Spectral equation S_1 for minimax estimators of parameters in linear models

V. L. GIRKO

Department of Probability and Statistics, Michigan State University
East Lansing, Michigan 48824

In many cases the estimation of parameters of simultaneous equations amounts to the search for the minimum value of maximum eigenvalues on a certain set of numbers. There is a wide range of published papers dealing with the estimation problem by means of the spectral theory of linear operators [1,2,6,12]. For this reason such estimators will be called spectral or S-estimators. In this section a set of S-estimators for parameters of linear systems, denoted by S_1 is suggested.

Assume that the linear regression model

$$\vec{y} = X\vec{c} + \vec{\varepsilon}$$

is given, where $\vec{c}$ is an unknown m-dimensional vector, $\vec{y}$ is an n -dimensional vector of observations,

$$X_{n\times m} = (x_{ij})\,,\ j = 1, ..., m;\ i = 1, ..., n,\ n \geq m$$

is a matrix, and ε is an n-dimensional random vector of unobservable perturbations such that

$$\mathbf{E}\,\vec{\varepsilon} = 0, \quad \mathbf{E}\,\vec{\varepsilon}\vec{\varepsilon}^T = R_{n\times n}.$$

Let $\vec{c}^T D_{m\times m}\vec{c} \leq \alpha$, where $D_{m\times m}$ is a positive definite matrix, $0 < \alpha < \infty$. We will find a linear transformation of the vector $\vec{y}$:

$$T_{m\times n}\vec{y}_n + \vec{t}_m$$

such that the maximum loss function

$$\varphi := \max_{\vec{c}:\vec{c}^T D_{m\times m}\vec{c}\leq a} \mathbf{E}\,\left(T\vec{y} + \vec{t} - \vec{c}\right)^T V_{m\times m}\left(T\vec{y} + \vec{t} - \vec{c}\right),$$

where $V_{m\times m}$ is a nonnegative definite symmetric matrix, is minimal. Let this minimum be attained for $T = \hat{T}_{m\times n}$; $\vec{t} = \hat{\vec{t}}_m$. The optimal vector

$$\hat{\vec{c}} = \hat{T}\vec{y} + \hat{\vec{t}}$$

is called the S_1-estimator (or minimax estimator) of the vector $\vec{c}$. Let

$$Y = R^{-1/2}X,\ \ B = D^{-1/2}VD^{-1/2} = U\Gamma^2U^T,$$

where $U_{m\times m}$ is the orthogonal matrix of eigenvectors, and $\Gamma_{m\times m} = \left(\gamma_i^{1/2}\delta_{ij}\right)$ is the diagonal matrix. Denote by

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_s > 0,\ \gamma_{s+1} = \cdots = \gamma_m = 0$$

the eigenvalues of the matrix B, where s is an integer,

$$Z = YD^{-1/2}U = HW^{1/2},\ H = Z\left(Z^TZ\right)^{-1/2},\ W = Z^TZ.$$

Let $L_{s\times m}$ be the set of real matrices of the size $s\times m$, and let K_m be the set of real vectors of dimension m.

THEOREM 5.1. *If the matrices D, X^TX and R are nondegenerate, then*

$$\min_{T_{m\times n}\in L_{m\times n},\,\vec{t}_m\in K_m}\ \max_{\vec{c}\in K_m:\vec{c}^TD\vec{c}\leq a}\mathbf{E}\ \left(\vec{c}-T_{m\times n}\vec{y}\pm\vec{t}_m\right)^TV\ \left(\vec{c}-T_{m\times n}\vec{y}\pm\vec{t}_m\right)$$
$$=\min_{\hat{T}\in L_{m\times n}}\left\{a\lambda_{\max}\left[D^{-1/2}\left(I-\hat{T}X\right)^TV\left(I-\hat{T}X\right)D^{-1/2}\right]+Tr\hat{T}^TV\hat{T}R\right\}$$
$$=\min_{A_{s\times m}\in L_{s\times m}}\left\{a\,\lambda_{\max}\left[A^T_{s\times m}A_{s\times m}\right]+Tr\,W^{\ \pm1}\left(A_{s\times m}+\Gamma_{s\times m}\right)^T\left(A_{s\times m}+\Gamma_{s\times m}\right)\right\},$$

where

$$V^{1/2}\hat{\vec{t}}=\vec{0},$$

$\lambda_{\max}$ is the maximum eigenvalue of a matrix, $I_{m\times m}$ is the identity matrix,

$$\Gamma_{s\times m}=\left[\Gamma_{s\times s},0_{s\times m-s}\right],\ \Gamma_{s\times s}=\left(\delta_{kl}\sqrt{\gamma_k}\right)^s_{k,l=1},\ 0_{s\times(m-s)}$$

is the matrix with zero entries,

$$\hat{T}=D^{-1/2}U\left[\begin{array}{c}\left(\Gamma^{-1}_{s\times s}A_{s\times m}+I_{s\times m}\right)U^TD^{1/2}\left(X^TR^{-1}X\right)^{-1}X^TR^{-1}\\ A_{(m-s)\times n}\end{array}\right],$$

and $A_{(m-s)\times n}$ are arbitrary matrices from the set $L_{(m-s)\times n},\ I_{s\times m}=[I_{s\times s},0_{s\times m-s}]$.

Proof. We transform the criterion of estimates quality by substituting the value of the vector and calculate the expectation

$$\varphi:=\max_{\vec{c}\in K_m:\vec{c}^TD\vec{c}\leq a}\left\{\vec{c}^T\left(TX-I\right)^TV\left(TX-I\right)\vec{c}\right.$$
$$\left.+\ 2\left|\vec{t}^TV\left(TX-I\right)\vec{c}\right|\right\}+\vec{t}^TV\vec{t}+\mathrm{Tr}\,R^{1/2}T^TVTR^{1/2}.$$

If we find a minimum for some $\vec{t}\in K_m$ then it is easy to see that $\vec{t}$ satisfies the equation $V^{1/2}\hat{\vec{t}}=\vec{0}$. Let us make the change of variables in the expression for φ :

$$T=\tilde{T}R^{-1/2},\ \tilde{T}\in L_{m\times n},\ \vec{c}=D^{-1/2}\tilde{\vec{c}},\ \tilde{\vec{c}}\in K_m.$$

Then by Rayleigh's formula, we obtain

$$\begin{aligned}\varphi &= a\lambda_{\max}\left\{D^{-1/2}\left(\tilde{T}Y-I\right)^T V\left(\tilde{T}Y-I\right)D^{-1/2}\right\}+\operatorname{Tr}\tilde{T}^T V\tilde{T}\\ &= a\lambda_{\max}\left\{\left(D^{1/2}\tilde{T}YD^{-1/2}-I\right)^T B\left(D^{1/2}\tilde{T}YD^{-1/2}-I\right)\right\}\\ &\quad+\operatorname{Tr}\tilde{T}^T D^{1/2}BD^{1/2}V\tilde{T}.\end{aligned}$$

From this formula, applying the transformation

$$B = U\Gamma^2 U^T,$$

we get

$$\varphi = a\lambda_{\max}[(T_1Z-I)^T\Gamma^2(T_1Z-I)]]+TrT_1^T\Gamma^2T_1, \tag{6.1}$$

where

$$T_1 = U^T D^{1/2}\tilde{T}.$$

Since the matrix $Z = YD^{-1/2}U$ can always be represented in the form

$$Z = HW^{1/2},\ H = Z\left(ZZ^T\right)^{-1/2},\ W = ZZ^T,$$

a simple manipulation yields:

$$\varphi = a\lambda_{\max}[(T_1\tilde{H}\tilde{W}-I_{m\times n})^T\Gamma^2(T_1\tilde{H}\tilde{W}-I_{m\times n})]]+\operatorname{Tr}T_1^T\Gamma^2T_1, \tag{6.2}$$

where $\tilde{H} = \left[H_{n\times m}, Q_{n\times(n-m)}\right]$, a real matrix Q is chosen so that the matrix $\left[H_{n\times m}, Q_{n\times(n-m)}\right]$ is a square orthogonal matrix; $I_{m\times n} = [I_{m\times m}, 0_{m\times n-m}]$,

$$\tilde{W}_{n\times n} = \begin{bmatrix} W^{1/2} & 0\\ 0 & 0\end{bmatrix},$$

where the matrix $W^{1/2}$ is augmented by zeros so that it has dimension $n\times n$. It is not hard to ascertain by multiplying the matrices that expressions (6.1) and (6.2) coincide. Now we can make in (6.2) the change of variables $T_1 = T_2\tilde{H}^T$ where $T_2 \in L_{m\times(n-m)}$, which is a one-to-one transformation, since $\tilde{H}$ is a square orthogonal matrix. After the change, expression (6.2) takes the form

$$\varphi = a\lambda_{\max}\left\{\left(T_2\tilde{W}-I_{m\times n}\right)^T\Gamma^2\left(T_2\tilde{W}-I_{m\times n}\right)\right\}+\operatorname{Tr}T_2^T\Gamma^2T_2. \tag{6.3}$$

Considering that the matrix $T_2\tilde{W}$ does not depend on the columns of matrix T_2 beginning with the $(m+1)$-st column and that

$$T_2 = \left[T_{2,(m\times m)}, T_{2,m\times(n-m)}\right].$$

from expression (6.3) we get

$$\varphi = a\lambda_{\max}\left[\left(T_{2,(m\times m)}W^{1/2} - I_{m\times m}\right)^T \Gamma^2 \left(T_{2,(m\times m)}W^{1/2} - I_{m\times m}\right)\right] \\ + \mathrm{Tr}\ \left[T_{2,(m\times m)}, T_{2,m\times(n-m)}\right]^T \Gamma^2 \left[T_{2,(m\times m)}, T_{2,m\times(n-m)}\right]. \tag{6.4}$$

Now, prior to searching for the minimum over all matrices $T_{2,m\times m}$, we may find the minimum over all matrices $T_{2,m\times(n-m)}$. Evidently, the matrix to be found is a solution to the equation

$$\Gamma\hat{T}_{2,m\times(n-m)} = 0_{m\times(n-m)}. \tag{6.5}$$

Turning to expression (6.4) and taking into account (6.5), we get

$$\min_{T_2\in L_{m\times(n-m)}}\varphi = \min_{T_2\in L_{m\times(n-m)}}\left\{a\lambda_{\max}\left[\left(T_{2,(m\times m)}W^{1/2} - I_{m\times m}\right)^T \Gamma^2 \right.\right. \\ \left.\left. \times\left(T_{2,(m\times m)}W^{1/2} - I_{m\times m}\right)\right] + \mathrm{Tr}\ T^T_{2,(m\times m)}\Gamma^2 T_{2,(m\times m)}\right\}. \tag{6.6}$$

Let us make the change of variables

$$T_{2,m\times m}W^{1/2} - I_{m\times m} = A_{m\times m}$$

in (6.6)

$$\min_{T_2\in L_{m\times(n-m)}}\varphi = \min_{A\in L_{s\times m}}\left\{a\lambda_{\max}\left[A^T\Gamma^2 A\right] + TrW^{-1}\left(A^T + I_{s\times m}\right)\Gamma^2\left(A + I_{s\times m}\right)\right\} \\ = \min_{\tilde{A}\in L_{s\times m}}\left\{a\lambda_{\max}\left(\tilde{A}^T_{s\times m}\tilde{A}_{s\times m}\right) + TrW^{-1}\left(\tilde{A}_{s\times m} + \Gamma_{s\times m}\right)^T\left(\tilde{A}_{s\times m} + \Gamma_{s\times m}\right)\right\}, \tag{6.7}$$

where $\tilde{A}_{s\times m} = \Gamma_{s\times s}A_{s\times m},\ \Gamma_{s\times m} = \left\{\Gamma_{s\times s},\ 0_{s\times(m-s)}\right\}.$

Consequently, to complete the proof of Theorem 6.1, we must find the matrix $\hat{T}$. Written in tandem are the necessary transformations of this matrix:

$$T = \tilde{T}R^{-1/2},\ \ T_1 = U^T D^{1/2}\tilde{T},\ \ T_1 = T_{2,m\times m}H^T + T_{2,m\times(n-m)}Q^T,$$

$$\Gamma T_{2,m\times(n-m)} = 0_{m\times(n-m)},$$

$$T_{2,m\times m} = \left[\left\{\begin{array}{c}\Gamma^{-1}_{s\times s}\tilde{A}_{s\times m}\\ A_{(m-s)\times m}\end{array}\right\} + I_{m\times m}\right]W^{-1/2}.$$

Considering that

$$T_2\,\tilde{H}^T = T_{2,m\times m}H^T + T_{2,m\times(n-m)}Q^T_{n\times(n-m)},$$

we get

$$T_{n\times m} = D^{-1/2}U\left\{\begin{bmatrix}\Gamma^{-1}_{s\times s}\tilde{A}_{s\times m}\\ A_{(m-s)\times m}\end{bmatrix} + I_{m\times m}\right] W^{-1/2}H^{T}_{n\times m} + T_{2,m\times(n-m)}Q^{T}_{n\times(n-m)}\right\} R^{-1/2}.$$

Now we can easily derive the expression for the matrix $\hat{T}$, defined in Theorem 6.1. This concludes the proof of the theorem. Thus we have somewhat simplified the search for estimators of the vector and have reduced it to finding

$$\min_{A\in L_{s\times m}} \left\{a\lambda_{\max}\left(A^TA\right) + \operatorname{Tr} W^{-1}\left(A+\Gamma_{s\times m}\right)^T\left(A+\Gamma_{s\times m}\right)\right\}. \tag{6.8}$$

As is easily verified,

$$\lambda_{\max}(A^TA) = \lambda_{\max}(AA^T).$$

For this reason we may assume in what follows that in expression (6.8) $\lambda_{\max}(AA^T)$ is replaced by $\lambda_{\max}(A^TA)$. In looking for the minimum of expression (6.8), we run into the main difficulty arising from the fact that the eigenvalue $\lambda_{\max}(AA^T)$ of the sought-for matrix $\hat{A}$ may turn out to be multiple, thus preventing us from utilizing well-known perturbation formulas for eigenvalues of multiplicity 1. To surmount this difficulty, we apply the method proposed in [15]. Consider the function

$$\varphi_\varepsilon(A) := a\mathbf{E}\,\lambda_{\max}\left(AA^T + \varepsilon\Xi_{s\times s}\right) + \operatorname{Tr} W^{-1}\left(A+\Gamma_{s\times m}\right)^T\left(A+\Gamma_{s\times m}\right),$$

where $\Xi_{s\times s} = (\xi_{ij})$ is a symmetric random matrix, whose entries $\xi_{ij},\ i\geq j;\ i,j=1,...,s$ are independent and are distributed in accordance with the normal law $N\left[0,(1+\delta_{ij})2^{-1}\right]$, $\varepsilon\neq 0$ is a real number. Since the eigenvalues of a square matrix K are continuous functions of the coefficients of its characteristic equation

$$\det\left(Iz-K\right)=0,$$

for all matrices $A\in L_{s\times m}$ we have

$$\lim_{\varepsilon\to 0}\sup_{A\in L_{s\times m}}|\varphi(A)-\varphi_\varepsilon(A)| = 0 \tag{6.9}$$

We prove that the function ε is strictly convex. Reasoning as in [10] and applying the Cauchy-Bunyakovskii inequality, we see that for all $A,B\in L_{s\times m},\ 0<\alpha,\ \beta;\ \ \alpha+\beta=1$

$$\begin{aligned}\varphi\left(\alpha A+\beta B\right) = a &\max_{\vec{c}\in K_m:\vec{c}^T\vec{c}\leq 1}\vec{c}^T\left(\alpha A+\beta B\right)^T\left(\alpha A+\beta B\right)\vec{c} + \operatorname{Tr} W^{-1}\\ &\times\left(\alpha A+\beta B+\Gamma_{s\times m}\right)^T\left(\alpha A+\beta B+\Gamma_{s\times m}\right)\\ &< a\alpha\lambda_{\max}\left(A^TA\right) + a\beta\lambda_{\max}\left(A^TA\right)\\ &+\alpha\operatorname{Tr} W^{-1}(A+\Gamma_{s\times m})^T(A+\Gamma_{s\times m})\\ &+\beta\operatorname{Tr} W^{-1}(B+\Gamma_{s\times m})^T(B+\Gamma_{s\times m}).\end{aligned}$$

A similar argument makes it possible to conclude that the function $\varphi_\varepsilon(A)$ is likewise strictly convex. Therefore, either of the two functions φ and $\varphi_\varepsilon(A)$ has unique points of minimums, $\hat{A}$ and $\hat{A}_\varepsilon$, respectively. But then it follows from (6.9) that

$$\lim_{\varepsilon\to 0}\hat{A}_\varepsilon=\hat{A}. \tag{6.10}$$

Let us represent the matrix $\hat{A}$:

$$\hat{A}_{s\times m}=\sum\nolimits_{k=1}^{s}\lambda_k\vec{u}_k\vec{v}_k^T,\quad \lambda_1=\lambda_2=\cdots=\lambda_j>\lambda_{j+1}\geq\cdots\geq\lambda_s>0,$$

where $\vec{u}_k,\ k=1,...,s$ are s-dimensional orthogonal vectors, $\vec{v}_k,\ k=1,...,s$ are m-dimensional orthogonal vectors and j is a number such that $1\leq j\leq s$.

THEOREM 6.2. *The numbers λ_k and the vectors $\vec{u}_k,\ \vec{v}_k$ satisfy the S_1 -equation*

$$Wal_1\sum_{k=1}^{j}\vec{\nu}_k\vec{u}_k^Tp_k+\lambda_1\sum_{k=1}^{j}\vec{\nu}_k\vec{u}_k^T+\sum_{k=j+1}^{s}\lambda_k\vec{\nu}_k\vec{u}_k^T+\Gamma_{s\times m}^T=0, \tag{6.11}$$

where

$$p_k>0,\ \sum\nolimits_{k=1}^{j}p_k=1,\ Y=R^{-1/2}X,\ \ B=D^{-1/2}VD^{-1/2}=U\Gamma^2U^T,$$

$U_{m\times m}$ is the orthogonal matrix of eigenvectors, and $\Gamma_{m\times m}=\left(\gamma_i^{1/2}\delta_{ij}\right)$ is the diagonal matrix,

$$\gamma_1\geq\gamma_2\geq\cdots\geq\gamma_s>0,\ \gamma_{s+1}=\cdots=\gamma_m=0$$

are the eigenvalues of the matrix B, s is an integer, and

$$Z=YD^{-1/2}U=HW^{1/2},\ H=Z\left(Z^TZ\right)^{-1/2},\ W=Z^TZ.$$

Proof. Let us first prove that there exists the derivative

$$(\partial/\partial\gamma)\varphi_\varepsilon\left(A+\gamma\Theta\right)|_{\gamma=0},\ \ \Theta\in L_{s\times m}.$$

To do this, we consider the expression

$$\begin{aligned}\lim_{\gamma\downarrow 0}\gamma^{-1}\left[\varphi_\varepsilon\left(A+\gamma\Theta\right)-\varphi_\varepsilon\left(A\right)\right]=\lim_{\gamma\downarrow 0}a\gamma^{-1}\mathbf{E}\ \Big[\lambda_{\max}\left\{\left(A+\gamma\Theta\right)\left(A+\gamma\Theta\right)^T+\varepsilon\Xi\right\}\\ -\lambda_{\max}\left(AA^T+\varepsilon\Xi\right)\Big]\left(\chi\left(B\right)+\chi\left(\bar{B}\right)\right)+2\mathrm{Tr}\,W^{\pm1}\left(A+\Gamma_{sxm}\right)^T\Theta,\end{aligned} \tag{6.12}$$

where B is the following random event:

$$B=\left\{\omega:\left|\lambda_{\max}\left(AA^T+\varepsilon\Xi\right)-\lambda_i\left(AA^T+\varepsilon\Xi\right)\right|>\delta,i\geq 2\right\},\ \delta>0,$$

$\chi(B)$ is the indicator for the event B. The density of eigenvalues $\nu_1 \geq \cdots \geq \nu_s$ of matrix $(A+\gamma\Theta)(A+\gamma\Theta)^T + \varepsilon\Xi$ is

$$p(y_1, ..., y_s) := c_\varepsilon \int\limits_G \exp\left\{-\frac{\mathrm{Tr}\,\left[(A+\gamma\Theta)(A+\gamma\Theta)^T - HYH^T\right]^2}{2\varepsilon^2}\right\}\mu(\mathrm{d}H)\prod_{i>j}|y_i - y_j|, \tag{6.13}$$

where H is an orthogonal matrix of the s-th order, μ is the Haar measure on the group G of orthogonal matrices $H = (h_{ij})$, c_ε is the normalizing factor, and $y_1 > \cdots > y_s$.

Using this density and the Schwarz inequality we get

$$\begin{aligned}
&\mathbf{E}\,\left|\lambda_{\max}\left\{(A+\gamma\Theta)(A+\gamma\Theta)^T + \varepsilon\Xi\right\} - \lambda_{\max}\left\{AA^T + \varepsilon\Xi\right\}\right|\chi(\bar{B}) \\
&\leq \sqrt{2}\left|\mathbf{E}\,\lambda^2_{\max}\left\{(A+\gamma\Theta)(A+\gamma\Theta)^T + \varepsilon\Xi\right\} + \mathbf{E}\,\lambda^2_{\max}\left\{AA^T + \varepsilon\Xi\right\}\right|^{1/2}\mathbf{E}\,\chi(\bar{B}) \\
&\leq c\sum_{i=2}^{s}\mathbf{P}\left\{|\nu_1 - \nu_i| \leq \delta\right\} \leq c_1\sum_{i=2}^{s}\int\limits_{|y_1-y_i|<\delta} p\left(y_1, \cdots, y_s\right)\prod_{k=1}^{s}\mathrm{d}y_k \leq c_2\int\limits_{|u|<\delta}|u|\,\mathrm{d}u \leq c_3\delta^2,
\end{aligned}$$

where c_i are constants. By virtue of the perturbation formulas for simple eigenvalues, (6.12) and (6.14), we get

$$\begin{aligned}
\left.\frac{\partial}{\partial\gamma}\varphi_\varepsilon(A+\gamma\Theta)\right|_{\gamma=0} &= \lim_{\gamma\downarrow 0}\gamma^{-1}\left\{a\mathbf{E}\,\vec{\psi}^T_{1\varepsilon}\left[(A+\gamma\Theta)\left(A+\gamma\Theta\right)^T - AA^T\right]\vec{\psi}_{1\varepsilon} + \nu(\delta)\right\} \\
&\quad + 2\mathrm{Tr}\,W^{-1}\left(A+\Gamma_{s\times m}\right)^T\Theta,
\end{aligned}$$

where $\vec{\psi}_{1\varepsilon}$ is an eigenvector associated with the eigenvalue ν_1,

$$\begin{aligned}
|\nu(\delta)| \leq c_1\Bigg[&\delta^2 + \frac{1}{\delta}\left\|(A+\gamma\Theta)(A+\gamma\Theta)^T - AA^T\right\|^2 \\
&\times\left(1 - \frac{1}{\delta}\left\|(A+\gamma\Theta)(A+\gamma\Theta)^T - AA^T\right\|\right)^{-1}\Bigg].
\end{aligned}$$

Choosing δ in such a way that

$$\lim_{\gamma\downarrow 0}[\gamma^{-1}\delta^2 + \delta^{-1}\gamma] = 0$$

we obtain

$$\left.\frac{\partial}{\partial\gamma}\varphi_\varepsilon(A+\gamma\Theta)\right|_{\gamma=0} = 2\mathbf{E}\,\vec{\psi}^T_{1\varepsilon}\Theta A^T\vec{\psi}_{1\varepsilon} + 2\mathrm{Tr}\,W^{-1}\left(A+\Gamma_{s\times m}\right)^T\Theta. \tag{6.15}$$

Insofar as the function $\varphi_\varepsilon(A)$ has a unique minimum, $\hat{A}_\varepsilon$, and is strictly convex, for all $\Theta \in L_{s\times s}$ we have

$$(\partial/\partial\gamma)\,\varphi_\varepsilon\left(\hat{A}_\varepsilon+\gamma\Theta\right)\Big|_{\gamma=0}=0.$$

From this equality, (6.15) and the fact that $\Theta\in L_{s\times m}$ is an arbitrary matrix, it follows that the unknown matrix $\hat{A}_\varepsilon$ is given by the equation

$$\mathbf{E}\,a\hat{A}_\varepsilon^T\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^T+W^{-1}\left(\hat{A}_\varepsilon^T+\Gamma_{s\times m}^T\right)=0. \tag{6.16}$$

Moreover this equation always has a unique solution. Since $\hat{A}_\varepsilon^T=H_\varepsilon\left(\hat{A}_\varepsilon\hat{A}_\varepsilon^T\right)^{1/2}$, where $H_\varepsilon=\hat{A}_\varepsilon^T\left(\hat{A}_\varepsilon\hat{A}_\varepsilon^T\right)^{-1/2}$ is an orthogonal matrix, then for small enough ε equation (6.16) is equivalent to

$$\begin{aligned}\mathbf{E}\,\Big\{&aH_\varepsilon\nu_1^{1/2}\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^T+W^{-1}\left(\hat{A}_\varepsilon^T+\Gamma_{sxm}^T\right)+H_\varepsilon a\Big[\left(\hat{A}_\varepsilon\hat{A}_\varepsilon^T\right)^{1/2}\\&-\left(\hat{A}_\varepsilon\hat{A}_\varepsilon^T+\varepsilon\Xi\right)^{1/2}\Big]\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^T\Big\}\chi\left(\nu_i\geq 0,\ i=1,...,s\right)=0.\end{aligned} \tag{6.17}$$

Let us prove the auxiliary assertion.

LEMMA 6.1. *For a certain subsequence $\varepsilon\to 0$*

$$\lim_{\varepsilon\downarrow 0}\mathbf{E}\,\nu_1^{1/2}H_\varepsilon\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^T=\lambda_1\sum\nolimits_{k=1}^{j}p_k\vec{v}_k\vec{u}_k^T, \tag{6.18}$$

where

$$p_k>0,\ \ \sum\nolimits_{k=1}^{j}p_k=1.$$

Proof. According to (6.13) we get

$$\begin{aligned}\mathbf{E}\,\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^T=c_\varepsilon\int\limits_{y_1>\cdots>y_s}&\vec{h}_1\vec{h}_1^T\exp\left\{-(2\varepsilon^2)^{-1}\mathrm{Tr}\left(\hat{A}_\varepsilon\hat{A}_\varepsilon^T-HYH^T\right)^2\right\}\mu(\mathrm{d}H)\\&\times\prod_{i>j}|y_i-y_j|\prod_{k=1}^{s}\mathrm{d}y_k,\end{aligned} \tag{6.19}$$

where $\vec{h}_1$ is the first column vector of the matrix H. Since the matrix $\hat{A}\,\hat{A}_\varepsilon^T$ can always be represented as $\hat{A}\,\hat{A}_\varepsilon^T=U_\varepsilon\Lambda_\varepsilon U_\varepsilon^T$, where U_ε is the orthogonal matrix of eigenvectors $\vec{u}_{i\varepsilon},\ \ i=1,...,s$ and

$$\Lambda_\varepsilon=\left(\lambda_{i\,\varepsilon}\delta_{ij}\right),\ \ \lambda_{1\varepsilon}\geq\cdots\geq\lambda_{s\,\varepsilon}$$

is the diagonal matrix of eigenvalues, making the change of variables $H=U_\varepsilon\tilde{H},\ \tilde{H}\in G$ and invoking the invariance of the Haar measure, from formula (6.19) we obtain

$$\mathbf{E}\,\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^{T}=c_{\varepsilon}U_{\varepsilon}\int_{y_1>\ldots.>y_s}\tilde{\vec{h}}_1\tilde{\vec{h}}_1^{T}U_{\varepsilon}^{T}\exp\left\{-\left(2\varepsilon^2\right)^{-1}\mathrm{Tr}\,\left(\Lambda_{\varepsilon}\ -\tilde{H}Y\tilde{H}^{T}\right)^2\right\}$$

$$\times\,\mu(d\tilde{H})\prod_{i>j}|y_i-y_j|\prod_{k=1}^{s}\mathrm{d}y_k \tag{6.20}$$

$$=U_{\varepsilon}\left\{\mathbf{E}\,\psi_{i1,\varepsilon}^2(\Lambda_{\varepsilon}+\varepsilon\Xi)\delta_{ij}\right\}_{i,j=1}^{s}U_{\varepsilon}^{T},$$

where $\vec{\psi}_1(\Lambda_{\varepsilon}+\varepsilon\Xi)$ is the eigenvector of matrix $\Lambda_{\varepsilon}+\varepsilon\Xi$ corresponding to the maximum eigenvalue. Let us represent the matrix $\hat{A}_{\varepsilon}$:

$$\hat{A}_{\varepsilon}=\sum\nolimits_{k=1}^{s}\lambda_{k\varepsilon}\vec{u}_{k\varepsilon}\vec{v}_{k\varepsilon}^{T},\ \ \lambda_{1\varepsilon}\geq\cdots\geq\lambda_{s\varepsilon}>0,$$

where $\vec{u}_{k\varepsilon},\ k=1,...,s$ are s-dimensional orthogonal vectors, $\vec{v}_{k\varepsilon},\ k=1,...,s$ are m-dimensional orthogonal vectors and j is a number such that $1\leq j\leq s$.

The perturbation formulas for eigenvalues imply that

$$\lim_{\varepsilon\to 0}\lambda_{k\varepsilon}=\lambda_1,\ k=1,...,j,\ \ \lim_{\varepsilon\to 0}\lambda_{q\,\varepsilon}=\lambda_q,\ q=j+1,...,s, \tag{6.21}$$

$$\lim_{\varepsilon\to 0}\vec{u}_{k\varepsilon}=\vec{u}_k,\ \ \lim_{\varepsilon\to 0}\vec{v}_{k\varepsilon}=\vec{v}_k,\ \ k=1,...,s,$$

and since $\lambda_q\neq\lambda_1, q=j+1,...,s$, then we have

$$\lim_{\varepsilon\to 0}\sum\nolimits_{q=j+1}^{s}\vec{\psi}_{q\,\varepsilon}\vec{\psi}_{q\,\varepsilon}^{T}=\begin{bmatrix}0_{j\times j} & 0_{j\times(s-j)}\\ 0_{(s-j)\times j} & I_{(s-j)\times(s-j)}\end{bmatrix},$$

where I is a square identity matrix of order $s-j$. But in this case $\psi_{1i}^2(\Lambda_{\varepsilon}+\varepsilon\Xi)\to 0$, as $\varepsilon\to 0$ for all $i=j+1,...,s$, and

$$\lim_{\varepsilon\to 0}\sum\nolimits_{q=j+1}^{s}\vec{\psi}_{q\,\varepsilon}\vec{\psi}_{q\,\varepsilon}^{T}\neq\begin{bmatrix}0_{j\times j} & 0_{j\times(s-j)}\\ 0_{(s-j)\times j} & I_{(s-j)\times(s-j)}\end{bmatrix},$$

otherwise. Therefore, setting

$$p_{i\,\varepsilon}=\mathbf{E}\,\psi_{i1,\varepsilon}^2\left(\Lambda_{\varepsilon}+\varepsilon\Xi\right)$$

and utilizing (6.21) from formula (6.20) we obtain that for a certain subsequence $\varepsilon'\to 0$

$$\lim_{\varepsilon\downarrow 0}\mathbf{E}\,\nu_1^{1/2}H_{\varepsilon}\vec{\psi}_{1\varepsilon}\vec{\psi}_{1\varepsilon}^{T}=\lambda_1\lim_{\varepsilon\downarrow 0}H_{\varepsilon}U_{\varepsilon}p_{i\,\varepsilon}\delta_{ij}{}_{i,j=1}^{s}U_{\varepsilon}^{T}=\lambda_1\sum\nolimits_{k=1}^{j}p_k\vec{v}_k\vec{u}_k^{T}.$$

Hence Lemma 6.1 is proved.

Lemma 6.1 implies that

$$\lim_{\varepsilon'\to 0}\hat{A}_{\varepsilon'}=\sum\nolimits_{k=1}^{s}\lambda_k\vec{u}_k\vec{v}_k^{T}$$

(see (6.10)). Then, passing to the limit as $\varepsilon' \to 0$ in equality (6.16) and utilizing (6.18), we get the S_1 -equation. Theorem 6.2 is proved.

Let us consider several corollaries.

COROLLARY 6.1. *If in addition to the conditions of Theorem 6.1 $s = m$, $\Gamma_s = I_{m\times m}$ then $j = m$ and equation (6.10) has the unique solution*

$$\hat{A} = -I_{m\times m} \mathrm{Tr}\, W^{-1} \left[a + \mathrm{Tr}\, W^{-1}\right]^{-1},$$

where $W = Z^T Z,\ Z = R^{-1/2} X D^{-1/2}$.

Proof. We represent the matrix W in the form $W = HBH^T$, where H is an orthogonal matrix and $B = (b_i \delta_{ij})$ is a diagonal matrix. Then from equation (6.11) we have

$$Ba\lambda_1 \sum_{k=1}^{j} \tilde{\vec{\nu}}_k \tilde{\vec{u}}_k^T p_k + \lambda_1 \sum_{k=1}^{j} \tilde{\vec{\nu}}_k \tilde{\vec{u}}_k^T + \sum_{k=j+1}^{s} \lambda_k \tilde{\vec{\nu}}_k \tilde{\vec{u}}_k^T + I = 0,$$

where $\tilde{\vec{v}}_k = H^T \vec{v}_k\,,\quad \tilde{\vec{u}}_k \ = H^T \vec{u}_k$.

Multiplying this equation from the right by $\tilde{\vec{u}}_k$ we get the system of equations

$$l_1 \left(Ba\, p_k + I\right) \tilde{\vec{v}}_k = -\tilde{\vec{u}}_k;\ k = 1, ..., j;\ \ \lambda_q \tilde{\vec{v}}_q = -\tilde{\vec{u}}_q;\ \ q = j+1, ..., m.$$

Hence $\lambda_q \equiv 1;\ \ q = j+1, ..., m$. But from the first j equalities, it follows that

$$l_1 \le \left[\tilde{\vec{v}}_k^T \left(Ba\, p_k + I\right) \tilde{\vec{v}}_k\right]^{-1} < 1.$$

Therefore $j = m$. Then

$$Ba\, l_1 \sum\nolimits_{k=1}^{m} \tilde{\vec{v}}_k \tilde{\vec{u}}_k^T p_k + \lambda_1 \sum\nolimits_{k=1}^{m} \tilde{\vec{v}}_k \tilde{\vec{u}}_k^T = -I.$$

Denoting

$$U = \left(\tilde{\vec{u}}_k, k = 1, ..., m\right), V = \left(\tilde{\vec{v}}_k, k = 1, ..., m\right), P = (p_k \delta_{kl})$$

from this equation we have

$$(aBVP + V)U^T = -l_1^{-1} I.$$

Hence

$$(aBVP + V)(aBVP + V)^T = l_1^{-2} I$$

and $(aK + B^{-1})^2 = l_1^{-2} B^{-2}$, where $K = VPV^T$. It is easy to see that

$$aK + B^{-1} = l_1^{-1} B^{-1}$$

and

$$l_1 = \mathrm{Tr}\, B^{-1} (a + TrB^{-1})^{-1},\ \ VU^T = -I.$$

Corollary 6.1 is proved.

Corollary 6.2. If in addition to the conditions of Theorem 6.1 $s = m$, matrix V is nondegenerate, $W = (\omega_i \delta_{ij})$ is a diagonal matrix, then

$$\hat{A}_{m\times m} = -\left(\lambda_k \delta_{kj}\right)_{k,j=1}^{m}$$

where

$$\lambda_i = \sum_{k=1}^{j} \gamma_k^{1/2} \omega_k^{-1} \left[a + \sum_{k=1}^{j} \omega_k^{-1} \right]^{-1}, \quad i = 1, ..., j; \quad \lambda_q = \gamma_q^{1/2}, q = j+1, ..., m,$$

the number j satisfies the inequality

$$\sum\nolimits_{k=1}^{j} \gamma_k^{1/2} \omega_k^{-1} \left[a + \sum\nolimits_{k=1}^{j} \omega_k^{-1} \right]^{-1} > \gamma_{j+1}^{1/2}, \quad \gamma_{m+1} \equiv 0,$$

and $\gamma_1 \geq \cdots \geq \gamma_m > 0$ are the eigenvalues of the matrix $B = D^{-1/2} V D^{-1/2}$.
Proof. In this case equation (6.11) has the form

$$W a \lambda_1 \sum_{k=1}^{j} \vec{\nu}_k \vec{u}_k^T p_k + \lambda_1 \sum_{k=1}^{j} \vec{\nu}_k \vec{u}_k^T + \sum_{k=j+1}^{m} \lambda_k \vec{\nu}_k \vec{u}_k^T + \Gamma_{m\times m} = 0.$$

Multiplying this equation by $\Gamma_{m\times m}^{-1}$, $\vec{v}_k^T$ and $\vec{u}_k$ we get

$$\vec{v}_k^T \lambda_1 \Gamma^{-1} (aWp_k + 1) \vec{v}_k = -\vec{v}_k^T \vec{u}_k; \;\; \vec{v}_q^T \lambda_q \Gamma^{-1} \vec{v}_q = -\vec{v}_q^T \vec{u}_q.$$

In our case

$$\Gamma^{-1}(aWp_k + 1)$$

is a symmetric matrix. Therefore, $V = U$,

$$\{\lambda_1 (aWp_k + I - \Gamma)\} \vec{v}_k = 0, k = 1, ..., j; \;\; (\Gamma - \lambda_q I) \vec{v}_q = 0; \;\; q = j+1, ..., m.$$

From this equation we get the system of equations

$$a\omega_k \lambda_1 p_k + \lambda_1 = \gamma_k^{1/2}, \;\; k = 1, ..., j; \;\; \lambda_q = \gamma_q^{1/2}, \;\; q = j+1, ..., m.$$

From these equations we obtain the assertion of Corollary 6.2.

REFERENCES

1. V.L. Girko, S.I. Lyascko and A.G. Nakonechny. On Minimax Regulator for Evolution Equations in Hilbert Space Under Conditions of Uncertainty. *Cybernetics*, N.1, 1987, 67–68 p.

2. V.L. Girko. Spectral Equations S_2 for Minimax Estimations of Solutions of Some Linear Systems*. Third International Workshop on Model Oriented Data Analysis (Moda-3) St. Petersburg, Petrodvorets, Russia 25–30 May 1992 p.11.
3. V.L. Girko. Spectral Equations S_1 and S_2 for Minimax Estimations of Solutions of some Linear Systems Linear Minimax Estimation - Theory and Practice 3–4 August 1992, Oldenburg.
4. V.L. Girko. Spectral Equation for Minimax Estimator of Parameters of Linear Systems*. *Calculating and Applied Mathematics*, N.63, 1987, 114–115 p. Translation in *J. Soviet Math.* **66** 2221–2222 (1993).
5. V.L. Girko and N. Christopeit. Minimax Estimators for Linear Models with Nonrandom Disturbances. *Random Operators and Stochastic Equations* **3**, N.4, 1995, 361–377 p.
6. V.L. Girko. Spectral Theory of Minimax Estimation, *Acta Applicandae Mathematicae* **43**, 1996, 59–69 p.
7. V.L. Girko. **Theory of Linear Algebraic Equations with Random Coefficients**, (monograph), Allerton Press, Inc. New York 1996, 320 p.