EigenSystems

$\underline{A}\,\underline{x} = \lambda \underline{x}$ ← eigenvector $\underline{x}$ is scaled by λ when you apply $\underline{A}$.

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

To compute analytically, find all λ such that

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

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

You could do numeric root finding, but that is typically not stable, and you still need to get $\underline{x}$!

=> Need iterative solvers for the eigenproblem.

Two classes of solvers:

1) Those that find the largest / smallest λ.

2) Those that find the spectrum (or a portion of it)

Largest Eigenvalue

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

- Rayleigh Quotient

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

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

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

$$\lambda = \frac{\underline{x}^T \underline{A}\,\underline{x}}{\underline{x}^T \underline{x}}$$ ← Given $\underline{x}$ & $\underline{A}$, find λ

- Power Iteration

Let $\underline{v}_0$ be any vector such that $\|\underline{v}_0\| = 1$ and $\underline{v}_0$ is <u>not</u> an eigenvector.

Let $q_1, q_2, \ldots, q_n$ be the orthonormal set of eigenvectors,

Then $\underline{v}_0 = a_1 q_1 + a_2 q_2 + \cdots + a_n q_n$

Look at $\underline{A}\,\underline{v}_0$

$$\underline{A}\underline{v}_0 = \underline{A}(a_1 q_1 + \dots + a_n q_n)$$

$$= a_1 \underline{A} q_1 + a_2 \underline{A} q_2 + \dots + a_n \underline{A} q_n$$

$$= a_1 \lambda_1 q_1 + a_2 \lambda_2 q_2 + \dots + a_n \lambda_n q_n$$

$$= \lambda_1 (a_1 q_1 + a_2 \lambda_2/\lambda_1\, q_2 + \dots + a_n \lambda_n/\lambda_1\, q_n)$$

$$\underline{A}^2 \underline{v}_0 = \underline{A}(\underline{A}\underline{v}_0) = \underline{A}(\lambda_1 (a_1 q_1 + \dots + a_n \lambda_n/\lambda_1\, q_n))$$

$$= \lambda_1 (a_1 \underline{A} q_1 + \dots + a_n \lambda_n/\lambda_1\, \underline{A} q_n)$$

$$= \lambda_1 (a_1 \lambda_1 q_1 + \dots + a_n \lambda_n^2/\lambda_1\, q_n)$$

$$= \lambda_1^2 (a_1 q_1 + \dots + a_n (\lambda_n/\lambda_1)^2 q_n)$$

$$\Rightarrow \underline{A}^n \underline{v}_0 = \lambda_1^n (a_1 q_1 + \dots + a_n (\lambda_n/\lambda_1)^n q_n)$$

Let $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$

Then $\lim\limits_{n\to\infty} \left(\dfrac{\lambda_j}{\lambda_1}\right)^n = 0$ for $j \neq 1$

$$\Rightarrow \lim_{n\to\infty} \underline{A}^n \underline{v}_0 = a_1 \lambda_1^n q_1 \qquad \text{w/ } a_1 = q_1^T \underline{v}_0$$

Combine w/ the Rayleigh Quotient!

Algorithm: Power Algorithm

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

for $k = 1, 2, \ldots,$

$\underline{w} = \underline{A}\,\underline{v}_{k-1}$

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

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

This converges at a rate of:

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

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

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

Try an inverse iteration w/ a shift

Let $\mu \in \mathbb{R}$ that is <u>not</u> an eigenvalue of $\underline{A}$.

$\underline{A} - \mu\underline{I}$ has the same eigenvectors of $\underline{A}$ w/ eigenvalues of $\lambda_j - \mu$

extension: Eigenvectors of $(\underline{A}-\mu \underline{I})^{-1}$ are the same as $\underline{A}$, and the eigenvalues of $\underline{A}$ 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 w/ shift

Let $\underline{v}_0$ = some vector w/ $\|\underline{v}_0\|=1$, choose $\mu>0$

for $k=1,2,\dots$

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

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

$\lambda(k) = \underline{v}_k^T \underline{A}\, \underline{v}_k$ 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!

$\underline{v}_0$ = 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 $\underline{w}$

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

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

This method has a convergence of

$$\|\underline{v}_{k+1} - (\pm q_J)\| = O(\|\underline{v}_k - (\pm q_J)\|^3)$$

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

Cubic order of convergence for the eigenvector q_J closest to $\underline{v}_0$

Look at lecture 27 of Trefethen.

Spectrum Calculations

Try to find all or a subset of the spectrum,

Recall that any square matrix has the Schur Decomposition:

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

Eigenvalue computations try to find this.

Note: If $\underline{A}$ is symmetric & Real then

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

$\uparrow$ Diagonal.

Looks similar to QR: $\underline{A} = \underline{Q}\,\underline{R}$ $\leftarrow$ upper

Recall Householder:

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

For this, though, we need $\underline{Q}_1^T \underline{A} \underline{Q}_1$

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

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

Look at $\underline{Q}_1^T \underline{A} \underline{Q}_1$

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

The original Householder will $\underline{\text{not}}$ work.

$\Rightarrow$ Not possible to get a Schur Decomposition $\underline{\text{Directly}}$

Instead, do 2 steps:

1) Reduce to Upper Hessenberg form
   $\rightarrow$ Close to upper Triangular
2) Iterate until upper triangular

1) Upper Hessenberg Matrix: A matrix w/ zeros below the first sub-diagonal.

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

Diagonal

1st Sub diagonal

Let Q_1^T be a unitary matrix ($Q_1^T Q_1 = I$) that zeros out values below the first subdiagonal of the first column, but does not touch the first row values.

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

Use a Householder Reflector!

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

Algorithm: Householder Reduction to Upper Hessenberg form

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

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

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

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

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

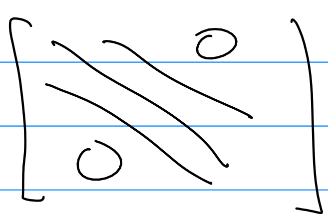
$\Rightarrow$ $\underline{A}$ then converts to upper Hessenberg

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

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

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

and the result is tri-diagonal



Part 2: Iterate to Upper triangular.

Focus on Real Symmetric matrices.

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

$\underline{A}$ could be any matrix or the result of Part 1 (upper Hessenberg)

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

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 $\underline{A}_k$,

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

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 eigen problems

Algorithm: QR for Eigen problems

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,

Converse to some tolerance,

Result will be upper triangular matrix $\underline{T}$.

To show why this converges look at the power method applied to matrices:

Original: chose $\underline{v}_0$, then $\underline{A}^k \underline{v}_0$ approaches q_1

Now, let $\{ \underline{v}_0^{(0)}, \underline{v}_0^{(1)}, \ldots, \underline{v}_0^{(n)} \}$ be a set of n linearly independent vectors close to the n largest eigenvalues:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > |\lambda_{n+1}| \geq \cdots \geq |\lambda_m|$$

then, it should be expected that

$$\{\underline{A}^k \underline{v}_0^{(0)}, \underline{A}^k \underline{v}_0^{(1)}, \ldots, \underline{A}^k \underline{v}_0^{(n)}\}$$

will approach $\{q_1, \ldots, q_n\}$ as $k \to \infty$.

Let $\underline{V}_0 = [\underline{v}_0^{(0)} \mid \underline{v}_0^{(1)} \mid \cdots \mid \underline{v}_0^{(n)}]$,

then $\underline{V}_k = \underline{A}^k \underline{V}_0$

and as $k \to \infty$, then the QR of $\underline{V}_k$

$\underline{V}_k = \underline{Q}_k \underline{R}_k$ converges to the eigenvectors

if $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > |\lambda_{n+1}| \geq \cdots$ holds &

$\hat{\underline{Q}}^T \underline{V}_0$ is non-singular

The true matrix of eigenvalues, not the QR decomposition.

Problem: 1) Converges linearly

2) Stability

To fix stability, orthogonalize every iteration:

Algorithm: Simultaneous Iteration

Let $\underline{A}$ be any matrix (or the result of upper Hessenberg)

Let $\underline{\hat{Q}}_0 \in \mathbb{R}^{m \times n}$ w/ orthonormal columns

for $k = 1, 2, \ldots$

$\underline{Z} = \underline{A}\,\underline{\hat{Q}}_{k-1}$ — make it orthonormal

$\underline{\hat{Q}}_k \underline{\hat{R}}_k = \underline{Z}$ — $\underline{Q}\,\underline{R}$ of $\underline{Z}$

You can show that if $\underline{\hat{Q}}_0 = \underline{I}$, then this $\underline{\text{is}}$ the $\underline{QR}$ method

Look in Trefethen, lectures 26–28