AMS526: Numerical Analysis I
(Numerical Linear Algebra)
Lecture 19: Other Eigenvalue Algorithms;
Computing SVD; Sensitivity of Eigenvalues

Xiangmin Jiao

Stony Brook University
Outline

1. Other Eigenvalue Algorithms (NLA§30)

2. Computing SVD (NLA§31)

3. Generalized Eigenvalue Problems

4. Sensitivity of Eigenvalues (MC§7.2)
Three Alternative Algorithms

- Jacobi algorithm: earliest known method
- Bisection method: standard way for finding few eigenvalues
- Divide-and-conquer: faster than QR and amenable to parallelization
- We only cover Jacobi algorithm here
The Jacobi Algorithm

- Diagonalize $2 \times 2$ real symmetric matrix by *Jacobi rotation*

$$J^T \begin{bmatrix} a & d \\ d & b \end{bmatrix} J = \begin{bmatrix} \neq 0 & 0 \\ 0 & \neq 0 \end{bmatrix}$$

where $J = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$, and $\tan(2\theta) = 2d/(b - a)$

- What are its similarity and differences with Givens rotation?
- Iteratively apply transformation to two rows and two corresponding columns of $A \in \mathbb{R}^{m \times m}$
- Need not tridiagonalize first, but loop over all pairs of rows and columns by choosing greedily or cyclically
- Magnitude of nonzeros shrink steadily, converging quadratic
- In each iteration, $O(m^2)$ Jacobi rotation, $O(m)$ operations per rotation, leading to $O(m^3 \log(\| \log \epsilon_{\text{machine}} \|))$ flops total
- Jacobi method is easy to parallelize (QR algorithm does not scale well), delivers better accuracy than QR algorithm, but far slower than QR algorithm
Outline

1. Other Eigenvalue Algorithms (NLA§30)

2. Computing SVD (NLA§31)

3. Generalized Eigenvalue Problems

4. Sensitivity of Eigenvalues (MC§7.2)
Intuitive idea for computing SVD of $A \in \mathbb{R}^{m \times n}$:

- Form $A^* A$ and compute its eigenvalue decomposition $A^* A = V \Lambda V^*$
- Let $\Sigma = \sqrt{\Lambda}$, i.e., $\text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$
- Solve system $U \Sigma = AV$ to obtain $U$

This method can be very efficient if $m \gg n$.

However, it is not very stable, especially for smaller singular values because of the squaring of the condition number

- For SVD of $A$, $|\tilde{\sigma}_k - \sigma_k| = O(\epsilon_{\text{machine}} \| A \|)$, where $\tilde{\sigma}_k$ and $\sigma_k$ denote the computed and exact $k$th singular value
- If computed from eigenvalue decomposition of $A^* A$, $|\tilde{\sigma}_k - \sigma_k| = O(\epsilon_{\text{machine}} \| A \|^2 / \sigma_k)$, which is problematic if $\sigma_k \ll \| A \|$

If one is interested in only relatively large singular values, then using eigenvalue decomposition is not a problem. For general situations, a more stable algorithm is desired.
Computing the SVD

- Typical algorithm for computing SVD are similar to computation of eigenvalues
- Consider $A \in \mathbb{C}^{m \times m}$, then hermitian matrix $H = \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$ has eigenvalue decomposition

$$H \begin{bmatrix} V & V \\ U & -U \end{bmatrix} = \begin{bmatrix} V & V \\ U & -U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix},$$

where $A = U\Sigma V^*$ gives the SVD. This approach is stable.
- In practice, such a reduction is done implicitly without forming the large matrix
- Typically done in two or more stages:
  - First, reduce to bidiagonal form by applying different orthogonal transformations on left and right,
  - Second, reduce to diagonal form using a variant of QR algorithm or divide-and-conquer algorithm
Outline

1. Other Eigenvalue Algorithms (NLA§30)
2. Computing SVD (NLA§31)
3. Generalized Eigenvalue Problems
4. Sensitivity of Eigenvalues (MC§7.2)
Generalized Eigenvalue Problem

- Generalized eigenvalue problem has the form

\[ Ax = \lambda Bx, \]

where \( A \) and \( B \) are \( m \times m \) matrices.

- For example, in structural vibration problems, \( A \) represents the stiffness matrix, \( B \) the mass matrix, and eigenvalues and eigenvectors determine natural frequencies and modes of vibration of structures.

- If \( A \) or \( B \) is nonsingular, then it can be converted into standard eigenvalue problem \( (B^{-1}A)x = \lambda x \) or \( (A^{-1}B)x = (1/\lambda)x \).

- If \( A \) and \( B \) are both symmetric, preceding transformation loses symmetry and in turn may lose orthogonality of generalized eigenvectors. If \( B \) is positive definite, alternative transformation is

\[ (L^{-1}AL^{-T})y = \lambda y, \text{ where } B = LL^T \text{ and } y = L^Tx. \]

- If \( A \) and \( B \) are both singular or indefinite, then use QZ algorithm to reduce \( A \) and \( B \) into triangular matrices simultaneously by orthogonal transformation (see Golub and van Loan for detail).
Outline

1. Other Eigenvalue Algorithms (NLA§30)
2. Computing SVD (NLA§31)
3. Generalized Eigenvalue Problems
4. Sensitivity of Eigenvalues (MC§7.2)
Sensitivity of Eigenvalues

- Condition number of matrix $X$ determines sensitivity of eigenvalues

**Theorem**

Let $A \in \mathbb{C}^{n \times n}$ be a nondefective matrix, and suppose $A = X \Lambda X^{-1}$, where $X$ is nonsingular and $\Lambda$ is diagonal. Let $\delta A \in \mathbb{C}^{n \times n}$ be some perturbation of $A$, and let $\mu$ be an eigenvalue of $A + \delta A$. Then $A$ has an eigenvalue $\lambda$ such that

$$|\mu - \lambda| \leq \kappa_p(X)\|\delta A\|_p$$

for $1 \leq p \leq \infty$.

- $\kappa_p(X)$ measures how far eigenvectors are from linear dependence
- For normal matrices, condition number $\kappa_2(X) = 1$ and $\kappa_p(X) = O(1)$, so eigenvalues of normal matrices are always well-conditioned
Sensitivity of Eigenvalues

Proof.

Let \( \delta \Lambda = X^{-1} (\delta A) X \). Then

\[
\| \delta \Lambda \|_p \leq \| X^{-1} \|_p \| \delta A \|_p \| X \|_p = \kappa_p(X) \| \delta A \|_p.
\]

Let \( y \) be an eigenvector of \( \Lambda + \delta \Lambda \) associated with \( \mu \). Suppose \( \mu \) is not an eigenvalue of \( A \), so \( \mu I - \Lambda \) is nonsingular.

\[
(\Lambda + \delta \Lambda)y = \mu y \Rightarrow (\mu I - \Lambda)y = (\delta \Lambda)y \Rightarrow y = (\mu I - \Lambda)^{-1} (\delta \Lambda)y.
\]

Thus

\[
\| (\mu I - \Lambda)^{-1} \|_p^{-1} \leq \| \delta \Lambda \|_p.
\]

\[
\| (\mu I - \Lambda)^{-1} \|_p = |\mu - \lambda|^{-1}, \text{ where } \lambda \text{ is the eigenvalue of } A \text{ closest to } \mu.
\]

Thus,

\[
|\mu - \lambda| \leq \| \delta \Lambda \|_p \leq \kappa_p(X) \| \delta A \|_p.
\]
Left and Right Eigenvectors

To analyze sensitivity of individual eigenvalues, we need to define left and right eigenvectors

1. $Ax = \lambda x$ for nonzero $x$ then $x$ is \textit{right eigenvector} associated with $\lambda$
2. $y^* A = \lambda y^*$ for nonzero $y$, then $y$ is \textit{left eigenvector} associated with $\lambda$

Left eigenvectors of $A$ are right eigenvectors of $A^*$

Theorem

Let $A \in \mathbb{C}^{n \times n}$ have distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with associated linearly independent right eigenvectors $x_1, \ldots, x_n$ and left eigenvectors $y_1, \ldots, y_n$. Then $y_j^* x_i \neq 0$ if $i = j$ and $y_j^* x_i = 0$ if $i \neq j$.

Proof.

If $i \neq j$, $y_j^* Ax_i = \lambda_i y_j^* x_i$ and $y_j^* Ax_i = \lambda_j y_j^* x_i$. Since $\lambda_i \neq \lambda_j$, $y_j^* x_i = 0$. If $i = j$, since $\{x_i\}$ form a basis for $\mathbb{C}^n$, $y_i^* x_i = 0$ together with $y_i^* x_j = 0$ would imply that $y_i = 0$. This leads to a contradiction.
We analyze sensitivity of individual eigenvalues that are distinct.

**Theorem**

Let $A \in \mathbb{C}^{n \times n}$ have $n$ distinct eigenvalues. Let $\lambda$ be an eigenvalue with associated right and left eigenvectors $x$ and $y$, respectively, normalized so that $\|x\|_2 = \|y\|_2 = 1$. Let $\delta A$ be a small perturbation satisfying $\|\delta A\|_2 = \epsilon$, and let $\lambda + \delta \lambda$ be the eigenvalue of $A + \delta A$ that approximates $\lambda$. Then

$$|\delta \lambda| \leq \frac{1}{|y^*x|} \epsilon + O(\epsilon^2).$$

- $\kappa = 1/|y^*x|$ is condition number for eigenvalue $\lambda$
- A simple eigenvalue is sensitive if its associated right and left eigenvectors are nearly orthogonal.
Sensitivity of Individual Eigenvalues

Proof.

We know that $|\delta \lambda| \leq \kappa_p(X)\epsilon = O(\epsilon)$. In addition, $\delta x = O(\epsilon)$ when $\lambda$ is a simple eigenvalue (proof omitted). Because

$$(A + \delta A)(x + \delta x) = (\lambda + \delta \lambda)(x + \delta x),$$

thus

$$(\delta A) x + A(\delta x) + O(\epsilon^2) = (\delta \lambda)x + \lambda(\delta x) + O(\epsilon^2).$$

Left multiplying by $y^*$ and using equation $y^* A = \lambda y^*$, we obtain

$$y^* (\delta A)x + O(\epsilon^2) = (\delta \lambda) y^* x + O(\epsilon^2)$$

and hence

$$\delta \lambda = \frac{y^* (\delta A)x}{y^* x} + O(\epsilon^2).$$

Since $|y^* (\delta A)x| \leq \|y\|_2 \| (\delta A)\|_2 \|x\|_2 = \epsilon$, $|\delta \lambda| \leq \frac{1}{|y^* x|} \epsilon + O(\epsilon^2)$.
Sensitivity of Multiple Eigenvalues and Eigenvectors

- Sensitivity of multiple eigenvalues is more complicated
  - For multiple eigenvalues, left and right eigenvectors can be orthogonal, hence very ill-conditioned
  - In general, multiple or close eigenvalues can be poorly conditioned, especially if matrix is defective

- Condition numbers of eigenvectors are also difficult to analyze
  - If matrix has well-conditioned and well-separated eigenvalues, then eigenvectors are well-conditioned
  - If eigenvalues are ill-conditioned or closely clustered, then eigenvectors may be poorly conditioned