Matrix norms & Iterative methods

Conditioning

Vineet Sahula

sahula@ieee.org

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

- Eigen values

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

- Eigen values
- Vector Norms

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

- Eigen values
- Vector Norms
- Matrix norms

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

- Eigen values
- Vector Norms
- Matrix norms
- Conditioning

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors

Determinants & Eigenvalues:

- Eigen values
- Vector Norms
- Matrix norms
- Conditioning
- Iterative correction from residual vectors

Overview

Matrix iterative Methods: Conditioning and Iterative correction from residual vectors
Determinants & Eigenvalues:

- Eigen values
- Vector Norms
- Matrix norms
- Conditioning
- Iterative correction from residual vectors
- Iterative methods for matrix eigen-values computation

Eigen Values & vector

Eigen Values & vector

- $det\mathbf{A} = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}A_{ij}$

Eigen Values & vector

- $det\mathbf{A} = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}A_{ij}$
- $\mathbf{Ax} = \lambda\mathbf{x} \qquad \Rightarrow \qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$

Similarity transformation

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar, while $\mathbf{T}$ is non-singular

Similarity transformation

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar, while $\mathbf{T}$ is non-singular
  - $\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1} \quad \Rightarrow \quad \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$

Similarity transformation

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar, while $\mathbf{T}$ is non-singular
  - $\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1} \qquad \Rightarrow \qquad \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$
- Similarity matrices have identical Eigen values

Similarity transformation

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar, while $\mathbf{T}$ is non-singular
  - $\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1} \quad \Rightarrow \quad \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$
- Similarity matrices have identical Eigen values
  - replacing $\mathbf{A}$: $\mathbf{T}\mathbf{B}\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{x}$
  - $\mathbf{B}\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{T}^{-1}\mathbf{x}$

Similarity transformation

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be similar, while $\mathbf{T}$ is non-singular
  - $\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1} \quad \Rightarrow \quad \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$
- Similarity matrices have identical Eigen values
  - replacing $\mathbf{A}$: $\mathbf{T}\mathbf{B}\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{x}$
  - $\mathbf{B}\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{T}^{-1}\mathbf{x}$
    - Hence $\mathbf{A}$ and $\mathbf{B}$ have identical Eigen values, i.e. λ

Simiarity Transformation using Diagonal Matrix

- Similarity matrix may be a diagonal matrix $\mathbf{\Lambda}$
  - $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$
  - $\mathbf{A}\mathbf{T} = \mathbf{T}\mathbf{\Lambda}$

Similarity Transformation using Diagonal Matrix

- Similarity matrix may be a diagonal matrix $\mathbf{\Lambda}$
  - $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$
  - $\mathbf{AT} = \mathbf{T}\mathbf{\Lambda}$
- Only elements of $\mathbf{\Lambda}$ are along diagonal, precisely $\lambda_1, \lambda_2, ..., \lambda_n$
  - $\mathbf{A}\mathbf{x_j} = \lambda_\mathbf{j}\mathbf{x_j}$

Similarity Transformation using Diagonal Matrix

- Similarity matrix may be a diagonal matrix $\mathbf{\Lambda}$
  - $\mathbf{A} = \mathbf{T \Lambda T}^{-1}$
  - $\mathbf{AT} = \mathbf{T \Lambda}$
- Only elements of $\mathbf{\Lambda}$ are along diagonal, precisely $\lambda_1, \lambda_2, ..., \lambda_n$
  - $\mathbf{Ax_j} = \lambda_\mathbf{j} \mathbf{x_j}$
- Here, $\mathbf{x}_j$ are j^{th} column of $\mathbf{T}$
  - Thus columns of $\mathbf{T}$ consists of eigen-vectors of $\mathbf{A}$
  - We need be able to choose n lineraly independent eigen-vectors, for $\mathbf{T}$ to be non-singular

Similarity Transformation...

- If $\mathbf{A}$ is symmetric matrix then
  - $\mathbf{T}$ may be chosen to be orthogonal, $\mathbf{T}^{-1} = \mathbf{T}^T$
  - and, $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^T$

Norms

- Vector p-norm
  - $\|\mathbf{x}\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{\frac{1}{p}}$
  - for $p = 2$, $\|\mathbf{x}\|_2 = \left[\mathbf{x}^{\mathbf{T}}\mathbf{x}\right]^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$
  - for $p = \infty$, $\|\mathbf{x}\|_\infty = \underset{1\leq i\leq n}{max} |x_i|$

Norms...

- Matrix norms
  - $\|\mathbf{A}\| = \sup_{\|x\|=1} \|\mathbf{Ax}\| = \max_{\|x\|=1} \|\mathbf{Ax}\|$
  - Also, $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{a}\| + \|\mathbf{B}\|$
  - and, $\|\mathbf{AB}\| \leq \|\mathbf{a}\| \cdot \|\mathbf{B}\|$

Norms...

- Computing matrix norm of order one

  - .
$$
\begin{aligned}
\|\mathbf{Ax}\|_1 &= \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \\
&\leq \sum_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}||x_j| \\
&\leqslant \sum_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}| \right) |x_j| \\
&\leq \sum_{j=1}^{n} \left(\max_{1\leq k\leq n} \sum_{i=1}^{n} |a_{ik}| \right) |x_j| \\
&\leq \left(\max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| \right) (|x_j|)
\end{aligned}
$$
  - The last factor is unity, if $\|\mathbf{x}\| = 1$, and thus
    - $\|\mathbf{A}\|_1 = \max_{\|x\|-1=1} \|\mathbf{Ax}\|_1 \leqslant \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$

Conditioning

- Definition- Condition number of $n \times n$ non-singular matrix $\mathbf{A}$ for the norm $\| \cdot \|_p$ is
  - $k_p(\mathbf{A}) = \|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p$
  - Perturbing $\mathbf{A}$,
    - $(\mathbf{A} + \delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$
    - $\delta\mathbf{x} = -\mathbf{A}^{-1}\delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) + \mathbf{A}^{-1}\delta\mathbf{b}$
    - $$\|\delta\mathbf{x}\| \leqslant \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\|(\|\mathbf{x}\| + \|\delta\mathbf{x}\|) + \|\mathbf{A}^{-1}\| \cdot$$

Conditioning

- $$\|\delta\mathbf{x}\| \leqslant \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\|(\|\mathbf{x}\| + \|\delta\mathbf{x}\|) + \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{b}\|$$

Conditioning

- $$\|\delta \mathbf{x}\| \leqslant \|\mathbf{A}^{-1}\| \cdot \|\delta \mathbf{A}\|(\|\mathbf{x}\| + \|\delta \mathbf{x}\|) + \|\mathbf{A}^{-1}\| \cdot \|\delta \mathbf{b}\|$$
- Computing two terms on right,
  - $\|\mathbf{A}^{-1}\|\|\delta \mathbf{b}\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\|\mathbf{x}\|\frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$
  - $\|\mathbf{A}^{-1}\|\|\delta \mathbf{A}\| = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\frac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|} = ke$

Conditioning

- $$\|\delta\mathbf{x}\| \leqslant \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\|(\|\mathbf{x}\| + \|\delta\mathbf{x}\|) + \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{b}\|$$
- Computing two terms on right,
  - $\|\mathbf{A}^{-1}\|\|\delta\mathbf{b}\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\|\mathbf{x}\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$
  - $\|\mathbf{A}^{-1}\|\|\delta\mathbf{A}\| = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|} = ke$
- $\|\delta\mathbf{x}\|(1 - ke) \leqslant ke\|\mathbf{x}\| + \mathbf{k}\|\mathbf{x}\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$

Conditioning

- $$\|\delta\mathbf{x}\| \leqslant \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\|(\|\mathbf{x}\| + \|\delta\mathbf{x}\|) + \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{b}\|$$
- Computing two terms on right,
  - $\|\mathbf{A}^{-1}\|\|\delta\mathbf{b}\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\|\mathbf{x}\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$
  - $\|\mathbf{A}^{-1}\|\|\delta\mathbf{A}\| = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|} = ke$
- $\|\delta\mathbf{x}\|(1 - ke) \leqslant ke\|\mathbf{x}\| + \mathbf{k}\|\mathbf{x}\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$
- $\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \frac{ke}{(1-ke)}\left(e + \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}\right)$

Iterative Methods

- Let's assume $\mathbf{x_0}$ be an initial approximation to solution of $\mathbf{Ax} = \mathbf{b}$

Iterative Methods

- Let's assume $\mathbf{x_0}$ be an initial approximation to solution of $\mathbf{Ax} = \mathbf{b}$
- Corresponding to $\mathbf{x_0}$, there is a *residual vector* $\mathbf{r_0}$ given by
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b}$

Iterative Methods

- Let's assume $\mathbf{x_0}$ be an initial approximation to solution of $\mathbf{Ax} = \mathbf{b}$
- Corresponding to $\mathbf{x_0}$, there is a *residual vector* $\mathbf{r_0}$ given by
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b}$
- If the bound for $k(\mathbf{A})$ is known, the residual vector may be used as a guide to accuracy
  - $\delta\mathbf{x} = \mathbf{x} - \mathbf{x_0}$
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b} = \mathbf{A}(\mathbf{x} - \delta\mathbf{x}) - \mathbf{b} = -\mathbf{A}\delta\mathbf{x}$

Iterative Methods

- Let's assume $\mathbf{x_0}$ be an initial approximation to solution of $\mathbf{Ax} = \mathbf{b}$
- Corresponding to $\mathbf{x_0}$, there is a *residual vector* $\mathbf{r_0}$ given by
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b}$
- If the bound for $k(\mathbf{A})$ is known, the residual vector may be used as a guide to accuracy
  - $\delta\mathbf{x} = \mathbf{x} - \mathbf{x_0}$
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b} = \mathbf{A}(\mathbf{x} - \delta\mathbf{x}) - \mathbf{b} = -\mathbf{A}\delta\mathbf{x}$
- So that $\delta\mathbf{x}$ is th solution of the system
  - $\mathbf{A}\delta\mathbf{x} = -\mathbf{r_0}$

Iterative Methods

- Let's assume $\mathbf{x_0}$ be an initial approximation to solution of $\mathbf{Ax} = \mathbf{b}$
- Corresponding to $\mathbf{x_0}$, there is a *residual vector* $\mathbf{r_0}$ given by
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b}$
- If the bound for $k(\mathbf{A})$ is known, the residual vector may be used as a guide to accuracy
  - $\delta\mathbf{x} = \mathbf{x} - \mathbf{x_0}$
  - $\mathbf{r_0} = \mathbf{Ax_0} - \mathbf{b} = \mathbf{A}(\mathbf{x} - \delta\mathbf{x}) - \mathbf{b} = -\mathbf{A}\delta\mathbf{x}$
- So that $\delta\mathbf{x}$ is th solution of the system
  - $\mathbf{A}\delta\mathbf{x} = -\mathbf{r_0}$
- We attempt to solve, $\mathbf{x} = \mathbf{x_0} + \delta\mathbf{x}$

Iterative Methods...

- Let's consider solution of $\mathbf{Ax} = \mathbf{b}$

Iterative Methods...

- Let's consider solution of $\mathbf{Ax} = \mathbf{b}$
- If $\mathbf{E}$ and are $n \times n$ matrices, such that $\mathbf{A} = \mathbf{E} - \mathbf{F}$, then
  - $\mathbf{Ex} = \mathbf{Fx} + \mathbf{b}$

Iterative Methods...

- Let's consider solution of $\mathbf{Ax} = \mathbf{b}$
- If $\mathbf{E}$ and are $n \times n$ matrices, such that $\mathbf{A} = \mathbf{E} - \mathbf{F}$, then
  - $\mathbf{Ex} = \mathbf{Fx} + \mathbf{b}$
- This suggests an iterative procedure
  - $\mathbf{Ex_{m+1}} = \mathbf{Fx_m} + \mathbf{b}$ for arbitrary $\mathbf{x_0}$
  - $\mathbf{x_{m+1}} = \mathbf{E}^{-1}\mathbf{Fx_m} + \mathbf{E}^{-1}\mathbf{b}$
  - The sequence $(\mathbf{x_m})_{\mathbf{m=0}}^{\infty}$ converges, if $\|\mathbf{E}^{-1}\mathbf{F}\| < 1$

Iterative Methods...

- Let's consider solution of $\mathbf{Ax} = \mathbf{b}$
- If $\mathbf{E}$ and are $n \times n$ matrices, such that $\mathbf{A} = \mathbf{E} - \mathbf{F}$, then
  - $\mathbf{Ex} = \mathbf{Fx} + \mathbf{b}$
- This suggests an iterative procedure
  - $\mathbf{Ex_{m+1}} = \mathbf{Fx_m} + \mathbf{b}$ for arbitrary $\mathbf{x_0}$
  - $\mathbf{x_{m+1}} = \mathbf{E}^{-1}\mathbf{Fx_m} + \mathbf{E}^{-1}\mathbf{b}$
  - The sequence $(\mathbf{x_m})_{\mathbf{m=0}}^{\infty}$ converges, if $\|\mathbf{E}^{-1}\mathbf{F}\| < 1$
- *Jacobi's* and *Gauss-Seidel* methods

Matrix Eigen-values computation

- Realtion between matrix norms and eigen-values; Gerschgorin theorems

Matrix Eigen-values computation

- Realtion between matrix norms and eigen-values; Gerschgorin theorems
- Simple & inverse iterative method;*Rayleigh* quotient

Matrix Eigen-values computation

- Realtion between matrix norms and eigen-values; Gerschgorin theorems
- Simple & inverse iterative method;*Rayleigh* quotient
- Sturm sequence method

Matrix Eigen-values computation

- Realtion between matrix norms and eigen-values; Gerschgorin theorems
- Simple & inverse iterative method;*Rayleigh* quotient
- Sturm sequence method
- The *QR* algorithm

Matrix Eigen-values computation

- Realtion between matrix norms and eigen-values; Gerschgorin theorems
- Simple & inverse iterative method;*Rayleigh* quotient
- Sturm sequence method
- The *QR* algorithm
- Reduction to tri-diagoonal form: *Householder's* method