

A Proof of the Sylvester Criterion for Quadratic Forms via Optimality Conditions for Quadratic Functions

Giorgio Giorgi

Correspondence: Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, Italy.

Received: January 17, 2022 Accepted: February 19, 2022 Online Published: March 3, 2022

doi:10.5539/jmr.v14n2p1

URL: <https://doi.org/10.5539/jmr.v14n2p1>

Abstract

We give a proof of the so-called Sylvester criterion for quadratic forms (for real symmetric matrices), based on elementary optimality properties of quadratic functions.

Keywords: Sylvester criterion, positive definite matrices, positive definite quadratic forms

AMS Subject Classification: 15A63

1. Introduction

Some time ago we have surveyed on this journal (Giorgi (2017)) several proofs of the so-called Sylvester criterion for quadratic forms. In the present paper we give a new proof of the said criterion, proof based on basic optimality properties of quadratic functions. A similar approach has been considered also by Hestenes (1966, 1975), however we think that our treatment is simpler and more suitable for didactic purposes. In our proof we shall make no reference to other properties of quadratic forms or symmetric matrices.

We consider, without loss of generality, a real symmetric matrix A of order n and $x \in \mathbb{R}^n$ as a column vector. We recall that the expression

$$Q(x) = x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

is a *quadratic form* associated to the matrix A . We recall that $Q(x)$ (or its associated symmetric matrix A) is

- *positive definite*, if $x^T A x > 0$ for all nonzero $x \in \mathbb{R}^n$;
- *negative definite*, if $x^T A x < 0$ for all nonzero $x \in \mathbb{R}^n$;
- *positive semidefinite*, if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$;
- *negative semidefinite*, if $x^T A x \leq 0$ for all $x \in \mathbb{R}^n$.

The reason for choosing a symmetric matrix is that, if A is not symmetric, then $\frac{1}{2}x^T(A + A^T)x = x^T A x$ for any $x \in \mathbb{R}^n$ and obviously $\frac{1}{2}(A + A^T)$ is symmetric.

For simplicity we shall obtain the Sylvester criterion only for definite quadratic forms. For the semidefinite case the reader may consult Chabrilac and Crouzeix (1984), Debreu (1952), Gantmacher (1959) and Takayama (1985).

We recall that the *k-th order leading principal minor* (or the *k-th order NW-principal minor* or the *k-th order successive principal minor*) of the square matrix A , of order n , not necessarily symmetric, denoted by Δ_k , $k = 1, \dots, n$, is the determinant of the square submatrix of A , of order k , consisting of the *first k rows* and the *first k columns* of A :

$$\Delta_1 = a_{11}, \Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \Delta_3 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \dots, \Delta_n = \det(A).$$

The Sylvester criterion for establishing the *sign* of $Q(x)$ (or of its associated symmetric matrix A) is the following one.

Theorem 1. Let be given the symmetric matrix A , of order n .

1. $Q(x)$ (or A) is positive definite if and only if all its *k-th order leading principal minors* Δ_k , $k = 1, \dots, n$, are positive.

2. $Q(x)$ (or A) is negative definite if and only if all its k -th order leading principal minors $\Delta_k, k = 1, \dots, n$, have the sign of $(-1)^k$, i. e.

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots, (-1)^n \Delta_n > 0.$$

2. The Main Results

Let us consider the following *quadratic function* $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$f(x) = x^T A x + 2x^T b + c, \tag{1}$$

where $x \in \mathbb{R}^n$, A a real and symmetric matrix of order n , with $\det(A) \neq 0$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. We have the following results, as a consequence of basic optimality properties of (1).

Theorem 2. The stationary point of (1) $x^0 = -A^{-1}b$ has the following properties.

1. x^0 is the unique stationary point of the quadratic function (1).
2. The following equality holds:

$$f(x^0) = \frac{\det(B)}{\det(A)},$$

where

$$B = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \\ b_1 & \cdots & b_n & c \end{bmatrix}.$$

3. The following equality holds:

$$f(x) = f(x^0) + (x - x^0)^T A (x - x^0), \quad \forall x \in \mathbb{R}^n. \tag{2}$$

Proof.

1. As $\nabla f(x) = 2Ax + 2b$, $x \in \mathbb{R}^n$, it holds $\nabla f(x) = 0$ if and only if $x = x^0 = -A^{-1}b$.
2. We remark that it holds

$$f(x^0) = (x^0)^T (Ax^0 + b) + (x^0)^T b + c = (x^0)^T b + c.$$

On the grounds of this relation and by a well-known property of the determinants, we have

$$\begin{aligned} \det(B) &= \begin{vmatrix} a_{11} & \cdots & a_{1n} & \sum_{j=1}^n a_{1j}x_j^0 + b_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & \sum_{j=1}^n a_{nj}x_j^0 + b_n \\ b_1 & \cdots & b_n & (x^0)^T b + c \end{vmatrix} = \\ &= \begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 \\ b_1 & \cdots & b_n & f(x^0) \end{vmatrix} = f(x^0) \det(A), \end{aligned}$$

from which we get

$$f(x^0) = \frac{\det(B)}{\det(A)}.$$

3. Being $\nabla^2 f(x) = 2A$, $x \in \mathbb{R}^n$, by Taylor's expansion formula we obtain relation (2). □

Theorem 3. Let be given the real symmetric matrix A , of order n and with $\det(A) \neq 0$. Then, the following conditions are equivalent.

- i) A is positive definite.
- ii) For every $b \in \mathbb{R}^n$ and every $c \in \mathbb{R}$, the point $x^0 = -A^{-1}b$ is a strict minimizer of f , defined by relation (1).
- iii) For every $b \in \mathbb{R}^n$ and every $c \in \mathbb{R}$, the point $x^0 = -A^{-1}b$ is a minimizer of f , defined by relation (1).

iv) A is positive semidefinite.

Proof. The implications $i) \implies ii)$ and $iii) \implies iv)$ result from relation (2). The implication $ii) \implies iii)$ is trivial. It remains to prove the implication $iv) \implies i)$. This implication is well known in the theory of quadratic forms and says, in other words, that the class of semidefinite quadratic forms, with $\det(A) \neq 0$, which are not definite, is empty. Let A be positive semidefinite and suppose that there exists a vector $y \in \mathbb{R}^n \setminus \{0\}$ such that $y^T A y = 0$. The vector y is therefore a local minimum point of the function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $g(x) = x^T A x$, $x \in \mathbb{R}^n$. We have therefore $\nabla g(y) = 0$, i. e.

$$2 \sum_{j=1}^n a_{ij} y_j = 0, \quad \forall i = 1, \dots, n.$$

Being $y \neq 0$, it results $\det(A) = 0$, in contradiction with the assumptions. Therefore A is positive definite. □

We are now ready to prove Theorem 1. We begin by proving the point 1.

Proof of Theorem 1, point 1.

a) *Necessity.* Let be $j \in \{1, \dots, n\}$. As A is positive definite, it is easy to see that also the submatrix

$$A_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & \dots & \vdots \\ a_{j1} & \cdots & a_{jj} \end{bmatrix}$$

is positive definite: being

$$Q(x_1, \dots, x_j) = Q(x_1, \dots, x_n),$$

with $(x_1, \dots, x_n) = (x_1, \dots, x_j, 0, \dots, 0)$, it results that $Q(x_1, \dots, x_j)$ is definite positive. We want to show that $\Delta_j > 0$, $j = 1, \dots, n$. Let be $t \in [0, 1]$. The matrix $(1-t)I + tA_j$, with I identity matrix of order j , is positive definite, as, for every $x \in \mathbb{R}^n$ we have

$$x^T [(1-t)I + tA_j] x = (1-t) \|x\|^2 + tx^T A_j x.$$

We put $P(t) = \det[(1-t)I + tA_j]$ and observe that $P(t) \neq 0$: indeed, if $P(t) = 0$, then the following system, defined on \mathbb{R}^j ,

$$[(1-t)I + tA_j] x = 0$$

admits solutions $y \neq 0$. For these solutions we have

$$y^T [(1-t)I + tA_j] y = 0,$$

which contradicts the fact that $[(1-t)I + tA_j]$ is positive definite. Being t an arbitrary number of the interval $[0, 1]$, it holds $P(t) \neq 0, \forall t \in [0, 1]$. In particular $P(0) = \det(I) > 0$ and, being $P(t)$ a continuous function on the interval $[0, 1]$, it will be $P(t) > 0, \forall t \in [0, 1]$. In particular, we have $P(1) > 0$, i. e. $\det(A_j) > 0, j = 1, \dots, n$. □

b) *Sufficiency.* This part of the proof will be performed by induction. For $n = 1$ the assertion is satisfied. Let us suppose that the assertion is true for $n = k$; we prove that the same assertion is true for $n = k + 1$.

Let us consider the real symmetric matrix $B = [a_{ij}]$, of order $(k + 1)$, such that it holds

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & \dots & \vdots \\ a_{j1} & \cdots & a_{jj} \end{pmatrix} > 0, \quad j \in \{1, \dots, k + 1\}.$$

Let us put

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \dots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix},$$

$$b = [a_{1,k+1}, \dots, a_{k,k+1}]^T, \quad c = a_{k+1,k+1}.$$

By the induction assumption A is positive definite. By Theorem 3 the point $x^0 = -A^{-1}b$ is a global minimum point of the function $f : \mathbb{R}^k \rightarrow \mathbb{R}$ defined by (1). Taking Theorem 2, point 2, into account, we have

$$f(x) \geq f(x^0) = \frac{\det(B)}{\det(A)} > 0, \quad \forall x \in \mathbb{R}^k. \tag{3}$$

Let be $y = [y_1, \dots, y_{k+1}]^T \in \mathbb{R}^{k+1} \setminus \{0\}$ and put $z = [y_1, \dots, y_k]^T$. If $y_{k+1} = 0$, then we have $z \neq 0$ and hence $y^T B y = z^T A z > 0$. If $y_{k+1} \neq 0$, then the point

$$\bar{z} = \frac{1}{y_{k+1}} z$$

satisfies, by relation (3), the following relation

$$y^T B y = z^T A z + 2y_{k+1} z^T b + c(y_{k+1})^2 = (y_{k+1})^2 f(\bar{z}) > 0.$$

Therefore $y^T B y > 0, \forall y \in \mathbb{R}^{k+1} \setminus \{0\}$, i. e. B is positive definite. □

The proof of Theorem 1, part 2, is very easy: the symmetric matrix A is negative definite if and only if $-A$ is positive definite and the k -th leading principal minor of $-A$ is $(-1)^k$ times the corresponding leading principal minor of A .

Remark 1. The necessity part of Theorem 1 can be proved in a more direct way if we make reference to (well known!) properties of quadratic forms. For example, if we recall that A (symmetric) is definite positive if and only if all its (real!) eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are positive and that

$$\det(A) = \lambda_1 \lambda_2 \dots \lambda_n > 0,$$

we have at once that $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$.

Another classical result establishes that the symmetric matrix A is definite positive if and only if there exists a non-singular matrix P , such that

$$A = P^T P.$$

Then we have $\det(A) = \det(P^T P) = [\det(P)]^2 > 0$.

Remark 2. Also the equivalence $iv) \iff i)$ of Theorem 3 is immediate if we make reference to the characterization of the sign of $Q(x)$ by means of eigenvalues of A : if A is positive semidefinite, without being positive definite, one at least of its eigenvalues is zero, but then $\det(A) = 0$, in contradiction with the assumptions of Theorem 3.

Remark 3. It is well known that the examination of the sign of the leading principal minors is *not* sufficient to check if a symmetric matrix is positive (negative) semidefinite. For example the matrix

$$\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$$

has nonnegative leading principal minors, but it is not positive semidefinite. Indeed, it is negative semidefinite. See, e. g., Chabrilac and Crouzeix (1984), Debreu (1952).

For other considerations on the relations between quadratic forms and quadratic functions, the reader may consult Han and Mangasarian (1984).

References

Bhatia, R. (1997). *Matrix Analysis*, New York, Springer-Verlag. <https://doi.org/10.1007/978-1-4612-0653-8>

Bhatia, R. (2007). *Positive Definite Matrices*. Princeton, N. J., Princeton University Press.

Chabrilac, Y., & Crouzeix, J. P. (1984). Definiteness and semidefiniteness of quadratic forms revisited. *Linear Algebra and Its Applications*, 63, 283-292. [https://doi.org/10.1016/0024-3795\(84\)90150-2](https://doi.org/10.1016/0024-3795(84)90150-2)

Debreu, G. (1952). Definite and semidefinite quadratic forms. *Econometrica*, 20, 295-300. <https://doi.org/10.2307/1907852>

Gantmacher, F. R. (1959). *Matrix Theory, Volume I*. New York, Chelsea.

Giorgi, G. (2017). Various proofs of the Sylvester criterion for quadratic forms. *Journal of Mathematics Research*, 9, N. 6, 55-66. <https://doi.org/10.5539/jmr.v9n6p55>

- Han, S. P., & Mangasarian, O. L. (1984). Characterization of positive definite and semidefinite matrices via quadratic programming duality. *SIAM J. Alg. Disc. Meth.*, 5, 26-32. [https://doi.org/10.1016/0024-3795\(84\)90116-2](https://doi.org/10.1016/0024-3795(84)90116-2)
- Hestenes, M. R. (1966). *Calculus of Variations and Optimal Control Theory*. New York, John Wiley & Sons.
- Hestenes, M. R. (1975). *Optimization Theory. The Finite Dimensional Case*. New York, John Wiley & Sons. Reprinted by R. E. Krieger, Huntington, New York, 1981.
- Rahma, A. A., Burqan, A., & Özer, O. (2021). New results for arithmetic-geometric mean inequality and singular values of matrices. *WSEAS Transactions on Mathematics*, 20, 625-637. <https://doi.org/10.37394/23206.2021.20.66>
- Takayama, A. (1985). *Mathematical Economics*. Cambridge, Cambridge University Press.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).