CS 450: Numerical Analysis¹
Eigenvalue Problems

University of Illinois at Urbana-Champaign

¹These 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

**Invertible similarity transformations** \( \mathbf{Y} = \mathbf{X} \mathbf{A} \mathbf{X}^{-1} \)

<table>
<thead>
<tr>
<th>matrix (A)</th>
<th>reduced form (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>arbitrary</td>
<td></td>
</tr>
<tr>
<td>diagonalizable</td>
<td></td>
</tr>
</tbody>
</table>

**Unitary similarity transformations** \( \mathbf{Y} = \mathbf{U} \mathbf{A} \mathbf{U}^H \)

<table>
<thead>
<tr>
<th>matrix (A)</th>
<th>reduced form (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>arbitrary</td>
<td></td>
</tr>
<tr>
<td>normal</td>
<td></td>
</tr>
<tr>
<td>Hermitian</td>
<td></td>
</tr>
</tbody>
</table>

**Orthogonal similarity transformations** \( \mathbf{Y} = \mathbf{Q} \mathbf{A} \mathbf{Q}^T \)

<table>
<thead>
<tr>
<th>matrix (A)</th>
<th>reduced form (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>real</td>
<td></td>
</tr>
<tr>
<td>real symmetric</td>
<td></td>
</tr>
<tr>
<td>SPD</td>
<td></td>
</tr>
</tbody>
</table>
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*:

- Real matrices are *orthogonally similar* to a block-triangular real matrix with $1 \times 1$ or $2 \times 2$ blocks (real 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 $T$:
Rayleigh Quotient

- For any vector $x$, the *Rayleigh quotient* provides an estimate for some eigenvalue of $A$: 
For non-defective $A = XD X^{-1}$, the eigenvalues of $A + \delta A = \hat{X}(D + \delta D) X^{-1}$ satisfy $\|\delta D\| \leq \kappa(X)\|\delta A\|$. 

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 $y$, $\lambda = y^H A x / y^H x$

- 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 $\hat{y} = y + \delta y$),
Google’s PageRank

A well-known application of eigenproblems is the problem of ranking \( n \) web-pages.
Power Iteration

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

  - The error of power iteration decreases at each step by the ratio of the largest eigenvalues:
If the error at the $k$th step with respect to the desired solution is $e_k$, $r$th order convergence implies that $\lim_{k \to \infty} \frac{\|e_k\|}{\|e_{k-1}\|^r} \leq C$.
Inverse and Rayleigh Quotient Iteration

- **Inverse iteration** uses LU/QR/SVD of $A$ to run power iteration on $A^{-1}$

- **Rayleigh quotient iteration** provides rapid convergence to an eigenpair
Deflation

- Power, inverse, and Rayleigh-quotient iteration compute a single eigenpair. To obtain further eigenpairs, can perform *deflation*
Direct Matrix Reductions

- We can always compute an orthogonal similarity transformation to reduce a general matrix to *upper-Hessenberg* (upper-triangular plus the first subdiagonal) matrix $H$, i.e. $A = QHQ^T$:

- In the symmetric case, Hessenberg form implies tridiagonal:
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
Orthogonal Iteration Convergence

If $A$ has distinct eigenvalues and $R_i$ has positive decreasing diagonal, the $j$th column of $Q_i$ converges to the $j$th Schur vector of $A$ linearly with rate $\max(\frac{|\lambda_{j+1}/\lambda_j|}{\lambda_j}, \frac{|\lambda_j/\lambda_{j-1}|}{\lambda_{j-1}})$. 
QR Iteration

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

- If orthogonal iteration starts with \( \hat{Q}_1 = Q_0 \), then \( \hat{Q}_i = \prod_{j=0}^{i-1} Q_j \),

- QR iteration converges to triangular \( A_i \) if the eigenvalues are distinct in modulus, and in general converges to block-triangular form with a block for each set of eigenvalues of equal modulus.
QR Iteration with Shift

- QR iteration can be accelerated using shifting:

  - The shift is selected to accelerate convergence to an eigenvalue (pair):
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:
Krylov subspace methods work with information contained in the $n \times k$ matrix

$$K_k = [x_0 \ A x_0 \ \cdots \ A^{k-1} x_0]$$

Assuming $K_n$ is invertible, the matrix $K_n^{-1} A K_n$ is a companion matrix $C$:
Krylov Subspaces

★ Given $Q_k R_k = K_k$, we obtain an orthonormal basis for the Krylov subspace,

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

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 $H_k = Q_k^T AQ_k$ minimizes $\|AQ_k - Q_k H_k\|_2$:

- $H_k$ is Hessenberg, because the companion matrix $C_k$ is Hessenberg:
Rayleigh-Ritz Procedure

- The eigenvalues/eigenvectors of $H_k$ are the \textit{Ritz values/vectors}:

- The Ritz vectors and values are the \textit{ideal approximations} of the actual eigenvalues and eigenvectors based on only $H_k$ and $Q_k$: 

\textbf{Demo: Arnoldi vs Power Iteration}
Arnoldi iteration computes the $i$th column of $H_n$, $h_i$ and the $i$th column of $Q_n$ directly using the recurrence $Aq_i = Q_n h_i = \sum_{j=1}^{i+1} h_{ij} q_j$.
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:
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:

Specialized canonical forms and methods exist for the generalized eigenproblem with fewer constraints on $B$ and better cost/stability.