POSITIVE DEFINITE REAL SYMMETRIC MATRICES

K. N. RAGHAVAN
FOR IST AT IITGN, JULY 2017

An $n \times n$ real symmetric matrix A is said to be *positive definite* if, for every $\underline{v} \in \mathbb{R}^n$, we have $\underline{v}^t A \underline{v} \geq 0$ and equality holds only if $\underline{v} = 0$. Fix notation as follows:

- RSM = real symmetric matrix
- $\mathcal{S}_n :=$ the set of all $n \times n$ real symmetric matrices
- $\mathcal{P}_n :=$ the set of all $n \times n$ real symmetric positive definite matrices
- $GL_n :=$ the set of all $n \times n$ real invertible matrices
- $\mathcal{O}_n :=$ the set of all $n \times n$ real orthogonal matrices
- $\mathcal{D}_n :=$ the set of all $n \times n$ real diagonal matrices
- $\mathcal{D}_n^+ :=$ the set of all elements of $\mathcal{D}_n$ with positive diagonal entries
- $\mathcal{U}_n :=$ the set of all $n \times n$ real upper triangular matrices with positive diagonal entries
- $\mathcal{L}_n :=$ the set of all $n \times n$ real lower triangular matrices with positive diagonal entries

Observe the following:

- GL_n, $\mathcal{D}_n^+$, $\mathcal{U}_n$, and $\mathcal{L}_n$ are all groups with respect to the usual matrix multiplication.
- The transpose map sets up an anti-isomorphism between the groups $\mathcal{U}_n$ and $\mathcal{L}_n$.
- $\mathcal{U}_n \cap \mathcal{L}_n = \mathcal{D}_n^+$.
- The only element of $\mathcal{D}_n^+$ that equals is its own inverse is the identity.

1. Spectral theorem for RSMs and positive definiteness

Here are two simple observations:

(1) $\mathcal{D}_n \cap \mathcal{P}_n = \mathcal{D}_n^+$.
(2) There is an action of the group GL_n on $\mathcal{S}_n$, the action map $GL_n \times \mathcal{S}_n \to \mathcal{S}_n$ being given by $(g, A) \mapsto gAg^t$. The subset $\mathcal{P}_n$ of $\mathcal{S}_n$ is GL_n-invariant.

Recall the spectral theorem for RSMs:

For A in $\mathcal{S}_n$, there exists g in $\mathcal{O}_n$ such that gAg^t belongs to $\mathcal{D}_n$.

Combining it with the two observations above, we obtain easily the following:

Theorem 1. $\boxed{\text{A matrix } A \text{ in } \mathcal{S}_n \text{ belongs to } \mathcal{P}_n \text{ iff } \exists\, g \text{ in } \mathcal{O}_n \text{ such that } gAg^t \in \mathcal{D}_n^+.}$

Corollary 2. *A real symmetric matrix A is positive definite if and only if its eigenvalues (which are all real by the spectral theorem) are all positive.*

Corollary 3. $\det A > 0$ *for $A \in \mathcal{P}_n$.*

These are notes for lectures at the NCM sponsored IST at IITGN held in July 2017. Corrections, comments, criticisms, etc. are welcome. Please email them to the author at `knr@imsc.res.in` or `knr.imsc@gmail.com`.

An analogous treatment holds for Hermitian positive definite matrices.

Remark 4. Theorem 1 could be rephrased as follows: given an inner product B on $\mathbb{R}^n$ (in other words, a positive definite RSM) there exists an orthonormal basis $\underline{v}_1, \ldots, \underline{v}_n$ with respect to the standard inner product such that $B(\underline{v}_i, \underline{v}_j) = 0$ for $1 \leq i \neq j \leq n$ and $B(\underline{v}_i, \underline{v}_i) > 0$ for $1 \leq i \leq n$.

2. The Gram-Schmidt process and positive definiteness

Let V be a finite dimensional real inner product space and U a subspace of V. Given an ordered basis $\underline{u}_1, \ldots, \underline{u}_m$ of U, the Gram-Schmidt process produces an ordered orthonormal basis $\underline{u}'_1, \ldots, \underline{u}'_m$ of U such that

$$(\underline{u}'_1, \ldots, \underline{u}'_m) = T(\underline{u}_1, \ldots, \underline{u}_m) \quad \text{with } T \text{ in } \mathcal{U}_n \tag{1}$$

Let us apply this to the following special case. Fix a matrix B in $\mathcal{P}_n$. Choose V to be $\mathbb{R}^n$ with inner product I given by $I(\underline{u}, \underline{v}) := \underline{u}^t B \underline{v}$. Choose U to be the whole space $\mathbb{R}^n$ (so that $\dim U = m = n$) and the ordered basis $\underline{u}_1, \ldots, \underline{u}_m$ to be the standard basis.

The right hand side of (1) in this special case becomes just T (since $\underline{u}_1, \ldots, \underline{u}_m$ is the standard basis of $\mathbb{R}^n$, it follows that $(\underline{u}_1, \ldots, \underline{u}_m)$ is the $n \times n$ identity matrix), and so we get

$$(\underline{u}'_1, \ldots, \underline{u}'_m) = T \tag{2}$$

The fact that $\underline{u}'_1, \ldots, \underline{u}'_m$ is an orthonormal basis with respect to the inner product I on $\mathbb{R}^n$ may be expressed as follows (we have written T in place of $(\underline{u}'_1, \ldots, \underline{u}'_m)$, thanks to (2)):

$$T^t \cdot B \cdot T = \text{identity}_{n \times n} \tag{3}$$

Put $Z^t = T^{-1}$. Observe that Z belongs to $\mathcal{L}_n$. We may rewrite (3) as follows:

$$B = Z \cdot Z^t \tag{4}$$

We have thus proved the existence part of the following theorem:

Theorem 5. $\boxed{\textit{For } B \in \mathcal{P}_n, \textit{ there exists unique expression } B = Z \cdot Z^t \textit{ with } Z \in \mathcal{L}_n.}$

To prove uniqueness, suppose that $Z \cdot Z^t = Y \cdot Y^t$ with both Z and Y in $\mathcal{L}_n$. This may be rewritten as $Y^{-1}Z = (Z^{-1}Y)^t$. Putting $W = Z^{-1}Y$, we may rewrite this once again as $W^{-1} = W^t$. Since $\mathcal{L}_n$ is a group and both Z and Y belong to it, it follows that W belongs to it, and so also W^{-1}. Observe that W^t belongs to $\mathcal{U}_n$.

Since $W^{-1} = W^t$, it follows that W^t belongs to $\mathcal{U}_n \cap \mathcal{L}_n$. But this intersection is precisely $\mathcal{D}_n^+$, and so W^t belongs to it. Since elements of $\mathcal{D}_n^+$ are all symmetric, it follows that $W^{-1} = W = W^t$. The only element of $\mathcal{D}_n^+$ that equals its own inverse is evidently the identity, so $W = \text{identity}_{n \times n}$. This means $Z = Y$, and the uniqueness assertion in the theorem is proved.

Corollary 6. *The action of the group GL_n on $\mathcal{P}_n$ is transitive. The stabilizer at the identity being evidently $\mathcal{O}_n$, we have* $\boxed{\mathcal{P}_n \simeq GL_n/\mathcal{O}_n}$.

Corollary 7. *The restriction to $\mathcal{L}_n$ of the action of GL_n on $\mathcal{P}_n$ is simply transitive. We thus have an identification* $\boxed{\mathcal{L}_n \simeq \mathcal{P}_n}$ *given by $Z \leftrightarrow Z \cdot Z^t$.*

3. Positivity of principal minors and positive definiteness

Let A be an $n \times n$ matrix. Let $\underline{i}$ be a subset of cardinality m of $\{1, 2, \ldots, n\}$. We write $\underline{i} = \{i_1, \ldots, i_m\}$ with $1 \leq i_1 < \ldots < i_m \leq n$. Let $A_{\underline{i}}$ denote the $m \times m$ submatrix of A whose entry in position (p, q) is A_{i_p, i_q}. Let the i^{th} *principal minor* of A, denoted $p_{\underline{i}}(A)$, be the determinant of the matrix $A_{\underline{i}}$. For example:

$$p_{\underline{i}}(A) = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{13}a_{31} \quad \text{for } \underline{i} = \{1, 3\}.$$

Proposition 8. $\boxed{\textit{For } A \textit{ in } \mathcal{P}_n, \textit{ we have } p_{\underline{i}}(A) > 0 \textit{ for all } \underline{i} \subseteq \{1, 2, \ldots, n\}.}$

Proof: Consider the inner product on $\mathbb{R}^n$ defined by A: $(\underline{u}, \underline{v})_A := \underline{u}^t A \underline{v}$. Let $W_{\underline{i}}$ be the subspace of $\mathbb{R}^n$ spanned by $\{e_{i_1}, \ldots, e_{i_m}\}$ (where $\underline{i} = \{i_1, \ldots, i_m\}$ with $1 \leq i_1 < \ldots < i_m \leq n$). The restriction to $W_{\underline{i}}$ of $(\ ,\ )_A$ is an inner product (and so positive definite). Observe that $A_{\underline{i}}$ is the matrix of this restriction with respect to the basis $e_{i_1}, \ldots, e_{i_m}$ of $W_{\underline{i}}$. The result now follows from Corollary 3. $\square$

In fact, we can say more. Let $\underline{u}_1, \ldots, \underline{u}_m$ be any set of linearly independent vectors in $\mathbb{R}^n$. Consider the $m \times m$ matrix B whose entry in position (i, j) is given by $\underline{u}_i^t A \underline{u}_j$. Then B is symmetric positive definite, for it is the matrix of the restriction to the span of $\underline{u}_1, \ldots, \underline{u}_m$ of $(\ ,\ )_A$ (which is defined as in the proof of the proposition above) with respect to $\underline{u}_1, \ldots, \underline{u}_m$. In particular $\det B > 0$ by Corollary 3.

Theorem 9. $\boxed{\textit{A matrix } A \textit{ in } \mathcal{S}_n \textit{ belongs to } \mathcal{P}_n \textit{ iff } p_{\underline{i}}(A) > 0 \textit{ for } \underline{i} = \{1, \ldots, j\}\ \forall\ 1 \leq j \leq n.}$

Proof: The only if part follows from the proposition above. For the if part, proceed by induction on n. The case $n = 1$ being clear, we assume $n \geq 2$. By induction, the $n-1 \times n-1$ top left corner of A is positive definite. By the spectral theorem, there exists orthogonal g of size $n-1 \times n-1$ such that conjugating A by the matrix $\operatorname{diag}(g, 1)$ we may assume that the $n-1 \times n-1$ top left corner of A is diagonal with positive entries, say $a_1, \ldots, a_{n-1}$. Let $v_1, \ldots, v_{n-1}, v$ be the entries in the last row of A; they are also the entries in the last column of A.

The determinant of A is

$$a_1 \cdots a_{n-1} v - \sum_{i=1}^{n-1} \frac{v_i^2}{a_i} a_1 \cdots a_{n-1}$$

Since it is positive (by hypothesis) and so also $a_1, \ldots a_{n-1}$, we get $v > \sum_{i=1}^{n-1} \frac{v_i^2}{a_i}$.

If $\underline{x}^t = (x_1, \ldots, x_n)$ is a general vector in $\mathbb{R}^n$, then, as an easy calculation shows, we have

$$\underline{x}^t A \underline{x} = v x_n^2 + \sum_{i=1}^{n-1} \left(a_i x_i^2 + 2 v_i x_i x_n \right)$$

Since $v > \sum_{i=1}^{n-1} \frac{v_i^2}{a_i}$, we have

$$\underline{x}^t A \underline{x} \geq \sum_{i=1}^{n-1} \left(a_i x_i^2 + 2 v_i x_i x_n + \frac{v_i^2 x_n^2}{a_i} \right) = \sum_{i=1}^{n-1} \left(\sqrt{a_i} x_i + \frac{v_i x_n}{\sqrt{a_i}} \right)^2 \geq 0$$

The first inequality is strict except when $x_n = 0$. Assuming $x_n = 0$, the second inequality is strict, except when x_i, $1 \leq i \leq n-1$ are all zero, in other words except when $\underline{x} = 0$, and we are done. $\square$

Exercises

(1) Let A be a real $n \times n$ symmetric positive definite matrix. Show from first principles that any real eigenvalue of A must be positive.

(2) Apply the Gram-Schmidt process to the ordered basis $(1,1,1)^t$, $(1,2,3)^t$, $(1,4,9)^t$ of $\mathbb{R}^3$ to produce an orthonormal basis of $\mathbb{R}^3$ with respect to the standard inner product.

(3) Given below is a 3×3 real symmetric matrix. Is it positive definite? If it is, write it as ZZ^t, where Z is a 3×3 real lower triangular matrix with positive entries on the diagonal:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

(4) Let A be a 2×2 invertible real matrix. When can we write A as LU, with L real lower triangular and U real upper triangular 2×2 matrices? Generalize to 3×3 and $n \times n$ matrices.

(5) Let A be an $n \times n$ real symmetric matrix. It is called *positive semi-definite* if $\underline{v}^t A \underline{v} \geq 0$ for all $\underline{v}$ in $\mathbb{R}^n$. Denote by $\mathcal{P}_n^0$ the set of all $n \times n$ real symmetric positive semi-definite matrices.

   (a) Observe that the action of GL_n on $\mathcal{S}_n$ preserves $\mathcal{P}_n^0$. In fact, if A is positive semi-definite, then so is the $m \times m$ matrix $B^t A B$, where B is any $m \times n$ real matrix.

   (b) Show that A belongs to $\mathcal{P}_n^0$ if and only if its eigenvalues (which are all real by the spectral theorem) are all non-negative.

   (c) Show that the principal minors of A are all non-negative for A in $\mathcal{P}_n^0$. Observe that it follows in particular that the principal submatrices of a positive semi-definite matrix are themselves positive semi-definite.

   (d) Prove or disprove the semi-definite analogue of the statement in Theorem 9: If $p_{\underline{i}}(A) \geq 0$ for all $\underline{i} = \{1, \ldots, j\}$ for all j, $1 \leq j \leq n$, then A is positive semi-definite.

   (e) Every element of $\mathcal{P}_n^0$ has a unique k^{th} root in $\mathcal{P}_n^0$, for every positive integer k.

(6) Find a positive definite real symmetric matrix 3×3 matrix A such that A^2 is the matrix in item (3) above.

(7) Prove or disprove: every positive definite real symmetric $n \times n$ matrix can be written uniquely as $Z \cdot Z^t$ for some real upper triangular $n \times n$ matrix Z with positive diagonal entries.

The Institute of Mathematical Sciences
E-mail address: knr@imsc.res.in, knr.imsc@gmail.com