Linear Analysis
Lecture 11

Unitary Equivalence

Recall that similar matrices represent the same linear transformation in different bases. We now focus on a very important class of similarity transformations.

Definition. We say that $A, B \in \mathbb{C}^{n\times n}$ are **unitarily equivalent** (or unitarily similar) if there is a unitary matrix $U \in \mathbb{C}^{n\times n}$,

$$U^H = U^{-1}, \quad \text{such that} \quad B = U^H A U.$$

Unitary equivalence is important for several reasons. Consider the following computational reasons.
(1) The inverse of a unitary matrix is easy to compute!
(2) Unitary matrices are perfectly conditioned in inner product spaces under the inner product operator norm.

$$(\forall\, x \in \mathbb{C}^n)\ \|Ux\| = \|x\| = \|U^H x\| \ \Rightarrow\ \|U\| = \|U^H\| = \|U^{-1}\| = 1\ .$$

(3) Unitary similarity transformations preserve the condition number (again in the inner product operator norm).

$$\kappa(U^H A U) = \|U^H A U\| \, \|U^H A^{-1} U\| \le \|A\| \, \|A^{-1}\| = \kappa(A) \; ;$$

similarly $\kappa(A) \le \kappa(U^H A U)$.

In general, a similarity transformation may degrade the condition of a matrix.

$$\begin{aligned} \kappa(S^{-1}AS) &= \|S^{-1}AS\| \cdot \|S^{-1}A^{-1}S\| \\ &\le \|S^{-1}\|^2 \|A\| \cdot \|A^{-1}\| \cdot \|S\|^2 = \kappa(S)^2 \kappa(A) \end{aligned}$$

and, similarly,

$$\kappa(A) \le \kappa(S)^2 \kappa(S^{-1}AS).$$

Preservation of Norms: Isometry

Proposition: Let $U \in \mathbb{C}^{n\times n}$ be unitary, and $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times k}$.

(1) In the operator norms induced by the Euclidean norms,

$$\|AU\| = \|A\| \quad \text{and} \quad \|UB\| = \|B\|.$$

(2) In the Frobenius norms,

$$\|AU\|_F = \|A\|_F \quad \text{and} \quad \|UB\|_F = \|B\|_F.$$

So multiplication by a unitary matrix preserves $\|\cdot\|$ and $\|\cdot\|_F$.

Proof Sketch.

(1) $$(\forall\, x \in \mathbb{C}^k)\ \|UBx\| = \|Bx\|, \quad \text{so} \quad \|UB\| = \|B\|.$$

Likewise, since U^H is also unitary,

$$\|AU\| = \|(AU)^H\| = \|U^H A^H\| = \|A^H\| = \|A\|.$$

(2) Let $b_1, \ldots, b_k$ be the columns of B. Then

$$\|UB\|_F^2 = \sum_{j=1}^{k} \|Ub_j\|_2^2 = \sum_{j=1}^{k} \|b_j\|_2^2 = \|B\|_F^2.$$

Likewise, since U^H is also unitary,

$$\|AU\|_F = \|U^H A^H\|_F = \|A^H\|_F = \|A\|_F.$$

□

Schur Unitary Triangularization Theorem

Every $A \in \mathbb{C}^{n\times n}$ is unitarily equivalent to an upper triangular T.
If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A in any prescribed order, then one can choose a unitary similarity transformation so that the diagonal entries of T are $\lambda_1, \ldots, \lambda_n$ in that order.

Proof Sketch: Induction on n: Obvious for $n = 1$. Assume true for $n-1$. Given $A \in \mathbb{C}^{n\times n}$ and an ordering $\lambda_1, \ldots, \lambda_n$ of its eigenvalues, choose an eigenvector x for λ_1 with Euclidean norm $\|x\| = 1$.
Extend $\{x\}$ to a basis of $\mathbb{C}^n$ and apply the Gram-Schmidt procedure to obtain an orthonormal basis $\{x, u_2, \ldots, u_n\}$ of $\mathbb{C}^n$. Then

$$U_1 = [x,\ u_2, \cdots,\ u_n] \in \mathbb{C}^{n\times n} \qquad \text{is unitary.}$$

Since $Ax = \lambda_1 x$,

$$U_1^H A U_1 = \begin{bmatrix} \lambda_1 & y_1^H \\ 0 & B \end{bmatrix} \quad \text{for some } y_1 \in \mathbb{C}^{n-1}, \quad B \in \mathbb{C}^{(n-1)\times(n-1)}.$$

Since similar matrices have the same characteristic polynomial,

$$p_A(t) = \det\left(tI - \begin{bmatrix} \lambda_1 & y_1^H \\ 0 & B \end{bmatrix}\right) = (t-\lambda_1)\det(tI-B) = (t-\lambda_1)p_B(t),$$

so the eigenvalues of B are $\lambda_2, \ldots, \lambda_n$.

Schur Unitary Triangularization Theorem

By the induction hypothesis, there is a unitary $\widetilde{U} \in \mathbb{C}^{(n-1)\times(n-1)}$ and upper triangular $\widetilde{T} \in \mathbb{C}^{(n-1)\times(n-1)}$ such that $\widetilde{U}^H B \widetilde{U} = \widetilde{T}$ and the diagonal entries on $\widetilde{T}$ are $\lambda_2, \ldots, \lambda_n$ in that order.

Let

$$U_2 = \begin{bmatrix} 1 & 0 \\ 0 & \widetilde{U} \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

Then U_2 is unitary, and

$$U_2^H U_1^H A U_1 U_2 = \begin{bmatrix} \lambda_1 & y_1^H \widetilde{U} \\ 0 & \widetilde{U}^H B \widetilde{U} \end{bmatrix} = \begin{bmatrix} \lambda_1 & y_1^H \widetilde{U} \\ 0 & \widetilde{T} \end{bmatrix} \equiv T.$$

Since $U \equiv U_1 U_2$ is unitary and $U^H A U = T$, the statement is true for n as well. □

Note: The basic iterative step that reduces the dimension by 1 is called a deflation. Deflation is used to derive many of important matrix factorizations.

Spectral Theorem for Normal Operators

Unitary equivalence preserves the classes of Hermitian, skew-Hermitian, and normal matrices.

Spectral Theorem: *Let $A \in \mathbb{C}^{n\times n}$ be normal. Then A is unitarily diagonalizable. That is, A is unitarily similar to a diagonal matrix; so there is an orthonormal basis of eigenvectors.*

Proof Sketch: By Schur Triangularization Theorem, there is a unitary $U \in \mathbb{C}^{n\times n}$ such that $U^H AU = T$ is upper triangular. Since A is normal, T is normal: $T^H T = U^H A^H AU = U^H AA^H U = TT^H$.

By equating diagonal entries of $T^H T$ and TT^H, we show T is diagonal.

The $(1,1)$ entries of $T^H T$ and TT^H are

$$|t_{11}|^2 \quad \text{and} \quad \sum_{j=1}^{n} |t_{1j}|^2, \quad \text{resp.ly.}$$

Hence $t_{1j} = 0 \quad \text{for} \quad j \geq 2$. Now the $(2,2)$ entries of $T^H T$ and TT^H are

$$|t_{22}|^2 \quad \text{and} \quad \sum_{j=2}^{n} |t_{2j}|^2, \quad \text{resp.ly.}$$

Hence, again,

$$t_{2j} = 0 \quad \text{for} \quad j \geq 3.$$

Continuing with the remaining rows yields the result. □

The Spectral Radius and Normal Operators

A useful tool in the analysis of operators is the **spectral radius**.

Given $L \in \mathcal{L}(V)$ with $\dim(V) < \infty$, the spectral radius of L is

$$\rho(L) = \max_{\lambda \in \sigma(L)} |\lambda| \ .$$

The spectral radius and the operator norm in a Euclidean space have an important relationship. This relationship is studied more deeply later, but for now, we have the following.

Fact: For any $A \in (\mathbb{C}^{n \times n}, \|\cdot\|)$ and any operator norm $\|\!|\cdot|\!\|$

$$\rho(A) \ \leq \|\!|A|\!\| \, ,$$

with equality if A is normal and $\|\!|\cdot|\!\| = \|\!|\cdot|\!\|_2$.

Proof: Let $\lambda \in \Sigma(A)$ be such that $|\lambda| = \rho(A)$, and let v be an associated eigenvector with $\|v\| = 1$. Then

$$\|\!|A|\!\| = \sup_{\|x\|=1} \|Ax\| \geq \|Av\| = |\lambda| = \rho(A) \ .$$

The Spectral Radius and Normal Operators

Next suppose A is normal with eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. The Spectral Theorem $\Rightarrow$ $\exists$ unitary $U \in \mathbb{C}^{n\times n}$ such that $U^H AU = D$, where $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.
Given $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$, set $v = U^H x$ so $\|v\|_2 = \|x\|_2 = 1$. Then

$$\begin{aligned}
\|Ax\|_2^2 &= \|U^H AUU^H x\|_2^2 = \|Dv\|_2^2 \\
&= \sum_{k=1}^{n} |\lambda_k v_i|^2 = \sum_{k=1}^{n} |\lambda_k|^2 |v_i|^2 \\
&\leq \left(\max_{k=1,\ldots,n} |\lambda_k|^2 \right) \sum_{k=1}^{n} |v_i|^2 \\
&= \rho^2(A) \ .
\end{aligned}$$

Therefore, $\|\|A\|\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2 \leq \rho(A)$, so $\|\|A\|\|_2 = \rho(A)$. $\square$

The Cayley-Hamilton Theorem

Theorem: (Cayley-Hamilton) *Every matrix* $A \in \mathbb{C}^{n\times n}$ *satisfies its characteristic polynomial, i.e.* $p_A(A) = 0$.

Proof: By Schur, $\exists$ unitary $U \in \mathbb{C}^{n\times n}$ and upper triangular $T \in \mathbb{C}^{n\times n}$ such that $U^H AU = T$, where the diagonal entries of T are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A (in some order). Since

$$A = UTU^H \quad \text{and} \quad A^k = UT^kU^H,$$

we have $p_A(A) = Up_A(T)U^H$.

Writing $p_A(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$ gives

$$p_A(T) = (T - \lambda_1 I)(T - \lambda_2 I)\cdots(T - \lambda_n I).$$

Since $T - \lambda_j I$ is upper triangular with its jj entry being zero, it follows that $p_A(T) = 0$. To see this, accumulate the product from the left, in which case one shows by induction on k that the first k columns of

$$(T - \lambda_1 I)\cdots(T - \lambda_k I)$$

are zero. $\square$