Isometries of $\mathbb{R}^n$ and Sylvester Criterion for Positive Definite Matrices

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

1 Introduction

This article is based on a lecture given in a Workshop for College Teachers, organized by Bombay Mathematics Colloquium and Bhavan's College on September 15, 2004. I was requested to speak on two topics: one is the isometries of $\mathbb{R}^3$ and the other is the criterion for a symmetric matrix to be positive definite. I first review some basic facts on isometries of $\mathbb{R}^n$ and $\mathbb{R}^2$ and then end up with the study of isometries of $\mathbb{R}^3$. Readers with a good background can go directly to the Subsection 2.1.

I thank Professor Dhvanita Rao for the invitation and the audience for an enthusiastic response.

2 Isometries of $\mathbb{R}^n$

Let (X, d) and (y, d) be metric spaces. A map $f : X \to Y$ is said to be an isometry if $d(f(x_1), f(x_2)) = d(x_1, x_2)$ for all $x_1, x_2 \in X$. Note that an isometry is always one-one but in general not onto. For instance, consider $f : [1, \infty) \to [1, \infty)$ given by $f(x) = x + 1$. If f and g are isometries of X to itself, then $g \circ f$ is also an isometry. The set of surjective isometries of a metric space form a group under the composition.

We consider $\mathbb{R}^n$ with the Euclidean inner product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$. We then have the notion of norm or length of a vector $\| x \| := \sqrt{\langle x, x\rangle}$. It is well-known that $d(x, y) := \| x - y \|$ defines a metric on X. Thus $(\mathbb{R}^n, d)$ becomes a metric space. The aim of this article is to give a complete description of all isometries of $\mathbb{R}^n$ and look a little more geometrically into the isometries of $\mathbb{R}^3$.

First a bit of convention: We consider $\mathbb{R}^n$ as the vector space of column vectors, that is, $n \times 1$ real matrices. 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.

Let us first of all look at some examples of isometries. For a fixed $v \in \mathbb{R}^n$, consider the translation $T_v\colon \mathbb{R}^n \to \mathbb{R}^n$ given by $T_v(x) := x + v$. Then T_v is an isometry of $\mathbb{R}^n$:

$$d(T_v x, T_v y) = \| T_v x - T_v y \| = \| (x + v) - (y + v) \| = \| x - y \| = d(x, y).$$

We now describe another important class of isometries of $\mathbb{R}^n$. A linear map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is said to be orthogonal if it preserves inner products: for every pair $x, y \in \mathbb{R}^n$, we have

$$\langle f(x), f(y) \rangle = \langle x, y \rangle .$$

The following result is well-known. For a proof, I refer the reader to my book on linear algebra.

Theorem 1. *For a linear map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ the following are equivalent:*

1. *f is orthogonal.*
2. *$\| f(x) \| = \| x \|$ for all $x \in \mathbb{R}^n$.*
3. *f takes an orthonormal basis of $\mathbb{R}^n$ to an orthonormal basis.* □

We now look at a very special class of orthogonal maps.

Fix a unit vector $u \in \mathbb{R}^n$. Let $W := (\mathbb{R}u)^\perp := \{x \in \mathbb{R}^n : \langle x, u \rangle = 0\}$. Then W is a vector subspace of dimension $n - 1$. This can be seen as follows. The map $f_u\colon \mathbb{R}^n \to \mathbb{R}$ given by $f_u(x) := \langle x, u \rangle$ is linear. It is nonzero, since $f_u(u) = 1$. Hence the image of f_u is a nonzero vector subspace and hence all of $\mathbb{R}$. Also, we observe that $W = \ker f_u$. Hence by the rank-nullity theorem,

$$n = \dim \mathbb{R}^n = \dim \ker f_u + \dim \operatorname{Im} f_u.$$

It follows that W is an $n - 1$ dimensional vector subspace of $\mathbb{R}^n$. We thus have an orthogonal decomposition $\mathbb{R}^n = W \oplus \mathbb{R}u$. We use W as a mirror and reflect across it. Thus any vector in W is mapped to itself whereas the vector u is mapped to $-u$. Thus the reflection $\rho_W\colon \mathbb{R}^n \to \mathbb{R}^n$ is given by $\rho(x) = w - tu$, where $x = w + tu$, $t \in \mathbb{R}$. This map is clearly an orthogonal linear map. For, if $\{w_1, \ldots, w_{n-1}\}$ is an orthonormal basis of W, then $\{w_1, \ldots, w_{n-1}, u\}$ is an orthonormal basis of $\mathbb{R}^n$. The linear map ρ_W carries this orthonormal basis to the orthonormal basis $\{w_1, \ldots, w_{n-1}, -u\}$ and hence is orthogonal. We have another description of this map as follows:

$$\rho_W(x) := x - 2 \langle x, u \rangle u. \tag{1}$$

The advantage of this expression is that it is basis-free. We also observe that the expression remains the same if we replace u by $-u$.

We shall look at the case when $n = 2$ in detail. We shall derive a matrix representation of ρ_W from (1). Consider the one dimensional subspace $\mathbb{R}\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$. Let $u = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$. Then

$$\begin{aligned}
\rho_W(e_1) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2 \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \\
&= \begin{pmatrix} 1 - 2\sin^2 t \\ 2 \sin t \cos t \end{pmatrix} = \begin{pmatrix} \cos 2t \\ \sin 2t \end{pmatrix} \\
\rho_W(e_2) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 2 \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \\
&= \begin{pmatrix} \sin 2t \\ -\cos 2t \end{pmatrix}.
\end{aligned}$$

Thus the reflection about the line is given by $\rho_W = \begin{pmatrix} \cos 2t & \sin 2t \\ \sin 2t & -\cos 2t \end{pmatrix}$.

Remark 2. Reflections generate the group of orthogonal linear maps. (We shall not prove this.) Thus, they are the building blocks of orthogonal linear maps. If we observe that $\rho^2 = 1$ for any reflection, then the analogy between the transpositions in a symmetric group and reflections is striking.

We now give a complete list of all orthogonal maps of $\mathbb{R}^2$. Let $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ be an orthogonal. If we represent A with respect to the standard orthonormal basis $\{e_1, e_2\}$ as a matrix, then A is either of the form

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \text{ or of the form } \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}.$$

In the first case, we say that it is a rotation by an angle t in the anticlockwise direction. In the latter case, it is reflection with respect to the line $\mathbb{R}\begin{pmatrix} \cos(t/2) \\ \sin(t/2) \end{pmatrix}$, as seen earlier.

Thus we have understood all the orthogonal maps of $\mathbb{R}^2$. We wish to do the same in $\mathbb{R}^3$. See Subsection 2.1.

A surprising result (Theorem 5) is that any isometry of $\mathbb{R}^n$ that fixes the zero vector must be linear and orthogonal. A proof of this can also be found in my book. However, I indicate a slightly different proof.

Lemma 3. *Let $x, y \in \mathbb{R}$. Assume that $\langle x, x\rangle = \langle x, y\rangle = \langle y, y\rangle$. Then $x = y$.*

Proof. We compute the length square of $x - y$:

$$\langle x - y, x - y\rangle = \langle x, x\rangle - 2\langle x, y\rangle + \langle y, y\rangle = 0,$$

in view of our hypothesis. Hence $x = y$. □

Lemma 4. *An isometry of $\mathbb{R}^n$ that fixes the origin preserves the inner products.*

Proof. Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be an isometry such that $f(0) = 0$. We need to prove that $\langle f(x), f(y)\rangle = \langle x, y\rangle$ for all $x, y \in \mathbb{R}^n$. Since $d(f(x), f(y)) = d(x, y)$, we have $\| f(x) - f(y) \| = \| x - y \|$. Taking squares, we get

$$\langle f(x) - f(y), f(x) - f(y)\rangle = \langle x - y, x - y\rangle \text{, for all } x, y \in \mathbb{R}^n. \tag{2}$$

Since $f(0) = 0$, setting $y = 0$ in (2), we get $\langle f(x), f(x)\rangle = \langle x, x\rangle$. Similarly, $\langle f(y), f(y)\rangle = \langle y, y\rangle$. Now, if we expand both sides of (2) and cancelling equal terms such as $\langle x, x\rangle$ and $\langle y, y\rangle$, we get the desired result. □

Theorem 5. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be an isometry with $f(0) = 0$. Then f is an orthogonal linear map.* □

Proof. In view of Lemma 4, we need only show that f is linear. Let $x, y \in \mathbb{R}^n$. Let $z = x+y$. We shall show first that $f(z) = f(x) + f(y)$. In view of Lemma 3, it suffices to show that

$$\langle f(z), f(z)\rangle = \langle f(z), f(x) + f(y)\rangle = \langle f(x) + f(y), f(x) + f(y)\rangle .$$

Expanding the last two terms, we need to show that

$$\langle f(z), f(z)\rangle = \langle f(z), f(x)\rangle + \langle f(z), f(y)\rangle = \langle f(x), f(x)\rangle + 2\langle f(x), f(y)\rangle + \langle f(y), f(y)\rangle . \quad (3)$$

Since f preserves inner products, (3) holds same iff

$$\langle z, z\rangle = \langle z, x\rangle + \langle z, y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle , \quad (4)$$

holds. Since $z = x + y$, (4) is clearly true!

Similarly, if we take $y = ax$, $a \in \mathbb{R}$, then we need to show that $f(y) = af(x)$. In view of Lemma 3, it is enough to show that

$$\langle f(ax), f(ax)\rangle = \langle f(ax), af(x)\rangle = \langle af(x), af(x)\rangle .$$

Since f is inner product preserving, it suffices to show that

$$\langle ax, ax\rangle = a\langle ax, x\rangle = a^2\langle x, x\rangle ,$$

which is true. □

As immediate consequences of Theorem 5, we have

Corollary 6. *(i) Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be an isometry. Then $f = T_v \circ A$ where $v = f(0)$.*
(ii) Any isometry of $\mathbb{R}^n$ is onto.

Proof. Consider $g := T_{-v} \circ f$ where $v = f(0)$. Then clearly, g is an isometry of $\mathbb{R}^n$ such that $g(0) = 0$. Hence g is an orthogonal linear map, say, A. Hence $f = T_v \circ A$. Thus (i) is proved.

Since any translation and any orthogonal linear map are onto, so is an isometry. □

2.1 Isometries of $\mathbb{R}^3$

Definition 7. Given a two dimensional vector subspace W of $\mathbb{R}^3$, we define a rotation in the plane W as follows. Fix an orthonormal basis $\{w_1, w_2\}$ of $\mathbb{R}^3$. Fix a unit vector u such that $u \perp W$. We consider the linear map given by

$$\begin{aligned} R_{W,t}(w_1) &= \cos t\, w_1 + \sin t\, w_2 \\ R_{w,t}(w_2) &= -\sin t\, w_1 + \cos t\, w_2 \\ R_{w,t}(u) &= u. \end{aligned}$$

Clearly, $R_{W,t}$ is orthogonal. Also, with respect to the orthonormal basis $\{w_1, w_2, u\}$, it is represented by

$$\begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

We shall refer to $R_{W,t}$ as the rotation in the plane W with $\mathbb{R}u$ as the axis of rotation. Note that $\det(R_{W,t}) = 1$.

Proposition 8. (i) *Let A be a 3×3 orthogonal matrix. Assume that* $\det A = 1$. *Then* 1 *is an eigenvalue of A.*
(ii) *Let A be a 3×3 orthogonal matrix. Assume that* $\det A = 1$. *Then A is a rotation.*

Proof. (i). Consider the following chain of equations:

$$\begin{aligned} \det(A - I) &= \det(A - AA^t) = \det(A(I - A^t)) \\ &= \det(A)\det(I - A^t) = \det((I - A^t)^t) = \det(I - A). \end{aligned}$$

For any $n \times n$ matrix B, we have $\det(-B) = (-1)^n \det B$. Hence, we see that $\det(A - I) = \det(I - A) = -\det(A - I)$. We conclude that $\det(A - I) = 0$. From this, (i) follows.

We now prove (ii). By (i), there exists a unit vector $u \in \mathbb{R}^3$ such that $Au = u$. Let $W := (\mathbb{R}u)^\perp$. Then W is a two dimensional vector subspace. We claim that $Aw \in W$ for any $w \in W$. This is seen as follows:

$$\langle Aw, u \rangle = \langle Aw, Au \rangle = \langle w, u \rangle = 0.$$

Thus $A\,|_W : W \to W$ is an orthogonal linear map. If we choose an orthonormal basis $\{w_1, w_2\}$ of W, then A can be represented with respect to the orthonormal basis $\{w_1, w_2, u\}$ either as

$$\begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ or as } \begin{pmatrix} \cos t & \sin t & 0 \\ \sin t & -\cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The second case cannot occur as the determinant is -1, contradicting our hypothesis. Hence (ii) follows. $\square$

Theorem 9. *Let A be any 3×3 orthogonal matrix. Then either A is a rotation or is a rotation followed by a reflection.*

Proof. If $\det A = 1$, then we know that it is a rotation. So we assume that $\det A = -1$. Then consider the reflection matrix corresponding to the xy plane: $\rho = \operatorname{diag}(1\ 1\ -1)$. Then $B = \rho A$ has determinant 1 and hence is a rotation. Since $A = \rho B$, the theorem follows. $\square$

Remark 10. Let A be an orthogonal 3×3 matrix with $\det A = 1$. How can we decide whether A is a pure reflection or is a rotation followed by a reflection? If it is a pure reflection, then its eigenvalues are $+1, +1, -1$. In the other case, either all eigenvalues are -1 or it has only one real eigenvalue, namely -1.

3 Sylvester Criterion for Positive Definiteness

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:

$$q(x_1, x_2) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2. \tag{5}$$

We choose a vector so that $x_1 + \frac{a_{12}}{a_{11}}x_2 = 0$ with $x_2 \neq 0$. It follows from (5) 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:

$$\begin{aligned} q(x) &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}}\right)x_3^2 \\ &\quad + 2\left(a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right)x_2x_3. \end{aligned} \tag{6}$$

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 11. *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 $0 < x^tAx = Ax \cdot x = \lambda x \cdot x = \lambda$.

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 $x = \sum_{i=1}^{n} x_iv_i$ where $x_i = x \cdot v_i$. We compute

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

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

Lemma 12. *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*

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

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

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

The result follows. □

Lemma 13. *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 $\{v_k : 1 \le k \le n\}$ 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 $1 \le k \le m$ and that $\lambda_k \le 0$ for $k > m$. If $m < \dim W$, then, by Lemma 12, there is a nonzero vector $v \in W$ such that $w = a_{m+1}v_{m+1} + \cdots + a_n v_n$. We compute

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

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

Definition 14. Let $A := (a_{ij})$ be an $n \times n$ matrix. Then the matrix $(a_{ij})_{1 \le i,j \le k}$ is called the k-th *principal submatrix* and determinant is known as the k-th *principal minor*.

Theorem 15 (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 12, 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. □

Reference

S. Kumaresan, *Linear Algebra—A Geometric Approach*, Prentice-Hall of India, 2000.