MA 341: Matrix Analysis and Positivity

Naman Mishra

August 2025

Contents

Lectures

The Course

Resources:

Lecture 1. Tuesday August 5

(1) A. Khare. Matrix analysis and entrywise positivity preservers, 2022.

(2) Rajendra Bhatia. Matrix analysis.

(3) Rajendra Bhatia. Positive definite matrices.

Grading:

(50%) Homework + Midterm

(50%) Final presentation

Various notions of matrix positivity and maps preserving these structures.

We'll briefly see how GPS triangulation works, and also Heron's formula in n dimensions.

Over this course, we will only work over $\mathbb{R}$, and "positive" will often mean "non-negative".

Chapter I

Positivity

One easy way to generalize positivity (non-negativity) of real numbers to matrices is to consider diagonal matrices with non-negative entries.

More generally, we

{thm:spectral}

Theorem I.1 (Spectral Theorem)**.** *Each symmetric matrix $A \in \mathbb{R}^{n\times n}$ has an orthonormal eigenbasis, that is, there exist orthonormal vectors $u_1, \dots, u_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ such that $Au_i = \lambda_i u_i$.*

We could write

$$A\begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} = \begin{pmatrix} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{pmatrix} = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Calling the matrix of eigenvectors U and the diagonal matrix of eigenvalues Λ, we can write $AU = U\Lambda$, so that $A = U\Lambda U^T$. Note that $U^{-1} = U^T$ since

$$\begin{pmatrix} u_1^T \\ \vdots \\ u_n^T \end{pmatrix}\begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} = (u_i^T u_j)_{ij} = \mathbb{I}_n$$

iff the u_i are orthonormal.

We call matrices of this form *positive semidefinite*. That is, a matrix A is positive semidefinite iff it can be expressed as $U\Lambda U^T$ where $U \in O(n)$ and Λ is a diagonal matrix with $\lambda_{ii} \geq 0$. The formal definition is different.

{def:psd}

Definition I.2 (Positive Semidefinite)**.** Let $A \in \mathbb{R}^{n\times n}$ and let $\kappa_A \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the associated bilinear form $\kappa_A(v, w) = v^T A w$. A is *positive semidefinite* iff $A = A^T$ and $\kappa_A(v, v) \geq 0$ for all $v \in \mathbb{R}^n$.

Remark. If $U = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}$ and $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ is a positive diagonal matrix, then

$$U\Lambda U^T = \sum_i \lambda_i u_i u_i^T$$

so $x^T(U\Lambda U^T)x = \sum_i \lambda_i \langle u_i,\, x\rangle^2 \geq 0$.

Thus the definition we discussed earliear is equivalent to the one in definition I.2.

A *principal submatrix* of $A \in \mathbb{R}^{n\times n}$ is the matrix obtained by choosing a non-empty subset $S \subseteq [n]$ and keeping precisely those rows and columns of A indexed by S.

We will over the course of the course prove most of the following.

Theorem I.3. *The following are equivalent for $A \in \mathbb{R}^{n\times n}$.*

(1) A is PSD.

(2) All eigenvalues of A are in $[0, \infty)$.

(3) $A = B^T B$ for some $B \in \mathbb{R}^{n\times n}$.

(4) All principal submatrices of A have non-negative determinants.

(5) A is a Gram matrix from $\mathbb{R}^n$.

(6) A is the covariance matrix of some data.

(7) A is the Cayley-Menger matrix of an $(n+1)$-point Euclidean metrix space $X \subseteq (\mathbb{R}^n, \|\cdot\|_2)$.

PSDs show up in many places, such as

- Nevanlinna-Pick condition,
- classifying Dynkin diagrams of something something,
- much earlier, using Hessians to find local minima.

Traditionally, authors like Rajendra Bhatia consider functions of the form $f(A) = f(U\Lambda U^T) = Uf(\Lambda)U^T$. Why?

- $A^2 = U\Lambda^2 U^T$. $A^3 = U\Lambda^3 U^T$, et cetera.
- For a polynomial $p(x) = \sum_{i=0}^d a_i x^i$, then $p(A) = Up(\Lambda)U^T$.

As long as the eigenvalues are compactly supported, polynomials give good approximations to all continuous functions. This study is called (holomorphic) functional calculus.

In this course, we will work with functions acting on the *entries* of the matrix, that is, functions that look like

$$f[A] = (f(a_{ij}))_{ij}.$$

For example, entrywise matrix multiplication of two PSD matrices is PSD, as we will see later. Entrywise calculus is not as well-developed.

Examples.

- [(Toeplitz) cosine matrices] Let $\theta_1, \dots, \theta_2 \in \mathbb{R}$ and set $a_{ij} = \cos(\theta_i - \theta_j)$. Then $A = uu^T + vv^T$ where

  $$u = \begin{pmatrix} \cos\theta_1 \\ \dots \\ \cos\theta_n \end{pmatrix}, \qquad v = \begin{pmatrix} \sin\theta_1 \\ \dots \\ \sin\theta_n \end{pmatrix}.$$

  Thus A is symmetric and $x^T Ax \geq 0$ for all x.

  A matrix is *Toeplitz* if a_{ij} depends only on $i - j$. That is, A looks like

  $$\begin{pmatrix} a_0 & a_{-1} & a_{-2} & a_{-3} \\ a_1 & a_0 & a_{-1} & a_{-2} \\ a_2 & a_1 & a_0 & a_{-1} \\ a_3 & a_2 & a_1 & a_0 \end{pmatrix}$$

  If $\theta_1, \dots, \theta_n$ are in arithmetic progression, then the cosine matrix discussed above is Toeplitz.

- A *Hankel* matrix is similar to a Toeplitz matrix, but constant along the anti-diagonals: something like

  $$\begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 & a_4 \\ a_2 & a_3 & a_4 & a_5 \\ a_3 & a_4 & a_5 & a_6 \end{pmatrix}.$$

  Choose $u_0 \in \mathbb{R}$ and define

  $$H_{u_0} := \begin{pmatrix} 1 & u_0 & u_0^2 & u_0^3 & \cdots \\ u_0 & u_0^2 & u_0^3 & u_0^4 & \cdots \\ u_0^2 & u_0^3 & u_0^4 & u_0^5 & \cdots \\ u_0^3 & u_0^4 & u_0^5 & u_0^6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

  Claim. Any (finite) leading principal truncation of H_{u_0} is PSD.

Proof. Let the truncation have size $n+1$. Then

$$\begin{pmatrix} 1 & u_0 & \cdots & u_0^n \\ u_0 & u_0^2 & \cdots & u_0^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_0^n & u_0^{n+1} & \cdots & u_0^{2n} \end{pmatrix} = \begin{pmatrix} 1 \\ u_0 \\ \vdots \\ u_0^n \end{pmatrix} \begin{pmatrix} 1 & u_0 & \cdots & u_0^n \end{pmatrix}$$

and hence is PSD. ∎

The *Dirac* δ measure at $u_0 \in \mathbb{R}$, denoted by $\delta_{u_0}(x) = \delta_{u_0,x}$ satisfies

$$\int_{\mathbb{R}} f d\delta_{u_0} = f(u_0).$$

The *k-th moment* of a measure $\mu \geq 0$ on $\mathbb{R}$ is

$$s_k(\mu) \coloneqq \int_{\mathbb{R}} x^k d\mu(x).$$

For example,

$$\begin{aligned} \text{mass of } \mu = s_0(\mu) &= \int_{\mathbb{R}} d\mu, \\ \text{mean of } \mu &= s_1(\mu)/s_0(\mu) \ (\text{if } s_0(\mu) > 0), \\ \text{variance of } \mu &= s_2(\mu) - s_1(\mu)^2 \ (\text{if } s_0(\mu) = 1). \end{aligned}$$