LINEAR ALGEBRA FOR APPLICATIONS

FERNANDO LIU LOPEZ

1. Classes of Matrices

Definition 1. Let $A \in \mathrm{Mat}_n(\mathbb{C})$. We also use $\vec{a}_i$ to refer to the i-th column of A, and use $a_{i,j}$ or $(A)_{i,j}$ to refer to the (i,j)-th entry of A. We call A...

(1) **normal** if it commutes with its (conjugate) transpose: $AA^* = A^*A$;
(2) **unitary** if it's invertible and $A^{-1} = A^*$ (resp. **orthogonal** and $A^{-1} = A^\intercal$ over $\mathbb{R}$);
(3) **Hermitian** if $A = A^*$ (resp. **symmetric** and $A = A^\intercal$ over $\mathbb{R}$);
(4) **positive definite** if it's Hermitian and $\vec{x}^* A\vec{x} > 0$ for all nonzero $x \in \mathbb{C}^n$.
(5) **positive semi-definite** if it's Hermitian and $\vec{x}^* A\vec{x} \geq 0$ for all nonzero $x \in \mathbb{C}^n$.
(6) **upper triangular** if $a_{i,j} = 0$ for all $i > j$ (resp. **lower triangular** for $i < j$).

Notation. We use the following notation for classes of matrices:

$\mathrm{GL}_n(\Bbbk)$	invertible $n \times n$ matrices over $\Bbbk$.
$\mathrm{SL}_n(\Bbbk)$	invertible $n \times n$ matrices over $\Bbbk$ with determinant 1.
$\mathrm{Diag}_n(\Bbbk)$	diagonal $n \times n$ matrices over $\Bbbk$.
$\mathcal{O}_n(\mathbb{R})$	orthogonal matrices over $\mathbb{R}$.
$\mathcal{U}_n(\mathbb{C})$	unitary matrices over $\mathbb{C}$.
$\mathcal{S}_n(\mathbb{R})$	symmetric matrices over $\mathbb{R}$.
$\mathcal{H}_n(\mathbb{C})$	Hermitian matrices over $\mathbb{C}$.

2. Diagonalizability

Notation. For $A \in \mathrm{Mat}_n(\Bbbk)$, we denote its eigenspaces by E_λ or $E_\lambda(A)$.

Definition 2. A matrix $A \in \mathrm{Mat}_n(\Bbbk)$ is **diagonalizable** iff:

- $\exists D \in \mathrm{Diag}_n(\Bbbk) : \exists P \in \mathrm{GL}_n(\Bbbk)$: $A = PDP^{-1}$.
- $\Bbbk^n$ admits a basis of eigenvectors of A.
- The minimal polynomial of A splits in $\Bbbk[x]$ and has distinct roots.

In particular, the columns of P are eigenvectors of A, with corresponding eigenvalues given by the diagonal of D.

Definition 3. Matrices $\{A_i\}$ are **simultaneously diagonalizable** if there exists a single matrix $P \in \mathrm{GL}(\mathbb{C})$ making all PA_iP^{-1} diagonal.

Theorem 4. Sets of diagonalizable matrices are simultaneously diagonalizable if and only if they commute.

Proof. Suppose A_i are simultaneously diagonalizable, with $A_i = PD_iP^{-1}$ for $P \in \mathrm{GL}_n(\Bbbk)$. Then they must commute because diagonal matrices commute.

$$A_iA_j = PD_iD_jP^{-1} = PD_jD_iP^{-1} = A_jA_i.$$

Conversely, use induction on the number of matrices r. The base case is clear. For $r \geq 2$, write $A_1, A_2, \ldots, A_r =: B$. First note that eigenspaces of a matrix are invariant under the action of any matrix it commutes with.

$$AB = BA \text{ and } v \in E_\lambda(A) \text{ implies } Bv \in E_\lambda(A) \text{ because } A(Bv) = BAv = \lambda(Bv).$$

To begin, note $\Bbbk^n$ is a direct sum $\bigoplus_\lambda E_\lambda(B)$ of the eigenspaces of B. By the above, we have that for all $v \in E_\lambda(B)$: $A_iv \in E_\lambda(B)$. Thus each A_i restricts to a linear map on $E_\lambda(B)$.

We now have maps $A_1|_{E_\lambda(B)}, \ldots, A_{r-1}|_{E_\lambda(B)}$. These commute since they commute in the entire space. Each is diagonalizable because their minimal polynomials divide the minimal polynomial of the corresponding A_i, and thus must have distinct factors. By the induction hypothesis, there is a basis of $E_\lambda(B)$ of common eigenvectors of $A_1|_{E_\lambda(B)}, \ldots, A_{r-1}|_{E_\lambda(B)}$. Each is also an eigenvector for $B|_{E_\lambda(B)}$ by definition. By combining the bases for each $E_\lambda(B)$, we obtain a full basis of eigenvectors since $\Bbbk^n = \bigoplus_\lambda E_\lambda(B)$. □

Theorem 5 (Spectral Theorem). The following are equivalent:

- A is normal.
- A is unitarily diagonalizable: $\exists D \in \mathrm{Diag}_n(\mathbb{C}), \exists U \in \mathcal{U}_n(\mathbb{C}) : A = UDU^{-1}$.
- $\mathbb{C}^n$ admits an orthonormal basis of eigenvectors of A.

If $A \in \mathrm{Mat}_n(\mathbb{C})$ is normal, then using $A = UDU^*$, we may write $A = \sum_j \vec{u}_j\lambda_j\vec{u}_j^*$. We call this a **spectral decomposition** of A. The eigenvalues and orthonormal eigenbasis for A

can be read off its spectral decomposition, since:

$$A\vec{u}_i = \sum_j \vec{u}_j \lambda_j \vec{u}_j^*(\vec{u}_i) = \sum_j \delta_{i,j} \lambda_j \vec{u}_j = \lambda_i \vec{u}_i,$$

shows that $(\lambda_i, \vec{u}_i)$ are eigenpairs of A. These are linearly independent since U is invertible. The verification that $\vec{u}_i$ are orthonormal is delegated to the section on unitary matrices below.

Example 6. Orthogonal, unitary, (skew) symmetric, and (skew) Hermitian matrices are all normal.

Unitary Matrices:

Unitary matrices are unitarily diagonalizable by the Spectral Theorem. The row and columns of A form orthonormal bases of $\mathbb{C}^n$, since for columns a_i, a_j of A:

$$\langle a_j, a_i \rangle = a_i^* \cdot a_j = (A^*A)_{i,j} = \delta_{i,j}.$$

Each $A \in \mathcal{U}_n(\mathbb{C})$ acts on $\mathbb{C}^n$ by an isometry (rotations and/or reflections), since:

$$\langle A\vec{x}, A\vec{y} \rangle = \langle \vec{x}, A^*A\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle,$$

for $\vec{x}, \vec{y} \in \mathbb{C}^n$. The eigenvalues (hence determinant) of A all have modulus 1, since:

$$|\lambda|^2 \langle \vec{v}, \vec{v} \rangle = \lambda\bar{\lambda} \langle \vec{v}, \vec{v} \rangle = \langle \lambda\vec{v}, \lambda\vec{v} \rangle = \langle A\vec{v}, A\vec{v} \rangle = \langle \vec{v}, \vec{v} \rangle,$$

for all eigenpairs $A\vec{v} = \lambda\vec{v}$. In particular, $\det(A) = \pm 1$ for orthogonal matrices.

Hermitian Matrices:

Hermitian matrices are unitarily diagonalizable by the Spectral Theorem. Each $A \in \mathcal{H}_n(\mathbb{C})$ defines a self-adjoint operator $\mathbb{C}^n$, since:

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^*\vec{y} \rangle = \langle \vec{x}, A\vec{y} \rangle.$$

All eigenvalues (hence determinant) of A are real, since for all eigenpairs $A\vec{v} = \lambda\vec{v}$:

$$\lambda \langle \vec{v}, \vec{v} \rangle = \langle \lambda\vec{v}, \vec{v} \rangle = \langle A\vec{v}, \vec{v} \rangle = \langle \vec{v}, A\vec{v} \rangle = \langle \vec{v}, \lambda\vec{v} \rangle = \bar{\lambda} \langle \vec{v}, \vec{v} \rangle.$$

Positive (Semi)-definite Matrices:

Let $A \in \mathcal{H}_n(\mathbb{C})$ have rank r. Then, the following are equivalent:

(1) A is positive definite (resp. semi-definite),

(2) all eigenvalues of A are positive (resp. non-negative),

(3) A factors as $A = B^*B$ for some $B \in \mathrm{GL}_n(\mathbb{C})$ (resp. $B \in \mathrm{Mat}_{r\times n}(\mathbb{C})$),

(4) the assignment $\langle \vec{x}, \vec{y} \rangle_A := \vec{y}^* A\vec{x}$ defines an inner product on $\mathbb{C}^n$ (resp. positive semi-definite Hermitian form).

Starting with (1) and (4), the assignment $\langle \vec{x}, \vec{y} \rangle_A$ is automatically sesquilinear and conjugate symmetric. It is positive definite if and only if A is.

For (1) and (2), if A is positive definite with eigenpair $(\lambda, \vec{v})$, then:

$$\lambda||\vec{v}||^2 = \lambda\langle \vec{v}, \vec{v}\rangle = \lambda(\vec{v}^* \cdot \vec{v}) = \vec{v}^*(\lambda\vec{v}) = \vec{v}^* A\vec{v} > 0,$$

shows that λ must be positive. Conversely, if all eigenvalues are positive, then for all nonzero $\vec{x} \in \mathbb{C}^n$, the spectral decomposition of A gives:

$$\vec{x}^* A\vec{x} = \vec{x}^* U^* DU\vec{x} = (U\vec{x})^* D(U\vec{x}).$$

The latter is a weighted version of $\langle U\vec{x}, U\vec{x}\rangle = ||U\vec{x}||^2$, with weights given by the eigenvalues of A. Since the eigenvalues are positive and $||U\vec{x}||^2 > 0$ because U^* is invertible, we get $\vec{x}^* A\vec{x} > 0$.

Finally for (1) and (3), if $A = B^*B$, then $\vec{x}^* B^* B\vec{x} = (B\vec{x})^*(B\vec{x}) = ||B\vec{x}||^2 > 0$ for all nonzero $\vec{x}$ since B is invertible. Conversely, if A is positive definite, let $A = U^*DU$ be a spectral decomposition. Let $\sqrt{D}$ denote the diagonal matrix obtained by taking the positive square roots of the entries of D. This exists and satisfies $(\sqrt{D})^* = \sqrt{D}$ because all eigenvalues of A are positive real numbers. Let $B = \sqrt{D}U$. Then:

$$A = U^*DU = U^*\sqrt{D}\sqrt{D}U = U^*(\sqrt{D})^*\sqrt{D}U = (\sqrt{D}U)^*\sqrt{D}U = B^*B.$$

3. Matrix Decompositions

3.1. **Singular Value Decomposition.** For any $A \in \mathrm{Mat}_{m\times n}(\mathbb{C})$, notice that A^*A and AA^* are Hermitian, since $(A^*A)^* = A^*(A^*)^* = A^*A$.

Definition 7. We call $\sigma > 0$ a **singular value** of $A \in \mathrm{Mat}_{m\times n}(\mathbb{C})$ if there exists unit vectors $\vec{u} \in \mathbb{C}^m$ and $\vec{v} \in \mathbb{C}^n$ satisfying $A\vec{v} = \sigma\vec{u}$ and $A^*\vec{u} = \sigma\vec{v}$. We call $\vec{u}$ a **left singular vector** and $\vec{v}$ a **right singular vector**.

Proposition 8. The singular values of A are precisely the square root of the eigenvalues of A^*A or AA^*.

Proof. Suppose σ is a singular value of A with left/right singular vectors $\vec{u}, \vec{v}$. Then $\vec{v} \in \mathbb{C}^n$ is an eigenvector of A^*A and $\vec{u} \in \mathbb{C}^m$ is an eigenvector of AA^*, since:

$$A^*A\vec{v} = A^*(\sigma\vec{u}) = \sigma(A^*\vec{u}) = \sigma^2\vec{v}.$$

and:

$$AA^*\vec{u} = A(\sigma\vec{v}) = \sigma(A\vec{v}) = \sigma^2\vec{u}.$$

□

Decomposition	**Formula**	**Description**
Rank Factorization	$A = CR$	$C \in \mathrm{Mat}_{m\times r}$ with full column rank $R \in \mathrm{Mat}_{r\times n}$ with full row rank
QR-Decomposition	$A = QR$	$A \in \mathrm{Mat}_{m\times n}$ with linearly independent columns (necessarily $m \geq n$ for A to be full column rank) $Q \in \mathcal{U}_m$: unitary/orthogonal $R \in \mathrm{Mat}_{m\times n}$: upper triangular
	$A = LQ$	$L \in \mathrm{Mat}_{m\times n}$: lower triangular $Q \in \mathcal{U}_n$: unitary/orthogonal
SVD	$A = U\Sigma V^*$	$U \in \mathcal{U}_m$: left singular vectors, ON eigenbasis of AA^* $V \in \mathcal{U}_n$: right singular vectors, ON eigenbasis of A^*A $\Sigma \in \mathrm{Diag}_{m\times n}$: singular values, $\sqrt{\text{EV}}$'s of A^*A

4. Square Matrix Decompositions

Decomposition	Formula	Description
LU-Decomposition	$A = LU$	L: lower triangular U: upper triangular
	$A = LDU$	D: diagonal L, U: unitriangular (1's in diagonal)
	$PAQ = LU$	P, Q: permutation matrices. use to avoid divergence/zero division
Jordan Normal Form	$A = PJP^{-1}$	$P \in \mathrm{GL}_n$: invertible, generalized eigenvectors J: block diagonal $J = \mathrm{diag}(J_1, \ldots, J_p)$ J_k: diagonals correspond to eigenvalue λ_i, with super diagonal of 1's for non 1×1 blocks.
Schur Decomposition	$A = QUQ^*$	$Q \in \mathcal{U}_n$: unitary $U \in \mathrm{Mat}_n$: upper triangular with eigenvalues of A in the diagonal. If A is normal, U is diagonal and this reduces to spectral decomposition.
Polar Decomposition	$A = UP$ $= P'U$	U: unitary P, P': positive semi-definite Hermitian

5. Symmetric/Hermitian Matrix Decompositions

Decomposition	Formula	Description
Cholesky Decomposition	$A = LL^*$	A: positive (semi)definite L: (non)unique lower triangular with positive (non-negative) diagonal entries
	$A = U^*U$	U: (non)unique upper triangular with positive (non-negative) diagonal entries
	$A = LDL^*$ $= U^*DU$	D: diagonal L, U: (non)unique unitriangular (algorithms avoid square root compuations)

Email address: fllfernando95@gmail.com

URL: https://sites.google.com/rice.edu/fernando-liu-lopez/