CS 450: Numerical Anlaysis[1]

Eigenvalue Problems

University of Illinois at Urbana-Champaign

[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 $\boldsymbol{A}$ has eigenvector-eigenvalue pair (eigenpair) $(\lambda, \boldsymbol{x})$ if

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$$

  - *For any scalar α, $\alpha\boldsymbol{x}$ is also an eigenvector of $\boldsymbol{A}$ with eigenvalue λ*
  - *Generally, an eigenvalue λ is associated with an eigenspace $\mathcal{X} \subseteq \mathbb{C}^n$ such that each $\boldsymbol{x} \in \mathcal{X}$ is an eigenvector of $\boldsymbol{A}$ with eigenvalue λ.*
  - *The dimensionality of an eigenspace is at most the multiplicity of an eigenvalue (when less, matrix is defective, otherwise matrix is diagonalizable).*
- Each $n \times n$ matrix has up to n eigenvalues, which are either real or complex
  - *The conjugate of any complex eigenvalue of a real matrix is also an eigenvalue.*
  - *The dimensionalities of all the eigenspaces (multiplicity associated with each eigenvalue) sum up to n for a diagonalizable matrix.*
  - *If the matrix is real, real eigenvalues are associated with real eigenvectors, but complex eigenvalues may not be.*

Eigenvalue Decomposition

- If a matrix $\boldsymbol{A}$ is diagonalizable, it has an *eigenvalue decomposition*

  $$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$$

  where $\boldsymbol{X}$ are the right eigenvectors, $\boldsymbol{X}^{-1}$ are the left eigenvectors and $\boldsymbol{D}$ are eigenvalues

  $$\boldsymbol{A}\boldsymbol{X} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{x}_1 & \cdots \boldsymbol{A}\boldsymbol{x}_n \end{bmatrix} = \boldsymbol{X}\boldsymbol{D} = \begin{bmatrix} d_{11}\boldsymbol{x}_1 & \cdots & d_{nn}\boldsymbol{x}_n \end{bmatrix}.$$

  - *If $\boldsymbol{A}$ is symmetric, its right and left singular vectors are the same, and consequently are its eigenvectors.*
  - *More generally, any normal matrix, $\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^H$, has unitary eigenvectors.*
- $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar*, if there exist $\boldsymbol{Z}$ such that $\boldsymbol{A} = \boldsymbol{Z}\boldsymbol{B}\boldsymbol{Z}^{-1}$
  - *Normal matrices are unitarily similar ($\boldsymbol{Z}^{-1} = \boldsymbol{Z}^H$) to diagonal matrices*
  - *Symmetric real matrices are orthogonally similar ($\boldsymbol{Z}^{-1} = \boldsymbol{Z}^T$) to real diagonal matrices*
  - *Hermitian matrices are unitarily similar to real diagonal matrices*

Similarity of Matrices

matrix	*similarity*	*reduced form*
SPD	orthogonal	real positive diagonal
real symmetric	orthogonal	real tridiagonal real diagonal
Hermitian	unitary	real diagonal
normal	unitary	diagonal
real	orthogonal	real Hessenberg
diagonalizable	invertible	diagonal
arbitrary	unitary invertible	triangular bidiagonal

Canonical Forms

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

$$\boldsymbol{A} = \boldsymbol{X} \begin{bmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_k \end{bmatrix} \boldsymbol{X}^{-1}, \quad \forall i, \quad \boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

  the Jordan form is unique modulo ordering of the diagonal Jordan blocks.

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

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^H$$

  where $\boldsymbol{T}$ is upper-triangular, so the eigenvalues of $\boldsymbol{A}$ is the diagonal of $\boldsymbol{T}$. Columns of $\boldsymbol{Q}$ are the Schur vectors.

Eigenvectors from Schur Form

- Given the eigenvectors of one matrix, we seek those of a similar matrix:
  Suppose that $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}$ and $\boldsymbol{B} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$ where $\boldsymbol{D}$ is diagonal,
  - *The eigenvalues of $\boldsymbol{A}$ are $\{d_{11}, \dots, d_{nn}\}$*
  - $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}\boldsymbol{S}^{-1}$ *so $\boldsymbol{S}\boldsymbol{X}$ are the eigenvectors of $\boldsymbol{A}$*
- Its easy to obtain eigenvectors of triangular matrix $\boldsymbol{T}$:
  - *One eigenvector is simply the first elementary vector.*
  - *The eigenvector associated with any diagonal entry (eigenvalue λ) may be obtaining by observing that*

$$\boldsymbol{0} = (\boldsymbol{T} - \lambda \boldsymbol{I})\boldsymbol{x} = \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{u} & \boldsymbol{T}_{13} \\ & 0 & \boldsymbol{v}^T \\ & & \boldsymbol{U}_{33} \end{bmatrix} \begin{bmatrix} -\boldsymbol{U}_{11}^{-1}\boldsymbol{u} \\ 1 \\ \boldsymbol{0} \end{bmatrix},$$

so it suffices to solve $\boldsymbol{U}_{11}\boldsymbol{y} = -\boldsymbol{u}$ to obtain eigenvector $\boldsymbol{x}$.

Rayleigh Quotient

- For any vector $\boldsymbol{x}$, the *Rayleigh quotient* provides an estimate for some eigenvalue of $\boldsymbol{A}$:
$$\rho_{\boldsymbol{A}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^H \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^H \boldsymbol{x}}.$$
  - *If $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$, then $\rho_{\boldsymbol{A}}(\boldsymbol{x})$ is the associated eigenvalue.*
  - *Moreover, for $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, the Rayleigh quotient is the best possible eigenvalue estimate given $\boldsymbol{x}$ and $\boldsymbol{y}$, as it is the solution to $\boldsymbol{x}\alpha \cong \boldsymbol{y}$.*
  - *The normal equations for this scalar-output least squares problem are*
$$\boldsymbol{x}^T \boldsymbol{x}\alpha = \boldsymbol{x}^T \boldsymbol{y} \quad \Rightarrow \quad \alpha = \frac{\boldsymbol{x}^T \boldsymbol{y}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

Perturbation Analysis of Eigenvalue Problems

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

  Note that the eigenvalues of $\boldsymbol{X}^{-1}(\boldsymbol{A} + \boldsymbol{\delta A})\boldsymbol{X} = \boldsymbol{D} + \boldsymbol{X}^{-1}\boldsymbol{\delta A}\boldsymbol{X}$ are also $\boldsymbol{D} + \boldsymbol{\delta D}$. So if we have perturbation to the matrix $||\boldsymbol{\delta A}||_F$, its effect on the eigenvalues corresponds to a (non-diagonal/arbitrary) perturbation $\boldsymbol{\delta}\hat{\boldsymbol{A}} = \boldsymbol{X}^{-1}\boldsymbol{\delta A}\boldsymbol{X}$ of a diagonal matrix of eigenvalues $\boldsymbol{D}$, with norm

  $$||\boldsymbol{\delta}\hat{\boldsymbol{A}}||_F \leq ||\boldsymbol{X}^{-1}||_2||\boldsymbol{\delta A}||_F||\boldsymbol{X}||_2 = \kappa(\boldsymbol{X})||\boldsymbol{\delta A}||_F.$$

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

  Given a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, let $r_i = \sum_{j\neq i}|a_{ij}|$, define the Gershgorin disks as

  $$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}.$$

  The eigenvalues $\lambda_1, \ldots, \lambda_n$ of any matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ are contained in the union of the Gershgorin disks, $\forall i \in \{1, \ldots, n\}, \lambda_i \in \bigcup_{j=1}^n D_j$.