ISSN: 2281-1346



UNIVERSITÀ DI PAVIA
Department of Economics and Management

DEM Working Paper Series

Again on Constrained Quadratic Forms

Giorgio Giorgi
(Università degli Studi di Pavia)

 # 234 (10-26)

Via San Felice, 5
I-27100 Pavia

https://economiaemanagement.dip.unipv.it/it

Again on Constrained Quadratic Forms

Giorgio Giorgi [1]

[1] Department of Economics and Management. University of Pavia. Via S. Felice, 5 - 27100 PAVIA (Italy). E-mail: giorgio.giorgi@unipv.it

Abstract. We consider a quadratic form under constraints, which are a generalization of the usual homogeneous linear constraints. We study the sign of the said constrained quadratic form by means of the sign of the roots of a polynomial equation. As a corollary we obtain the corresponding criterion for the sign of a quadratic form under homogeneous linear constraints. In the Appendix we give an English translation (due to M. El-Hodiri) of a paper of R. Shostak, published in Russian on *Uspekhi Mat. Nauk*, 1954, Volume 9, Issue 2 (60), pp. 199-206.

Keywords. Constrained quadratic forms. Positive definite quadratic forms. Negative definite quadratic forms.

1. Introduction.

Positive definiteness and positive semidefiniteness of the restriction of a quadratic form to a linear subspace given by the solutions of a system of homogeneous linear equations, have been investigated since a long time. See, e. g., the classical paper of Debreu (1952), Bellman (1970), Chabrillac and Crouzeix (1984), Farebrother (1977), Ibramov and Saveleva (2000). As it is known, it is possible to determine the sign of a linearly constrained quadratic form by means of a suitable bordered matrix. See the above references and Giorgi (2017), Mann (1943), Välihao (1984). It seems that in the paper of Mann there is a slip in the statement of negative semidefiniteness conditions.

More precisely, we have to study the sign of the quadratic form

$$Q(x) = x^\top Ax$$

subject to the constraints

$$Bx = [0]\,,$$

where $A = [a_{ij}]$ is a symmetric matrix of order n and $B = [b_{ij}]$ is a matrix of order (m, n), with $rank(B) = m < n$. Let us denote by

$$A_r = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{bmatrix}$$

the *leading principal submatrix of* A *of order* r and by

$$B_{mr} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mr} \end{bmatrix}$$

the submatrix of B, of order (m, r), made with the first r columns of B. Furthermore, define the following determinants

$$|C_r| = \det \begin{bmatrix} [0] & B_{mr} \\ B_{mr}^{\top} & A_r \end{bmatrix}, \quad r = m+1,\ m+2, ..., n.$$

Then, we have the following classical theorem (however, not always reported in a correct way in some textbooks).

Theorem 1.1. Without loss of generality, let be B_{mm} nonsingular (i. e. the rank of B is given by the first m columns). Then:

(i) $x^{\top}Ax < 0$, $\forall x \neq [0]$ such that $Bx = [0]$ if and only if

$$(-1)^r |C_r| > 0,\ r = m+1, m+2, ..., n.$$

(ii) $x^{\top}Ax > 0$, $\forall x \neq [0]$ such that $Bx = [0]$ if and only if

$$(-1)^m |C_r| > 0,\ r = m+1, m+2, ..., n.$$

Conditions (i) and (ii) of Theorem 1.1 can be rewritten in the form, perhaps more useful:

(i) $x^{\top}Ax < 0$, $\forall x \neq [0]$ such that $Bx = [0]$ if and only if the leading principal minors of the bordered matrix

$$C_n = \begin{bmatrix} [0] & B \\ B^{\top} & A \end{bmatrix}$$

of order $2m+1, ..., m+n$, alternate in sign, beginning with the sign of $(-1)^{m+1}$.

(ii) $x^{\top}Ax > 0$, $\forall x \neq [0]$ such that $Bx = [0]$ if and only if the leading principal minors of C_n, of order $2m+1, ..., m+n$, have all the sign of $(-1)^m$.

If we consider the boredered matrix

$$\tilde{C}_n = \begin{bmatrix} A & B^{\top} \\ B & [0] \end{bmatrix},$$

we have the equivalences

$$\begin{aligned} (-1)^r |C_r| &> 0 \text{ if and only if } (-1)^r \left|\tilde{C}_r\right| > 0,\ r = m+1, ..., n. \\ (-1)^m |C_r| &> 0 \text{ if and only if } (-1)^m \left|\tilde{C}_r\right| > 0,\ r = m+1, ..., n. \end{aligned}$$

Remark 1.2. Note that the assumption (not restrictive) that the rank of B is given by the first m columns, is essential in order to obtain necessary and sufficient conditions. Otherwise, we have only sufficient conditions. Consider, for instance, the following example:

$$Q(x) = (x_1)^2 + (x_2)^2 - (x_3)^2,$$

which is obviously positive definite under the constraint $x_3 = 0$ (hence here $b^\top = [0, 0, 1]$). Yet we have

$$\begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0.$$

Besides the above conditions, there is in the literature another condition, based on the roots of a bordered determinant and used in Economic Analysis, for instance by Samuelson (1947). This condition can be found also in the book of Hancock (1917, reprinted in 1960), who attributes it to Richelot (1858) and to Zajaczkowski (*Annals of the Scientific Society of Cracow, 12*, 1867). Debreu (1952) attributes this condition to Richelot (1858) and subsequently to Hotelling (1935) and Samuelson (1947). See also Chabrillac and Crouzeix (1984).

Theorem 1.2. Let be given the quadratic form

$$Q(x) = x^\top Ax,$$

where A is symmetric of order n, subject to the constraints

$$Bx = [0],$$

with B of order (m, n) and rank $m < n$. Consider the determinantal equation, which has only real roots,

$$f(\lambda) = \begin{vmatrix} A - \lambda I & B^\top \\ B & [0] \end{vmatrix} = 0.$$

Then $Q(x)$ is positive definite, under the constraints $Bx = [0]$ if and only if all the roots of the equation $f(\lambda) = 0$ are positive. $Q(x)$ is negative definite, under the constraints $Bx = [0]$, if and only if all the roots of the equation $f(\lambda) = 0$ are negative.

In Section 2 we prove in a simple way a generalization of the above theorem and in the Appendix we give a translation into English (due to M. El-Hodiri) of a paper in Russian of R. Ya. Shostak, published in *Uspekhi Mat. Nauk, 9, 1954,* 199-206, and whose translation, as far as we are aware, never has been published. This paper gives an algebraic proof of the criterion described in Theorem 1.1 and it is also concerned with the case of non homogeneous constraints, i. e. of the type $Bx = c$.

2. A generalization of the criterion of Theorem 1.2.

We consider the following

Theorem 2.1. Let be given the quadratic form

$$Q(x) = x^\top Ax$$

defined on the set

$$X = \left\{x \in \mathbb{R}^n : x^\top Cx = 1,\ Bx = [0]\right\},$$

where A and C are symmetric matrices, of order n, C is positive definite and B is a matrix of order (m, n), with $rank(B) = m < n$. Then:

(i) The equation, of degree $(n - m)$ in λ

$$P(\lambda) = \begin{vmatrix} A - \lambda C & B^\top \\ B & [0] \end{vmatrix} = 0$$

has only real roots.

(ii) If λ_* and λ^* are, respectively, the smallest and the greatest roots of the determinantal equation sub (i), then

$$\begin{aligned} \min_x \left\{x^\top Ax : x^\top Cx = 1,\ Bx = [0]\right\} &= \lambda_*; \\ \max_x \left\{x^\top Ax : x^\top Cx = 1,\ Bx = [0]\right\} &= \lambda^*. \end{aligned}$$

Moreover, $x \in \mathbb{R}^n$ (with $x^\top Cx = 1$) is a global minimizer (resp. global maximizer) of $Q(x)$ over X if and only if there exists $u \in \mathbb{R}^n$ such that (x, u) solves the system

$$\begin{cases} (A - \lambda C)x + B^\top u = [0] \\ Bx = [0] \end{cases}$$

for $\lambda = \lambda_*$ (respectively, for $\lambda = \lambda^*$).

(iii) In particular, for $C = I$, we deduce that a necessary and sufficient condition such that $Q(x) = x^\top Ax$ be positive definite (resp. positive semidefinite) over the space of the solutions of $Bx = [0]$, is that the determinantal equation

$$f(\lambda) = \begin{vmatrix} A - \lambda I & B^\top \\ B & [0] \end{vmatrix} = 0$$

has all positive roots (resp. all nonnegative roots). That is, the thesis of Theorem 1.2.

Proof.

(*i*) Let λ be a complex root and $(x,u) \neq [0]$ a (complex) solution of the system

$$(A - \lambda C)x + B^\top u = [0] \tag{1}$$

$$Bx = [0]\,. \tag{2}$$

By taking the conjugates of λ, x and u, we obtain $B\bar{x} = [0]$ and

$$A\bar{x} + B^\top \bar{u} = \bar{\lambda} C\bar{x}. \tag{3}$$

(The conjugate of a vector is intended as the vector given by the conjugate of each component). From (1) and (3), we obtain, respectively,

$$\begin{aligned} x^\top A\bar{x} + u^\top B\bar{x} &= \lambda x^\top C\bar{x}; \\ x^\top A\bar{x} + x^\top B^\top \bar{u} &= \bar{\lambda} x^\top C\bar{x}, \end{aligned}$$

which, by (2) and $B\bar{x} = [0]$, by subtraction, give

$$(\lambda - \bar{\lambda}) x^\top C\bar{x} = 0,$$

from which we have $\lambda = \bar{\lambda}$: hence λ is real. We have $x^\top C\bar{x} \neq 0$: if $0 = x^\top C\bar{x} = (x^1 + ix^2)^\top C(x^1 - ix^2) = (x^1)^\top Cx^1 + (x^2)^\top Cx^2$ and being C positive definite, we have $(x^i)^\top Cx^i = 0$, $i = 1, 2$, hence $x^1 = x^2 = [0]$, that is $x = [0]$; therefore from (1) and the fact that $B^\top$ has linearly independent columns, it holds $u = [0]$, against the assumption that $(x,u) \neq [0]$.

(*ii*) As $Q(x)$ is continuous on the compact set X, there exists a global minimizer $x^0 \in X$ (and also a global maximizer). We have $rank(B^\top, Cx^0) = m + 1$. If $B^\top v + v_0 Cx^0 = [0]$, it results $0 = (x^0)^\top B^\top v + v_0 (x^0)^\top Cx^0 = v_0$ and hence $B^\top v = [0]$, from which, since $B^\top$ has linearly independent columns, $v = [0]$. Then, there exist multipliers $\lambda_0 \in \mathbb{R}$ and $u^0 \in \mathbb{R}^m$ such that for the Lagrangian function

$$\mathcal{L}(x, \lambda, u) = x^\top Ax - \lambda(x^\top Cx - 1) + 2u^\top Bx,$$

we have

$$\nabla_x \mathcal{L}(x^0, \lambda, u^0) = 2\left[(A - \lambda_0 C)x^0 + B^\top u^0\right] = [0]\,.$$

As the system (1)-(2) has non trivial solutions, it results that λ_0 is a root of the equation $P(\lambda) = 0$ and in particular, from (1) we get $Q(x^0) = \lambda_0$ and x^0 global minimizer of $Q(x)$ over X. Moreover, λ_0 is necessarily the smallest root of the equation $P(\lambda) = 0$, that is $\lambda_0 = \lambda_*$. We have proved that each global minimizer of $Q(x)$ is the component x of a solution of the system (1)-(2), with $\lambda = \lambda_*$. Conversely, if (x,u) satisfies (1)-(2), and $x^\top Cx = 1$, then from (1) we get $Q(x) = \lambda_*$ and hence x is a global minimizer of $Q(x)$ over X.

(*iii*) Trivial. We notice only that λ^0 is a negative root of $f(\lambda) = 0$ if and only if $-\lambda^0$ is a positive root of $f(-\lambda) = 0$. □

Remark 2.1. If we consider the example of Remark 1.2, we have, by applying Theorem 1.2 (i. e. (iii) of Theorem 2.1):

$$\begin{vmatrix} A-\lambda I & b^\top \\ b & 0 \end{vmatrix} = 0 \Longrightarrow \begin{vmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & -(1-\lambda) & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix} = 0 \Longrightarrow$$

$$\Longrightarrow (-1)(1-\lambda)^2 = 0 \Longrightarrow \lambda_1 = \lambda_2 = 1,$$

which correctly states that the constrained quadratic form of this example is positive definite, under the given constraint.

Remark 2.2. If, as in Chabrillac and Crouzeix (1984), instead of the function $f(\lambda)$, we consider the function

$$g(\lambda) = \begin{vmatrix} A+\lambda I & B^\top \\ B & [0] \end{vmatrix},$$

then the conditions of Theorem 1.2 are modified as follows:

(i) $Q(x)$ is positive definite under the constraints $Bx = [0]$ if and only if the roots of the equation $g(\lambda) = 0$ are *negative.*

(ii) $Q(x)$ is negative definite under the constraints $Bx = [0]$ if and only if the roots of the equation $g(\lambda) = 0$ are *positive.*

Remark 2.3. Summing up: $Q(x)$ is positive definite under the constraints $Bx = [0]$ if and only if the roots of $f(\lambda) = 0$ are all positive; it is positive semidefinite, under the same constraints, if and only if the roots of $f(\lambda) = 0$ are all nonnegative. It is negative definite, under the same constraints, if and only if the roots of $f(\lambda) = 0$ are all negative; it is negative semidefinite, under the same constraints, if and only if the roots of $f(\lambda) = 0$ are all nonpositive. It is indefinite, under the same constraints, if and only if the roots of $f(\lambda) = 0$ present a variation of sign. If $Q(x)$ is semidefinite, under the same constraints, and $f(0) \neq 0$, that is

$$\begin{vmatrix} A & B^\top \\ B & [0] \end{vmatrix} \neq 0,$$

then $Q(x)$ is actually definite, under the same constraints. Indeed, suppose that $Q(x)$ is positive semidefinite under the constraints $Bx = [0]$ with $rank(B) = m < n$ and that $\bar{x}^\top A\bar{x} = 0$, $B\bar{x} = [0]$. Then $\bar{x}$ minimizes $Q(x)$ under the constraints and since the constraint qualifications are satisfied (the contraints are linear functions), there will exist a vector $\lambda \in \mathbb{R}^m$ such that

$$A\bar{x} + B^\top \lambda = [0].$$

Thus

$$\begin{bmatrix} A & B^\top \\ B & [0] \end{bmatrix} \begin{bmatrix} \bar{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} A\bar{x} + B^\top \lambda \\ B\bar{x} \end{bmatrix} = \begin{bmatrix} [0] \\ [0] \end{bmatrix}.$$

Since

$$\begin{vmatrix} A & B^\top \\ B & [0] \end{vmatrix} \neq 0,$$

we see that $\bar{x} = [0]$ and $\lambda = [0]$. Thus $Bx = [0]$ and $\bar{x} \neq [0]$ imply $x^\top Ax > 0$. On the other hand, if $f(0) \neq 0$, obviously 0 is not a root of $f(\lambda) = 0$.

Remark 2.4. The problem with applying the criterion of Theorem 1.2, is that it does not provide a simple formula for the coefficients of the polynomial. Even without applying the celebrated Descates' Rule of Signs, the following facts are easy to see: if all nonzero coeffcients of the polynomial $f(\lambda)$ have the same sign, then f has no strictly positive roots and hence A is negative semidefinite under constraints. If the coefficients of the polynomial $f(\lambda)$ alternate in sign, then A is positive semidefinite under constraints. Here we must consider the zero coefficients to be alternating in sign. For the reader's convenience we recall the Descates' Rule: let

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_0$$

be a polynomial with real coefficients and with real roots (just such as $f(\lambda)$). Then, the number of positive roots of $P(x) = 0$, counted with their respective algebraic multiplicity, is equal to the number of variations of the sequence of the coefficients of $P(x)$.

Appendix

R. Ya. Shostak

On a criterion of conditional definiteness of a quadratic form of n variables subject to linear relations, and on a sufficient condition for a constrained extremum of a function of n variables (in Russian),
Uspekhi Mat. Nauk, 1954, Volume 9, Issue 2 (60), 199-206.
Translated by M. El-Hodiri

1.

Formulation of the problem. Consider the following quadratic form in n variables

$$U(x_1, x_2, ..., x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji}). \tag{4}$$

The variables $x_1, ..., x_n$ satisfy the following k linear homogeneous equations:

$$\psi_i(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} b_{ij} x_j = 0, \ (i = 1, 2, ..., k). \tag{5}$$

The rank of the matrix $[b_{ij}]$ is assumed to equal the number of equations, i. e. equals k. We shall find necessary and sufficient conditions for the form (4) to be positive definite subject to (5).

2. To solve the above problem, we assume that the matrix $[b_{ij}]$ has a square submatrix of order k with a non-zero determinant. Without loss of generality, we may assume this submatrix to consist of the first k columns of $[b_{ij}]$. Let B denote the determinant of that submatrix, i. e.

$$B = \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{vmatrix} \neq 0. \tag{6}$$

In what follows B_j^r denotes the determinant obtained by replacing the j-th column of b by the r-th column of $[b_{ij}]$, i. e.

$$B_j^r = \begin{vmatrix} b_{11} & \cdots & b_{1,j-1} & b_{1r} & b_{1,j+1} & \cdots & b_{1k} \\ b_{21} & \cdots & b_{2,j-1} & b_{2r} & b_{2,j+1} & \cdots & b2k \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ b_{k1} & \cdots & b_{k,j-1} & b_{kr} & b_{k,j+1} & \cdots & b_{kk} \end{vmatrix} \quad (j = 1, 2, ..., k;\ r = k{+}1, k{+}2, ..., n). \tag{7}$$

3. Solving (5) for the k "dependent" variables $x_1, ..., x_k$, we have:

$$x_i = -\frac{1}{B} \sum_{r=k+1}^{n} B_i^r x_r \ (i = 1, 2, ..., k). \tag{8}$$

Substituting for $x_1, ...x_k$ in (4) from (8), the form $U(x_1, ..., x_n)$ becomes:

$$W(x_{k+1}, x_{k+2}, ..., x_n) = \frac{1}{B^2} \sum_{r,s=k+1}^{n} A_{rs} x_r x_s, \tag{9}$$

where

$$A_{r,s} = a_{rs}B^2 - \sum_{i=1}^{k} a_{is} B B_i^r - \sum_{j=1}^{k} a_{rj} B B_j^s + \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} B_i^r B_j^s. \tag{10}$$

The positive definiteness of the form (4) subject to (5) is equivalent to the unconstrained positive definiteness of the form (9) and they have the same necessary and sufficient conditions. These are: $M_i > 0$, $i = 1, 2, ..., n-k$, where M_i are the principal diagonal minors of the determinants of W, i. e. of:

$$D = \begin{vmatrix} A_{k+1,k+1} & A_{k+1.k+2} & \cdots & A_{k+1,n} \\ A_{k+2,k+1} & A_{k+2,k+2} & \cdots & A_{k+2,n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n,k+1} & A_{n,k+2} & \cdots & A_{nn} \end{vmatrix}. \tag{11}$$

But such conditions are not easy to verify; thus we introduce a determinant which is equivalent, for our purpose, to the determinant of (9). Such a determinant, as we shall show, has the form:

$$\Delta = (-1)^k \begin{vmatrix} 0 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & 0 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & b_{k1} & b_{k2} & \cdots & b_{kn} \\ b_{11} & b_{21} & \cdots & b_{k1} & a_{11} & a_{12} & \cdots & a_{1n} \\ b_{12} & b_{22} & \cdots & b_{k2} & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ b_{1n} & b_{2n\cdots} & & b_{kn} & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{12}$$

We shall show that the determinant Δ is a positive multiple of D and that the principal diagonal minors of Δ, μ_{2k+j}, $j = 1, 2, ..., n-k$, are positive multiples of the corresponding principal diagonal minors, M_j, of D (all principal diagonal minors of order less than or equal to $2k-1$ are zeros and $\mu_{2k} = B^2$).

To prove the above assertions we introduce three transformations of Δ.

4. First Transformation:

We multiply the determinant Δ by the determinant β (of order $n+k$) where

$$\beta = \begin{vmatrix} B_{11} & B_{12} & \cdots & B_{1k} & 0 & 0 & \cdots & 0 \\ B_{21} & B_{22} & \cdots & B_{2k} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ B_{k1} & B_{k2} & \cdots & B_{kk} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{vmatrix}. \tag{13}$$

Here B_{ij} denotes the cofactor of b_{ij} in B. We first multiply the columns of β by the columns of Δ, obtaining a determinant $\beta\Delta$. We then multiply $\beta\Delta$ by β, this time multiply the rows of $\beta\Delta$ by the rows of β. Since $\beta = B^{k-1}$, the transformation yields Δ multiplied by the positive factor B^{2k-2}. In view of the identity

$$\sum_{i=1}^{k} b_{ir} B_{ij} = \begin{cases} 0 \text{ for } j \neq r,\ r \leqq k, \\ B \text{ for } j = r, \\ B_j^r \text{ for } r > k, \end{cases}$$

we have

$$B^{2k-2}\Delta=(-1)^k\begin{vmatrix}
0 & 0 & \cdots & 0 & B & 0 & \cdots & 0 & B_1^{k+1} & \cdots & B_1^n\\
0 & 0 & \cdots & 0 & 0 & B & \cdots & 0 & B_2^{k+1} & \cdots & B_2^n\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & B & B_k^{k+1} & \cdots & B_k^n\\
B & 0 & \cdots & 0 & a_{11} & a_{12} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n}\\
0 & B & \cdots & 0 & a_{21} & a_{22} & \cdots & a_{2k} & a_{2,k+1} & \cdots & a_{2n}\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
0 & 0 & \cdots & B & a_{k1} & a_{k2} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn}\\
B_1^{k+1} & B_2^{k+1} & \cdots & B_k^{k+1} & a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n}\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
B_1^n & B_2^n & \cdots & B_k^n & a_{n1} & a_{n2} & \cdots & a_{nk} & a_{n.k+1} & \cdots & a_{nn}
\end{vmatrix}. \tag{14}$$

5. Second Transformation:

In (14) we set at the intersection of each of the first k rows and the last $n-k$ columns equal to zero. Also set equal to zero, the elements of (14) at the intersection of the first k columns with the last $n-k$ rows.

To achieve this, multiply the elements of the $(2k+m)$-th column ($m=1,...,n-k$) by B and subtract from them the sum of the products of the elements of the $(k+1)$-st,...,$2k$-th columns, respectively, by $B_1^{k+m}, B_2^{k+m},..., B_k^{k+m}$. Then operate analogously on the last $n-k$ rows of (14), utilizing the elements of the preceding k rows. The determinant (14), thus, gets multiplied by a positive number B^{2n-2k} and we have:

$$B^{2n-2}\Delta=(-1)^k\begin{vmatrix}
0 & 0 & \cdots & 0 & B & 0 & \cdots & 0 & 0 & \cdots & 0\\
0 & 0 & \cdots & 0 & 0 & B & \cdots & 0 & 0 & \cdots & 0\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & B & 0 & \cdots & 0\\
B & 0 & \cdots & 0 & a_{11} & a_{12} & \cdots & a_{1k} & c_{1,k+1} & \cdots & c_{1n}\\
0 & B & \cdots & 0 & a_{21} & a_{22} & \cdots & a_{2k} & c_{2,k+1} & \cdots & c_{2n}\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
0 & 0 & \cdots & B & a_{k1} & a_{k2} & \cdots & a_{kk} & c_{k,k+1} & \cdots & c_{kn}\\
0 & 0 & \cdots & 0 & c'_{k+1,1} & c'_{k+1,2} & \cdots & c'_{k+1,k} & d_{k+1,k+1} & \cdots & d_{k+1,n}\\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
0 & 0 & \cdots & 0 & c'_{n1} & c'_{n2} & \cdots & c'_{nk} & d_{n.k+1} & \cdots & d_{nn}
\end{vmatrix}. \tag{15}$$

In (15) we have:

$$\left.\begin{array}{c} c_{rs} = a_{rs}B - \sum_{j=1}^{k} a_{rj}B_j^s \ (r = 1,2,...,n; s = k+1, k+2, ..., n), \\ c'_{rs} = a_{rs}B - \sum_{i=1}^{k} a_{is}B_i^r \ (r = k+1,\ k+2, ..., n;\ s = 1,2,...,k), \\ d_{rs} = c_{rs}B - \sum_{i=1}^{k} c_{is}B_i^r = a_{rs}B^2 - \sum_{j=1}^{k} a_{rj}BB_j^s - \sum_{i=1}^{k} a_{is}BB_i^r + \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}B_i^r B_j^s \end{array}\right\}. \tag{16}$$

From (16) and (10) we note that $d_{rs} = A_{rs}$.

6. Third Transformation:

In (15) let us place the columns numbered $k+1$ to $2k$ in place of the first k columns after multiplying them by $(-1)^k$. Noting that $d_{rs} = A_{rs}$ and substituting, we get:

$$B^{2n-2}\Delta = (-1)^{2k}\begin{vmatrix} B & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & B & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & B & 0 & \cdots & 0 & 0 & \cdots & 0 \\ a_{11} & a_{12} & \cdots & a_{1k} & B & \cdots & 0 & c_{1,k+1} & \cdots & c_{1n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & 0 & \cdots & B & c_{k,k+1} & \cdots & c_{kn} \\ c'_{k+1,1} & c'_{k+1,2} & \cdots & c'_{k+1,k} & 0 & \cdots & 0 & A_{k+1,k+1} & \cdots & A_{k+1,n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ c'_{n1} & c'_{n2} & \cdots & c'_{nk} & 0 & \cdots & 0 & A_{n,k+1} & \cdots & A_{nn} \end{vmatrix}. \tag{17}$$

Observe that exchanging columns in (16) does change the principal diagonal minors from the $(2k+1)$-st on.

But (17) may be immediately simplified. The value of the left hand corner subdeterminant of order k is B^{2k}. The value of the determinant (17) is given by

$$B^{2n-2}\Delta = B^{2k}D. \tag{18}$$

7. We have now shown that D and Δ are positive multiplies of one another. By similar transformations we can show the same for diagonal principal minors, i. e.

$$\mu_{2k+j} = \gamma^2 M_j \quad (j = 1,2,...,n-k). \tag{19}$$

Thus we have proved the following theorem:

• The quadratic form $U(x_1,...,x_n) = \sum_{i,j=1}^{n} a_{ij}x_i x_j$ subject to $\sum_{j=1}^{b} b_{ij}x_j = 0 \ \ (i = 1,...,k)$ is positive definite if and only if all the principal diagonal minors, from the $(2k+1)$-st on, of Δ (given by (12)), are positive. For negative definiteness, the conditions is that the principal diagonal minors, starting with the $(2k+1)$-st, alternate in sign with the first one bering negative. This assumes $[b_{ij}]$ has rank k, given by the first k columns of $[b_{ij}]$.

8. We now apply the conditions to finding the extrema of a function of n-variables subject to constraints, i. e. finding the extrema of $f(x_1, x_2, ..., x_n)$ subject to $\phi_i(x_1, ..., x_n) = 0$, $i = 1, 2, ..., k < n$. It is well known that a necessary condition for a point to be an extremum is that it be a critical point, i. e. there exist λ_i, $i = 1, ..., k$, such that

$$\begin{aligned} \frac{\partial V}{\partial x_j} &= 0 \quad (j = 1, 2, ..., n), \\ \phi_i(x_1, ..., x_n) &= 0, i = 1, 2, ..., k, \end{aligned} \tag{20}$$

where $V = f + \sum \lambda_i \phi_i$.

9. Let $\bar{x}_1, ..., \bar{x}_n$ be the coordinates of a critical point, i. e. of a point satisfying (20) with parameters $\bar{\lambda}_i$. We assume that f and all ϕ_i are twice differentiable and that, at $\bar{x}$, the matrix

$$A = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \frac{\partial \phi_1}{\partial x_2} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \frac{\partial \phi_2}{\partial x_1} & \frac{\partial \phi_2}{\partial x_2} & \cdots & \frac{\partial \phi_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \phi_k}{\partial x_1} & \frac{\partial \phi_k}{\partial x_2} & \cdots & \frac{\partial \phi_k}{\partial x_n} \end{bmatrix} \tag{21}$$

has rank k. Without loss of generality the last condition implies that the first k columns of A constitute a matrix with a nonzero determinant. Utilizing the implicit function theorem, we solve the ϕ_i 's for the first k components of x as functions of $(x_{k+1}, ..., x_n)$. Let z denote f after substituting for $x_1, ..., x_k$ as functions of $x_{k+1}, ..., x_{n-k}$. It suffices to study the extrema of z. The second differential of z is:

$$d^2 z = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j + \sum_{i=1}^{k} \frac{\partial f}{\partial x_i} d^2 x_i. \tag{22}$$

But, taking the first and the second derivatives of $\phi_i = 0$, we must have:

$$\sum_{j=1}^{n} \frac{\partial \phi_r}{\partial x_j} dx_j = 0 \quad (r = 1, 2, ..., k), \tag{23}$$

$$\sum_{i,j=1}^{n} \frac{\partial^2 \phi_r}{\partial x_i \partial x_j} dx_i dx_j + \sum_{i=1}^{k} \frac{\partial \phi_r}{\partial x_i} d^2 x_i = 0 \quad (r = 1, 2, ..., k). \tag{24}$$

Multiply each of the equations (24) by the corresponding $\bar{\lambda}_r$, sum from $r = 1$ to $r = k$ and add to (22). We then get

$$d^2 z = \sum_{i,j=1}^{n} \left\{ \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{r=1}^{k} \bar{\lambda}_r \frac{\partial^2 \phi_r}{\partial x_i \partial x_j} \right\} dx_i dx_j + \sum_{i=1}^{k} \left\{ \frac{\partial f}{\partial x_i} + \sum_{i=1}^{k} \bar{\lambda}_r \frac{\partial \phi_r}{\partial x_i} \right\} d^2 x_i. \tag{25}$$

Since $\bar{x}$ is a critical point, (20) holds, the coefficients of d^2x_i in (25) vanish, and

$$d^2z = d^2V. \tag{26}$$

Let us introduce the following notation for the second partial derivatives of V and for the first derivatives of ϕ_r :

$$\frac{\partial^2 V}{\partial x_i \partial x_j} = a_{ij},\ (i,j = 1,2,...,n);\ \frac{\partial \phi_r}{\partial x_j} = \beta_{rj}\ (r = 1,2,...,k; j = 1,2,...,n), \tag{27}$$

and denote dx_i by h_i.

A sufficient test for an extremum of z is determined by the value of $d^2z = \sum_{i,j} a_{ij}h_ih_j$, where h_i are subject to:

$$\sum_{j=1}^{n} \beta_{rj}h_j = 0, \quad (r = 1,2,...,k).$$

The conditions for this were given in **7.**

10. In case the linear constraints are not homogeneous, we may use our result to obtain appropriate conditions. Let us consider the quadratic form:

$$U(x_1,...,x_n) = \sum_{i,j}^{n} a_{ij}x_ix_j \tag{28}$$

subject to the equations

$$\psi_i(x_1,x_2,...,x_n) = \sum_{j=1}^{n} b_{ij}x_j = e_i \quad (i = 1,2,...,k < n), \tag{29}$$

where $[b_{ij}]$ has rank k and where the determinant of its first k columns is non-zero.

We transform the variables in (28) and (29) according to:

$$x_j = \frac{y_j}{y_{n+1}} \quad (j = 1,2,...,n). \tag{30}$$

Then form (28) becomes

$$U'(y_1,y_2,...,y_{n+1}) = \frac{1}{y_{n+1}^2} \sum_{i,j}^{n} a_{ij}y_iy_j, \tag{31}$$

and form (29) becomes

$$\sum_{j=1}^{n} b_{ij}y_j - e_iy_{n+1} = 0 \quad (i = 1,2,...,k). \tag{32}$$

The problem now is to characterize the definiteness of (31) subject to (32) and the solution depends on the study of the signs of the principal diagonal minors, starting with the $(2k+1)$-st of:

$$\Delta' = (-1)^k \begin{vmatrix} 0 & 0 & \cdots & 0 & b_{11} & \cdots & b_{1n} & e_1 \\ 0 & 0 & \cdots & 0 & b_{21} & \cdots & b_{2n} & e_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & b_{k1} & \cdots & b_{kn} & e_k \\ b_{11} & b_{21} & \cdots & b_{k1} & a_{11} & \cdots & a_{1n} & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ b_{1n} & b_{2n} & \cdots & b_{kn} & a_{n1} & \cdots & a_{nn} & 0 \\ e_1 & e_2 & \cdots & e_k & 0 & \cdots & 0 & 0 \end{vmatrix}.$$

(I thank prof. M. El-Hodiri who sent me his translation from Russian of the paper of R. Ya. Shostak.)

References

R. E. BELLMAN (1970), Introduction to Matrix Analysis, McGraw-Hill, New York.

Y. CHABRILLAC and J.-P. CROUZEIX (1984), Definiteness and semidefiniteness of quadratic forms revisited, *Linear Algebra and Its Appl.,6*, 283-292.

G. DEBREU (1952), Definite and semidefinite quadratic forms, *Econometrica, 20,* 295-300.

R. W. FAREBROTHER (1977), Necessary and sufficient conditions for a quadratic form to be positive whenever a set of homogeneous linear constraints are satisfied, *Linear Algebra and Its Appl., 16,* 39-42.

G. GIORGI (2017), Various proofs of the Sylvester criterion for quadratic forms, *Journal of Mathematics Research, 9,* 55-66.

M. HANCOCK (1917), Theory of Maxima and Minima, Ginn, Boston. Reprinted by Dover, New York, 1960.

H. HOTELLING (1935), Demand functions with limited budgets, *Econometrica, 3,* 66-78.

KH. D. IBRAMOV and N. V. SAVELEVA (2000), Conditionally definite matrices, *Journal of Mathematical Sciences, 98,* 1-50.

H. B. MANN (1943), Quadratic forms with linear constraints, *American Math. Monthly, 50,* 430-433.

F. J. RICHELOT (1858), Bemerkungen zur Theorie der Maxima and Minima, *Astronomische Nachrichten, 48,* 273-286.

P. A. SAMUELSON (1947), Foundations of Economic Analysis, Harvard Univ. Press, Cambridge, Mass.

R. YA. SHOSTAK (1954), On a criterion of conditional definiteness of a quadratic form of n variables, subject to linear relations, and on a sufficient

condition for a conditional extremum of a function of n variables, *Uspekhi Mat. Nauk, 9,* 198-206.

H. VÄLIAHO (1982), On the definity of a quadratic form subject to linear constraints, *J. Optim. Theory Appl., 38,* 143-145.