Modified on Monday 5th October, 2020, 11:41:1711:41

Oct 5, Symmetric matrices and diagonalization. [8.2, 8.3]

- **Equivalence relation**
  Let X be a nonempty set. Let R be a relation on X, i.e., $R \subset X \times X$, and let xRy denote $(x, y) \in R$. The relation R is an *equivalence relation*, if it satisfies the following three properties:
  - For any $x \in X$, xRx. (R is reflexive)
  - For any $x, y \in X$, if xRy then yRx. (R is symmetric)
  - For any $x, y, z \in X$, if xRy and yRz then xRz. (R is transitive)
- Row equivalence relation
  Let $\sim_r$ denote the relation on $m \times n$ matrices such that $A \sim_r B$ if and only if B can be obtained from A by applying a sequence of elementary row operations. This relation is an equivalence relation.
- Definition
  Two $n \times n$ matrices A and B are said to be similar, denoted by $A \sim B$, if there exists an $n \times n$ invertible matrix P such that $B = P^{-1}AP$. This relation is an equivalence relation.
- Recall: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator. $\mathcal{S}$: the standard basis. $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$: an ordered basis for $\mathbb{R}^n$.
$$P = [\mathrm{id}]_{\mathcal{S},\mathcal{B}} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$$
$$[T]_{\mathcal{S}} = P[T]_{\mathcal{B}}P^{-1}$$
$$[T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{S}}P$$
- Two square matrices are similar if and only if there are bases with respect to which the matrices represent the same linear operator.
- Any propery that is shared by similar matrices is said to be a **similarity invariant**. Let A and B be $n \times n$ similar matrices.
  - $\det(A) = \det(B)$
  - $\mathrm{tr}(A) = \mathrm{tr}(B)$
  - $p_A(t) = p_B(t)$
  - A and B have the same eigenvalues λ. The dimension of the eigenspace corresponding to eigenvalue λ is the same for A and B.

  Does $\mathrm{tr}(A) = \mathrm{tr}(B)$ imply $A \sim B$? Does $p_A(t) = p_B(t)$ imply $A \sim B$?
- Algebraic multiplicity, geometric multiplicity
  Let λ be an eigenvalue of an $n \times n$ matrix A. The **algebraic multiplicity** of λ is the greatest integer d such that $(t - \lambda)^d$ divides $p_A(t) = \det(tI_n - A)$, and the **geometric multiplicity** of λ is the dimension of the eigenspace $\{\mathbf{v} \in \mathbb{R}^n : A\mathbf{v} = \lambda\mathbf{v}\}$. For any eigenvalue λ, the geometric multiplicity is at least one and is at most the algebraic multiplicity.
- Diagonal matrix and diagonalizable matrix
  An $n \times n$ matrix $A = (a_{ij})$ is called a **diagonal matrix** if $a_{ij} = 0$ for all $i \neq j$, and is said to be **diagonalizable** if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.
- If P diagonalizes A, $P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, then for each i, the i-th column of P is an eigenvector of A corresponding to eigenvalue λ_i.
- Theorem
  If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of A of size $n \times n$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent.

Proof. There are several interesting proofs.

- Proof 1: Suppose $\sum_{i=1}^{k} c_i\mathbf{v}_i = 0$. We will show that $c_1 = c_2 = \cdots = c_k = 0$. For every $j = 0, 1, \ldots, k-1$, we also have $\sum_{i=1}^{k} c_i A^j \mathbf{v}_i = \sum_{i=1}^{k} c_i \lambda_i^j \mathbf{v}_i = 0$.

$$\begin{bmatrix} c_1\mathbf{v}_1 & c_2\mathbf{v}_2 & \cdots & c_k\mathbf{v}_k \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{k-1} \end{bmatrix} = \mathbf{0}_{n\times k}$$

$$\begin{bmatrix} c_1\mathbf{v}_1 & c_2\mathbf{v}_2 & \cdots & c_k\mathbf{v}_k \end{bmatrix} = \mathbf{0}_{n\times k} \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{k-1} \end{bmatrix}^{-1} = \mathbf{0}_{n\times k}$$

  It follows that $c_1 = c_2 = \cdots = c_k = 0$.

- Proof 2: Suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly dependent. Choose a linear dependence relation of the set involving the smallest number of $\mathbf{v}_i$'s, say $c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$ with $m > 1$ and $c_i \neq 0$ for all $i = 1, \ldots, m$. Then we have $c_1\lambda_1\mathbf{v}_1 + \cdots + c_m\lambda_m\mathbf{v}_m = \mathbf{0}$ and

$$c_1(\lambda_1 - \lambda_m)\mathbf{v}_1 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{v}_{m-1} = \mathbf{0}.$$

  Since λ_i's are distinct, the latter is a shorter linear dependence relation with $c_i(\lambda_i - \lambda_m) \neq 0$ for all $i = 1, \ldots, m-1$, contradicting our choice of m.

- Theorem
  If A is an $n \times n$ matrix, then the following are equivalent.

  1. A is diagonalizable.
  2. A has n linearly independent eigenvectors.
  3. $\mathbb{R}^n$ has a basis consisting of eigenvectors of A.
  4. The geometric multiplicity of each eigenvalue of A is the same as the algebraic multiplicity.

- If an $n \times n$ matrix A has n distinct eigenvalues, then it is diagonalizable.
- Orthogonally diagonalizable matrix and symmetric matrix.
- An $n \times n$ matrix $A = (a_{ij})$ is said to be **orthogonally diagonalizable**, if there exists an orthogonal matrix P such that P^TAP is a diagonal matrix, i.e., $A = PDP^T$ for some diagonal matrix D.
- Two matrices are orthogonally similar if and only if there exist orthonormal bases with respect to which the matrices represent the same linear operator.
- Any orthogonally diagonalizable matrix A is (real) symmetric, i.e., $A^T = A$. Does the converse hold? Yes, any real symmetric matrix is orthogonally diagonalizable.
- If P orthogonally diagonalizes A, $P^TAP = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then columns of P are eigenvectors of A and form an orthonormal basis for $\mathbb{R}^n$. Let P=$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$.

$$AP = P\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix}$$

$$A = P\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)P^T = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}$$

The spectral decomposition of A (or the eigenvalue decomposition of A).

$$A = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T$$

$$A^k = \lambda_1^k \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2^k \mathbf{v}_2 \mathbf{v}_2^T + \cdots + \lambda_n^k \mathbf{v}_n \mathbf{v}_n^T$$

- If A is symmetric and $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors of A corresponding to distinct eigenvalues λ_1, λ_2 respectively, then $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal, i.e., $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

  Proof. $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = A\mathbf{v}_1 \cdot \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 = \mathbf{v}_1^T A^T \mathbf{v}_2 = \mathbf{v}_1^T A \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$. Since $\lambda_1 \neq \lambda_2$, we have $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

- **Cayley-Hamilton theorem**
  For any $n \times n$ matrix A, $p_A(A) = \mathbf{0}_{n\times n}$, where $p_A(t) = \det(tI_n - A)$.