MA431 Spectral Graph Theory: Lecture 0

Ahmad Abdi Neil Olver

January 13, 2022

Linear Algebra: A Brief Review

Let A be an $n \times n$ matrix over the complex numbers. The *characteristic polynomial of* A is $p_A(x) := \det(xI - A)$, where I is the $n \times n$ identity matrix. Notice that $p_A(x)$ is a polynomial of degree n, and that its distinct roots are precisely the distinct eigenvalues of A.

The *algebraic multiplicity* of an eigenvalue $\lambda \in \mathbb{C}$ is the largest integer d such that $(x - \lambda)^d$ is a factor of $p_A(x)$. As a consequence, the sum of the algebraic multiplicities of the distinct eigenvalues of A is equal to n. The *geometric multiplicity* of an eigenvalue $\lambda \in \mathbb{C}$ is the dimension of its eigenspace. It is known that the geometric multiplicity is always less than or equal to the algebraic multiplicity.

Two $n \times n$ matrices A, B are *similar* if for an invertible matrix P, we have $A = P^{-1}BP$. It is known that similar matrices have the same characteristic polynomial.

Suppose A is a real symmetric matrix, an assumption that is often made in this course. Then the eigenvalues of A are real numbers (this is a nice exercise). Moreover, A is diagonalizable, that is, it is similar to a diagonal matrix D (this follows from the theorem below). In this case, the geometric and algebraic multiplicities of every eigenvalue coincide, so we may speak of *the* multiplicity of an eigenvalue. Moreover, the diagonal entries of D are the eigenvalues of A, repeated according to their multiplicity.

Theorem 0.1 (Spectral Decomposition Theorem)**.** Let A be an $n \times n$ real symmetric matrix. Then the following statements hold:

1. There exists an orthonormal basis $u_1, u_2, \dots, u_n \in \mathbb{R}^n$ such that each u_i is an eigenvector for A.

For each i, let λ_i be the eigenvalue corresponding to u_i. Let D be the diagonal matrix whose diagonal entries are $\lambda_1, \dots, \lambda_n$. Define the $n \times n$ orthogonal matrix $P := [u_1, u_2, \dots, u_n]$. Then

2. $A = PDP^{-1} = PDP^\top$. That is,
$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top.$$

For each eigenvalue λ of A, let $E_\lambda := \sum \left(u_i u_i^\top : \lambda_i = \lambda \right)$. Then

3. E_λ is the matrix of projection onto the λ-eigenspace. In particular, $E_\lambda^2 = E_\lambda$.

4. Given that $ev(A)$ denotes the set of distinct eigenvalues of A, we have the following:

$$I = \sum_{\lambda \in ev(A)} E_\lambda,$$
$$A = \sum_{\lambda \in ev(A)} \lambda E_\lambda.$$

Proof. Exercise. □

As a consequence, we get the following theorem:

Theorem 0.2 (Courant-Fischer Theorem)**.** Let A be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Then the following statements hold:

1. $\min\{x^\top Ax : x^\top x = 1\} = \lambda_n$. Moreover, equality is achieved only by vectors in the λ_n-eigenspace.

2. $\max\{x^\top Ax : x^\top x = 1\} = \lambda_1$. Moreover, equality is achieved only by vectors in the λ_1-eigenspace.

Let $u_1, \ldots, u_n$ be an orthogonal basis of eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. For each $j \in \{1, \ldots, n-1\}$, let U_j be the subspace spanned by $u_1, \ldots, u_j$. Then we have the following *Rayleigh equalities*:

3. $\min\{x^\top Ax : x^\top x = 1, x \in U_j\} = \lambda_j$. Moreover, equality is achieved only by vectors in the λ_j-eigenspace.

4. $\max\{x^\top Ax : x^\top x = 1, x \in U_j^\perp\} = \lambda_{j+1}$. Moreover, equality is achieved only by vectors in the λ_{j+1}-eigenspace.

Moreover, for any subspace U of dimension $j \in \{1, \ldots, n-1\}$, we have the following *Rayleigh inequalities*:

5. $\min\{x^\top Ax : x^\top x = 1, x \in U\} \leq \lambda_j$. Moreover, if equality holds, then U contains an eigenvector with eigenvalue λ_j.

6. $\max\{x^\top Ax : x^\top x = 1, x \in U^\perp\} \geq \lambda_{j+1}$. Moreover, if equality holds, then $U^\perp$ contains an eigenvector with eigenvalue λ_{j+1}.

Proof. Exercise. □

There is a subtle difference between (3)-(4) and (5)-(6). Notice that in (3)-(4), we have a characterization of when equality holds, whereas in (5)-(6), we do not. The reason for this difference is made clear when one proves the statements; we invite the reader to do so.

Let us present one final application of the Spectral Decomposition Theorem. A real symmetric matrix A is *positive semidefinite (PSD)*, denoted as $A \succcurlyeq \mathbf{0}$, if $x^\top Ax \geq 0$ for all vectors x, and it is *positive definite* if $x^\top Ax > 0$ for all nonzero vectors x. We have the following characterization of PSD matrices:

Theorem 0.3 (Characterization of PSD Matrices)**.** Let A be an $n \times n$ real symmetric matrix. Then the following statements are equivalent:

1. A is positive semidefinite (resp. positive definite),
2. every eigenvalue of A is nonnegative (resp. strictly positive),
3. $A = B^\top B$ for a real matrix B (resp. nonsingular real matrix B).

Proof. Exercise. □

Given real symmetric matrices A, B of the same dimension, we write $A \succcurlyeq B$ if $A - B$ is a positive semidefinite matrix. The relation $\succcurlyeq$ defines a partial order, called the *Loewner order*, on the space of symmetric matrices, that is, the following three properties are satisfied (the proof of which is left as an exercise):

- $A \succcurlyeq A$ (reflexivity),
- if $A \succcurlyeq B$ and $B \succcurlyeq A$, then $A = B$ (antisymmetry),
- if $A \succcurlyeq B$ and $B \succcurlyeq C$, then $A \succcurlyeq C$ (transitivity).

The pseudoinverse. The *Moore-Penrose pseudoinverse* is a generalization of the inverse to all matrices.

Theorem 0.4. Let A be an $m \times n$ real matrix. Then there is a unique $n \times m$ matrix A^+, called the pseudoinverse of A, satisfying the following:

(i) $AA^+A = A$ and $A^+AA^+ = A^+$;

(ii) AA^+ and A^+A are symmetric.

Further, A^+ satisfies the following.

1. If A is square and invertible, then $A^+ = A^{-1}$.
2. $(A^+)^+ = A$.
3. AA^+ is the orthogonal projection onto the range of A, and A^+A is the orthogonal projection onto the range of A^T.

We will primarily be interested in square symmetric matrices, in which case we have a more straightforward interpretation. Consider the diagonalization $A = PDP^\top$ of a real symmetric matrix A, where D is diagonal and P orthogonal. Then we must have that $A^+ = PD^+P^\top$ (you can easily check that this satisfies the requirements to be the pseudoinverse of A; for example, $AA^+A = PDD^+DP^\top = PDP^\top = A$). But the pseudoinverse of D is straightforward to see: it is diagonal, with $D^+_{ii} = 0$ if $D_{ii} = 0$, and $D^+_{ii} = 1/D_{ii}$ otherwise. Again, you can easily check that this indeed satisfies the requirements to be the pseudoinverse of D. Hence we have the following.

Lemma 0.5. If A is a symmetric $n \times n$ real matrix, with spectral decomposition $A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$ (with $\{u_1, \ldots, u_n\}$ an orthonormal basis of eigenvectors), then

$$A^+ = \sum_{i:\lambda_i \neq 0} \tfrac{1}{\lambda_i} u_i u_i^\top.$$

Geometrically, we can view the pseudoinverse in the symmetric case as follows. View A as a self-adjoint operator from $\mathbb{R}^n \to \mathbb{R}^n$. Let W be the range of A. Note that W is an A-invariant subspace, and furthermore, the map $A' : W \to W$ obtained by restricting to W is a bijection (this is a consequence of the self-adjointness of A). Consider the map obtained by first projecting orthogonally onto W, and then applying the inverse of A' (viewing the result as a vector in the ambient space $\mathbb{R}^n$). This is precisely the pseudoinverse.

(A similar geometric construction applies for general linear operators, but it's a bit more complicated. It is no longer true that A is a bijection from W to W. Instead, one defines A_x^+ to be the minimum norm point y such that Ay is equal to the orthogonal projection of x onto W.)

Exercises

1. Let A be an $n \times n$ matrix. Recall that $p_A(x) = \det(xI - A)$. Choose $\sigma_0(A), \sigma_1(A), \ldots, \sigma_n(A)$ such that

   $$p_A(x) = \sum_{k=0}^{n} (-1)^k \sigma_k(A) x^{n-k}.$$

   Prove the following statements for each k:

   (a) $\sigma_k(A)$ is the sum of the product of any k eigenvalues, counted according to their algebraic multiplicity. That is, if $\lambda_1, \ldots, \lambda_n$ are the n eigenvalues of A, repeated according to the algebraic multiplicity of the eigenvalues, then

   $$\sigma_k(A) = \sum_{S \subseteq [n], |S|=k} \prod_{i \in S} \lambda_i.$$

   (b) $\sigma_k(A)$ is the sum of the determinants of all principal $k \times k$ submatrices. That is,

   $$\sigma_k(A) = \sum \left(\det(B) : B \text{ is a } k \times k \text{ principal submatrix of } A\right).$$

2. Let A be an $n \times n$ matrix. Prove that the trace of A is equal to the sum of its eigenvalues, respecting their algebraic multiplicities.

3. Let A be an $n \times n$ matrix, and take an integer $\ell \geq 1$. Prove that the eigenvalues of A^ℓ are precisely the eigenvalues of A raised to the power ℓ, preserving algebraic multiplicities.

4. Let A be an $n \times n$ real symmetric matrix. A subspace $U \subseteq \mathbb{R}^n$ is A-*invariant* if $Ax \in U$ for all $x \in U$.

   (a) Prove that if U is A-invariant, then so is $U^\perp$.

   (b) Prove that any A-invariant subspace of dimension at least one contains an eigenvector of A.

(c) Prove that for any integer $1 \leq m < n$, any orthogonal set of m eigenvectors can be extended to orthogonal set of $m+1$ eigenvectors.

(d) Prove Theorem 0.1.

5. Prove Theorem 0.2 parts (1), (3) and (5). Then apply those parts to $-A$ to prove parts (2), (4) and (6).

6. Let A be an $n \times n$ real symmetric matrix, and denote by $ev(A)$ the set of distinct eigenvalues of A. Prove that for any polynomial p,
$$p(A) = \sum_{\lambda \in ev(A)} p(\lambda) E_\lambda.$$
Then prove that the vector space of all the polynomials in A has dimension equal to the number of distinct eigenvalues of A.

7. Prove Theorem 0.3.

8. Prove from the definition of the pseudoinverse that it must be unique (in general, no restriction to the symmetric case).

9. Let A be a symmetric matrix, and suppose $x = A^+ b$. Show that x minimizes $\|Ax - b\|$, and moreover, that amongst all such minimizers, x has minimum norm.

10. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times m}$. Show that AB and BA have the same nonzero eigenvalues (with multiplicities), and hence give a relationship between the characteristic polynomials of AB and of BA.