1. Eigenvalues and eigenvectors

Definition 1 设 A 是复数域 $\mathbb{C}$ 上的 n 阶矩阵，若存在数 $\lambda \in \mathbb{C}$ 和非零的 n 维向量 x 使得：

$$Ax = \lambda x$$

就称 λ 为矩阵 A 的特征值，x 是 A 对应 λ 的特征向量

Definition 2 设 n 阶矩阵 $A = (a_{ij})$ 则

$$f(\lambda) = \det(\lambda I - A)$$

$$= \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为 A 的特征多项式，$\lambda I - A$ 称为 A 的特征矩阵，

$$\det(\lambda I - A) = 0$$

称为 A 的特征方程

Theorem 1 (Rank-Nullity Theorem 秩-零化度定理) 设 $A \in \mathbb{F}^{m \times n}$ 定义零空间 $N(A) = \{x \in \mathbb{F}^n \mid Ax = 0\}$，则有：

$$\operatorname{rank}(A) + \dim(N(A)) = n$$

Proof: ① A 初等行变换为：

$$R = \begin{bmatrix} E_r & F \\ 0 & 0 \end{bmatrix}$$

$$\therefore \operatorname{rank}(A) = r(R) = r(E_r) = r$$

② $F \in \mathbb{F}^{r\times(n-r)}$

构造 R 的零空间矩阵 $N(R) = N(A)$

$$P = \begin{bmatrix} -F \\ E_{n-r} \end{bmatrix}$$

验证：

$$RP = \begin{bmatrix} E_r & F \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -F \\ E_{n-r} \end{bmatrix} = \begin{bmatrix} -F+F \\ 0 \end{bmatrix} = 0$$

显然 $r(P) = n-r$ 即列向量线性无关

③ 下证：$Rx = 0$ 的解 x 可由 P 的列向量表出

对 $\forall\ x = [x_1, x_2]^T \quad x_1 \in \mathbb{F}^r \quad x_2 \in \mathbb{F}^{n-r}$

使 $Rx = 0$ 则

$$Rx = \begin{bmatrix} E_r & F \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + Fx_2 \\ 0 \end{bmatrix} = 0$$

$\therefore \quad x_1 = -Fx_2$

即 $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -Fx_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -F \\ E_{n-r} \end{bmatrix} x_2 = Px_2$ □

也即 $N(R) = \{x \in \mathbb{F}^r \mid x \in P\}$

④ 综上：

$\operatorname{rank}(A) = r$

$\dim(N(A)) = \dim(N(R)) = \dim(C(P)) = r(P) = n-r$

$\therefore \quad \operatorname{rank}(A) + \dim(N(A)) = n$ □

Theorem 2 下面两式等价：

(a). λ 是 A 的特征值，α 是 A 关于 λ 的特征向量

(b). λ 满足 $|\lambda I - A| = 0$，α 满足 $\alpha \neq 0 \quad (\lambda I - A)\alpha = 0$

Proof: (a) $\Rightarrow$ (b)

$A\alpha = \lambda\alpha \quad \alpha \neq 0$

即 $(\lambda I - A)\alpha = 0$ 存在非零解

$\therefore \operatorname{rank}(\lambda I - A) < n$

$\Rightarrow |\lambda I - A| = 0 \quad \square$

(b) $\Rightarrow$ (a)

定义显然 $\square$

<u>Proposition 1 (性质)</u> λ 是 A 的特征值，$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A 对应 λ 的 m 个特征向量：

(a) $\forall a,b \in \mathbb{F}$，$b\lambda + a$ 是 $bA + aI$ 的特征值

$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 $bA+aI$ 关于 $b\lambda+a$ 的特征向量

(b) $\forall k \in \mathbb{N}_+$，$\lambda^k$ 是 A^k 的特征值，$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A^k 关于 λ^k 的特征向量

(c) 如果 A 可逆，$\lambda \neq 0$．则 λ^{-1} 是 A^{-1} 的特征值 $\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A^{-1} 关于 λ^{-1} 的特征向量

(d). λ 也是 A^T 的特征值

(e). 任意 α_i 的非零线性组合都是 A 关于 λ 的特征向量

Proof (a) $(b\lambda+a)\alpha = b\lambda\alpha + a\alpha$

$= bA\alpha + a\alpha$

$= (bA + aI)\alpha$ □

(b) $\lambda^k\alpha = \lambda^{k-1}(\lambda\alpha) = \lambda^{k-2}\cdot A\alpha$

$= A\cdot\lambda^{k-2}\cdot\alpha = A^2\cdot\lambda^{k-3}\cdot\alpha$

$= \cdots = A^k\alpha$ □

(c). $\lambda^{-1}\alpha = A^{-1}\alpha$

$\Leftrightarrow A\alpha = \lambda\alpha$ □

(d) $|\lambda I - A| = 0$

$\Leftrightarrow |(\lambda I - A)^T| = 0$

$\Leftrightarrow |\lambda I - A^T| = 0$ □

(e). $A\alpha_i = \lambda\alpha_i \Rightarrow A k_i\alpha_i = \lambda k_i\alpha_i$

$\Rightarrow A(\Sigma k_i\alpha_i) = \lambda(\Sigma k_i\alpha_i)$ □

Theorem 3 (Linear Independence of eigenvectors) $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是矩阵 A 的 s 个不同的特征值，$\{\alpha_{ij} \mid j = 1, 2, \cdots, r_i\}$ 是 A 关于 λ_i 的 r_i 个线性无关的特征向量。则整体集 $\{\alpha_{ij} \mid j = 1, 2, \cdots, r_i, i = 1, 2, \cdots, s\}$ 也是线性无关的。

Proof: ① 先证简单情形：$r_1 = r_2 = \cdots = r_s = 1$

(i) 当 $s=1$ 显然，

(ii) 假设 $s=m-1$ 有 $\{\alpha_1, \alpha_2, \cdots, \alpha_{m-1}\}$ 线性无关

(iii) 当 $s=m$ 时

若 $\exists \{k_1, k_2, \cdots, k_m\}$ 使

$$\sum_{i=1}^{m} k_i \alpha_i = 0 \qquad (1)$$

(1) 同乘 A：

$$\sum_{i=1}^{m} k_i (A\alpha_i) = 0 \qquad (2)$$

$\because A\alpha_i = \lambda_i \alpha_i$ 故，

$$\sum_{i=1}^{m} k_i \lambda_i \alpha_i = 0 \qquad (3)$$

(1) 同乘 λ_m：

$$\sum_{i=1}^{m} k_i \lambda_m \alpha_i = 0 \qquad (4)$$

用 (3) − (4)：

$$\sum_{i=1}^{m} k_i (\lambda_i - \lambda_m) \alpha_i = \sum_{i=1}^{m-1} k_i (\lambda_i - \lambda_m) \alpha_i = 0 \qquad (5)$$

$\because s=m-1$ 有 $\{\alpha_1, \cdots, \alpha_{m-1}\}$ 线性无关

故式 (5) 有：

$$k_i(\lambda_i - \lambda_m) = 0 \quad \forall i = 1, 2, \cdots, m-1 \qquad (6)$$

$\because \{\lambda_1, \lambda_2, \cdots \lambda_m\}$ 互不相同

$\therefore k_1 = k_2 = \cdots = k_{m-1} = 0$

代回 (1) 有

$$k_m \alpha_m = 0 \qquad (7)$$

$\because \alpha_m \neq 0 \quad \therefore k_m = 0$

同一个 λ_i，它的特征向量必定线性独立

即 $\{k_i\} = 0$ 故 $\{\alpha_1, \cdots, \alpha_m\}$ 线性无关 $s=m$ 成立 □

② 一般情形：

若 $\exists \{k_{ij} \mid j=1,2,\cdots,r_i\ ;\ i=1,2,\cdots,s\}$ 使

$$\sum_{i=1}^{s}\sum_{j=1}^{r_i} k_{ij}\alpha_{ij} = 0 \tag{8}$$

记 $\xi_i = \sum_{j=1}^{r_i} k_{ij}\alpha_{ij}$，则 (8) 变为：

$$\sum_{i=1}^{s} \xi_i = 0 \tag{9}$$

$\because \forall$ 给定 i 有 $\{\alpha_{ij} \mid j=1,2,\cdots,r_i\}$ 线性无关

故若 $\xi_i = 0$ 有 $k_{i1} = k_{i2} = \cdots = k_{ir_i} = 0$

$\therefore$ 原问题 $\Longleftrightarrow$ $\forall i$ 有 $\xi_i = 0$

(反证) 记非空集合 $I \subset \{1,2,\cdots,s\}$ $I \neq \phi$ 使得集合

$\{\xi_i \neq 0 \mid i \in I\}$，$\{\xi_i = 0 \mid i \notin I\}$

原式 (9) 有：

$$\sum_{i\in I} \xi_i = 0 \tag{10}$$

于是 $0 \neq \xi_i = \sum_{j=1}^{r_i} k_{ij}\alpha_{ij}$ 线性组合 $\Rightarrow$ ξ_i 也是 A 关于 $\lambda_i\ (i\in I)$

的特征值。但在①中，我们已证：

每个特征值对应的自己的特征向量一定线性独立

(ξ_i 是 λ_i 的特征向量，但 $\sum 1\cdot\xi_i = 0$ 却表明 ξ_i 之间不是独立，)

这与 (10) $\sum_{i\in I}\xi_i = 0$ 矛盾！

故 $\xi_i = 0$ $\square$

2. Diagonalization of Square Matrices

Theorem 4　A可以被对角化 当且仅当 $A \in \mathbb{F}^{n\times n}$，且A有 n个线性无关的特征向量

$\Leftrightarrow$

Proof (1) $\Rightarrow$

已有 $A = PDP^{-1}$　其中 $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$　$P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$

且P是可逆（即满秩）

$\therefore \quad AP = PD$

即 $A(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n) \cdot diag(\lambda_1, \cdots, \lambda_n)$

$\therefore (A\alpha_1, \cdots, A\alpha_n) = (\lambda_1\alpha_1, \cdots, \lambda_n\alpha_n)$

$\therefore \quad A\alpha_i = \lambda_i \alpha_i$

$\therefore$ α_i 是 A 关于 λ_i 的特征向量，

又 $\because$ P 满秩 $\Rightarrow$ $\alpha_i \neq 0$ 且线性独立　$(i = 1, 2, \cdots, n)$ $\square$

(2) $\Leftarrow$

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 A 的 n 个线性独立的特征向量（关于 λ_i）

$\therefore \quad A\alpha_i = \lambda_i \alpha_i$

构造 $P = (\alpha_1, \cdots, \alpha_n)$ 是可逆的

构造 $D = diag(\lambda_1, \cdots, \lambda_n)$ 是对角阵

有 $AP = A(\alpha_1, \cdots, \alpha_n) = (A\alpha_1, \cdots, A\alpha_n)$

$$= (\lambda_1\alpha_1, \cdots, \lambda_n\alpha_n) = (\alpha_1, \cdots, \alpha_n) \cdot diag(\lambda_1, \cdots, \lambda_n)$$

$$= PD$$

$\therefore \quad A = PDP^{-1}$ $\square$

Theorem 5 已知 $\{\lambda_i : i=1,2,\cdots,s\}$ 是 A 的 s 个不同的特征值，用

$$V_i = \{x \in \mathbb{F}^n \mid (\lambda_i I - A)x = 0\}$$

表示 λ_i 的特征空间. $\dim(V_i)$ 称为 λ_i $(i=1,2,\cdots,s)$ 的几何重数 (geometric multiplicity). A 可对角化，当且仅当 $\sum_{i=1}^{s}\dim(V_i) = n$

$$\boxed{A \text{ 可对角化} \iff \sum_{i=1}^{s}\dim(V_i) = n}$$

Proof:

根据 Theorem 4，我们去证：

$$\sum_{i=1}^{s}\dim(V_i) = n \iff A \text{ 有 } n \text{ 个线性独立的特征向量}$$

$\Rightarrow$：

可以从每个 V_i 中取 $\dim(V_i)$ 个线性独立的特征向量

又 $\sum_{i=1}^{s}\dim(V_i) = n$，加之 Theorem 3

故找到了 n 个线性独立的特征向量 □

$\Leftarrow$：

A 有 n 个线性独立的特征向量

而 A 的所有特征向量在所有特征空间 V_i 中 $(i=1,2,\cdots,s)$

$\therefore\ n \le \sum_{i=1}^{s}\dim(V_i)$

另一方面：x：$(\lambda I - A)x = 0$ 其中 $(\lambda I - A) \in \mathbb{F}^{n\times n}$

$\therefore$ 解空间 $\le n$，即

至多只有找到 n 个线性独立的特征向量

$\therefore \quad n \geq \sum_{i=1}^{s} \dim(V_i)$

故 $\quad n = \sum_{i=1}^{s} \dim(V_i) \quad \square$

Theorem 6 设不同的特征值 $\lambda_i \ (i=1,2,\cdots,s)$，则特征多项式（characteristic polynomial）$|\lambda I - A|$ 可表示为：

$$|\lambda I - A| = \prod_{i=1}^{s} (\lambda - \lambda_i)^{r_i} \qquad \left(\sum r_i = n\right)$$

我们称 r_i 为 λ_i 的代数重数（algebraic multiplicity）

(a) 每个特征值对应的几何重数 ≤ 代数重数

$$\dim(V_i) \leq r_i \qquad (i=1,2,\cdots,s)$$

(b) A可对角化，当且仅当 几何重数 = 代数重数

A is diagonalizable

$$\Longleftrightarrow \quad \forall i \quad \dim(V_i) = r_i$$

(c) 如果A是可对角化的，取A的每个 λ_i 的特征空间的基，它们共同构成 F^n 的一组基

Proof (a) 略

(b) 根据 Theorem 5： A可对角化 $\Longleftrightarrow \sum_{i=1}^{s} \dim(V_i) = n$

我们只须证： $\sum_{i=1}^{s} \dim(V_i) = n \Longleftrightarrow \dim(V_i) = r_i$

① $\Leftarrow$ ：

$\because \sum_{i=1}^{s} r_i = n$ 而 $\dim(V_i) = r_i$

故 $\sum_{i=1}^{s} \dim(V_i) = n \quad \square$

② $\Rightarrow$:

由(a)知 $\dim(V_i) \le r_i$

$\therefore \sum_{i=1}^{s} \dim(V_i) \le \sum_{i=1}^{s} r_i = n$

但 $\sum_{i=1}^{s} \dim(V_i) = n$ 等号成立

故 $\forall \dim(V_i) = r_i$ $\square$

(c) A 可对角化 $\Leftrightarrow$ $\sum_{i=1}^{s} \dim(V_i) = n$

又 V_i 为 λ_i 的特征空间

λ_i 的基张成了 V_i，和 $\dim(V_i)$ 求和为 n

故这些基共同张成 $\mathbb{F}^n$ $\square$

Corollary 2 A 可对角化

$\Leftrightarrow$ $r_i + \operatorname{rank}(\lambda_i I - A) = n \quad i = 1, 2, \cdots, s$

Proof: A 可对角化 $\Leftrightarrow$ $\dim(V_i) = r_i$

V_i 是 $(\lambda_i I - A)x = 0$ 的解空间

$\therefore \dim(V_i) + \operatorname{rank}(\lambda_i I - A) = n$

$\therefore r_i + \operatorname{rank}(\lambda_i I - A) = n$ $\square$

3. Eigenvalues and eigenvectors of real sysmmetric matrices

下面，我们仅考虑，实矩阵 $A \in \mathbb{R}^{n\times n}$

Lemma 1 已知 λ 是 A 的特征值，α 是 A 关于 λ 的特征向量

则 $\bar{\lambda}$ 也是 A 的特征值，$\bar{\alpha}$ 是 A 关于 $\bar{\lambda}$ 的特征向量

Proof: $A\alpha = \lambda\alpha$ $\quad$ ($\because A \in \mathbb{R}^{n\times n}$ $\therefore \bar{A} = A$

$\therefore A\bar{\alpha} = \bar{A}\bar{\alpha} = \overline{A\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$

$\because \alpha \neq 0 \quad \therefore \bar{\alpha} \neq 0$

$\therefore$ $\bar{\lambda}$ 也是 A 的特征值，$\bar{\alpha}$ 是 A 关于 $\bar{\lambda}$ 的特征向量 □

Theorem 7 已知 A 是实对称矩阵（RSM: real symmetric matrix）

则 A 的所有特征值均为实数.

$$A \in \mathbb{R}^{n\times n} \text{ and } A^T = A \quad \Rightarrow \quad \lambda_i \in \mathbb{R}$$

Proof: $\lambda \in \mathbb{C}$ 是 A 的特征值，特征向量 $\alpha = (a_1+b_1 i, a_2+b_2 i, \cdots, a_n+b_n i)^T$

$$A\alpha = \lambda\alpha \qquad \alpha \neq 0$$

由 Lemma 1：$A\bar{\alpha} = \bar{\lambda}\bar{\alpha}$

$$\Rightarrow (A\bar{\alpha})^T = (\bar{\lambda}\bar{\alpha})^T \Rightarrow \bar{\alpha}^T A^T = \bar{\alpha}^T A = \bar{\lambda}\bar{\alpha}^T$$

右乘 α：$\bar{\lambda}\bar{\alpha}^T\alpha = \bar{\alpha}^T A\alpha = \bar{\alpha}^T(\lambda\alpha) = \lambda\bar{\alpha}^T\alpha$

$$\therefore 0 = (\lambda-\bar{\lambda})\bar{\alpha}^T\alpha = (\lambda-\bar{\lambda})\cdot(a_1-b_1 i, \cdots, a_n-b_n i)(a_1+b_1 i, \cdots, a_n+b_n i)^T$$

$$= (\lambda-\bar{\lambda})\sum_{j=1}^{n}(a_j^2+b_j^2)$$

$\because \quad 0 \neq \alpha = (a_j + b_j i)^T \quad \Rightarrow \quad \sum (a_j^2 + b_j^2) \neq 0$

$\therefore \quad \lambda - \bar{\lambda} = 0 \quad \Rightarrow \quad \lambda \in \mathbb{R} \qquad \square$

Theorem 8　已知 A 是实对称矩阵，则 A 的不同特征根的特征向量是相互正交的 (orthogonal)

Proof: α_1, α_2 是 A 关于 λ_1, λ_2 的特征向量 ($\lambda_1 \neq \lambda_2$)

$$A\alpha_1 = \lambda_1 \alpha_1 \qquad A\alpha_2 = \lambda_2 \alpha_2$$

$$\begin{aligned} \therefore \quad \lambda_1 \alpha_1^T \alpha_2 &= (\lambda_1 \alpha_1)^T \alpha_2 = (A\alpha_1)^T \alpha_2 \\ &= \alpha_1^T A^T \alpha_2 = \alpha_1^T (A^T \alpha_2) \\ &= \alpha_1^T (A\alpha_2) = \alpha_1^T (\lambda_2 \alpha_2) \\ &= \lambda_2 \alpha_1^T \alpha_2 \end{aligned}$$

$\therefore \quad (\lambda_1 - \lambda_2)\alpha_1^T \alpha_2 = 0 \qquad (\lambda_1 \neq \lambda_2)$

$\therefore \quad \alpha_1^T \alpha_2 = 0$

(注：正交必独立，易证)

4. Orthogonal Diagonalization of RSM

☆ **Theorem 9** A 是实对称矩阵，等价于存在正交阵 Q，使得 Q^TAQ 是一个对角阵。我们称 A 可正交对角化：

$$A \text{ 是实对称阵}$$
$$\Leftrightarrow \quad A = QDQ^T$$

其中 Q 为正交阵（$QQ^T = I$），D 为对角阵

Proof:

If $D = Q^{\mathrm{T}}AQ$ is diagonal, the symmetry of A follows from the representation $A = QDQ^{\mathrm{T}}$. We now prove the converse by the induction method:

(i) Clearly, the conclusion holds when $n = 1$.

(ii) Assume the conclusion holds when $n = m - 1$.

(iii) Consider the case with $n = m$. Suppose λ_1 is an eigenvalue of A with an eigenvector α_1. Since $c\,\alpha_1$ is also an eigenvector of A belonging to α_1 for any scalar $c \neq 0$, we assume $\|\alpha_1\| = 1$ without loss of generality. Expand α_1 to an orthonormal basis $\{\alpha_i : i = 1, \ldots, m\}$ of $\mathbb{R}^m$, where all the vectors are unit vectors and orthogonal to each other, using the Gram-Schmidt process as described in Theorem 11. Denote $P = (\alpha_2, \ldots, \alpha_m)$, $Q_1 = (\alpha_1, P)$ and $A_2 = P^{\mathrm{T}}AP$. Then Q_1 is orthogonal. Since $\{\alpha_i : i = 1, \ldots, m\}$ satisfy $\alpha_1^{\mathrm{T}}\alpha_1 = 1$ and $\alpha_i^{\mathrm{T}}\alpha_1 = 0$ for $i = 2, \ldots, m$, we have $\alpha_1^{\mathrm{T}}A\alpha_1 = \lambda_1\alpha_1^{\mathrm{T}}\alpha_1 = \lambda_1$ and $P^{\mathrm{T}}A\alpha_1 = \lambda_1 P^{\mathrm{T}}\alpha_1 = \lambda_1(\alpha_2^{\mathrm{T}}\alpha_1, \ldots, \alpha_m^{\mathrm{T}}\alpha_1)^{\mathrm{T}} = 0$. It follows that

$$Q_1^{\mathrm{T}}AQ_1 = (\alpha_1, P)^{\mathrm{T}}A(\alpha_1, P) = \begin{pmatrix} \alpha_1^{\mathrm{T}}A\alpha_1 & \alpha_1^{\mathrm{T}}AP \\ P^{\mathrm{T}}A\alpha_1 & A_2 \end{pmatrix} = \mathrm{diag}(\lambda_1, A_2). \tag{10}$$

Since $A_2 \in \mathbb{R}^{(m-1)\times(m-1)}$ is symmetric, assumption (ii) indicates there exist some orthogonal matrix $Q_2 \in \mathbb{R}^{(m-1)\times(m-1)}$ and diagonal matrix $D_2 \in \mathbb{R}^{(m-1)\times(m-1)}$, such that $Q_2^{\mathrm{T}}A_2Q_2 = D_2$. Let $Q = Q_1\mathrm{diag}(1, Q_2) \in \mathbb{R}^{m\times m}$. Since

$$Q^{\mathrm{T}}Q = \mathrm{diag}(1, Q_2^{\mathrm{T}})Q_1^{\mathrm{T}}Q_1\mathrm{diag}(1, Q_2) = \mathrm{diag}(1, Q_2^{\mathrm{T}})\mathrm{diag}(1, Q_2) = I,$$

we know Q is orthogonal. Then, it follows from (10) that

$$\begin{aligned} Q^{\mathrm{T}}AQ &= \operatorname{diag}(1, Q_2^{\mathrm{T}})Q_1^{\mathrm{T}}AQ_1\operatorname{diag}(1, Q_2) = \operatorname{diag}(1, Q_2^{\mathrm{T}})\operatorname{diag}(\lambda_1, A_2)\operatorname{diag}(1, Q_2) \\ &= \operatorname{diag}(\lambda_1, Q_2^{\mathrm{T}}A_2Q_2) = \operatorname{diag}(\lambda_1, D_2), \end{aligned}$$

which is a diagonal matrix. This completes the proof.

Proposition 3 A是实对称阵. 记 λ_{min} 和 λ_{max} 为 A 的最小和最大的特征值。对任意,模长为1的向量 $x \in \mathbb{R}^n$ ($\|x\| = 1$), 有:

$$\lambda_{min} \leq x^T A x \leq \lambda_{max}$$

Proof: $\because$ A 为实对称矩阵

$\therefore$ $A = QDQ^T \quad QQ^T = I \quad D = diag(\lambda_1, \cdots, \lambda_n)$

对 $\forall\ x \in \mathbb{R}^n \quad \|x\| = 1$

令 $y = Q^T x = (y_1, \cdots, y_n)^T$

有 $y^T y = (Q^T x)^T Q^T x = x^T Q Q^T x = x^T x = \|x\| = 1$

$\therefore \|y\| = 1$

计算 $x^T A x = x^T Q D Q^T x$

$$= (Q^T x)^T D (Q^T x)$$

$$= y^T D y$$

$$= \sum_{i=1}^n \lambda_i y_i^2$$

对 $\forall\ i$ 有 $\lambda_{min} y_i^2 \leq \lambda_i y_i^2 \leq \lambda_{max} y_i^2$

$\therefore \lambda_{min} \sum y_i^2 \leq \sum \lambda_i y_i^2 \leq \lambda_{max} \sum y_i^2$

又 $\|y\| = \sum y_i^2 = 1$

$\therefore \lambda_{min} \le \sum \lambda_i y_i^2 = x^TAx \le \lambda_{max}$

取等时：(以 $x^TAx = \lambda_{max}$ 为例)

当 $Ax = \lambda_{max} x$ 时 $x^TAx = x^T\lambda_{max} x = \lambda_{max}\|x\| = \lambda_{max}$

即 x 为 A 关于 λ_{max} 的特征向量

Theorem 12

实对称阵是正定的，当且仅当 它的特征值均正

实对称阵是半正定的，当且仅当 它的特征值均非负

Proof: $\Leftarrow$

由 $y = Q^Tx$ $A = QDQ^T$ $D = diag(\lambda_1, \cdots, \lambda_n)$

有 $x^TAx = x^TQDQ^Tx = (Q^Tx)^TD(Q^Tx)$

$= y^TDy$

$= \sum_{i=1}^n \lambda_i y_i^2$

① $\forall \lambda_i > 0$ 有 $x^TAx = \sum \lambda_i y_i^2 > 0$ A正定 □

② $\forall \lambda_i \ge 0$ 有 $x^TAx = \sum \lambda_i y_i^2 \ge 0$ A半正定 □

$\Rightarrow$ (以第①为例)

由 $\forall x$ 有 $x^TAx > 0$

而解 $A = QDQ^T$ $D = diag(\lambda_1, \cdots, \lambda_n)$

取 x 为 $\forall \lambda_i$ 对应的特征向量 α_i

$\therefore \quad x^T A x = \alpha_i^T A \alpha_i = \alpha_i^T \lambda_i \alpha_i = \lambda_i \|\alpha_i\| > 0$

$\therefore \quad \lambda_i > 0 \quad \square$