Turkish J. Ineq., 6 (1) (2022), Pages 1 –6 .

Turkish Journal of
INEQUALITIES

Available online at www.tjinequality.com

A STUDY ON LEAST SQUARES ESTIMATION WITH INEQUALITY CONSTRAINT

FATMA BUĞLEM YALÇIN[1]

Abstract. In this paper, the least-squares estimators are obtained by imposing inequality constraints on parameter vector in classical regression model and using genaralized inverses of matrices. Additionally, the Kantorovich type inequality for a special matrix is expressed.

1. Introduction

In many fields, the data collected through the application is examined and it is desired to find a function that models this data. It may not always be possible to find the function that fits the data exactly. The regression analysis is a method of finding the function that best fits these datas [15].

The ordinary least squares (OLS) method is one of most commonly used methods in regression analysis [14]. The famous mathematician C. F. Gauss used firstly this method to determine the orbit of the Cres asteroid [4]. The OLS method is one of the methods used in determining the relationships between variables in many science such as mathematics, sociology, engineering, medicine and agriculture [6]. The OLS method has become a subject that mathematicians and other scientists have been work on (see. [1,2,8,10,12,16]).

The OLS method is the optimal method according to the Gauss-Markov Theorem, since it aims to minimize the sum of squares error (SSE). When some assumptions are provided for the data, the method provides reliable estimates [7]. However, it can sometimes produce misleading results. Therefore, it may be necessary to impose a linear equality or an inequality constraint on the parameter. It is aimed to minimize the SSE again [5,11,13].

In this study, our aim is to find the inequality constrained least squares (ICLS) estimation in linear models by means of the Moore-Penrose g-inverse, which is a special case of the generalized inverse of matrices. And it is to prove the Kantorovich type inequality for the matrix consisting of the difference between the variance of the ICLS estimator and

Key words and phrases. Constrained Least Squares, Moore-Penrose Generalized Inverse, Kantorovich type inequalities, Parameter Estimation.

2010 *Mathematics Subject Classification.* Primary: 93E24. Secondary: 47A05, 15A42, 15A39.

Received: 26/12/2021 *Accepted*: 04/05/2022.

Cited this article as: F. B. Yalçın, A study on least squares estimation with inequality constraint, Turkish Journal of Inequalities, 6(1) (2022), 1-6.

the variance of the OLS estimator. Firstly, let's obtain the ICLS estimator by using the Lagrange method.

2. Inequality Constrained Least Squares Estimation

Consider a general linear model

$$\omega = \varphi\theta + \varepsilon \tag{2.1}$$

where φ is an $s \times t$ matrix of known constant and $rank(\varphi) = t \leq s$, ω is an $s \times 1$ observable random vector, θ is a $t \times 1$ vector of parameters to be estimated, ε is a random error vector with $E(\varepsilon) = 0, cov(\varepsilon) = \sigma^2 I$ and σ^2 is an unknown positive parameter. As is known, the OLS estimator $\hat{\theta}$ is obtained by minimizing with the regard to θ:

$$\hat{\theta} = (\varphi^{'}\varphi)^{-1}\varphi^{'}\omega \tag{2.2}$$

and

$$\hat{\sigma}^2 = \frac{(\omega - \varphi\hat{\theta})^{'}(\omega - \varphi\hat{\theta})}{s - rank(\varphi)} \tag{2.3}$$

is an unbiased estimator of σ^2 where the estimated error vector is $\hat{\varepsilon} = \omega - \varphi\hat{\theta}$. $E(\hat{\theta}) = \theta$ and $var(\hat{\theta}) = \sigma^2(\varphi^{'}\varphi)^{-1}$ is obtained.

Consider q-linear inequality constraints

$$\xi\theta \geq \eta \tag{2.4}$$

imposed on coefficients where ξ is a matrix of known constant, $rank(\xi_{q\times t}) = q$, and η is a $q \times 1$ known vector. We can write 2.4 as

$$\xi\theta - \mu = \eta \tag{2.5}$$

where μ is $q \times 1$ surplus vector, θ isn't otherwise constrained. The ICLS estimation is obtained by minimizing the following objective function

$$\varepsilon^{'}\varepsilon = (\omega - \varphi\theta)^{'}(\omega - \varphi\theta) \tag{2.6}$$

subject to $\xi\theta - \mu = \eta$. Constructed the Lagrange function

$$\Im = \omega^{'}\omega - 2\theta^{'}\varphi^{'}\omega + \theta^{'}\varphi^{'}\varphi\theta - \alpha^{'}(\eta - \xi\theta + \delta) \tag{2.7}$$

where $\alpha^{'}$ is an $q \times 1$ vector of Lagrange multipliers. By using the result in [17], differentiation of the Lagrange function with respect to θ and α yields the following conditions

$$-2\varphi^{'}\omega + 2\varphi^{'}\varphi\theta + \xi^{'}\alpha, \tag{2.8}$$

$$\xi\theta - \mu - \eta = 0. \tag{2.9}$$

Now pre-multiplying equation 2.8 by $\xi(\varphi^{'}\varphi)^{-1}$ gives that

$$-2\xi(\varphi'\varphi)^{-1}\varphi'\omega+2\xi\theta+\xi(\varphi'\varphi)^{-1}\xi'\alpha=0. \tag{2.10}$$

Since the matrix $\xi(\varphi'\varphi)^{-1}\xi'$ is positive definite, by using $\hat{\theta}=(\varphi'\varphi)^{-1}\varphi'\omega$, we get

$$\alpha=-2(\xi(\varphi'\varphi)^{-1}\xi')^{-1}(\eta+\mu-\xi\hat{\theta}) \tag{2.11}$$

By substituting this value in 2.8, we obtain the ICLS estimator $\hat{\theta}_{icls}$ of θ:

$$\hat{\theta}_{icls}=\hat{\theta}+(\varphi'\varphi)^{-1}\xi'(\xi(\varphi'\varphi)^{-1}\xi')^{-1}(\eta+\delta-\xi\hat{\theta}). \tag{2.12}$$

$\hat{\theta}_{icls}$ is unbiased for θ. This expression is the same as in [3]. As to the variance-covariance matrix of $\hat{\theta}_{icls}$, we have

$$var(\hat{\theta}_{icls})=E[(\hat{\theta}_{icls}-\theta)(\hat{\theta}_{icls}-\theta)']=\sigma^2 M_{icls}(\varphi'\varphi)^{-1}(M_{icls})' \tag{2.13}$$

where the matrix $M_{icls}=[I-(\varphi'\varphi)^{-1}\xi'(\xi(\varphi'\varphi)^{-1}\xi')^{-1}\xi]$ is idempotent but not symmetric. Now let's get the ICLS estimator by another method. We will use the Moore-Penreose generalized inverse in matrices for this.

3. The Inequality Constrained Least Squares Estimation via The Moore-Penrose Generalized Inverse

Under the constraints 2.5 and the condition $rank(\xi_{q\times t})=q$, by using Moore-Penrose generalized inverse $\xi^{+}=\xi'(\xi\xi')^{-1}$ of ξ (ξ^{+} is unique), we get the solution

$$\hat{\theta}_{icls}=\xi^{+}(\eta+\mu)+(I-\xi^{+}\xi)\hat{\theta}. \tag{3.1}$$

If we use $(\varphi'\varphi)^{-1}\xi'[\xi(\varphi'\varphi)^{-1}\xi']^{-1}$ in place of ξ^{+}, then this g-inverse satisfies the three properties of Moore-Penrose generalized inverse. Therefore, we get

$$\begin{aligned}\hat{\theta}_{icls}&=\{(\varphi'\varphi)^{-1}\xi'[\xi(\varphi'\varphi)^{-1}\xi']^{-1}\}(\eta+\mu)+(I-\{(\varphi'\varphi)^{-1}\xi'[\xi(\varphi'\varphi)^{-1}\xi']^{-1}\}\xi)\hat{\theta}\\&=\hat{\theta}+(\varphi'\varphi)^{-1}\xi'[\xi(\varphi'\varphi)^{-1}\xi']^{-1}(\eta+\mu-\xi\hat{\theta}).\end{aligned} \tag{3.2}$$

The covariance matrix of the ICLS estimator is obtained in the following forms:

$$var(\hat{\theta}_{icls})=\sigma^2(\varphi'\varphi)^{-1}+\sigma^2[-(\varphi'\varphi)^{-1}\xi^{+}\xi-\xi^{+}\xi(\varphi'\varphi)^{-1}+\xi^{+}\xi(\varphi'\varphi)^{-1}\xi^{+}\xi], \tag{3.3}$$

and

$$var(\hat{\theta})-var(\hat{\theta}_{icls})=\sigma^2[(\varphi'\varphi)^{-1}\xi^{+}\xi+\xi^{+}\xi(\varphi'\varphi)^{-1}-\xi^{+}\xi(\varphi'\varphi)^{-1}\xi^{+}\xi]. \tag{3.4}$$

Now let's analyse $\phi=(\varphi'\varphi)^{-1}\xi^{+}\xi+\xi^{+}\xi(\varphi'\varphi)^{-1}-\xi^{+}\xi(\varphi'\varphi)^{-1}\xi^{+}\xi$ matrix consisting of the difference of $var(\hat{\theta})$ and $var(\hat{\theta}_{icls})$.

Let be choose $x\epsilon C(\xi')=N(I-\xi'\xi'^{+})$ vector where $C(\xi')$ is column space of ξ' and N is null space. Therefore $x=\xi' y$, $y\epsilon R^q$ and $x\epsilon R^t$. Using Moore-Penrose g-inverse, we have

$$\begin{aligned}x'\phi x&=y'\xi[(\varphi'\varphi)^{-1}\xi^{+}\xi+\xi^{+}\xi(\varphi'\varphi)^{-1}-\xi^{+}\xi(\varphi'\varphi)^{-1}\xi^{+}\xi]\xi' y\\&=y'[\xi(\varphi'\varphi)^{-1}\xi^{+}\xi\xi'+\xi\xi^{+}\xi(\varphi'\varphi)^{-1}\xi'-\xi\xi^{+}\xi(\varphi'\varphi)^{-1}\xi^{+}\xi\xi']y\\&=y'[\xi(\varphi'\varphi)^{-1}\xi'+\xi(\varphi'\varphi)^{-1}\xi'-\xi(\varphi'\varphi)^{-1}\xi']y=y'[\xi(\varphi'\varphi)^{-1}\xi']\end{aligned}$$

for $\forall x \epsilon R^t$. Since $rank(\varphi) = t$, the matrices $(\varphi^{'}\varphi)$, $(\varphi^{'}\varphi)^{-1}$ and $\xi(\varphi^{'}\varphi)^{-1}\xi^{'}$ are positive definite. Then ϕ is a positive definite matrix. Thus, the variance of the ICLS estimator is the variance of the OLS estimator minus a positive definite matrix.

4. The Kantorovich Inequality for The Difference Matrix

Let us prove Kantorovich type inequality for the matrix which is the difference between the variance of the ICLS estimator and the variance of the OLS estimator (See [9] for Kantorovich type equality).

Theorem 4.1. *Let ϕ be an positive definite matrix which is the difference between the variance of the ICLS estimator and the variance of the OLS estimator, λ_1 and λ_t be ϕ's largest and smallest eigenvalues. If $\varphi^{'}$ is an $t \times s$ matrix such that $(\varphi^{'})^*\varphi^{'} = I_s$, then*

$$(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\varphi^{'}$$

$$\leq \frac{(\lambda_1+\lambda_t)^2}{4\lambda_1\lambda_t}[(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2, \tag{4.1}$$

$$\begin{aligned}&(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\varphi^{'}\\ &-[(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2\\ &\leq \frac{1}{4}(\lambda_1-\lambda_t)^2 I_s,\end{aligned} \tag{4.2}$$

$$\begin{aligned}&(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} -\\ &\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\varphi^{'} - [(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2\\ &\leq (\sqrt{\lambda_1}-\sqrt{\lambda_t})^2[(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]\end{aligned} \tag{4.3}$$

(Here φ^* is the conjugate transpose of φ)

Proof. Since $0 \leq (\lambda_1 I_t - (\varphi^{'}\varphi)^{-1}\xi^+\xi - \xi^+\xi(\varphi^{'}\varphi)^{-1} + \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi - \lambda_t I_t)$, we get

$$\begin{aligned}&[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\\ &\leq (\lambda_1+\lambda_t)[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi] - \lambda_1\lambda_t I_t\end{aligned}$$

and

$$(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\varphi^{'}$$

$$\leq (\lambda_1+\lambda_t)(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]\varphi^{'} - \lambda_1\lambda_t I_s. \tag{4.4}$$

We can write the right-hand side of 4.4 as

$$\begin{aligned}&(\lambda_1+\lambda_t)(\varphi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]\varphi^{'} - \lambda_1\lambda_t I_s\\ &= \tfrac{(\lambda_1+\lambda_t)^2}{4\lambda_1\lambda_t}[(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2\\ &-\{\sqrt{\lambda_1\lambda_t I_s} - \tfrac{\lambda_1+\lambda_t}{2\sqrt{\lambda_1\lambda_t}}[(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]\}^2\\ &= [(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2 + \tfrac{1}{4}(\lambda_1-\lambda_t)^2 I_s\\ &-\{(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'} - \tfrac{\lambda_1+\lambda_t}{2}I_s\}^2\\ &= [(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi + \xi^+\xi(\varphi^{'}\varphi)^{-1} - \xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2\end{aligned}$$

$$+(\sqrt{\lambda_1}-\sqrt{\lambda_t})^2(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}$$
$$-[\sqrt{\lambda_1\lambda_t I_s}-(\varphi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\varphi^{'}]^2$$

We have 4.1, 4.2 and 4.3. The proof is complete. □

Corollary 4.1. *Under the assumptions of 4.1, we have the following inequalities such that* $(\xi^{'})^*\xi^{'}=I_q$

$$(\xi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\xi^{'}$$
$$\leq\frac{(\lambda_1+\lambda_t)^2}{4\lambda_1\lambda_t}[(\xi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\xi^{'}]^2 \tag{4.5}$$

$$(\xi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\xi^{'}$$
$$-[(\xi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\xi^{'}]^2$$
$$\leq\frac{1}{4}(\lambda_1-\lambda_t)^2I_q \tag{4.6}$$

$$(\xi^{'})^*[(\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi]^2\xi^{'}$$
$$-[(\xi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\xi^{'}]^2$$
$$\leq(\sqrt{\lambda_1}-\sqrt{\lambda_t})^2[(\xi^{'})^*((\varphi^{'}\varphi)^{-1}\xi^+\xi+\xi^+\xi(\varphi^{'}\varphi)^{-1}-\xi^+\xi(\varphi^{'}\varphi)^{-1}\xi^+\xi)\xi^{'}] \tag{4.7}$$

Acknowledgements. The author would like to thank her deceased thesis advisor (Prof. Dr. Cemil YAPAR) for guiding her with his valuable information.

References

[1] S.T. Barratt and S.P. Boyd, *Least Squares Auto-Tuning,* Engineering Optimization, **53**(5) (2020), 789-810.
[2] G. Casella and R.L. Berger, *Statistical Inference,* Cengage Learning, USA, 2001.
[3] C.K. Liew, *Inequality Constrained Least-Squares Estimation,* Journal of the American Statistical Association, **71**(355) (1976), 746-751.
[4] D.D. Mazanek, R.G. Merrill, J.R. Brophy and R.P. Mueller, *Asteroid Redirect Mission concept: A bold approach for utilizing space resources,* Acta Astronautica, **117** (2015), 163-171.
[5] Q. Fan, C. Jia, J. Liu and Y. Luo, *Robust Recovery in 1-bit Compressive Sensing via ℓq-Constrained Least Squares,* Signal Processing, **23**(5) (2021), 2073-2088.
[6] J.J. Faraway, *Linear Models with R,* Chapman and Hall/CRC Texts in Statistical Science, USA, 2014.
[7] J. Fox, *Applied Regression Analysis: Linear Models and Related Methods,* Sage Publications, USA, 2002.
[8] F.A. Graybill, *Introduction to Matrices with Applications in Statistics,* Wadsworth Publishing, USA, 1969.
[9] G. He, *Equivalence of Some Matrix Inequalities,* Journal of Mathematical Inequalities, **8**(1) (2014), 153-158.
[10] M. Landreman and C. Zhu, *Calculation of Permanent Magnet Arrangements for Stellarators: A Linear Least-Squares Method,* Plasma Physics and Controlled Fusion, **63**(3) (2021), 035001.
[11] Q. Liu, C. Chen and Q. Zhang, *Perturbation analysis for total least squares problems with linear equality constraint,* Applied Numerical Mathematics, **161** (2021), 69-81.
[12] S.J. Miller, *The Method of Least Squares*, Mathematics Department Brown University, Brown University, 2006, pp. 1-7.
[13] T. Nikazad and M. Karimpour, *Column-Oriented Algebraic Iterative Methods for Nonnegative Constrained Least Squares Problems,* Numerical Algorithms, **86** (2021), 1265-1284.
[14] S.V. Carvalho de Souza and R.G. Junqueira, *A procedure to assess linearity by ordinary least squares method,* Analytica Chimica Acta, **552**(1-2) (2005), 25-35.

[15] S.H. Brown, *Multiple Linear Regression Analysis: A Matrix Approach with MATLAB,* Alabama Journal of Mathematics, Spring/Fall 2009, 1-3.

[16] S.A. Van de Geer, *Least Squares Estimation,* Encyclopedia of Statistics in Behavioral Science, **2** (2005), 1041-1045.

[17] P. Krottnerus, *On New Variance Approximations for Linear Models with Inequality Constraints,* Statica Neerlandica, **70** (2016), 26-46.

[1]Faculty of Science and Arts,
Department of Mathematics,
Ordu University,
Ordu, Turkey
E-mail address: yalcinfatmabuglem@gmail.com