PART ONE

PRELIMINARIES: ENTRYWISE POWERS PRESERVING POSITIVITY IN A FIXED DIMENSION

1

The Cone of Positive Semidefinite Matrices

A *kernel* is a function $K : X \times Y \to \mathbb{R}$. Broadly speaking, the goal of this text is to understand:

Which functions $F : \mathbb{R} \to \mathbb{R}$, when applied to kernels that are positive semidefinite, preserve that notion?

To do so, we first study the test sets of such kernels K themselves, and then the post-composition operators F that preserve these test sets. We begin by understanding such kernels when the domains X, Y are finite, i.e., matrices.

In this text, we will assume familiarity with linear algebra and a first course in calculus/analysis. To set notation: an uppercase letter with a two-integer subscript (such as $A_{m\times n}$) represents a matrix with m rows and n columns. If m, n are clear from context or unimportant, then they will be omitted. Three examples of real matrices are $0_{m\times n}, \mathbf{1}_{m\times n}, \mathrm{Id}_{n\times n}$, which are the (rectangular) matrix consisting of all zeros, all ones, and the identity matrix, respectively. The entries of a matrix A will be denoted a_{ij}, a_{jk}, etc. Vectors are denoted by lowercase letters (occasionally in bold) and are columnar in nature. All matrices, unless specified otherwise, are real; and similarly, all functions, unless specified otherwise, are defined on – and take values in – $\mathbb{R}^m$ for some $m \geq 1$. As is standard, we let $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}$ denote the complex numbers, reals, rationals, integers, and positive integers respectively. Given $S \subset \mathbb{R}$, let $S^{\geq 0} := S \cap [0, \infty)$.

1.1 Preliminaries

We begin with several basic definitions.

Definition 1.1 A matrix $A_{n\times n}$ is said to be *symmetric* if $a_{jk} = a_{kj}$ for all $1 \leq j, k \leq n$. A real symmetric matrix $A_{n\times n}$ is said to be *positive semidefinite*

Cambridge University Press
978-1-108-79204-2 — Matrix Analysis and Entrywise Positivity Preservers
Apoorva Khare
Excerpt
More Information

if the real number $x^T Ax$ is nonnegative for all $x \in \mathbb{R}^n$ – in other words, the quadratic form given by A is positive semidefinite. If, furthermore, $x^T Ax > 0$ for all $x \neq 0$ then A is said to be *positive definite*. Denote the set of (real symmetric) positive semidefinite matrices by $\mathbb{P}_n$.

We state the spectral theorem for symmetric (i.e., self-adjoint) operators without proof.

Theorem 1.2 (Spectral theorem for symmetric matrices) *For $A_{n\times n}$ a real symmetric matrix, $A = U^T DU$ for some orthogonal matrix U (i.e., $U^T U = \mathrm{Id}$) and real diagonal matrix D. D contains all the eigenvalues of A (counting multiplicities) along its diagonal.*

As a consequence, $A = \sum_{j=1}^n \lambda_j v_j v_j^T$, where each v_j is an eigenvector for A with *real* eigenvalue λ_j, and the v_j (which are the columns of U^T) form an orthonormal basis of $\mathbb{R}^n$.

We also have the following related results, stated here without proof: the spectral theorem for two commuting matrices, and the singular value decomposition.

Theorem 1.3 (Spectral theorem for commuting symmetric matrices) *Let $A_{n\times n}$ and $B_{n\times n}$ be two commuting real symmetric matrices. Then A and B are simultaneously diagonalizable, i.e., for some common orthogonal matrix U, $A = U^T D_1 U$ and $B = U^T D_2 U$ for D_1 and D_2 diagonal matrices (whose diagonal entries comprise the eigenvalues of A, B respectively).*

Theorem 1.4 (Singular value decomposition) *Every real matrix $A_{m\times n} \neq 0$ decomposes as $A = P_{m\times m} \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q_{n\times n}$, where P, Q are orthogonal and Σ_r is a diagonal matrix with positive eigenvalues. The entries of Σ_r are called the* singular values *of A and are the square roots of the nonzero eigenvalues of AA^T (or $A^T A$).*

1.2 Criteria for Positive (Semi)Definiteness

We write down several equivalent criteria for positive (semi)definiteness. There are three initial criteria which are easy to prove, and a final criterion which requires separate treatment.

Theorem 1.5 (Criteria for positive (semi)definiteness) *Given $A_{n\times n}$ a real symmetric matrix of rank $0 \leq r \leq n$, the following are equivalent:*

(1) *A is positive semidefinite (respectively, positive definite).*
(2) *All eigenvalues of A are nonnegative (respectively, positive).*
(3) *There exists a matrix $B \in \mathbb{R}^{r \times n}$ of rank r, such that $B^T B = A$. (In particular, if A is positive definite then B is square and nonsingular.)*

Proof We prove only the positive semidefinite statements; minor changes show the corresponding positive definite variants. If (1) holds and λ is an eigenvalue – for an eigenvector x – then $x^T Ax = \lambda \|x\|^2 \geq 0$. Hence, $\lambda \geq 0$, proving (2). Conversely, if (2) holds then by the spectral theorem, $A = \sum_j \lambda_j v_j v_j^T$ with all $\lambda_j \geq 0$, so A is positive semidefinite:

$$x^T Ax = \sum_j \lambda_j x^T v_j v_j^T x = \sum_j \lambda_j (x^T v_j)^2 \geq 0, \qquad \forall x \in \mathbb{R}^n.$$

Next, if (1) holds then write $A = U^T DU$ by the spectral theorem; note that $D = UAU^T$ has the same rank as A. Since D has nonnegative diagonal entries d_{jj}, it has a square root $\sqrt{D}$, which is a diagonal matrix with diagonal entries $\sqrt{d_{jj}}$. Write $D = \begin{pmatrix} D'_{r\times r} & 0 \\ 0 & 0_{(n-r)\times(n-r)} \end{pmatrix}$, where D' is a diagonal matrix with positive diagonal entries. Correspondingly, write $U = \begin{pmatrix} P_{r\times r} & Q \\ R & S_{(n-r)\times(n-r)} \end{pmatrix}$. If we set $B := (\sqrt{D'}P \mid \sqrt{D'}Q)_{r\times n}$, then it is easily verified that

$$B^T B = \begin{pmatrix} P^T D' P & P^T D' Q \\ Q^T D' P & Q^T D' Q \end{pmatrix} = U^T DU = A.$$

Hence, (1) $\implies$ (3). Conversely, if (3) holds then $x^T Ax = \|Bx\|^2 \geq 0$ for all $x \in \mathbb{R}^n$. Hence, A is positive semidefinite. Moreover, we claim that B and $B^T B$ have the same null space and hence the same rank. Indeed, if $Bx = 0$ then $B^T Bx = 0$, while

$$B^T Bx = 0 \implies x^T B^T Bx = 0 \implies \|Bx\|^2 = 0 \implies Bx = 0. \quad \square$$

Corollary 1.6 *For any real symmetric matrix $A_{n\times n}$, the matrix $A - \lambda_{\min} \mathrm{Id}_{n\times n}$ is positive semidefinite, where $\lambda_{\min}$ denotes the smallest eigenvalue of A.*

We now state Sylvester's criterion for positive (semi)definiteness. (Incidentally, Sylvester is believed to have first introduced the use of "matrix" in mathematics, in the nineteenth century.) This requires some additional notation.

Definition 1.7 Given an integer $n \geq 1$, define $[n] := \{1, \ldots, n\}$. Now given a matrix $A_{m\times n}$ and subsets $J \subset [m]$, $K \subset [n]$, define $A_{J\times K}$ to be the submatrix of A with entries a_{jk} for $j \in J, k \in K$ (always considered to be arranged in increasing order in this text). If J, K have the same size then $\det A_{J \times K}$ is called a *minor* of A. If A is square and $J = K$ then $A_{J \times K}$ is called a *principal submatrix* of A, and $\det A_{J \times K}$ is a *principal minor*. The principal submatrix (and principal minor) are *leading* if $J = K = \{1, \dots, m\}$ for some $1 \leq m \leq n$.

Theorem 1.8 (Sylvester's criterion) *A symmetric matrix is positive semidefinite (respectively, positive definite) if and only if all its principal minors are nonnegative (respectively, positive).*

We will show Theorem 1.8 with the help of a few preliminary results.

Lemma 1.9 *If $A_{n \times n}$ is a positive semidefinite (respectively, positive definite) matrix, then so are all principal submatrices of A.*

Proof Fix a subset $J \subset [n] = \{1, \dots, n\}$ (so $B := A_{J \times J}$ is the corresponding principal submatrix of A), and let $x \in \mathbb{R}^{|J|}$. Define $x' \in \mathbb{R}^n$ to be the vector, such that $x'_j = x_j$ for all $j \in J$ and 0 otherwise. It is easy to see that $x^T B x = (x')^T A x'$. Hence, B is positive (semi)definite if A is. □

As a corollary, all the principal minors of a positive semidefinite (positive definite) matrix are nonnegative (positive) since the corresponding principal submatrices have nonnegative (positive) eigenvalues and hence nonnegative (positive) determinants. So one direction of Sylvester's criterion holds trivially.

Lemma 1.10 *Sylvester's criterion is true for positive definite matrices.*

Proof We induct on the dimension of the matrix A. Suppose $n = 1$. Then A is just an ordinary real number, so its only principal minor is A itself, and so the result is trivial.

Now, suppose the result is true for matrices of dimension $\leq n - 1$. We claim that A has at least $n - 1$ positive eigenvalues. To see this, let $\lambda_1, \lambda_2 \leq 0$ be eigenvalues of A. Let W be the $n - 1$ dimensional subspace of $\mathbb{R}^n$ with last entry 0. If v_j are orthogonal eigenvectors for λ_j, $j = 1, 2$, then the span of the v_j must intersect W nontrivially, since the sum of dimensions of these two subspaces of $\mathbb{R}^n$ exceeds n. Define $u := c_1 v_1 + c_2 v_2 \in W$; then $u^T A u > 0$ by Lemma 1.9. However,

$$u^T A u = \left(c_1 v_1^T + c_2 v_2^T\right) A (c_1 v_1 + c_2 v_2) = c_1^2 \lambda_1 ||v_1||^2 + c_2^2 \lambda_2 ||v_2||^2 \leq 0,$$

thereby giving a contradiction and proving the claim.

Now since the determinant of A is positive (it is the minor corresponding to A itself), it follows that all eigenvalues are positive, completing the proof. □

We will now prove the Jacobi formula, an important result in its own right. A corollary of this result will be used, along with the previous result and the idea that positive semidefinite matrices can be expressed as entrywise limits of positive definite matrices, to prove Sylvester's criterion for all positive semidefinite matrices.

Theorem 1.11 (Jacobi formula) *Let $A_t : \mathbb{R} \to \mathbb{R}^{n\times n}$ be a matrix-valued differentiable function. Then,*

$$\frac{d}{dt}(\det A_t) = \operatorname{tr}\left(\operatorname{adj}(A_t)\frac{dA_t}{dt}\right), \tag{1.1}$$

where $\operatorname{adj}(A_t)$ *denotes the adjugate matrix of* A_t.

Proof The first step is to compute the differential of the determinant. We claim that for any $n \times n$ real matrices A, B,

$$d(\det)(A)(B) = \operatorname{tr}(\operatorname{adj}(A)B).$$

As a special case, at $A = \mathrm{Id}_{n\times n}$, the differential of the determinant is precisely the trace.

To show the claim, we need to compute the directional derivative

$$\lim_{\epsilon\to 0} \frac{\det(A + \epsilon B) - \det A}{\epsilon}.$$

The fraction is a polynomial in ϵ with vanishing constant term (e.g., set $\epsilon = 0$ to see this); and we need to compute the coefficient of the linear term. Expand $\det(A + \epsilon B)$ using the Laplace expansion as a sum over permutations $\sigma \in S_n$; now each individual summand $(-1)^\sigma \prod_{k=1}^n (a_{k\sigma(k)} + \epsilon b_{k\sigma(k)})$ splits as a sum of 2^n terms. (It may be illustrative to try and work out the $n = 3$ case by hand.) From these $2^n \cdot n!$ terms, choose the ones that are linear in ϵ. For each $1 \leq i, j \leq n$, there are precisely $(n-1)!$ terms corresponding to ϵb_{ij}; and added together, they equal the (i, j)th cofactor C_{ij} of A – which equals $\operatorname{adj}(A)_{ji}$. Thus, the coefficient of ϵ is

$$d(\det)(A)(B) = \sum_{i,j=1}^n C_{ij} b_{ij},$$

and this is precisely $\operatorname{tr}(\operatorname{adj}(A)B)$, as claimed.

More generally, the above argument shows that if $B(\epsilon)$ is any family of matrices, with limit $B(0)$ as $\epsilon \to 0$, then

$$\lim_{\epsilon\to 0} \frac{\det(A + \epsilon B(\epsilon)) - \det A}{\epsilon} = \operatorname{tr}(\operatorname{adj}(A)B(0)). \tag{1.2}$$

Returning to the proof of the theorem, for $\epsilon \in \mathbb{R}$ small and $t \in \mathbb{R}$ we write

$$A_{t+\epsilon} = A_t + \epsilon B(\epsilon),$$

where $B(\epsilon) \to B(0) := \frac{dA_t}{dt}$ as $\epsilon \to 0$, by definition. Now compute using (1.2)

$$\frac{d}{dt}(\det A_t) = \lim_{\epsilon \to 0} \frac{\det(A_t + \epsilon B(\epsilon)) - \det A_t}{\epsilon} = \operatorname{tr}\left(\operatorname{adj}(A_t)\frac{dA_t}{dt}\right). \qquad \square$$

With these results at hand, we can finish the proof of Sylvester's criterion for positive semidefinite matrices.

Proof of Theorem 1.8 For positive definite matrices, the result was proved in Lemma 1.10. Now suppose $A_{n\times n}$ is positive semidefinite. One direction follows by the remarks preceding Lemma 1.10. We show the converse by induction on n, with an easy argument for $n = 1$ similar to the positive definite case.

Now suppose the result holds for matrices of dimension $\leq n-1$ and let $A_{n\times n}$ have all principal minors nonnegative. Let B be any principal submatrix of A, and define $f(t) := \det(B + t\,\mathrm{Id}_{n\times n})$. Note that $f'(t) = \operatorname{tr}(\operatorname{adj}(B + t\,\mathrm{Id}_{n\times n}))$ by the Jacobi formula (1.1).

We claim that $f'(t) > 0\ \forall t > 0$. Indeed, each diagonal entry of $\operatorname{adj}(B + t\,\mathrm{Id}_{n\times n})$ is a proper principal minor of $A + t\,\mathrm{Id}_{n\times n}$, which is positive definite since $x^T(A + t\,\mathrm{Id}_{n\times n})x = x^T Ax + t\|x\|^2$ for $x \in \mathbb{R}^n$. The claim now follows using Lemma 1.9 and the induction hypothesis.

The claim implies: $f(t) > f(0) = \det B \geq 0\ \forall t > 0$. Thus, all principal minors of $A + tI$ are positive, and by Sylvester's criterion for positive definite matrices, $A + tI$ is positive definite for all $t > 0$. Now note that $x^T Ax = \lim_{t\to 0^+} x^T(A + t\,\mathrm{Id}_{n\times n})x$; therefore the nonnegativity of the right-hand side implies that of the left-hand side for all $x \in \mathbb{R}^n$, completing the proof. $\square$

1.3 Examples of Positive Semidefinite Matrices

We next discuss several examples of positive semidefinite matrices.

1.3.1 Gram Matrices

Definition 1.12 For any finite set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^m$, their *Gram matrix* is given by $\mathrm{Gram}((\mathbf{x}_j)_j) := (\langle \mathbf{x}_j, \mathbf{x}_k \rangle)_{1\leq j,k\leq n}$.

A *correlation matrix* is a positive semidefinite matrix with ones on the diagonal.

In fact, we need not use $\mathbb{R}^m$ here; any inner product space/Hilbert space is sufficient.

Proposition 1.13 *Given a real symmetric matrix $A_{n\times n}$, it is positive semidefinite if and only if there exist an integer $m > 0$ and vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^m$, such that $A = Gram((\mathbf{x}_j)_j)$.*

As a special case, correlation matrices precisely correspond to those Gram matrices for which the $\mathbf{x}_j$ are unit vectors. We also remark that a "continuous" version of this result is given by a well-known result of Mercer [172].

Proof If A is positive semidefinite, then by Theorem 1.5 we can write $A = B^T B$ for some matrix $B_{m\times n}$. It is now easy to check that A is the Gram matrix of the columns of B.

Conversely, if $A = \mathrm{Gram}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ with all $\mathbf{x}_j \in \mathbb{R}^m$, then to show that A is positive semidefinite, we compute for any $\mathbf{u} = (u_1, \dots, u_n)^T \in \mathbb{R}^n$

$$\mathbf{u}^T A \mathbf{u} = \sum_{j,k=1}^{n} u_j u_k \langle \mathbf{x}_j, \mathbf{x}_k \rangle = \left\| \sum_{j=1}^{n} u_j \mathbf{x}_j \right\|^2 \geq 0. \qquad \square$$

1.3.2 (Toeplitz) Cosine Matrices

Definition 1.14 A matrix $A = (a_{jk})$ is *Toeplitz* if a_{jk} depends only on $j-k$.

Lemma 1.15 *Let $\theta_1, \dots, \theta_n \in [0, 2\pi]$. Then the matrix $C := (\cos(\theta_j - \theta_k))_{j,k=1}^n$ is positive semidefinite, with rank at most 2. In particular, $\alpha \mathbf{1}_{n\times n} + \beta C$ has rank at most 3 (for scalars α, β), and it is positive semidefinite if $\alpha, \beta \geq 0$.*

Proof Define the vectors $u, v \in \mathbb{R}^n$ via: $u^T = (\cos\theta_1, \dots, \cos\theta_n)$, $v^T = (\sin\theta_1, \dots, \sin\theta_n)$. Then $C = uu^T + vv^T$ via the identity $\cos(a - b) = \cos a \cos b + \sin a \sin b$, and clearly the rank of C is at most 2. (For instance, it can have rank 1 if the θ_j are equal.) As a consequence,

$$\alpha \mathbf{1}_{n\times n} + \beta C = \alpha \mathbf{1}_n \mathbf{1}_n^T + \beta uu^T + \beta vv^T$$

has rank at most 3; the final assertion is straightforward. $\square$

As a special case, if $\theta_1, \dots, \theta_n$ are in arithmetic progression, i.e., $\theta_{j+1} - \theta_j = \theta \ \forall j$ for some θ, then we obtain a positive semidefinite *Toeplitz* matrix

$$C = \begin{pmatrix} 1 & \cos\theta & \cos 2\theta & \cdots & & & \\ \cos\theta & 1 & \cos\theta & \cos 2\theta & \cdots & & \\ \cos 2\theta & \cos\theta & 1 & \cos\theta & \cos 2\theta & \cdots & \\ & \cos 2\theta & \cos\theta & 1 & \cos\theta & \cos 2\theta & \cdots \\ \vdots & & \ddots & & \ddots & & \ddots \end{pmatrix}.$$

This family of Toeplitz matrices was used by Rudin in a 1959 paper [197] on entrywise positivity preservers; see Theorem 11.3 for his result.

1.3.3 Hankel Matrices

Definition 1.16 A matrix $A = (a_{jk})$ is *Hankel* if a_{jk} depends only on $j+k$.

Example 1.17 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is Hankel but not positive semidefinite.

Example 1.18 For $x \geq 0$, the matrix $\begin{pmatrix} 1 & x & x^2 \\ x & x^2 & x^3 \\ x^2 & x^3 & x^4 \end{pmatrix} = \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \begin{pmatrix} 1 & x & x^2 \end{pmatrix}$ is Hankel and positive semidefinite of rank 1.

A more general perspective is as follows. Define

$$H_x := \begin{pmatrix} 1 & x & x^2 & \cdots \\ x & x^2 & x^3 & \cdots \\ x^2 & x^3 & x^4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

and let δ_x be the Dirac measure at $x \in \mathbb{R}$. The moments of this measure are given by

$$s_k(\delta_x) := \int_{\mathbb{R}} y^k \, d\delta_x(y) = x^k, \qquad k \geq 0.$$

Thus, H_x is the "moment matrix" of δ_x. More generally, given any nonnegative measure μ supported on $\mathbb{R}$, with all moments finite, the corresponding *Hankel moment matrix* is the bi-infinite "matrix" given by

$$H_\mu := \begin{pmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & s_3 & \cdots \\ s_2 & s_3 & s_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \text{where } s_k = s_k(\mu) := \int_{\mathbb{R}} y^k \, d\mu(y). \quad (1.3)$$