Lecture 3: Outline

- Ch. 2: Unitary equiv, QR factorization, Schur's thm, Cayley-H., Normal matrices, Spectral thm, Singular value decomp.
- Ch. 3: Canonical forms: Jordan/Matrix factorizations

Unitary matrices

- A set of vectors $\{x_i\} \in \mathbf{C}^n$ are called
  - *orthogonal* if $x_i^* x_j = 0, \; \forall \, i \neq j$ and
  - *orthonormal* if they are orthogonal and $x_i^* x_i = 1, \; \forall i.$
- A matrix $U \in M_n$ is *unitary* if $U^* U = I$.
- A matrix $U \in M_n(\mathbf{R})$ is *real orthogonal* if $U^T U = I$.
- (A matrix $U \in M_n$ is *orthogonal* if $UU^T = I$.)
- If U, V are unitary then UV is unitary.
  - Unitary matrices form a group under multiplication.

Unitary matrices cont'd

The following are equiv.

1. U is unitary
2. U is nonsingular and $U^{-1} = U^*$
3. $UU^* = I$
4. U^* is unitary
5. the columns of U are orthonormal
6. the rows of U are orthonormal
7. for all $x \in \mathbf{C}^n$, the Euclidean length of $y = Ux$ equals that of x.

Def: A linear transformation $T : \mathbf{C}^n \to \mathbf{C}^m$ is a *Euclidean isometry* if $x^* x = (Tx)^*(Tx)$ for all $x \in \mathbf{C}^n$

Unitary U is an Euclidean isometry.

Euclidean isometry and Parseval's Theorem

1. F_N be the FFT (Fast Fourier Transform matrix) of dimension $N \times N$, i. e,
$$F_N(m,n) = \frac{1}{\sqrt{N}} e^{\frac{-2\pi(m-1)(n-1)}{N}}$$
2. F is a unitary matrix.
3. Let $y = F_N x$ i.e, y is the N point FFT of x.
   (a) Length of x = Length of y
   (b) $\sum_{j=1}^{N} |x(j)|^2 = \sum_{j=1}^{N} |y(j)|^2$: This is energy conservation or Parseval's Theorem in DSP.

Unitary equivalence

Def: A matrix $B \in M_n$ is *unitarily equivalent (or similar)* to $A \in M_n$ if $B = U^* A U$ for some unitary matrix U.

(i) $A \to S^{-1}AS$: similarity (Ch 1,3)

(ii) $A \to S^* AS$: *congruence (Ch 4)

(iii) $A \to S^* AS$ with S unitary : unitary similarity (Ch 2)

Since in (iii) $S^* = S^{-1}$, we have that (iii) is "included" in both (i) and (ii).

Theorem: If A and B are unitarily equivalent then

$$\|A\|_F^2 \triangleq \sum_{i,j} |a_{ij}|^2 = \sum_{i,j} |b_{ij}|^2 = \|B\|_F^2$$

Unitary matrices and Plane Rotations : 2-D case

- Consider rotating the $2-D$ Euclidean plane counter-clockwise by an angle θ.
- Resulting coordinates,

$$x' = x\cos\theta - y\sin\theta$$
$$y' = x\sin\theta + y\cos\theta$$

- Equivalently,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- Note that $U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is unitary.

Unitary matrices and Plane Rotations : General Case

$$U(\theta, 2, 4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 0 & 1 & 0 \\ 0 & \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

- $U(\theta, 2, 4)$ rotates the *second* and *fourth* axes in $\mathbf{R}^4$ counter clock-wise by θ.
- The other axes are not changed.
- Left multiplication by $U(\theta, 2, 4)$ affects only rows 2 and 4.
- Note that $U(\theta, 2, 4)$ is unitary.
- Such $U(\theta, m, n)$ are called Givens rotations.

Product of Givens rotations

- $U = U(\theta_1, 1, 3)U(\theta_2, 2, 4)$ rotates
  - *second* and *fourth* axes in $\mathbf{R}^4$ counter clock-wise by θ_2.
  - *first* and *third* axes in $\mathbf{R}^4$ counter clock-wise by θ_1.
- U is unitary $\Rightarrow$ product of Givens rotations is unitary.
- Such matrices are used in Least-Squares and eigenvalue computations.

Special Unitary matrices: Householder matrices

Let $w \in \mathbf{C}^n$ be a normalized ($w^*w = 1$) vector and define

$$U_w = I - 2ww^*$$

Properties:

1. U_w is unitary and Hermitian.
2. $U_w x = x, \forall\, x \perp w.$
3. $U_w w = -w$
4. There is a Householder matrix such that

$$y = U_w x$$

for any given *real* vectors x and y of the same length.

QR-factorization

Thm: If $A \in M_{n,m}$ and $n \geq m$, then

$$A = QR$$

with $Q \in M_{n,m}$ such that $Q^*Q = I$ and $R \in M_m$ is upper triangular.

- If $m = n$ then Q is unitary.
- If A is nonsingular, then the diagonal elements of R can be taken to be positive (Q and R are in this case unique).
- Gram Schmidt orthogonalization followed by book keeping.
- Useful in Least squares solutions, eigenvalue computations etc.

The QR-algorithm for eigenvalue computation

Initialization: Given $A \in M_n$, set $k = 0$, $A_k = A$,

1. $A_k = Q_k R_k$ (QR decomposition)
2. Obtain $A_{k+1} = R_k Q_k$.
3. Set $k = k + 1$ and go to Step 1.

Prove : A_k and A are unitarily equivalent for all k.

Under certain conditions A_k converges to an upper triangular matrix (as $k \to \infty$) whose diagonal elements correspond to the eigenvalues of A (see p. 114-115 in Horn and Johnson).

Schur's unitary triangularization thm

Theorem: Given $A \in M_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, there is a unitary matrix $U \in M_n$ such that

$$U^*AU = T = [t_{ij}]$$

is upper triangular with $t_{ii} = \lambda_i$ $(i = 1, \ldots, n)$ in any prescribed order. If $A \in M_n(\mathbf{R})$ and all λ_i are real, U may be chosen real and orthogonal.

Consequence: Any matrix in M_n is unitarily similar to an upper (or lower) triangular matrix. Note that $A = UTU^*$.

Uniqueness:

1. Neither U nor T is unique.
2. Eigenvalues can appear in any order
3. Two triangular matrices can be unitarily similar

SCHUR: THE GENERAL REAL CASE

Given $A \in M_n(\mathbf{R})$, there is a real orthogonal matrix $Q \in M_n(\mathbf{R})$ such that

$$Q^T A Q = \begin{bmatrix} A_1 & * & \dots & * \\ 0 & A_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & A_k \end{bmatrix} \in M_n(\mathbf{R})$$

where A_i $(i = 1, \dots, k)$ are real scalars or 2 by 2 blocks with a non-real pair of complex conjugate eigenvalues.

IMPLICATIONS OF THE SCHUR THEOREM

1. $\mathrm{tr} A = \sum_j \lambda_j(A)$
2. $\det A = \prod_j \lambda_j(A)$
3. Cayley-Hamilton theorem.
4. ...

CAYLEY-HAMILTON THEOREM

Let $p_A(t) = \det(tI - A)$ be the characteristic polynomial of $A \in M_n$. Then

$$p_A(A) = 0$$

Consequences:

- $A^{n+k} = q_k(A)$ $(k \geq 0)$ for some polynomials $q_k(t)$ of degrees $\leq n - 1$.
- If A is nonsingular: $A^{-1} = q(A)$ for some polynomial $q(t)$ of degree $\leq n - 1$.

Important : Note $p_A(C)$ is a matrix valued function.

NORMAL MATRICES

Def: A matrix $A \in M_n$ is *normal* if $A^* A = A A^*$.

Examples:

All unitary matrices are normal.

All Hermitian matrices are normal.

Def: $A \in M_n$ is *unitarily diagonalizable* if A is unitarily equivalent to a diagonal matrix.

Facts for normal matrices

The following are equivalent:

1. A is normal
2. A is *unitarily diagonalizable*
3. $\|A\|_F^2 \triangleq \sum_{i,j} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$
4. there is an orthonormal set of n eigenvectors of A

The equivalence of 1 and 2 is called "the *Spectral Theorem for Normal matrices*."

Important special case: Hermitian (sym) matrices

Spectral theorem for Hermitian matrices:
If $A \in M_n$ is Hermitian, then,

- all eigenvalues are real
- A is unitarily diagonalizable.

If $A \in M_n(\mathbf{R})$ is symmetric, then A is real orthogonally diagonalizable.

SVD: Singular Value Decomposition

Theorem: Any $A \in M_{m,n}$ can be decomposed as

$$A = V \Sigma W^*$$

- $V \in M_m$: Unitary with columns being eigenvectors of AA^*.
- $W \in M_n$: Unitary with columns being eigenvectors of A^*A.
- $\Sigma = [\sigma_{ij}] \in M_{m,n}$ has $\sigma_{ij} = 0, \ \forall\, i \neq j$

Suppose $\operatorname{rank}(A) = k$ and $q = \min\{m, n\}$, then

- $\sigma_{11} \geq \cdots \geq \sigma_{kk} > \sigma_{k+1,k+1} = \cdots = \sigma_{qq} = 0$
- $\sigma_{ii} \equiv \sigma_i$ square roots of non-zero eigenvalues of AA^* (or A^*A)
- Unique : σ_i, Non-unique : V, W

Canonical forms

- An equivalence relation partitions the domain.
- Simple to study equivalence if two objects in an equivalence class can be related to one *representative* object.
- Requirements of the *representatives*
  - Belong to the equivalence class.
  - One per class.
- Set of such *representatives* is a *Canonical form*
- We are interested in a canonical form for equivalence relation defined by similarity.

Canonical forms: Jordan form

Every equivalence class under similarity contains *essentially* only one, so called, Jordan matrix:

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{n_k}(\lambda_k) \end{bmatrix}$$

where each block $J_k(\lambda) \in M_k$ has the structure

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \lambda & 1 \\ 0 & & & & \lambda \end{bmatrix}$$

The Jordan form theorem

Note that the orders n_i or λ_i are generally not distinct.

Theorem: For a given matrix $A \in M_n$, there is a nonsingular matrix $S \in M_n$ such that $A = SJS^{-1}$ and $\sum_i n_i = n$. The Jordan matrix is unique up to permutations of the Jordan blocks.

The Jordan form may be numerically unstable to compute but it is of theoretical interest.

Jordan form cont'd

- The number k of Jordan blocks is the number of linearly independent eigenvectors. (Each block is associated with an eigenvector from the standard basis.)
- J is diagonalizable iff $k = n$.
- The number of blocks corresponding to the same eigenvalue is the geometric multiplicity of that eigenvalue.
- The sum of the orders (dimensions) of all blocks corresponding to the same eigenvalue equals the algebraic multiplicity of that eigenvalue.

Applications of the Jordan form

Linear systems:

$$\dot{x}(t) = Ax(t); \;\; x(0) = x_0$$

The solution may be "easily" obtained by changing state variables to the Jordan form.

Convergent matrices: If all elements of A^m tend to zero as $m \to \infty$, then A is a *convergent matrix*. Fact: A is convergent iff $\rho(A) < 1$ (that is, iff $|\lambda_i| < 1, \; \forall i$). This may be proved, e.g., by using the Jordan canonical form.

TRIANGULAR FACTORIZATIONS

Linear systems of equations are easy to solve if we can factorize the system matrix as $A = LU$ where L (U) is lower (upper) triangular.

Theorem: If $A \in M_n$, then there exist permutation matrices $P, Q \in M_n$ such that

$$A = PLUQ$$

(in some cases we can take $Q = I$ and/or $P = I$).

WHEN TO USE WHAT?

	Theoretical derivations	Practical implem.
Schur triangularization	☺	☹
QR factorization	☺	☺
Spectral dec.	☺	☺(?)
SVD	☺	☺
Jordan form	☺	☹!!