第七章　特征值问题的迭代解法

当矩阵规模很大时, 计算其所有的特征值和特征向量是非常困难的. 而在实际应用中, 我们通常也只对其中的某些特征值和特征向量感兴趣, 因此也没必要计算所有的特征值和特征向量.

本章讨论计算部分特征值和特征值向量的迭代解法. 这些算法的存储量要远小于 $O(n^2)$, 运算量也远小于 $O(n^3)$.

7.1 投影算法

最简单的特征值问题就是仅仅计算一个特征值, 如计算模最大的特征值. 这时我们可以使用幂迭代算法.

算法 7.1 幂迭代: 计算最大特征值

1: Given $x^{(0)}$
2: **for** $i = 1, 2, \ldots$, until converge **do**
3: $\quad y^{(i)} = Ax^{(i-1)}$
4: $\quad x^{(i)} = y^{(i)} / \|y^{(i)}\|_2$
5: **end for**

幂迭代所产生的迭代向量 $x^{(0)}, x^{(1)}, \ldots, x^{(m-1)}$ 生成一个 Krylov 子空间

$$\mathcal{K}_m(A, x^{(0)}) = \text{span}\left\{x^{(0)}, Ax^{(0)}, \ldots, A^{m-1}x^{(0)}\right\}.$$

在幂迭代中, 我们取 $x^{(m-1)}$ 为近似特征向量. 显然, 如果我们在 $\mathcal{K}_m(A, x^{(0)})$ 中找出 "最佳" 的近似特征向量, 则收敛速度就可能会大大加快.

下面我们讨论如何在 $\mathcal{K}_m = \mathcal{K}_m(A, x^{(0)})$ 中寻找 "最佳" 的近似特征向量. 设 $A \in \mathbb{R}^{n \times n}$, 并设 $\mathcal{K}_m$ 和 $\mathcal{L}_m$ 是 $\mathbb{R}^n$ 的两个 m 维子空间. 投影算法就是在寻找 A 的近似特征对 $(\tilde{\lambda}, \tilde{x})$, 满足下面的 Petrov-Galerkin 条件

$$\text{find } \tilde{\lambda} \in \mathbb{C} \text{ and } \tilde{x} \in \mathcal{K}_m \quad \text{such that} \quad A\tilde{x} - \tilde{\lambda}\tilde{x} \perp \mathcal{L}_m. \tag{7.1}$$

这样的算法我们称为**斜投影算法**. 如果我们取 $\mathcal{L}_m = \mathcal{K}_m$, 则上面的算法就是一个**正交投影算法**, 此时条件 (7.1) 称为 Galerkin 条件.

设 $v_0, v_1, \ldots, v_{m-1}$ 和 $w_0, w_1, \ldots, w_{m-1}$ 分别是 $\mathcal{K}_m$ 和 $\mathcal{L}_m$ 的一组标准正交基. 令 $V_m = [v_0, v_1, \ldots, v_{m-1}]$, $W_m = [w_0, w_1, \ldots, w_{m-1}]$. 则对任意 $\tilde{x} \in \mathcal{K}_m$, 存在向量 $y \in \mathbb{R}^n$ 使得 $\tilde{x} = V_m y$, 即 $\tilde{x}$ 可以由 $v_0, v_1, \ldots, v_{m-1}$ 线性表出. 根据条件 (7.1), 我们可得

$$(AV_m y - \tilde{\lambda} V_m y, w_i) = 0, \quad i = 1, 2, \ldots, m,$$

即

$$W_m^T A V_m y = \tilde{\lambda} W_m^T V_m y. \tag{7.2}$$

这是一个广义特征值问题. 如果我们取 $\mathcal{L}_m = \mathcal{K}_m$, 并令 $W_m = V_m$, 则 (7.2) 就化为

$$T_m y = \tilde{\lambda} y, \tag{7.3}$$

其中 $T_m = V_m^T A V_m \in \mathbb{R}^{m \times m}$. 这意味着 $(\tilde{\lambda}, y)$ 是矩阵 T_m 的一个特征对. 由于 m 通常比较小, 因此我们可以使用先前讨论的方法 (如 QR 迭代) 来计算 $(\tilde{\lambda}, y)$. 这样我们就可以计算出 A 的一个近似特征对 $(\tilde{\lambda}, \tilde{x})$, 其中 $\tilde{x} = V_m y$.

7.2 Rayleigh-Ritz 算法

事实上, 我们可以在 $\mathcal{K}_m(A, x^{(0)})$ 中找出 m 个最佳近似特征向量及相应的最佳近似特征值. 这些近似特征值和近似特征向量就是 Ritz 值 和 Ritz 向量.

定义 7.1 设 $\mathcal{K}_m$ 是 $\mathbb{R}^{n\times n}$ 的一个 m 维子空间, 它的一组标准正交基为 $v_0, v_1, \ldots, v_{m-1}$, 并令 $V_m = [v_0, v_1, \ldots, v_{m-1}]$. 记 $T_m = V_m^T A V_m$, 设 $(\tilde{\lambda}, y)$ 是 T_m 的一组特征对, 即 $T_m y = \tilde{\lambda} y$ 且 $\|y\|_2 = 1$. 则我们成 $\tilde{\lambda}$ 是 A 的一个 Ritz 值, $\tilde{x} = V_m y$ 是 A 的一个 Ritz 向量.

Rayleigh-Ritz 算法 就是用 Ritz 值和 Ritz 向量来近似 A 的特征值与特征向量.

算法 7.2 Rayleigh Ritz procedure

1: Compute an orthonormal basis of $\mathcal{K}$: $V_m = [v_0, v_1, \ldots, v_{m-1}]$.
2: Compute the eigenvalues of the matrix $T_m = V_m^T A V_m$, *i.e.*, the Ritz values of A.
3: Select k desired ones: $\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_k$ where $k \le m$.
4: Compute the eigenvectors y_i of T_m associated with $\tilde{\lambda}_i$, $i = 1, 2, \ldots, k$.
5: Approximate the eigenvectors of A by the Ritz vectors $\tilde{x}_i = V_m y_i$, $i = 1, 2, \ldots, k$.

7.2.1 对称矩阵

这里我们讨论 A 对称时, Rayleigh Ritz 算法的性质.

设 $V_u \in \mathbb{R}^{n\times n-m}$ 是一个列正交矩阵, 且使得 $V=[V_m, V_u] \in \mathbb{R}^{n\times n}$ 是一个正交矩阵. 于是我们有

$$V^TAV = [V_m, V_u]^T A [V_m, V_u] = \begin{bmatrix} V_m^TAV_m & V_m^TAV_u \\ V_u^TAV_m & V_u^TAV_u \end{bmatrix}.$$

由于 A 对称的, 故 $T_m = V_m^TAV_m \in \mathbb{R}^{m\times m}$ 也是对称的. 此时, Ritz 值和 Ritz 向量具有下面的最优性质.

定理 7.1 设 $A \in \mathbb{R}^{n\times n}$ 对称, 则对任意的对称矩阵 $R \in \mathbb{R}^{m\times m}$, 有

$$\|AV_m - V_mR\|_2 \geq \|AV_m - V_mT_m\|_2,$$

即 $\|AV_m - V_mR\|_2$ 在 $R = T_m$ 处取最小值, 此时 $\|AV_m - V_mR\|_2 = \|V_u^TAV_m\|_2$.

证明. Let $R = T_m + Z$ where $Z \in \mathbb{R}^{m\times m}$ is symmetric. Note that both A and T_m are symmetric and $T_m = V_m^TAV_m$, we have

$$\begin{aligned}
\|AV_m - V_mR\|_2^2 &= (AV_m - V_mR)^T(AV_m - V_mR) \\
&= \big(AV_m - V_m(T_m+Z)\big)^T\big(AV_m - V_m(T_m+Z)\big) \\
&= \|AV_m - V_mT_m\|_2^2 - V_m^TAV_mZ + T_mV_m^TV_mZ \\
&\quad - ZV_m^TAV_m + ZV_m^TV_mT_m + Z^TV_m^TV_mZ \\
&= \|AV_m - V_mT_m\|_2^2 + \|Z\|_2^2.
\end{aligned}$$

It follows that $\|AV_m - V_mR\|_2^2$ is minimized when $Z = 0$ and

$$\|AV_m - V_mT_m\|_2 = \left\|VV^TAV_m - V_mT_m\right\|_2$$

$$
\begin{aligned}
&= \left\| (V_m, V_u) \begin{bmatrix} V_m^T A V_m \\ V_u^T A V_m \end{bmatrix} - V_m T_m \right\|_2 \\
&= \left\| V_u (V_u^T A V_m) \right\|_2 \\
&= \left\| V_u^T A V_m \right\|_2 .
\end{aligned}
$$

□

注: 定理 7.1 中的 2-范数可以改成任意的酉不变范数, 如 F-范数.

定理 7.2 设 $A \in \mathbb{R}^{n\times n}$ 对称, 并设 $T_m = U\Lambda U^T$ 是 $T_m = V_m^T A V_m$ 的特征值分解. 设 $Q \in \mathbb{R}^{n\times m}$ 是满足 $\mathrm{span}(Q) = \mathcal{K}$ 的任意列正交矩阵, $D \in \mathbb{R}^{m\times m}$ 是任意对角矩阵. 我们有

$$
\|AQ - QD\|_2 \ge \|AV_m - V_m T_m\|_2,
$$

且当 $Q = V_m U$, $D = \Lambda$ 时等式成立.

证明. Because $\mathrm{span}(Q) = \mathcal{K} = \mathrm{span}(V_m)$, it follows that there exist a matrix $W \in \mathbb{R}^{m\times m}$ such that $Q = V_m W$. As Q is column orthogonal, we have

$$
I = Q^T Q = (V_m W)^T V_m W = W^T V_m^T V_m W = W^T W,
$$

which implies that W is orthogonal. Let $WDW^T = T_m + Z$. Then we have

$$
\|AQ - QD\|_2^2 = \|AV_m W - V_m W D\|_2^2
$$

$$
\begin{aligned}
&= \|AV_m - V_m W D W^T\|_2^2 \\
&= \|AV_m - V_m T_m\|_2^2 + \|Z\|_2^2 \\
&\geq \|AV_m - V_m T_m\|_2^2.
\end{aligned}
$$

If we choose $W = U$ and $D = \Lambda$, then $Z = WDW^T - T_m = U\Lambda U^T - T_m = 0$ and, hence, the above equality holds. $\square$

This result shows that the minimum of $\|AQ - QD\|_2$ over all n-by-k column orthogonal matrices Q with $\operatorname{span}(Q) = \mathcal{K}$ and over all diagonal matrices D is attained by $Q = V_m U$ and $D = \Lambda$.

定理 7.1 和定理 7.2 表明, 在 $\|AQ - QD\|_2$ 极小的意义下, Ritz 值是特征值的 “最佳” 近似. 所以我们用 Ritz 值作为特征值的近似是有道理的.

7.3 Lanczos 算法

设 $A \in \mathbb{R}^{n\times n}$ 是对称矩阵. Lanczos 就是利用 Lanczos 算法来计算 $\mathcal{K}_m$ 的基和 $T_m = V_m^T A V_m$, 然后计算 A 的 Ritz 值和 Ritz 向量.

算法 7.3 Lanczos Algorithm

```
Choose a vector v_0 such that ||v_0|| = 1, and set β_0 = 0
for j = 0, 1, ... do
    Compute w = A v_j − β_j v_{j−1}
    α_{j+1} = (w, v_j)
    w = w − α_{j+1} v_j
    β_{j+1} = ||w||_2
    if β_{j+1} = 0 then
        stop
    end if
    v_{j+1} = w / β_{j+1}
    Compute the eigenvalues and eigenvectors of T_j
    Check the convergence
end for
```

在 Lanczos 算法 ?? 中, 迭代 m 步后, 向量 $v_0, v_1, \ldots, v_{m-1}$ 构成子空间

$$\mathcal{K}_m(A, v_0) = \mathrm{span}\left\{v_0, Av_0, \ldots, A^{m-1}v_0\right\},$$

的一组基, 并且有

$$AV_m = V_m T_m + \beta_m v_m e_m^T,$$

其中 $e_m = [0, 0, \ldots, 0, 1]^T \in \mathbb{R}^m$,

$$T_m = V_m^T A V_m = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \beta_{m-1} & \alpha_m \end{bmatrix}.$$

设 $(\tilde{\lambda}, y)$ 是 T_m 的一个特征对, 则有

$$A(V_m y) = V_m T_m y + \beta_m v_m e_m^T y = \tilde{\lambda} V_m y + \beta_m (e_m^T y) v_m.$$

于是

$$Ax - \tilde{\lambda} x = \beta_m (e_m^T y) v_m,$$

即

$$\|Ax - \tilde{\lambda} x\|_2 = |\beta_m (e_m^T y)|,$$

其中 $x = V_m y$. 如果 $|\beta_m (e_m^T y)|$ 很小, 则我们就认为 $\tilde{\lambda}$ 是 A 的某个特征值的很好的近似. 事实上, 关于 Ritz 值 $\tilde{\lambda}$, 我们有下面的性质.

引理 7.1 设 $A \in \mathbb{R}^{n\times n}$ 是对称矩阵, 设 $r = Ax - \tilde{\lambda}x$, 其中 $x \neq 0$. 则

$$\min_{\lambda\in\sigma(A)} |\lambda - \tilde{\lambda}| \leq \frac{\|r\|_2}{\|x\|_2},$$

其中 $\sigma(A)$ 表示 A 的谱, 即所有特征值组成的集合.

证明. Let $A = U\Lambda U^T$ be the eigendecomposition of A. Then we have

$$r = (A - \tilde{\lambda}I)x = U(\Lambda - \tilde{\lambda}I)U^T x.$$

Denote $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. For any vector $z = [z_1, z_2, \ldots, z_n]^T \in \mathbb{R}^n$, it holds that

$$\begin{aligned}
\|(\Lambda - \tilde{\lambda}I)z\|_2^2 &= \sum_{i=1}^{n} (\lambda_i - \tilde{\lambda})^2 z_i^2 \\
&\geq \sum_{i=1}^{n} \min_{\lambda\in\sigma(A)} |\lambda - \tilde{\lambda}|^2 z_i^2 \\
&= \|z\|_2^2 \min_{\lambda\in\sigma(A)} |\lambda - \tilde{\lambda}|^2.
\end{aligned}$$

Therefore, we have

$$\|r\|_2 = \|U^T r\|_2 = \|(\Lambda - \tilde{\lambda}I)U^T x\|_2 \geq \|U^T x\|_2 \min_{\lambda\in\sigma(A)} |\lambda - \tilde{\lambda}| = \|x\|_2 \cdot \min_{\lambda\in\sigma(A)} |\lambda - \tilde{\lambda}|,$$

which complete the proof. □

由引理 7.1 可知, 存在 A 的某个特征值 λ, 使得

$$|\lambda - \tilde{\lambda}| \leq \frac{\|\beta_m(e_m^T y)v_m\|_2}{\|V_m y\|_2} = \frac{|\beta_m| \cdot |e_m^T y|}{\|y\|_2} = |\beta_m| \cdot |e_m^T y|. \tag{7.4}$$

注: 在前面的讨论中, 我们都没有考虑实际计算时可能的误差. 在实际计算中, 由于浮点运算的误差, 即使 m 很小 (如 $m = 10$ 或 $m = 20$), 也可能会导致向量 $\{v_i\}$ 失去正交性. 这时我们必须采取一些补救措施, 最简单的方法就是对它们重新来一次正交话, 即在算法 ?? 的第 5 步后加上一条语句

$$w = w - \sum_{i=1}^{j} (w, v_i) v_i.$$

这个过程就称为带全正交过程的 Lanczos 算法. 显然, 这个过程是非常费时的.

另外一个可行的方法就是选择性正交.

7.4 Arnoldi 算法

这里考虑非对称情形, 即计算非对称矩阵 A 的特征值. 与 Lanczos 算法的思想相类似, 我们可以使用 Arnoldi 算法, 即通过 Arnoldi 算法计算 $\mathcal{K}_m$ 的标准正交基 $v_0, v_1, \ldots, v_{m-1}$ 和上 Hessenberg 矩阵 $H_m = V_m^T A V_m$, 使得

$$AV_m = V_m H_m + h_{m+1,m} v_m e_m^T \quad \text{and} \quad V_m^T A V_m = H_m.$$

但此时 H_m 只是上 Hessenberg, 而不是对称三对角. 但我们同样可以通过计算 H_m 的特征值和特征向量来得到 A 的 Ritz 值的 Ritz 向量, 并用它们来近似 A 的特征值和特征向量.

设 $(\tilde{\lambda}, y)$ 是 H_m 的一个特征对, 其中 $\|y\|_2 = 1$, 则

$$A(V_m y) = V_m H_m y + h_{m+1,m} v_m e_m^T y = \tilde{\lambda} V_m y + h_{m+1,m}(e_m^T y) v_m,$$

所以

$$\|Ax - \tilde{\lambda} x\|_2 = \|h_{m+1,m}(e_m^T y) v_m\|_2 = |h_{m+1,m}| \cdot |e_m^T y|,$$

其中 $x = V_m y$. 若 $|h_{m+1,m}| \cdot |e_m^T y|$ 足够小, 我就认为 $(\tilde{\lambda}, x)$ 是 A 的某个特征对的近似.

算法 7.4 Arnoldi Algorithm

1: Choose a vector v_0 such that $\|v_0\|_2 = 1$
2: **for** $j = 0, 1, \ldots$ **do**
3:     Compute $w_{j+1} = Av_j$

for $i = 0, 1, \ldots, j$ **do**
  $h_{ij} = (w_{j+1}, v_i)$
  $w_{j+1} = w_{j+1} - h_{ij} v_i$
end for
$h_{j+1,j} = \|w_{j+1}\|_2$
if $h_{j+1,j} = 0$ **then**
  stop
end if
$v_{j+1} = w_{j+1}/h_{j+1,j}$
Compute the eigenvalues and eigenvectors of T_j
Check the convergence
end for

由于 A 是非对称的, 其特征值可能是复的, 或者是坏条件的, 此时 Lanczos 算法的一些最优性质就不再成立. 尽快如此, 目前还是存在一些有效 Arnoldi 算法的实现方式, 可参见 [22, 28, 29].

7.5 非对称 Lanczos 算法

非对称 Lanczos 算法就是 Lanczos 算法在非对称矩阵上的推广, 它是基于 Lanczos 双正交化过程.

设 v_0 和 w_0 是任意的非零向量, 并设

$$\mathcal{K}_m(A, v_0) = \mathrm{span}\{v_0, Av_0, \ldots, A^{m-1}v_0\}$$

和

$$\mathcal{K}_m(A^T, w_0) = \mathrm{span}\{w_0, A^T w_0, \ldots, (A^T)^{m-1}w_0\}.$$

Lanczos 双正交化过程就是计算 $\mathcal{K}_m(A, v_0)$ 和 $\mathcal{K}_m(A, w_0)$ 的基 $\{v_i\}$ 和 $\{w_i\}$, 满足 $\{v_i\}$ 和 $\{w_i\}$ 相互正交, 即

$$(v_i, w_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

所以

$$W_m^T V_m = I,$$

其中 $V_m = [v_0, v_1, \ldots, v_{m-1}], W_m = [w_0, w_1, \ldots, w_{m-1}]$.

注意, 通常 $\{v_i\}_{i=0}^m$ 或 $\{w_j\}_{j=0}^m$ 本身并不一定正交, 故 V_m 和 W_m 通常并不列正交.

根据 Lanczos 双正交化过程, 我们可得

$$\begin{aligned} AV_m &= V_m T_m + \gamma_m v_m e_m^T, \\ A^T W_m &= W_m T_m^T + \beta_m w_m e_m^T, \end{aligned}$$

其中 $e_m = [0, \ldots, 0, 1]^T \in \mathbb{R}^m$,

$$T_m = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \gamma_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \gamma_{m-1} & \alpha_m \end{bmatrix}.$$

所以

$$W_m^T A V_m = W_m^T V_m T_m + \gamma_m (W_m^T v_m) e_m^T = T_m.$$

设 $\tilde{\lambda}$ 是 T_m 的特征值, 其对应的右特征向量和左特征向量分别为 y 和 z, 且 $\|y\|_2 = \|z\|_2 = 1$, 即

$$T_m y = \tilde{\lambda} y \quad \text{and} \quad z^T T_m = \tilde{\lambda} z^T.$$

于是有

$$\begin{aligned} A V_m y &= V_m T_m y + \gamma_m v_m e_m^T y = \tilde{\lambda} V_m y + \gamma_m e_m^T y v_m, \\ (W_m z)^T A &= (A^T W_m z)^T = (W_m T_m^T z)^T + \beta_m e_m^T z w_m^T = \tilde{\lambda} (W_m z)^T + \beta_m e_m^T z w_m^T. \end{aligned}$$

若 $|\gamma_m (e_m^T y)|$ 和 $|\beta_m (e_m^T z)|$ 足够小, 我们就认为 $\tilde{\lambda}$ 是 A 的某个特征值的近似, 而 $V_m y$ 和 $W_m z$ 就是相应的右特征向量和左特征向量的近似.

算法 7.5 Nonsymmetric Lanczos Algorithm

1: Choose two vectors v_0 and w_0 such that $(v_0, w_0) = 1$

Set $\beta_0 = 0$ and $\gamma_0 = 0$
for $j = 0, 1, \ldots$ **do**
    Compute $\alpha_{j+1} = (Av_j, w_j)$
    $\tilde{v}_{j+1} = Av_j - \alpha_{j+1} v_j - \beta_j v_{j-1}$
    $\tilde{w}_{j+1} = A^T w_j - \alpha_{j+1} w_j - \gamma_j w_{j-1}$
    $\gamma_{j+1} = |(\tilde{v}_{j+1}, \tilde{w}_{j+1})|^{1/2}$
    if $\gamma_{j+1} = 0$ **then**
        stop
    end if
    $\beta_{j+1} = (\tilde{v}_{j+1}, \tilde{w}_{j+1}) / \gamma_{j+1}$
    $v_{j+1} = \tilde{v}_{j+1} / \gamma_{j+1}$
    $w_{j+1} = \tilde{w}_{j+1} / \beta_{j+1}$
    Compute the eigenvalues and eigenvectors of T_j
    Check the convergence
end for

非对称 Lanczos 算法的显著优点就是节省运算量, 缺点是更容易被中断.

参考文献

[1] J. O. Aasen, On the reduction of a symmetric matrix to tridiagonal form, BIT, 11 (1971), 233–242.

[2] R. Barrett, et.al, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, 1994. (http://www.netlib.org/templates/index.html)

[3] Åke Björck, Solving linear least square problems by Gram-Schmidt orthogonalization, BIT, 7 (1967), 1–21.

[4] Åke Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1996.

[5] J. J. M. Cuppen, A Divide and Conquer Method for the Symmetric Tridiagonal Eigenproblem, Numerische Mathematik, 36 (1981), 177–195.

[6] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997.

[7] Z. Drmač and K. Veselić, New fast and accurate jacobi SVD algorithm. I SIAM Journal on Matrix Analysis and Applications, 29 (2008), 1322–1342.

[8] Z. Drmač and K. Veselić, New fast and accurate jacobi SVD algorithm. II SIAM Journal on Matrix Analysis and Applications, 29 (2008), 1343–1362.

[9] K. Fernando and B. Parlett, Accurate singular values and differential qd algorithms, Numerische Mathematik, 67 (1994), 191–229.

[10] N. Gastinel, *Linear Numerical Analysis*, Kershaw Publishing, London, 1083.

[11] G. H. Golub, History of numerical linear algebra: A personal view, Stanford, 2007. Available at http://forum.stanford.edu/events/2007slides/plenary/history-revised-2007-03-19-golub.pdf

[12] G. H. Golub and W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM Journal on Numerical Analysis, Series B, 2 (1965), 205–224.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The 4th Editon, The Johns Hopkins University Press, Baltimore, MD, 2013.

[14] M. Gu and S. C. Eisenstat, A stable algorithm for the rank-1 modification of the symmetric eigenproblem, SIAM Journal on Matrix Analysis and Applications, 15 (1994), 1266–1276.

[15] M. Gu and S. C. Eisenstat, A Divide-and-Conquer algorithm for the bidiagonal SVD, SIAM Journal on Matrix Analysis and Applications, 16 (1995), 79–92.

[16] M. Gu and S. C. Eisenstat, A Divide-and-Conquer algorithm for the symmetric tridiagonal eigenproblem, SIAM Journal on Matrix Analysis and Applications, 16 (1995), 172–191.

[17] A. Hadjidimos, Accelerated overrelaxation method, Mathematics of Computation, 32 (1978), 149–157.

[18] Nicholas J. Higham, *Accuracy and Stability of Numerical Algorithms*, Second Edition, SIAM, Philadelphia, 2002.

[19] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[20] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[21] W. Kahan Numerical Linear Algebra, Canadian Math. Bull., 9 (1966), 757–801.

[22] R. Lehoucq, *Analysis and Implementation of an Implicitly Restarted Arnoldi Iteration*, Ph.D. thesis, Rice University, Houston, TX, 1995.

[23] E. H. Moore, On the reciprocal of the general algebraic matrix, Bull. Amer. Math. Soc., 26 (1920), 394–395.

[24] Christopher C. Paige, Miroslav Rozložník and Zdeněk Strakoš, Modified Gram–Schmidt (MGS), least squares, and backward stability of MGS-GMRES, SIAM Journal on Matrix Analysis and Applications, (28) 2006, 264--284.

[25] B.N. Parlett, *The Symmetric Eigenvalue Problem*, The 2nd Edition, SIAM, Philadelphia, PA, 1998.

[26] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc., 51 (1955), 406–413.

[27] J. Rutter, *A Serial Implementation of Cuppen'S Divide and Conquer Algorithm for the Symmetric Eigenvalue Problem*, Master's Thesis, University of California, 1994.

[28] Y. Saad, *Numerical Methods for Large Eigenvalue Problems: Theory and Algorithms*, Manchester University Press, Manchester, UK, 1992.

[29] D. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, *SIAM Journal on Matrix Analysis and Applications*, 13 (1992), 357–385.

[30] G. W. Stewart, *Matrix Algorithms, Vol I: Basic Decomposition*, SIAM, Philadelphia, PA, 1998.

[31] G. W. Stewart and Ji-guang Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[32] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997.

[33] L. N. Trefethen, Numerical Analysis, in *Princeton Companion to Mathematics*, Edited by T. Gowers, J. Barrow-Green and I. Leader, Princeton University Press, 2008.

[34] D. S. Watkins, *The Matrix Eigenvalue Problem: GR and KrylovSubsp ace Methods*, SIAM, Philadelphia, 2007.

[35] D. S. Watkins and L. Elsner, Convergence of algorithms of decomposition type for the eigenvalue problem, *Linear Algebra and its Applications*, 143 (1991), 19–47.

[36] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford University, Oxford, 1965.

[37] R.S. Varga, *Matrix Iterative Analysis*, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 2000.

[38] D.M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.

[39] 胡家赣, 线性代数方程组的迭代解法, 科学出版社, 1991.

[40] 蒋尔雄, 矩阵计算, 科学出版社, 2008.

[41] 李大明, 数值线性代数, 清华大学出版社.

[42] 孙继广, 矩阵扰动分析, 科学出版社, 北京, 2001.

[43] 魏木生, 广义最小二乘问题的理论与计算, 科学出版社, 北京, 2006.

[44] 徐树方, 矩阵计算的理论与方法, 北京大学出版社, 北京, 1995.