CS 450: Numerical Analysis\textsuperscript{1}
Eigenvalue Problems

University of Illinois at Urbana-Champaign

\textsuperscript{1}These slides have been drafted by Edgar Solomonik as lecture templates and supplementary material for the book “Scientific Computing: An Introductory Survey” by Michael T. Heath (slides).
Eigenvalues and Eigenvectors

- A matrix $A$ has eigenvector-eigenvalue pair (eigenpair) $(\lambda, x)$ if

- Each $n \times n$ matrix has up to $n$ eigenvalues, which are either real or complex
Eigenvalue Decomposition

- If a matrix $A$ is diagonalizable, it has an eigenvalue decomposition

- $A$ and $B$ are similar, if there exist $Z$ such that $A = ZBZ^{-1}$
Similarity of Matrices

<table>
<thead>
<tr>
<th>matrix</th>
<th>similarity</th>
<th>reduced form</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>real symmetric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hermitian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>real</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diagonalizable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>arbitrary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \begin{align*}
A^+ &= A & B &= MAM^{-1} \\
A^{H} &= A & \text{eigenvalues } (B) = \text{eigenvalues } (A) \\
A^H A &= A A^H & \text{diagonalizable} \\
\text{real} & & \text{real} \\
A &= QDU & \text{Schaar form } A = Q D Q^H \\
\text{arbitrary} & & \text{arbitrary} \\
\end{align*} \]
Canonical Forms

- Any matrix is *similar* to a bidiagonal matrix, giving its *Jordan form*:

- Any diagonalizable matrix is *unitarily similar* to a triangular matrix, giving its *Schur form*:
Eigenvectors from Schur Form

Given the eigenvectors of one matrix, we seek those of a similar matrix:

- Its easy to obtain eigenvectors of triangular matrix $T$: 

Activity: Calculating Eigenpairs of a Triangular Matrix
Rayleigh Quotient

For any vector \( x \), the \textit{Rayleigh quotient} provides an estimate for some eigenvalue of \( A \):

\[
\rho_A(x) = \frac{x^\top A x}{x^\top x}
\]
Perturbation Analysis of Eigenvalue Problems

Suppose we seek eigenvalues $D = X^{-1}AX$, but find those of a slightly perturbed matrix $D + \delta D = \hat{X}^{-1}(A + \delta A)\hat{X}$:

$\exists (X)$ \quad if \quad $A = A^T \quad \mathbb{R}(X) = 1$

Gershgorin’s theorem allows us to bound the effect of the perturbation on the eigenvalues of a (diagonal) matrix:

*Given a matrix $A \in \mathbb{R}^{n \times n}$, let $r_i = \sum_{j \neq i} |a_{ij}|$, define the Gershgorin disks as $D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$.***
Gershgorin Theorem Perturbation Visualization

- Top corresponds to Gershgorin disks on complex plane of 4-by-4 real matrix.
- Bottom part corresponds to bounds on Gershgorin disks of $X^{-1}(A + \delta A)X$, which contain the eigenvalues $D$ of $A$ and the perturbed eigenvalues $D + \delta D$ of $A + \delta A$ provided that $||\delta A||$ is sufficiently small.
Conditioning of Particular Eigenpairs

Consider the effect of a matrix perturbation on an eigenvalue $\lambda$ associated with a right eigenvector $x$ and a left eigenvector $y$, $\lambda = y^H A x / y^H x$

$|\tilde{\lambda} - \lambda| \leq \left| \begin{array}{c} y^H (A + \delta A) x \\ y^H x \end{array} \right| + \ldots$

A more accurate eigenvalue approximation than Rayleigh quotient for a normalized perturbed eigenvector (e.g., iterative guess) $\hat{x} = x + \delta x$, can be obtained with an estimate of both eigenvectors (also $\hat{y} = y + \delta y$),

$|\tilde{\lambda} - \lambda| \approx \left| \begin{array}{c} y^H A x + x^H A \delta x \\ y^H x \end{array} \right| \leq \|x\| \|A\| + \left( \|\delta x\| + \|y\| \|x\| \right)$

while $y \approx x$ then $\tilde{\lambda} = \lambda$... (ill-cond.)
Power Iteration

- **Power iteration** can be used to compute the largest eigenvalue of a real symmetric matrix $A$:

  $x^{(i)} = A x^{(i-1)}$

  if $\exists \alpha, \beta$ such that $\lambda = \beta \alpha$ for any other eigenvalue $\lambda'$ of $A$

  $\lim_{i \to \infty} \frac{\|x^{(i)}\|}{\|x^{(i)}\|}$

- The error of power iteration decreases at each step by the ratio of the largest eigenvalues:

  $A = U \Lambda U^H$ where $V^H = U^{-1}$

  $x^{(i)} = A^k x^{(0)} = (U \Lambda U^H)^k x^{(0)} = U \Lambda^k U^H x^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)}$

  $\lambda_1$ is dominant eigenvalue, then $|\alpha_1^{(i)k} / \alpha_1^{(i)k'}| \leq \left( \frac{\lambda_1}{\lambda_{1'}} \right)^k \cdot C$
\[ A^2 = XDX^{-1}XDX^{-1} \]

\[ \Delta A = XDX^{-1} \]

\[ A^k = XDX^{-1} \]
Inverse and Rayleigh Quotient Iteration

- **Inverse iteration** uses LU/QR/SVD of $A$ to run power iteration on $A^{-1}$.
  
  - Power iteration on $A^{-1}:
    - $x(0)$ = random
    - for $k = 0$ to
      - $x(k+1)$ = $A^{-1}x(k)$
      - solve for $x(k+1)$ in $(A^{-1} - \sigma I)x(k+1) = x(k)$
    - normalize $x(k+1)$

- **Rayleigh quotient iteration** provides rapid convergence to an eigenpair.
  
  - Rayleigh, quotient iterate (RQI):
    - $x^{(0)}$ = random
    - for $k = 0$ to...
      - solve for $x^{(k+1)}$ in $n(A - PA(x^{(k)}))x^{(k+1)} = x^{(k)}$
    - $O(n^3k)$ cubic convergence

Activity: Inverse Iteration with a Shift
Activity: Rayleigh Quotient Iteration
Deflation

- Power, inverse, and Rayleigh-quotient iteration compute a single eigenpair, to obtain further eigenpairs, can perform deflation

Given eigenvalue \( \lambda \), eigenvector \( x \), seek \( v \) so that

\[
B = A - \lambda I, \quad v^H B v = 0, \quad \lambda_1, \ldots, \lambda_n \text{ equals of } A
\]

\[ v = y_i \text{ where } y_i \text{ is a left eigenvector of } A, \text{ so } y_i^H A = \lambda y_i \]

For symmetric \( A \), we have that \( y_i = x_i \), \( y_i^H A = \lambda y_i \)

So

\[
B = A - \lambda I, x_i x_i^T
\]

For nonsymmetric, pick \( v = x_i \), observe Schur vectors are stable

\[
B = Q^T A Q - \lambda I, \quad Q^T Q = Q (T - \lambda I) Q^T Q = Q (T - \lambda I) Q
\]

\[ Q = Q (T - \lambda I, e_i e_i^T) \]
We can always compute an orthogonal similarity transformation to reduce a general matrix to \textit{upper-Hessenberg} (upper-triangular plus the first subdiagonal) matrix $H$, i.e. $A = QHQ^T$:

\[
\begin{align*}
A &= QH
\end{align*}
\]

$R = Q^TA$, $B = Q^TAQ = RQ$

In the symmetric case, Hessenberg form implies tridiagonal:

$O(n^3)$
Simultaneous and Orthogonal Iteration

- *Simultaneous iteration* provides the main idea for computing many eigenvectors at once:

- Orthogonal iteration performs QR at each step to ensure stability
QR Iteration

- QR iteration reformulates orthogonal iteration for \( n = k \) to reduce cost/step,

Using induction, we assume \( A_i = \hat{Q}_i^T A \hat{Q}_i \) and show that QR iteration obtains \( A_{i+1} = \hat{Q}_{i+1}^T A \hat{Q}_{i+1} \)
QR Iteration with Shift

- QR iteration can be accelerated using shifting:

- The shift is typically selected to accelerate convergence with respect to a particular eigenvalue:
QR Iteration Complexity

- QR iteration is accelerated by first reducing to upper-Hessenberg or tridiagonal form:
Solving Tridiagonal Symmetric Eigenproblems

A variety of methods exists for the tridiagonal eigenproblem:

▶ QR iteration

▶ Divide and conquer
Solving the Secular Equation for Divide and Conquer

To solve the eigenproblem at each step, the divide and conquer method needs to diagonalize a rank-1 perturbation of a diagonal matrix

\[ A = D + \alpha uu^T. \]
Introduction to Krylov Subspace Methods

- *Krylov subspace methods* work with information contained in the $n \times k$ matrix

$$K_k = \begin{bmatrix} x_0 & Ax_0 & \cdots & A^{k-1}x_0 \end{bmatrix}$$

- The matrix $K_n^{-1}AK_n$ is a *companion matrix* $C$:
Given $Q_k R_k = K_k$, we obtain an orthonormal basis for the Krylov subspace,

$$K_k(A, x_0) = \text{span}(Q_k) = \{ p(A) x_0 : \text{deg}(p) < k \},$$

where $p$ is any polynomial of degree less than $k$.

The Krylov subspace includes the $k - 1$ approximate dominant eigenvectors generated by $k - 1$ steps of power iteration:
Krylov Subspace Methods

The $k \times k$ matrix $H_k = Q_k^T A Q_k$ minimizes $\|A Q_k - Q_k H_k\|_2$:

$H_k$ is Hessenberg, because the companion matrix $C_k$ is Hessenberg:
The eigenvalues/eigenvectors of $H_k$ are the Ritz values/vectors:

The Ritz vectors and values are the *ideal approximations* of the actual eigenvalues and eigenvectors based on only $H_k$ and $Q_k$.
Arnoldi Iteration computes $H = H_n$ directly using the recurrence $q_i^T A q_j = h_{ij}$, where $q_l$ is the $l$th column of $Q_n$:

After each matrix-vector product, orthogonalization is done with respect to each previous vector:
Lanczos Iteration

- Lanczos iteration provides a method to reduce a symmetric matrix to a tridiagonal matrix:

- After each matrix-vector product, it suffices to orthogonalize with respect to two previous vectors:
Cost Krylov Subspace Methods

- The cost of matrix-vector multiplication when the matrix has $m$ nonzeros

- The cost of orthogonalization at the $k$th iteration of a Krylov subspace method is
Restarting Krylov Subspace Methods

- In finite precision, Lanczos generally loses orthogonality, while orthogonalization in Arnoldi can become prohibitively expensive:

- Consequently, in practice, low-dimensional Krylov subspace methods are constructed repeatedly using carefully selected new starting vectors:
Generalized Eigenvalue Problem

- A generalized eigenvalue problem has the form $Ax = \lambda Bx$,

- When $A$ and $B$ are symmetric and $B$ is SPD, we can perform Cholesky on $B$, multiply $A$ by the inverted factors, and diagonalize it:

- Alternative canonical forms and methods exist that are specialized to the generalized eigenproblem.