Lecture 18 Outline

Monday, October 28, 2019 1:07 PM

Review Session

Tuesday, November 12

3:30pm - 5pm, 805 Furnas

HW1 Grade Corrections

HW5 Updates

Matrix Diagonalization

Eigensystem Solutions

Matrix Diagonalization

Tuesday, October 22, 2019 5:19 PM

Now, let $\underline{A}$ be Normal (i.e. $\underline{A}\,\underline{A}^T = \underline{A}^T \underline{A}$)

$\Rightarrow$ All eigenvectors are orthonormal

$\Rightarrow$ $\underline{S}$ = unitary matrix, such that

$$\underline{S}^T \underline{S} = \underline{S}\,\underline{S}^T = \underline{I} \Rightarrow \underline{S}^{-1} = \underline{S}^T$$

and this $\underline{S}$ is denoted as $\underline{Q}$

If $\underline{A}$ is normal, then

$$\underline{A} = \underline{Q}\,\underline{\Lambda}\,\underline{Q}^T$$ unitary decomposition

Any unitary decomposition is the summation of rank 1 matrices

outer product

Rank 1 matrix has the form $\underline{u}\,\underline{v}^T$

$$\underline{A} = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Phi}^T = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \cdots & \underline{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix}$$

$$= \lambda_1 \underline{x}_1 \underline{x}_1^T + \lambda_2 \underline{x}_2 \underline{x}_2^T + \cdots + \lambda_n \underline{x}_n \underline{x}_n^T$$

or, from another perspective,

$\underline{x}_i \underline{x}_i^T$ is the projection onto the eigenspace with a basis given by $\{\underline{x}_i\}$

What if $\underline{A}$ is non-diagonalizable?

(e.g., $\underline{A}$ is defective)

Use Schur's Theorem, which states that every square matrix $\underline{A}$ can be written as

$\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}$, where $\underline{T}$ is an upper-triangular matrix and $\underline{Q}$ is unitary

Summary

All square matrices: $\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T$

If $\underline{A}$ is complete: $\underline{A} = \underline{S}\,\underline{\Lambda}\,\underline{S}^{-1}$

If $\underline{A}$ is normal: $\underline{A} = \underline{Q}\,\underline{\Lambda}\,\underline{Q}^T$

Now consider real, symmetric $\underline{A}$ with only positive eigenvalues $\rightarrow$ positive definite

One way to establish positive definiteness is to compute all of the eigenvalues

However, this is very expensive $\mathcal{O}(n^3)$

Another approach:

$$\underline{A}\,\underline{x} = \lambda\,\underline{x}$$

$$\underline{x}^T \underline{A}\,\underline{x} = \lambda\,\underline{x}^T \underline{x}$$

with $\underline{x}^T \underline{x} = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 > 0$

Therefore, with $\underline{x}^T \underline{A}\,\underline{x} > 0$, we must have $\lambda > 0$

Beyond this, it can be shown that if $\underline{x}^T \underline{A}\,\underline{x} > 0$ for any $\underline{x}$, then all the λ's will be positive

Here, $\underline{x}^T \underline{A} \, \underline{x}$ is the "energy" definition of positive definite

Eigensystem Solutions

Sunday, October 27, 2019 5:49 PM

$\underline{A}\,\underline{x} = \lambda \underline{x}$ ← Eigenvector $\underline{x}$ is scaled by λ when premultiplied by $\underline{A}$

$\underline{A}$ must be square (column & row space must be equal)

To solve analytically, find all λ, such that

$$|\underline{A} - \lambda \underline{\underline{I}}| = 0$$

Issue: No closed form solutions for polynomials of size ≥ 5

One could do numerical root finding, but that is typically not stable, and one would still need the eigenvectors $\underline{x}$ ⇒ Need iterative solvers for the eigenproblem

Two classes of solvers:

1) Finding largest (or smallest) λ

2) Finding the entire spectrum (or a portion of it)

Largest Eigenvalue

Restrict to real, symmetric $\underline{A}$

- Rayleigh Quotient

Let $\underline{x}$ be an eigenvector of $\underline{A}$, then

$$\underline{A}\,\underline{x} = \underline{\lambda}\,\underline{x}$$

and

$$\underline{x}^T \underline{A}\,\underline{x} = \lambda\, \underline{x}^T \underline{x}$$

$$\therefore \lambda = \frac{\underline{x}^T \underline{A}\,\underline{x}}{\underline{x}^T \underline{x}} \quad \leftarrow \text{Given } \underline{x} + \underline{A}\text{, find } \lambda$$

- Power Iteration

Let $\underline{v}_0$ be any vector such that

$\|\underline{v}_0\| = 1$ and $\underline{v}_0$ is not an eigenvector

Let $\underline{q}_1, \underline{q}_2, \dots, \underline{q}_n$ be the orthonormal set of eigenvectors,

then

$$\underline{v}_0 = a_1 \underline{q}_1 + a_2 \underline{q}_2 + \dots + a_n \underline{q}_n$$

Consider $\underline{A}\,\underline{v}_0$

$$\underline{A}\,\underline{v}_0 = \underline{A}\left(a_1 \underline{q}_1 + a_2 \underline{q}_2 + \dots + a_n \underline{q}_n\right)$$

$$= a_1 \underline{A}\,\underline{q}_1 + a_2 \underline{A}\,\underline{q}_2 + \dots + a_n \underline{A}\,\underline{q}_n$$

$$= a_1 \lambda_1 \underline{q}_1 + a_2 \lambda_2 \underline{q}_2 + \dots + a_n \lambda_n \underline{q}_n$$

$$= \lambda_1 \left(a_1 \underline{q}_1 + a_2 \frac{\lambda_2}{\lambda_1} \underline{q}_2 + \dots + a_n \frac{\lambda_n}{\lambda_1} \underline{q}_n\right)$$

$$\underline{A}^2 \underline{v}_0 = \underline{A}(\underline{A}\,\underline{v}_0) = \underline{A}\,\lambda_1 \left(a_1 \underline{q}_1 + \dots + a_n \frac{\lambda_n}{\lambda_1} \underline{q}_n\right)$$

$$\underline{A}^2 \underline{v}_0 = \underline{A}(\underline{A}\underline{v}_0) = \underline{A}\lambda_1\left(a_1 \underline{q}_1 + \dots + a_n \frac{\lambda_n}{\lambda_1}\underline{q}_n\right)$$

$$= \lambda_1\left(a_1 \underline{A}\,\underline{q}_1 + a_2 \underline{A}\frac{\lambda_2}{\lambda_1}\underline{q}_2 + \dots + a_n \frac{\lambda_n}{\lambda_1}\underline{A}\,\underline{q}_n\right)$$

$$= \lambda_1\left(a_1 \lambda_1 \underline{q}_1 + a_2 \frac{\lambda_2^2}{\lambda_1}\underline{q}_2 + \dots + a_n \frac{\lambda_n^2}{\lambda_1}\underline{q}_n\right)$$

$$= \lambda_1^2\left(a_1 \underline{q}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^2\underline{q}_2 + \dots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^2\underline{q}_n\right)$$

$$\underline{A}^p \underline{v}_0 = \lambda_1^p\left(a_1 \underline{q}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^p\underline{q}_2 + \dots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^p\underline{q}_n\right)$$

Let $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ (order eigenvalues)

Then

$$\lim_{p\to\infty}\left(\frac{\lambda_j}{\lambda_1}\right)^p = 0 \quad \text{for } j \neq 1$$

$$\Rightarrow \lim_{p\to\infty} \underline{A}^p \underline{v}_0 = a_1 \lambda_1^p \underline{q}_1 \quad \text{with } a_1 = \underline{q}_1^T \underline{v}_0$$

Combine this result with the Rayleigh Quotient

Algorithm: Power Algorithm

$\underline{v}_0 \Rightarrow$ Some vector with $\|\underline{v}_0\| = 1$

for $k = 1, 2, \ldots$

$$\underline{w} = \underline{A}\,\underline{v}_{k-1} \qquad \underline{A}^k \underline{v}_0 = \underline{A}\,\underline{A}\,\overbrace{\underline{A}\underbrace{(\underline{A}\,\underline{v}_0)}_{\underline{v}_1}}^{\underline{v}_2}$$

$$\underline{v}_k = \underline{w} / \|\underline{w}\| \quad \rightarrow \quad \underline{v}_k^T v_k = 1$$

$$\lambda_{(k)} = \underline{v}_k^T \underline{A}\,\underline{v}_k \quad \rightarrow \text{Rayleigh quotient}$$

end

This converges at a rate of

$$\|\underline{v}_k - (\pm \underline{q}_1)\| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

$$|\lambda_{(k)} - \lambda_1| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

This causes an issue if $\lambda_1 \sim \lambda_2$

In this case, try an inverse iteration w/shift

Let $\mu \in \mathbb{R}$, such that μ is not an eigenvalue of $\underline{A}$. Then, $\underline{A} - \mu \underline{I}$ has the same eigenvectors as $\underline{A}$ with eigenvalues $\lambda_j - \mu$

Extension: Eigenvectors of $(\underline{A} - \mu \underline{I})^{-1}$ are the same as those for $\underline{A}$, and the eigenvalues for $(\underline{A} - \mu \underline{I})^{-1}$ are $(\lambda_j - \mu)^{-1}$

Let μ be close to λ_1, then $|\lambda_1 - \mu|^{-1}$ will be much larger than $|\lambda_j - \mu|^{-1}$ for $j > 1$

Algorithm: Inverse iteration with shift

Let $\underline{v}_0$ be some vector with $\|\underline{v}_0\| = 1$,

choose $\mu > 0$

for $k = 1, 2, \ldots$

Solve $(\underline{A} - \mu \underline{I}) \underline{w} = \underline{v}_{k-1}$ for $\underline{w}$

$\underline{v}_k = \underline{w} / \|\underline{w}\|$ Normalize $\underline{v}_k$

$\lambda_{(k)} = \underline{v}_k^T \underline{A} \underline{v}_k$ Rayleigh Quotient

$\lambda_{(k)} = \underline{v}_k' \underline{A} \underline{v}_k$ Rayleigh Quotient

Convergence order of

$$\|\underline{v}_k - (\pm \underline{q}_1)\| = O\left(\left|\frac{\mu-\lambda_1}{\mu-\lambda_2}\right|^k\right)$$

$$|\lambda_{(k)} - \lambda_1| = O\left(\left|\frac{\mu-\lambda_1}{\mu-\lambda_2}\right|^{2k}\right)$$

Now combine to get the

Rayleigh Quotient Iteration:

$\underline{v}_0$ is some vector w/ $\|\underline{v}_0\| = 1$

$\lambda_{(0)} = \underline{v}_0^T \underline{A} \underline{v}_0$

for $k = 1, 2, \ldots$

Solve $(\underline{A} - \lambda_{(k-1)} \underline{I})\underline{w} = \underline{v}_{k-1}$ for $\underline{w}$

$\underline{v}_k = \underline{w} / \|\underline{w}\|$

$\lambda_{(k)} = \underline{v}_k^T \underline{A} \underline{v}_k$

$$\lambda_{(k)} = \underline{v}_k^T \underset{\sim}{A} \underline{v}_k$$

end

This method has a convergence of

$$\| \underline{v}_{k+1} - (\pm \underset{\sim}{q}_J) \| = O\left(\| \underline{v}_k - (\pm \underset{\sim}{q}_J) \|^3 \right)$$

$$| \lambda_{(k+1)} - \lambda_J | = O\left(| \lambda_{(k)} - \lambda_J |^3 \right)$$

Cubic order convergence of the eigenvector $\underline{q}_J$ closest to $\underset{\sim}{v}_0$

See Lecture 27 of Trefethan & Bau

Spectrum Calculations

Try to find all or a subset of the eigenvalue spectrum

Recall that any square matrix has the

Schur Decomposition

$$\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T, \text{ where } \underline{T} \text{ is upper triangular}$$

Eigenvalue computations can try the find this decomposition

Note: If $\underline{A}$ is symmetric and real, then

$$\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T = \underline{S}\,\underline{\Lambda}\,\underline{S}^{-1}$$

diagonal

The above looks similar to QR decomposition, where $\underline{A} = \underline{Q}\,\underline{R}$ upper triangular

Recall Householder reflections

$$\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \xrightarrow{\underline{Q}_1^T} \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix}$$

$\underline{A}$ $\qquad\qquad$ $\underline{Q}_1^T \underline{A}$

For the eigenproblem, we need $\underline{Q}_1^T \underline{A} \underline{Q}_1$

Then

$$\underbrace{\underline{Q}_n^T \underline{Q}_{n-1}^T \cdots \underline{Q}_1^T}_{\underline{Q}^T} \underline{A} \underbrace{\underline{Q}_1 \underline{Q}_2 \cdots \underline{Q}_n}_{\underline{Q}} = \underline{T}$$

and then $\underline{A} = \underline{Q}\, \underline{T}\, \underline{Q}^T$

Consider $\underline{Q}_1^T \underline{A} \underline{Q}_1$

$$\underbrace{\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}}_{\underline{Q}_1^T \underline{A}} \longrightarrow \underbrace{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}}_{\underline{Q}_1^T \underline{A} \underline{Q}_1}$$

Fill-in of the zeros

$\therefore$ The original Householder approach will not work

$\Rightarrow$ Not possible to get a Schur Decomposition directly

Instead, two steps are needed:

1) Reduce to upper Hessenberg form, which is nearly upper triangular

2) Iterate until upper triangular is obtained

Details of these two steps:

1) Upper Hessenberg Matrix: A matrix with zeros below the first sub-diagonal

$$\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \end{bmatrix}$$

$$\begin{bmatrix} 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

diagonal

first sub-diagonal

Let $\underline{Q}_1^T$ be a unitary matrix $(\underline{Q}_1^T \underline{Q}_1 = \underline{I})$ that zeros out values below the first subdiagonal of the first column, but does <u>not</u> touch the first row values

$$\underset{\underline{A}}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}} \rightarrow \underset{\underline{Q}_1^T \underline{A}}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}} \rightarrow \underset{\underline{Q}_1^T \underline{A}\, \underline{Q}_1}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}}$$

Use a Householder reflector to assure ortho-normality

$$\underline{Q} = \begin{bmatrix} \underline{I} & \underline{0} \\ \underline{0} & \underline{F} \end{bmatrix}$$

Algorithm: Householder Reduction to Upper Hessenberg Form

for $k = 1$ to $m-2$

$\quad \underline{x} = \underline{A}(k+1:m, k)$

$\quad \underline{v}_k = \text{sign}(x_1) \|\underline{x}\|_2 \underline{e}_1 + \underline{x}$

$\quad \underline{A}(k+1:m, k:m) = \underline{A}(k+1:m, k:m) - 2\underline{v}_k\left(\underline{v}_k^T \underline{A}(k+1:m, k:m)\right)$ $\qquad \underline{Q}^T\underline{A}$

$\quad \underline{A}(1:m, k+1:m) = \underline{A}(1:m, k+1:m) - 2\left(\underline{A}(1:m, k+1:m)\underline{v}_k\right) v_k^T$ $\qquad \underline{Q}^T\underline{A}\underline{Q}$

end

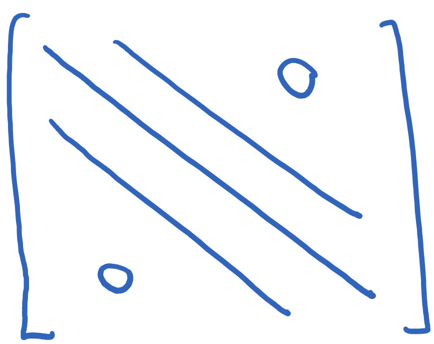
→ $\underline{A}$ then converts to upper Hessenberg

Note: $\underline{Q}$ is never formed

Cost: $\Theta\left(\frac{10}{3} m^3\right)$

If $\underline{A}$ is symmetric, then the cost reduces to $\Theta\left(\frac{4}{3} m^3\right)$

and the result is tri-diagonal



2) Iterate to Upper Triangular Form

Focus on real, symmetric matrices

Turn to $\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T$ (Schur decomposition)

$\underline{A}$ could be any matrix or the result of part 1) (i.e., upper Hessenberg)

$$\underline{T} = \underline{Q}^T \underline{A}\, \underline{Q}$$

Make this an iteration

Given $\underline{A}_k$, let $\underline{A}_{k+1} = \underline{Q}_k^T \underline{A}_k \underline{Q}_k$

Now, let $\underline{A}_k = \underline{Q}_k \underline{R}_k$ be the QR decomposition of $\underline{A}_k$

Then,

$$\underline{A}_{k+1} = \underline{Q}_k^T \underline{A}_k \underline{Q}_k = \underline{Q}_k^T (\underline{Q}_k \underline{R}_k) \underline{Q}_k$$

$$= \underline{I}\, \underline{R}_k \underline{Q}_k = \underline{R}_k \underline{Q}_k$$

Given $\underline{A}_k$, find $\underline{Q}_k \underline{R}_k$, then $\underline{A}_{k+1} = \underline{R}_k \underline{Q}_k$

$\Rightarrow$ This is the QR Algorithm for eigenproblems

Algorithm: QR for Eigenproblems

Let $\underline{A}_0 = \underline{A}$

for $k = 1, 2, \ldots$

$\underline{Q}_k \underline{R}_k = \underline{A}_{k-1}$ (QR of $\underline{A}_{k-1}$)

$\underline{A}_k = \underline{R}_k \underline{Q}_k$ Recombination in reverse

end

Converge to some tolerance,

result will be upper triangular matrix $\underline{T}$

To show why this converges consider the power method applied to matrices