mathbf{M}^{1/2}(\mathbb{I} - \mathbf{P}_L)\hat{\boldsymbol{\vartheta}}\right] \\ &= \mathbf{M}^{1/2}\mathbf{P}_L\mathbb{M}SE(\hat{\boldsymbol{\vartheta}})(\mathbb{I} - \mathbf{P}_L)\mathbf{M}^{1/2}.\end{aligned}$$

Projectors $\mathbf{P}_L$ and $\mathbb{I} - \mathbf{P}_L$ project onto orthogonal subspaces, which means that $\mathbf{P}_L(\mathbb{I} - \mathbf{P}_L) = 0$. Since by (7) $\mathbb{M}SE(\hat{\boldsymbol{\vartheta}}) = \mathbb{I}$, it holds:

$$E\left[(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta})\hat{\boldsymbol{e}}'\right] = 0. \tag{20}$$

For the mean squre error of $\hat{\boldsymbol{\theta}}$ we have:

$$\begin{aligned}\mathbf{M} = \mathbb{M}SE(\hat{\boldsymbol{\theta}}) &= \mathbb{E}\left[(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta} + \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_R)(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta} + \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_R)'\right] \\ &= \mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) + \mathbb{E}\left(\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}'\right) + \mathbb{E}\left[(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta})\hat{\boldsymbol{e}}'\right] + \mathbb{E}\left[\hat{\boldsymbol{e}}(\hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta})'\right].\end{aligned}$$

Using (20) we finally get:

$$\mathbb{M}SE(\hat{\boldsymbol{\theta}}) = \mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) + \mathbb{E}(\mathbf{e}\mathbf{e}'). \tag{21}$$

Since $\mathbb{E}(\mathbf{e}\mathbf{e}') \geq \mathbf{0}$ (nonnegative definite), we have $MSE(\hat{\boldsymbol{\theta}}) - \mathbb{M}SE(\hat{\boldsymbol{\theta}}_R) \geq \mathbf{0}$ from which it is clear that $\mathbb{M}SE(\hat{\boldsymbol{\theta}}) \geq \mathbb{M}SE(\hat{\boldsymbol{\theta}}_R)$ (in Löwner ordering). The last expression claims that the diagonal elements of the matrix on the right are not greater than the respective diagonal elements of the matrix on the left.

References

Knottnerus, P. (2003) *Sample Survey Theory. Some Pythagorean Perspectives.* Wiley, New York.

Sõstra, K. (2007) *Restriction estimator for domains. Doctoral dissertation.* Tartu University PRESS, Tartu.

[3]The MSE between two random vectors can be defined analogically to the covariance matrix as $\mathbb{M}SE(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\vartheta}}) = \mathbb{E}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta})'$. Then for the constant matrices $\mathbf{A}$, $\mathbf{B}$ and constant vectors $\mathbf{a}$ and $\mathbf{b}$ is hold:

$$\mathbb{M}SE(\mathbf{A}\hat{\boldsymbol{\theta}} + \mathbf{a}, \mathbf{B}\hat{\boldsymbol{\vartheta}} + \mathbf{b}) = \mathbf{A}\mathbb{M}SE(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\vartheta}})\mathbf{B}'.$$