Sylvester Criterion for Positive Definiteness

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

We shall consider $\mathbb{R}^n$ as the vector space of column vectors, that is, matrices of type $n \times 1$. The standard inner product or the dot product of two vectors $x, y \in \mathbb{R}^n$ is given by

$$\langle x, y \rangle = x \cdot y = y^t x,$$

where the 1×1 matrix is identified as a real number. Given an $n \times n$ matrix A, we have a linear map on $\mathbb{R}^n$ given by $x \mapsto Ax$. In the sequel, we shall not distinguish between the matrix A and the associated linear map. I am sure that the context will make it clear what we are referring to.

A quadratic form $q \colon \mathbb{R}^n \to \mathbb{R}$ is said to be positive definite iff $q(v) > 0$ for any nonzero $v \in \mathbb{R}^n$. We say that an $n \times n$ real symmetric matrix A is positive definite if the associated quadratic form $q \colon x \mapsto x^t Ax$ is positive definite.

Let us first look at lower dimensions to gain some insight. When $n = 1$, any quadratic from on $\mathbb{R}$ is of the form $q(x) = ax^2$. This is positive definite iff $a > 0$. Now, consider a form in two variables:

$$q(x_1, x_2) := a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

(We chose to represent the coordinates of vectors in $\mathbb{R}^2$ by x_1, x_2, in stead of x, y, which are easier to type and write, so that we can perceive how the higher dimensional case will go!) Assume that this is positive definite. Then for all vectors $(x_1, 0)$ with $x_1 \neq 0$, we must have $a_{11}x_1^2 > 0$. Hence we conclude that $a_{11} > 0$. We can rewrite the form as follows:

$$q(x_1, x_2) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2. \tag{1}$$

We choose a vector so that $x_1 + \frac{a_{12}}{a_{11}}x_2 = 0$ with $x_2 \neq 0$. It follows from (1) that $(a_{22} - \frac{a_{12}^2}{a_{11}}) > 0$. This is the same as saying that $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0$.

Let us now look at $n = 3$. Let the quadratic form be given by $q(x_1, x_2, x_3) = \sum_{i,j=1}^3 a_{ij}x_ix_j$. If this is positive definite, by taking vectors with $x_2 = x_3 = 0$, we see that $a_{11} > 0$. Hence we rewrite the quadratic form as follows:

$$\begin{aligned} q(x) &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}}\right)x_3^2 \\ &\quad + 2\left(a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right)x_2x_3. \end{aligned} \tag{2}$$

As analyzed earlier, we see that q is positive definite iff $a_{11} > 0$ and the quadratic form in the variables x_2, x_3 is positive definite. The latter entails in the conditions

$$a_{22} - \frac{a_{12}^2}{a_{11}} > 0 \text{ and } \det \begin{pmatrix} a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{pmatrix} > 0.$$

The second condition may be understood if we compute the determinant of $A = (a_{ij})$, suing an elementary operation, as follows:

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ 0 & a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{pmatrix}.$$

The above can be put in a more tractable form a follows. Let

$$y = x_1 + (a_{12}/a_{11})x_2 + (a_{13}/a_{11})x_3 \text{ and } z = y - x_1.$$

Then $q(x) = a_{11}y^2 - a_{11}z^2 + \sum_{i,j=2}^{3} a_{ij}x_ix_j$. Now it is clear how the general case will look like. Given $q(x) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$, we let

$$y = x_1 + (a_{12}/a_{11})x_2 + \cdots + (a_{1n}/a_{11})x_n \text{ and } z = y - x_1.$$

We check that $q(x) = a_{11}y^2 - a_{11}z^2 + \sum_{i,j=2}^{n} a_{ij}x_ix_j$. This suggests how to proceed by induction. We now define a quadratic form that depends on $n-1$ variables, namely, $x_2, \ldots, x_n$: $q'(x') = \sum_{i,j=2}^{n} a_{ij}x_ix_j - a_{11}z^2$. If we set $b_{ij} := a_{ij} - (a_{i1}a_{1j})/a_{11}$, we find that

$$q'(x,) = \sum_{i,j=2}^{n} b_{ij}x_ix_j.$$

The relation between determinants the symmetric matrix A of the quadratic form q and that of q' is given by

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & b_{22} & b_{23} & \ldots & b_{2n} \\ 0 & b_{23} & b_{33} & \ldots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & b_{n3} & \ldots & b_{nn} \end{pmatrix}.$$

Note that the matrix, say, B on the right side of the equation is obtained by an obvious elementary operation. If q is positive definite, then $a_{11} > 0$. Also, if we denote by $M_k(X)$ the k-th principal minor of a square matrix X, then $M_k(A) = M_k(B)$. It follows by induction that q' is positive definite and hence q is. This completes the classical proof of the criterion for the positive definiteness of a real symmetric matrix. Note that the proof carries through in the case of hermitian matrices also, with obvious modifications such as x_j^2 replaced by $z_j\overline{z}_j$ etc.

We now indicate a more conceptual and less computational proof which uses basic concepts from linear algebra.

Lemma 1. *A real symmetric matrix A is positive definite iff all its eigenvalues are positive.*

Proof. Let A be a real symmetric matrix of order n. If A is positive definite and λ is an eigenvalue of A with a unit eigenvector $x \in \mathbb{R}^n$, then $0 < x^t Ax = Ax \cdot x = \lambda x \cdot x = \lambda$.

Conversely, if A is symmetric with all its eigenvalues positive, by diagonalization theorem, there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors. Assume that $\{v_i : 1 \leq i \leq n\}$ be such a basis with $Av_i = \lambda_i v_i$. Then any $x \in \mathbb{R}^n$ can be written as $x = \sum_{i=1}^n x_i v_i$ where $x_i = x \cdot v_i$. We compute

$$Ax \cdot x = A\left(\sum_{i=1}^n x_i v_i\right)\left(\sum_{j=1}^n x_j v_j\right) = \sum_{i,j=1}^n \lambda_j x_i x_j v_i \cdot v_j = \sum_{i=1}^n \lambda_i x_i^2 > 0$$

if $x \neq 0$. Thus A is positive definite. □

Lemma 2. *Let $v_1, \ldots, v_n$ be a basis of a vector space V. Suppose that W is a vector subspace. If $\dim W > m$, then*

$$W \cap \operatorname{span}\{v_{m+1}, \ldots, v_n\} \neq (0).$$

Proof. Recall that if W_j, $j = 1, 2$, are vector subspaces, then $\dim(W_1 \cap W_2) = \dim W_1 + \dim W_2 - \dim(W_1 + W_2)$. Now, if $\dim W > m$, then

$$\begin{aligned} &\dim(W \cap \operatorname{span}\{v_{m+1}, \ldots, v_n\}) \\ &= \dim W + \dim(\operatorname{span}\{v_{m+1}, \ldots, v_n\}) - \dim(W + \operatorname{span}\{v_{m+1}, \ldots, v_n\}) \\ &> m + (n - m) - n = 0. \end{aligned}$$

The result follows. □

Lemma 3. *Let A be an $n \times n$ real symmetric matrix. If $\langle Aw, w\rangle > 0$ for all $w \in W$, then A has at least $\dim W$ positive eigenvalues (counted with multiplicity).*

Proof. Let $\dim W = r$. Let $\{v_k : 1 \leq k \leq n\}$ be an orthonormal eigen-basis of A on $\mathbb{R}^n$ such that $Av_k = \lambda_k v_k$ for all k. Let us assume, without loss of generality, that $\lambda_k > 0$ for $1 \leq k \leq m$ and that $\lambda_k \leq 0$ for $k > m$. If $m < \dim W$, then, by Lemma 2, there is a nonzero vector $v \in W$ such that $w = a_{m+1}v_{m+1} + \cdots + a_n v_n$. We compute

$$\langle Aw, w\rangle = \sum_{j,k=m+1}^n a_j a_k \langle Av_j, v_k\rangle = a_{m+1}^2 \lambda_{m+1} + \cdots + a_n^2 \lambda_n \leq 0,$$

a contradiction. Hence $m \geq \dim W$, as required. □

Definition 4. Let $A := (a_{ij})$ be an $n \times n$ matrix. Then the matrix $(a_{ij})_{1 \leq i,j \leq k}$ is called the k-th *principal submatrix* and determinant is known as the k-th *principal minor.*

Theorem 5 (Sylvester)**.** *A real symmetric $n \times n$ matrix is positive definite iff all its principal minors are positive.*

Proof. Let A be be a real positive definite $n \times n$ symmetric matrix. Since the eigenvalues of A are positive, it follows that $\det A$, being the product of the eigenvalues must be positive. Now the restriction A_k of A to the k dimensional vector subspace $\mathbb{R}^k := \{x \in \mathbb{R}^n : x_j = 0 \text{ for } j > k\}$

is also positive definite. Clearly the matrix of A_k is the k-th principal matrix and hence its determinant must be positive by the argument above.

Let A be be a real $n \times n$ symmetric matrix all of whose principal minors are positive. We prove, by induction that A is positive definite by showing that all its eigenvalues are positive. For $n = 1$, the result is trivial. Assume the sufficiency of positive principal minors for $(n-1) \times (n-1)$ real symmetric matrices. If A is an $n \times n$ real symmetric matrix, then its $(n-1)$-th principal submatrix is positive definite by induction. Let $W = \mathbb{R}^{n-1} \subset \mathbb{R}^n$ be the subspace whose last coordinate is 0. Then for any nonzero $w \in W$, we observe that$\langle Aw, w \rangle = \langle A_{n-1}x', x' \rangle$ where $x = (x', 0) \in \mathbb{R}^n$ and $x' \in \mathbb{R}^{n-1}$. Since A_{n-1} is positive definite by induction, we see that $\langle A_{n-1}x', x' \rangle > 0$ for $x' \in \mathbb{R}^{n-1}$. Hence $\langle Ax, x \rangle > 0$ for $x \in W$. By Lemma 2, A has at least $(n-1)$ positive eigenvalues. Now $\det A$ is the product of the eigenvalues of A and $(n-1)$ of these eigenvalues are positive. Hence, it follows that all the eigenvalues of A are positive. Hence A is positive definite. $\square$