MA-GY 7043: Linear Algebra II

Unitarily Equivalent Matrices
Schur Decomposition

Deane Yang

Courant Institute of Mathematical Sciences
New York University

February 27 2025

Outline I

Unitarily Equivalent Matrices

- $M_1, M_2 \in \mathrm{gl}(n, \mathbb{F})$ are **unitarily equivalent** if there exists a unitary matrix U such that

$$M_2 = UM_1U^*$$

- Since $U^* = U^{-1}$, unitarily equivalent implies similar

Unitary Equivalence to Diagonal Matrix

- **Recall:** A matrix D is diagonal if and only if the standard basis vectors are eigenvectors
- Given a matrix M, the following are equivalent:
  - M is similar to a diagonal matrix D, i.e., there exists an invertinble matrix S such that $M = SDS^{-1}$
  - There exists a basis of eigenvectors
- If this holds, then the columns of S are a basis of eigenvectors
- In particular, a matrix M is unitarily equivivalent to a diagonal matrix if and only if there is a unitary basis of eigenvectors
- For each $1 \leq k \leq n$, let $f_k = U(e_k)$
- For each $1 \leq j, k \leq n$,

$$(f_j, f_k) = (U(e_j), U(e_k)) = (e_j, e_k) = \delta_{jk}$$

- Moreover, if $De_k = \lambda_k e_k$, then

$$Mf_k = UDU^* Ue_k = UDe_k = U(\lambda_k e_k) = \lambda_k Ue_k = \lambda_k f_k$$

- Therefore, $(f_1, \ldots, f_n)$ is a unitary basis of eigenvectors
- Converse is even easier

Hermitian Inner Product With Respect To Basis

- Let V be a complex vector space and let $(b_1, \ldots, b_n)$ be a basis of V
- Any inner product on V is uniquely determined by the matrix A, where
$$A_{ij} = (b_i, b_j)$$
- The matrix A satisfies the following properties
  - Hermitian:
  $$A_{ij} = (b_i, b_j) = \overline{(b_j, b_i)} = \bar{A}_{ji}$$
  (In particular, since $A_{ii} = \bar{A}_{ii}$, it follows that $A_{ii} \in \mathbb{R}$)
  - Positive definite: For any nonzero $v = a^k b_k = Ba \in V$,
  $$0 < (v, v) = (a^j b_j, a^k b_k) = a^j \bar{a}^k (b_j, b_k) = a^T A \bar{a}$$
- Conversely, given the basis $(b_1, \ldots, b_n)$ of V, any positive definite Hermitian matrix A defines an inner product where
$$(b_i, b_j) = A_{ij}$$

Schur Decomposition of a Real Linear Map

- Let V be a finite dimensional real inner product space
- **Theorem:** Given any linear map $L : V \to V$ with only real eigenvalues, there exists an **orthonormal** basis $E = (e_1, \ldots, e_n)$ of V such that for each $1 \leq k \leq n$, $L(e_k)$ is a linear combination of $e_1, \ldots, e_k$,
$$L(e_k) = e_k M_k^k + \cdots + e_n M_k^n$$
- In other words, there exists an orthonormal basis E such that
$$L(E) = EM,$$
where M is a lower triangular matrix.
- **Corollary:** Given any real matrix M with only real eigenvalues, there is an orthogonal matrix O such that the matrix $O^t MO$ is lower triangular

Schur Decomposition of a Complex Linear Map

- **Theorem:**
  - Let V be an n-dimensional Hermitian vector space over $\mathbb{F}$
  - Let $L : V \to V$ be a linear map with n eigenvalues in $\mathbb{F}$, counting multiplicity

  Then there exists a **unitary** basis $U = (u_1, \ldots, u_n)$ of V such that for each $1 \leq k \leq n$, $L(u_k)$ is a linear combination of $u_1, \ldots, u_k$,
  $$L(u_k) = u_k M_k^k + \cdots + u_n M_k^n$$
- In other words, there exists a unitary basis U such that
  $$L(U) = UM,$$
  where M is a lower triangular matrix
- **Corollary:** Given any complex matrix M, there is a unitary matrix U such that the matrix $U^* MU$ is triangular

Proof (Part 1)

- Proof by induction
- Theorem holds when $\dim V = 1$
- Suppose theorem holds when $\dim V = n - 1$
- Consider a linear map $L : V \to V$, where $\dim V = n$ with eigenvalues $\lambda_1, \dots, \lambda_n$
- Let u_n be a unit eigenvector for the eigenvalue λ_n, i.e.,
$$\|u_n\| = 1 \text{ and } L(u_n) = \lambda_n u_n$$
- Let
$$u_n^\perp = \{v \in V \ : \ (v, u_n) = 0\}$$
- Recall that the orthogonal projection maps onto $[u_n$ and $u_n^\perp$ respectively are
$$\begin{aligned} \pi : V &\to [u_n] \\ v &\mapsto (v, u_n)u_n \\ \pi^\perp : V &\to u_n^\perp \\ v &\mapsto v - u_n(v, u_n) \end{aligned}$$

Proof (Part 2)

- If $(v_1, \ldots, v_{n-1})$ is a basis of $u_n^\perp$, then $(v_1, \ldots, v_{n-1}, u_n)$ is a basis of V
- Let M be the matrix such that for $1 \leq k \leq n-1$,

$$L(v_k) = v_1 M_k^1 + \cdots + v_{n-1} M_k^{n-1} + u_n M_k^n$$

and

$$L(u_n) = v_1 M_n^1 + \cdots + v_{n-1} M_n^{n-1} + u_n M_n^n$$

Proof (Part 3)

- Since $L(u_n) = \lambda_n u_n$,

$$M_n^1 = \cdots = M_n^{n-1} = 0 \text{ and } M_n^n = \lambda_n$$

and therefore M is of the form

$$M = \left[\begin{array}{ccc|c} M_1^1 & \cdots & M_{n-1}^1 & M_n^1 \\ \vdots & & \vdots & \vdots \\ M_1^{n-1} & \cdots & M_{n-1}^{n-1} & M_n^{n-1} \\ \hline 0 & \cdots & 0 & M_n^n \end{array}\right]$$

- Let

$$\widehat{M} = \begin{bmatrix} M_1^1 & \cdots & M_{n-1}^1 \\ \vdots & & \vdots \\ M_1^{n-1} & \cdots & M_{n-1}^{n-1} \end{bmatrix}$$

- It follows that

$$p_L(x) = \det(L - xI) = \det(M - xI = \det(\widehat{M} - xI)(M_n^n - x)$$

- In particular, the eigenvalues of $\widehat{M}$ are eigenvalues of M

Proof (Part 4)

- Let $L^{\perp} : u_n^{\perp} \to u_n^{\perp}$ be the linear transformation given by
$$L^{\perp}(v) = \pi^{\perp}\left(L|_{u_n^{\perp}}(v) \right)$$
- By assumption, $L^{\perp}$ has a Schur decomposition, i.e., a unitary basis $u_1, \ldots, u_{n-1}$ and a lower triangular matrix $\widehat{M}$ such that for each $1 \leq k \leq n-1$,
$$L^{\perp}(u_k) = u_k \widehat{M}_k^k + \cdots + u_{n-1} \widehat{M}_k^{n-1}$$
and therefore
$$L(u_k) = u_k \widehat{M}_k^k + \cdots + u_{n-1} \widehat{M}_k^{n-1} + u_n M_k^n$$
- Also,
$$L(u_n) = \lambda_n u_n$$
- Therefore,
$$L(u_k) = M_k^k u_k + \cdots + M_k u^{n-1} u_{n-1} + M_k^n u_n,\ 1 \leq k \leq n,$$
where $M_n^n = \lambda_n$
- This proves the theorem