MATH 311-504

Topics in Applied Mathematics

Lecture 3-9:
Symmetric and orthogonal matrices.

Definition. An $n \times n$ matrix A is called

- **symmetric** if $A^T = A$;
- **orthogonal** if $AA^T = A^T A = I$, that is, if $A^T = A^{-1}$;
- **normal** if $AA^T = A^T A$.

Proposition For any $n \times n$ matrix A and any column vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\boxed{A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^T \mathbf{y}.}$

Theorem If **x** and **y** are eigenvectors of a symmetric matrix A associated with different eigenvalues, then $\mathbf{x} \cdot \mathbf{y} = 0$.

Theorem If A is a normal matrix then
$\mathrm{Null}(A) = \mathrm{Null}(A^T)$ (that is, $A\mathbf{x} = \mathbf{0} \Longleftrightarrow A^T\mathbf{x} = \mathbf{0}$).

Proof: $A\mathbf{x} = \mathbf{0} \Longleftrightarrow A\mathbf{x} \cdot A\mathbf{x} = 0 \Longleftrightarrow \mathbf{x} \cdot A^T A\mathbf{x} = 0$
$\Longleftrightarrow \mathbf{x} \cdot AA^T\mathbf{x} = 0 \Longleftrightarrow A^T\mathbf{x} \cdot A^T\mathbf{x} = 0 \Longleftrightarrow A^T\mathbf{x} = \mathbf{0}$.

Proposition If a matrix A is normal, so are matrices $A - \lambda I$, $\lambda \in \mathbb{R}$.

Proof: Let $B = A - \lambda I$, where $\lambda \in \mathbb{R}$. Then
$B^T = (A - \lambda I)^T = A^T - (\lambda I)^T = A^T - \lambda I$.
We have $BB^T = (A - \lambda I)(A^T - \lambda I) = AA^T - \lambda A - \lambda A^T + \lambda^2 I$,
$B^T B = (A^T - \lambda I)(A - \lambda I) = A^T A - \lambda A - \lambda A^T + \lambda^2 I$.
Hence $AA^T = A^T A \implies BB^T = B^T B$.

Thus any normal matrix A shares with A^T all real eigenvalues and the corresponding eigenvectors.
How about complex eigenvalues?

Dot product of complex vectors

Dot product of real vectors
$\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$:

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Dot product of complex vectors
$\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{C}^n$:

$$\mathbf{x} \cdot \mathbf{y} = x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}.$$

If $z = r + it$ $(r, t \in \mathbb{R})$ then $\overline{z} = r - it$,
$z\overline{z} = r^2 + t^2 = |z|^2$.
Hence $\mathbf{x} \cdot \mathbf{x} = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \geq 0$.
Also, $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

Since $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z}\,\overline{w}$, it follows that $\mathbf{y} \cdot \mathbf{x} = \overline{\mathbf{x} \cdot \mathbf{y}}$.

Definition. Let V be a complex vector space. A function $\beta : V \times V \to \mathbb{C}$, denoted $\beta(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle$, is called an **inner product** on V if

(i) $\langle \mathbf{x}, \mathbf{y} \rangle \geq 0$, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ only for $\mathbf{x} = \mathbf{0}$ (positivity)

(ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ (conjugate symmetry)

(iii) $\langle r\mathbf{x}, \mathbf{y} \rangle = r\langle \mathbf{x}, \mathbf{y} \rangle$ (homogeneity)

(iv) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ (additivity)

$\langle \mathbf{x}, \mathbf{y} \rangle$ is complex-linear as a function of $\mathbf{x}$.
The dependence on the second argument is called *half-linearity*: $\langle \mathbf{x}, \lambda\mathbf{y} + \mu\mathbf{z} \rangle = \overline{\lambda}\langle \mathbf{x}, \mathbf{y} \rangle + \overline{\mu}\langle \mathbf{x}, \mathbf{z} \rangle$.

Example. $\langle f, g \rangle = \displaystyle\int_a^b f(x)\overline{g(x)}\, dx$,
$f, g \in C([a, b], \mathbb{C})$.

Proposition For any $n \times n$ matrix B with complex entries and any column vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$B\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \overline{B}^T \mathbf{y}.$$

If $B\overline{B}^T = \overline{B}^T B$ then $B\mathbf{x} = \mathbf{0} \iff \overline{B}^T \mathbf{x} = \mathbf{0}$.

Theorem Suppose A is a normal matrix. Then for any $\mathbf{x} \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ one has

$$A\mathbf{x} = \lambda \mathbf{x} \iff A^T \mathbf{x} = \overline{\lambda} \mathbf{x}.$$

Also, $A\mathbf{x} = \lambda \mathbf{x} \iff A^T \overline{\mathbf{x}} = \overline{\lambda}\, \overline{\mathbf{x}}$.

Corollary All eigenvalues of a symmetric matrix are real. Any eigenvalue λ of an orthogonal matrix satisfies $\overline{\lambda} = \lambda^{-1} \iff |\lambda| = 1$.

Theorem If $\mathbf{x}$ and $\mathbf{y}$ are eigenvectors of a normal matrix A associated with different eigenvalues, then $\mathbf{x} \cdot \mathbf{y} = 0$.

Theorem Let A be an $n \times n$ matrix with real entries. Then

(a) A is normal $\iff$ there exists an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors of A;

(b) A is symmetric $\iff$ there exists an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

Example. $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

- A is symmetric.
- A has three eigenvalues: 0, 2, and 3.
- Associated eigenvectors are $\mathbf{v}_1 = (-1, 0, 1)$, $\mathbf{v}_2 = (1, 0, 1)$, and $\mathbf{v}_3 = (0, 1, 0)$.
- Vectors $\frac{1}{\sqrt{2}}\mathbf{v}_1, \frac{1}{\sqrt{2}}\mathbf{v}_2, \mathbf{v}_3$ form an orthonormal basis for $\mathbb{R}^3$.

Example. $A_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$.

- $A_\phi A_\psi = A_{\phi+\psi}$
- $A_\phi^{-1} = A_{-\phi} = A_\phi^T$
- A_ϕ is orthogonal
- Columns of A_ϕ form an orthonormal basis.
- Rows of A_ϕ form an orthonormal basis.
- Eigenvalues: $\lambda_1 = \cos\phi + i\sin\phi = e^{i\phi}$, $\lambda_2 = \cos\phi - i\sin\phi = e^{-i\phi}$.
- Associated eigenvectors: $\mathbf{v}_1 = (1, -i)$, $\mathbf{v}_2 = (1, i)$.
- Vectors $\frac{1}{\sqrt{2}}\mathbf{v}_1$ and $\frac{1}{\sqrt{2}}\mathbf{v}_2$ form an orthonormal basis for $\mathbb{C}^2$.

Orthogonal matrices

Theorem Given an $n \times n$ matrix A, the following conditions are equivalent:
(i) A is orthogonal: $A^T = A^{-1}$;
(ii) columns of A form an orthonormal basis for $\mathbb{R}^n$;
(iii) rows of A form an orthonormal basis for $\mathbb{R}^n$.

Proof: Entries of the matrix A^TA are the dot products of columns of A. Entries of AA^T are the dot products of rows of A.

Thus an orthogonal matrix is the transition matrix from one orthonormal basis to another.

Consider a linear operator $L : \mathbb{R}^n \to \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$, where A is an $n \times n$ matrix.

Theorem The following conditions are equivalent:
(i) $|L(\mathbf{x})| = |\mathbf{x}|$ for all $\mathbf{x} \in \mathbb{R}^n$;
(ii) $L(\mathbf{x}) \cdot L(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;
(iii) the matrix A is orthogonal.

[(ii) $\Longrightarrow$ (iii): $L(\mathbf{e}_i) \cdot L(\mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j = 1$ if $i = j$, and 0 otherwise. But $L(\mathbf{e}_1), \dots, L(\mathbf{e}_n)$ are columns of A.]

Definition. A transformation $f : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if it preserves distances between points: $|f(\mathbf{x}) - f(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|$.

Theorem Any isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ is represented as $f(\mathbf{x}) = A\mathbf{x} + \mathbf{x}_0$, where $\mathbf{x}_0 \in \mathbb{R}^n$ and A is an orthogonal matrix.

Consider a linear operator $L : \mathbb{R}^n \to \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$, where A is an $n \times n$ orthogonal matrix.

Theorem There exists an orthonormal basis for $\mathbb{R}^n$ such that the matrix of L relative to this basis has a diagonal block structure

$$\begin{pmatrix} D_{\pm 1} & O & \dots & O \\ O & R_1 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & R_k \end{pmatrix},$$

where $D_{\pm 1}$ is a diagonal matrix whose diagonal entries are equal to 1 or -1, and

$$R_j = \begin{pmatrix} \cos\phi_j & -\sin\phi_j \\ \sin\phi_j & \cos\phi_j \end{pmatrix}, \quad \phi_j \in \mathbb{R}.$$

Classification of 2×2 orthogonal matrices:

	$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
	rotation about the origin	reflection in a line
Determinant:	1	-1
Eigenvalues:	$e^{i\phi}$ and $e^{-i\phi}$	-1 and 1

Classification of 3 × 3 orthogonal matrices:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}.$$

$A =$ rotation about a line; $B =$ reflection in a plane; $C =$ rotation about a line combined with reflection in the orthogonal plane.

$\det A = 1$, $\det B = \det C = -1$.

A has eigenvalues 1, $e^{i\phi}$, $e^{-i\phi}$. B has eigenvalues -1, 1, 1. C has eigenvalues -1, $e^{i\phi}$, $e^{-i\phi}$.