Lecture 20

Last time we studied a few basic properties of eigenvalues of a matrix.

Recall a number λ is said to be an eigenvalue of a matrix A if there exists $x \neq 0$ s.t $\underline{Ax = \lambda x}$

x is said to be eigenvector associated to λ.

characteristic polynomial

$$p(\lambda) = \det(\lambda I - A)$$

clearly λ is an eigenvalue of A iff λ is a root of $p(\lambda)$.

Def" A basis of n-vectors which consists entirely of eigenvectors of B is called a complete set of eigenvectors for B.

We proved that if a matrix A has n-distinct eigen ~~vectors~~ values then it has a complete set of eigenvectors.

If A does not have distinct eigenvalues then it may or may not have a basis consisting of eigenvectors of B.

I gave two examples { one has / one does not have }

Fact If A is a symmetric matrix then it has a complete set of eigenvectors.

We then defined similar matrices
If A and B are similar then eigenvalues of A and B are the same.

Thm Let A be a nxn - matrix

TFAE
1) A has a complete set of eigenvectors
2) A is similar to a diagonal matrix.

Localization results

① λ eigenvalue of A. Then

$|\lambda| \leq \|A\|$ for any matrix norm $\|\ \|$ (associated to a norm in $\mathbb{R}^n$)

Gershgorin discs

Let $A = (a_{ij})_{n \times n}$.

$$R_i = \left\{ z \in \mathbb{C} \mid |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right\}$$

Theorem Let λ be an eigenvalue of A then $\lambda \in R_i$ for some $i = 1, 2, \ldots, n$.

Extension of Gerschgorin Theorem

If p Gerschgorin disks form a set S that is disjoint from the $n-p$ other disks of a given matrix A, then S contains precisely p eigenvalues of A (counted with multiplicity)

Lecture 20

Eigen values of a symmetric matrix

Let A be a symmetric real matrix.

Fact (1) All eigen values of A are real

(2) $\exists$ an orthogonal matrix Q (i.e $Q^{-1} = Q^t$) such that $Q^t A Q$ is a diagonal matrix.

Remark 1 Let Q be an orthogonal matrix.

Let $Q = [v_1, v_2, \ldots, v_n]$ i.e v_i is the i^{th} column of Q.

$$Q^t Q = I \Rightarrow$$

$$v_i^t v_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Thus $\{v_1, v_2, \ldots, v_n\}$ form an orthonormal basis of $\mathbb{R}^n$.

Remark 2

If $D = Q^t A Q$

D a diagonal matrix

Q orthogonal matrix

Let $Q = [v_1, v_2, \ldots, v_n]$. Then $v_1, \ldots, v_n$ are eigenvectors of A

$$Av_i = d_{ii} v_i$$

Rayleigh Quotients

Let A be a symmetric matrix. Let $\bar{x}$ be a non-zero vector

Then $\dfrac{x^t A x}{x^t x}$

is called the Rayleigh Quotient (for x and A)

Theorem :- For a symmetric matrix A

both $\max_{x \neq 0} \frac{x^t A x}{x^t x}$

and $\min_{x \neq 0} \frac{x^t A x}{x^t x}$

are eigenvalues of A. Any other eigenvalue of A lies in between these two.

Proof :- We will use the fact that all eigenvalues of A are real.

Let $\lambda_1, \ldots, \lambda_n$ be eigenvalues of A.

Assume $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$

A symmetric. So $\exists$ an orthonormal basis $\{v_1, \ldots, v_n\}$ consisting of eigenvectors of A

Assume $Av_i = \lambda_i v_i$

Let $x \in \mathbb{R}^n$ be any non-zero vector

$$x = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \qquad x_i \in \mathbb{R}$$

$$x^t x = (x_1 v_1 + \cdots + x_n v_n)^t (x_1 v_1 + \cdots + x_n v_n)$$

$$= x_1^2 + \cdots + x_n^2 \qquad (\because v_i^t v_j = \delta_{ij})$$

$$Ax = \lambda_1 x_1 v_1 + \lambda_2 x_2 v_2 + \cdots + \lambda_n x_n v_n$$

$$x^t A x = (x_1 v_1 + \cdots + x_n v_n)^t (\lambda_1 x_1 v_1 + \cdots + \lambda_n x_n v_n)$$

$$= \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

So $\lambda_n (x_1^2 + \cdots + x_n^2) \leq x^t A x \leq \lambda_1 (x_1^2 + \cdots + x_n^2)$

Thus

$$\lambda_n \leq \frac{x^t A x}{x^t x} \leq \lambda_1.$$

$$\frac{x^t A x}{x^t x} \quad \text{for } x = v_1 \quad = \lambda_1$$

$$\frac{x^t A x}{x^t x} \quad \text{for } x = v_n \quad = \lambda_n$$

Thus $$\max_{x \neq 0} \frac{x^t A x}{x^t x} = \lambda_1$$

$$\min_{x \neq 0} \frac{x^t A x}{x^t x} = \lambda_n$$

The result follows. ∎

Remark Eigenvalues can be complex even if A is a real matrix.

Example $$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$|\lambda I - A| = \begin{vmatrix} \lambda & -1 \\ 1 & \lambda \end{vmatrix}$$

$$= \lambda^2 + 1$$

eigenvalues are $\pm i$

If the matrix A is not too large then one can find eigenvalues of A simply by finding roots of the characteristic polynomial

$$p(\lambda) = \det(\lambda I - A)$$

Since roots can be complex we cannot use Newton's method or secant-method.

We have to use "Muller's method".

For tridigonal symmetric matrices special methods are available. This is of importance since any real symmetric matrix can be transformed by similarity into a real tridigonal symmetric matrix.

Let A be a symmetric tridigonal matrix.

$$A = \begin{bmatrix} a_1 & b_1 & & & & \\ b_1 & a_2 & b_2 & & & \\ & b_2 & a_3 & b_3 & & \\ & & b_3 & \ddots & \ddots & \\ & & & \ddots & & b_{n-1} \\ & & & & b_{n-1} & a_n \end{bmatrix}$$

Its characteristic polynomial can be obtained as $P_n(\lambda) = \det(\lambda I - A)$ with

$$\begin{cases} P_0(\lambda) = 1 \\ P_1(\lambda) = \lambda - a_1 \\ P_j(\lambda) = (\lambda - a_j) P_{j-1}(\lambda) - b_{j-1}^2 P_{j-2}(\lambda) \\ \qquad\qquad\qquad\qquad j = 2, \cdots, n \end{cases}$$

Here $P_j(\lambda)$ = det of matrix formed by first j rows and columns of $\lambda I - A$.

This can be proved by expanding the determinant at the j^{th} row.

The recurrance (*) allows the evaluation of $P_n(\lambda)$ in about $3n$ operations.

Furthermore the recurrance is easily differentiated making it possible to calculate $P_n'(\lambda)$ by recurrance

As eigenvalues of A are real, this allows for the application of Newton's method to calculate the eigenvalues of A.

Recall a symmetric matrix A is said to be positive definite if $x^t A x > 0$ for each $x \neq 0$

Theorem Let A be a symmetric matrix. The following are equivalent

1) A is positive definite
2) All eigenvalues of A are +ve.

Proof Recall that for a symmetric matrix

$$\alpha = \max_{x \neq 0} \frac{x^t A x}{x^t x} \quad \text{and} \quad \min_{x \neq 0} \frac{x^t A x}{x^t x} = \beta$$

are all eigenvalues of A and all other eigenvalues are inbetween these values

1) $\Rightarrow$ 2) As A is +ve definite

$x^t A x > 0$ for $x \neq 0$

also $x^t x = \|x\|_2^2 > 0$ for $x \neq 0$

So $\frac{x^t A x}{x^t x} > 0 \quad \forall x \neq 0$

$\Rightarrow \beta > 0.$

~~Thus if~~

Thus if λ is any eigenvalue of A then

$\lambda \geq \beta > 0$.

So λ is +ve.

2) $\Rightarrow$ 1)

Conversely if all eigenvalues of A are positive

then $\beta > 0$

So $x^t A x > 0$ for any $x \neq 0$

(since $x^t x > 0$ for any $x \neq 0$)

Thus A is positive definite.

Power method to determine eigenvalue of a matrix A

In many practical problems the matrix A has a dominant eigen-value i.e $\exists \lambda_1$ an eigen value of A such that

$$|\lambda_1| > |\lambda_i|$$

for any other eigenvalue λ_i of A.

It is of interest to find this eigenvalue (i.e λ_1) and corresponding eigenvector

Idea of the method

We assume that A has a complete set of eigenvectors $v_1, v_2, \ldots, v_n$

Let $x \in \mathbb{R}^n$. So there exists constants $\beta_1, \beta_2, \ldots, \beta_n$ such that

$$x = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

$$Ax = \beta_1 A v_1 + \beta_2 A v_2 + \cdots + \beta_n A v_n$$

$$= \beta_1 \lambda_1 v_1 + \beta_2 \lambda_2 v_2 + \cdots + \beta_n \lambda_n v_n$$

$$A^2 x = \beta_1 \lambda_1^2 v_1 + \cdots + \beta_n \lambda_n^2 v_n$$

$$A^3 x = \beta_1 \lambda_1^3 v_1 + \cdots + \beta_n \lambda_n^3 v_n$$

In general

$$A^k x = \beta_1 \lambda_1^k v_1 + \beta_2 \lambda_2^k v_2 + \cdots + \beta_n \lambda_n^k v_n$$

If we factor λ_1^k we get

$$A^k x = \lambda_1^k \left(\beta_1 v_1 + \beta_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right)$$

we are assuming $|\lambda_1| > |\lambda_i|$ for $i \geq 2$

So $\lim_{k \to \infty} \left(\frac{\lambda_i}{\lambda}\right)^k = 0$

Thus

$$(*) \quad \lim_{k\to+\infty} A^k x = \lim_{k\to\infty} \lambda_1^k \beta_1 v_1$$

This sequence converges to 0 if $|\lambda_1| < 1$ and diverges if $|\lambda_1| > 1$ provided of course that $\beta_1 \neq 0$

Advantage can be made of the relationship expressed in $(*)$ by scaling powers of $A^k x$ in an appropriate manner to ensure that the limit is finite and non-zero

Recall if $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$$

We write v_j as $v^{(j)}$

Let $x^{(0)}$ be a unit vector w.r.t to $\|\ \|_\infty$

Choose $x^{(0)}_{p_0} = 1 = \|x^{(0)}\|_\infty$

Let $y^{(1)} = A x^{(0)}$ and define $\mu^{(1)} = y^{(1)}_{p_0}$

$$\mu^{(1)} = y^{(1)}_{p_0} = \frac{y^{(1)}_{p_0}}{x^0_{p_0}} = \frac{\beta_1 \lambda_1 v^{(1)}_{p_0} + \sum_{j=2}^{n} \beta_j \lambda_j v^{(j)}_{p_0}}{\beta_1 v^{(1)}_{p_0} + \sum_{j=2}^{n} \beta_j v^{j}_{p_0}}$$

$$= \lambda_1 \left[\frac{\beta_1 v^{(1)}_{p_0} + \sum_{j \geq 2} \beta_j \left(\frac{\lambda_j}{\lambda_1}\right) v^{j}_{p_0}}{\beta_1 v^{(1)}_{p_0} + \sum_{j=2}^{n} \beta_j v^{(j)}_{p_0}} \right]$$

Let p_1 be the least integer such that

$$1 y^{(1)}_{p_1} = \|y^{(1)}\|_\infty$$

$$x^{(1)} = \frac{1}{y^{(1)}_{p_1}} y^{(1)} = \frac{1}{y^{(1)}_{p_1}} A x^{(0)}$$

Then $x^{(1)}_{p_1} = 1 = \|x^{(1)}\|_\infty$

$$y^{(2)} = Ax^{(1)} = \frac{1}{y_{p_1}^{(1)}} A^2 x^{(0)}$$

$$\mu^2 = y_{p_1}^{(2)} = \frac{y_{p_1}^{(2)}}{x_{p_1}^{(1)}}$$

$$= \frac{\left(\beta_1 \lambda_1^2 v_{p_1}^{(1)} + \sum_{j \geq 2} \beta_j \lambda_j^2 v_{p_1}^{(j)}\right) / y_{p_1}^{(1)}}{\left(\beta_1 \lambda_1 v_{p_1}^{(1)} + \sum_{j \geq 2} \beta_j \lambda_j v_{p_1}^{(j)}\right) / y_{p_1}^{(1)}}$$

$$= \lambda_1 \left[\frac{\beta_1 v_{p_1}^{(1)} + \sum_{j \geq 2} \beta_j \left(\frac{\lambda_j}{\lambda_1}\right)^2 v_{p_1}^{(j)}}{\beta_1 v_{p_1}^{(1)} + \sum_{j \geq 2} \beta_j \left(\frac{\lambda_j}{\lambda_1}\right) v_{p_1}^{(j)}} \right]$$

Let p_2 be the smallest integer with

$$|y_{p_2}^{(2)}| = \|y^{(2)}\|_\infty$$ and define

$$x^{(2)} = \frac{1}{y_{p_2}^{(2)}} y^{(2)}$$

$$= \frac{1}{y_{p_2}^{(2)}} A x^{(1)}$$

$$= \frac{1}{y_{p_2}^{(2)} y_{p_1}^{(1)}} A^2 x^{(0)}$$

In a similar manner define sequences of vectors $\{x^{(m)}\}_{m=0}^{\infty}$ and $\{y^{(m)}\}_{m=1}^{\infty}$ and sequence of scalars $\{\mu^{(m)}\}_{m=1}^{\infty}$ inductively defined by

$$y^{(m)} = Ax^{(m-1)}$$

$$\mu^{(m)} = y^{(m)}_{p_{m-1}}$$

$$x^{(m)} = \frac{y^{(m)}}{y^{(n)}_{p_m}}$$

At each step p_m is used to represent smallest integer for which

$$|y^{(m)}_{p_m}| = \|y^{(m)}\|_\infty$$

Note that

$$\mu^{(m)} = \lambda_1 \left[\frac{\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j \geq 2} \left(\frac{\lambda_j}{\lambda_1}\right)^m \beta_j v_{p_{m-1}}^{(j)}}{\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j \geq 2} \left(\frac{\lambda_j}{\lambda_1}\right)^{m-1} \beta_j v_{p_{m-1}}^{(j)}} \right]$$

Since $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$ for $j \geq 2$

we get $\lim_{m \to \infty} \mu^{(m)} = \lambda_1$ provided $x^{(0)}$

is chosen such that $\beta_1 \neq 0$.

Moreover $\{x^{(m)}\}_{m=0}^{\infty}$ converges to an eigenvector associated with λ_1 that has ∞-norm one.

Example

$$A = \begin{pmatrix} 1 & 3 \\ 7 & 5 \end{pmatrix}$$

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -3 \\ -7 & \lambda - 5 \end{vmatrix}$$

$$= (\lambda - 1)(\lambda - 5) - 21$$

$$= \lambda^2 - 6\lambda + 5 - 21$$

$$= \lambda^2 - 6\lambda - 16$$

$$= (\lambda - 8)(\lambda + 2)$$

So eigen-values of A are $8, -2$

Let us compute the dominant eigen value numerically.

$$X^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$P_0 = 1$$

$$y^{(1)} = A x^{(0)} = \begin{pmatrix} 4 \\ 12 \end{pmatrix} \qquad \lambda' = y_1^{(1)} = 4$$

$P_1 = 2$

$$x^{(1)} = \frac{1}{12} y^{(1)} = \begin{pmatrix} 0.33 \\ 1.00 \end{pmatrix}$$

$$y^{(2)} = A x^{(1)} = \begin{pmatrix} 3.33 \\ 7.33 \end{pmatrix}$$

$$\mu^{(2)} = y_2^{(2)} = 7.33$$

$P_2 = 2$

$$x^{(2)} = \frac{1}{7.33} y^{(2)} = \begin{pmatrix} 0.45 \\ 1.00 \end{pmatrix}$$

$$y^{(3)} = A x^{(2)} = \begin{pmatrix} 3.45 \\ 8.18 \end{pmatrix}$$

$$\mu^3 = y_2^{(2)} = 8.18$$

$P_3 = 2$

$$x^{(3)} = \frac{1}{8.18} y^{(3)} = \begin{pmatrix} 0.42 \\ 1.00 \end{pmatrix}$$

$$y^4 = A x^{(3)} = \begin{pmatrix} 3.42 \\ 7.96 \end{pmatrix}$$

(2g)

$\mu^{(4)} = 7.96$

$p_4 = 2$

$$x^4 = \frac{1}{7.96}\begin{pmatrix} 3.42 \\ 7.96 \end{pmatrix} = \begin{pmatrix} 0.43 \\ 1.00 \end{pmatrix}$$

$$y^{(5)} = A x^{(4)} = \begin{pmatrix} 3.43 \\ 8.01 \end{pmatrix}$$

$\mu^{(5)} = 8.01$

$p_5 = 2$

$$x^{(5)} = \frac{1}{8.01}\begin{pmatrix} 3.43 \\ 8.01 \end{pmatrix} = \begin{pmatrix} 0.43 \\ 1.00 \end{pmatrix} = x^{(4)}$$

$y^{(6)} = y^{(5)}$ etc

So we get 8.01 as dominant eigen value of A with $\begin{pmatrix} 0.43 \\ 1.00 \end{pmatrix}$ as corresponding eigen-vector.

Convergence

The rate at which $\{\mu^{(m)}\}_{m=1}^{\infty}$ converges to λ_1 is determined by the ratios $\left|\frac{\lambda_j}{\lambda_1}\right|^m$ for $j = 2, 3, \ldots, n$.

If we assume $|\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$ it is determined by $\left|\frac{\lambda_2}{\lambda_1}\right|^m$.

The rate of convergence is $O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^m\right)$

So there is a constant k such that for large m

$$|\mu^{(m)} - \lambda_1| \approx k \left|\frac{\lambda_2}{\lambda_1}\right|^m$$

This implies that

$$\lim_{m \to \infty} \frac{|\mu^{(m+1)} - \lambda_1|}{|\mu^{(m)} - \lambda_1|} \approx \left|\frac{\lambda_2}{\lambda_1}\right| < 1$$

So the sequence $\{\mu^{(m)}\}$ converges linearly to λ_1.

Thus we can use Aitken Δ^2 method to speed up convergence

Power method when A is a symmetric matrix

Let A be a symmetric matrix. Let x_0 be a vector with $\|x_0\|_2 = 1$

Set $y_1 = Ax_0$

$$\mu_1 = x_0^t y_1$$

$$= x_0^t A x_0$$

$$x_0^t x_0$$

$\{v_1, \ldots, v_n\}$ orthonormal basis of eigenvectors

$$x_0 = \beta_1 v_1 + \ldots + \beta_n v_n$$

$$x_0^t A x_0 = \beta_1^2 \lambda_1 + \beta_2^2 \lambda_2 + \cdots + \beta_n^2 \lambda_n$$

$$\mu_1 = \frac{\lambda_1 \left(\beta_1^2 + \sum_{j \geq 2} \beta_j^2 \frac{\lambda_j}{\lambda_1} \right)}{\beta_1^2 + \sum_{j \geq 2} \beta_j^2}$$

$$x_1 = \frac{y_1}{\|y_1\|_2} = \frac{1}{\|y_1\|} A x_0 = \frac{1}{\|Ax_0\|} Ax_0$$

$$y_2 = A x_1 = \frac{1}{\|y_1\|} A y_1 = \frac{1}{\|Ax_0\|} A^2 x_0$$

$$\mu_2 = x_1^t y_2 = \frac{x_0^t A^3 x_0}{\|Ax_0\|^2} = \frac{x_0^t A^3 x_0}{x_0^t A^2 x_0}$$

$$\mu_2 = \frac{\beta_1^2 \lambda_1^3 + \beta_2^2 \lambda_2^3 + \cdots + \beta_n^2 \lambda_n^3}{\beta_1^2 \lambda_1^2 + \beta_2^2 \lambda_2^2 + \cdots + \beta_n^2 \lambda_n^2}$$

$$= \lambda_1 \left[\frac{\beta_1^2 + \sum_{j \geq 2} \beta_j^2 \left(\frac{\lambda_j}{\lambda_1} \right)^3}{\beta_1^2 + \sum_{j \geq 2} \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^2} \right]$$

$$x_2 = \frac{y_2}{\|y_2\|} = \frac{1}{\|y_2\|} A x_1 = \frac{1}{\|y_2\| \|y_1\|} A^2 x_0$$

$$y_2 = \frac{1}{\|y_1\|} A^2 x_0 \qquad \text{so } \|y_2\| = \frac{1}{\|y_1\|} \|A^2 x_0\|$$

$$\text{So} \quad x_2 = \frac{1}{\|A^2 x_0\|} A^2 x_0$$

$$y_3 = A x_2 = \frac{1}{\|A^2 x_0\|} A^3 x_0$$

$$\mu_3 = x_2^t y_3 = \frac{x_0^t A^5 x_0}{x_0^t A^4 x_0}$$

$$= \lambda_1 \left[\frac{\beta_1^2 + \sum_{j>1,2} \beta_j^2 \left(\frac{\lambda_j}{\lambda_1}\right)^5}{\beta_1^2 + \sum_{j>1,2} \beta_j^2 \left(\frac{\lambda_j}{\lambda_1}\right)^4} \right]$$

$$x_3 = \frac{y_3}{\|y_3\|} = \frac{A^3 x_0}{\|A^3 x_0\|}$$

$$y_4 = A x_3 = \frac{A^4 x_0}{\|A^3 x_0\|}$$

$$\mu_4 = x_3^t y_4 = \frac{x_0^t A^7 x_0}{x_0^t A^6 x_0}$$

In general

$$x_n = \frac{A^n x_0}{\|A^n x_0\|_2}$$

$$y_{n+1} = A x_n = \frac{1}{\|A^n x_0\|} A^{n+1} x_0$$

$$\nu_{n+1} = x_n^t y_{n+1}$$

$$= \frac{x_0^t A^{2n+1} x_0}{x_0^t A^{2n} x_0}$$

$$= \lambda_1 \left[\frac{\beta_1^2 + \sum_{j \geq 2} \beta_j^2 \left(\frac{\lambda_j}{\lambda_1}\right)^{2n+1}}{\beta_1^2 + \sum_{j \geq 2} \beta_j^2 \left(\frac{\lambda_j}{\lambda_1}\right)^{2n}} \right]$$

Thus $\lim_{n \to \infty} \nu_{n+1} = \lambda_1$ (if $\beta_1 \neq 0$)

Algorithm

Choose x with $\|x\|_2 = 1$

do ~~until~~ until satisfied

$$y = Ax$$

$$\nu = x^t y$$

$$x = \frac{y}{\|y\|_2}$$

Example

$$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{bmatrix}$$

A is symmetric with eigenvalues

$\lambda_1 = 6,\ \lambda_2 = 3,\ \lambda_3 = 1$

m	$x^{(m)}$	$\mu^{(m)}$
0	(1, 0, 0)	
1	(0.942, −0.235, 0.235)	4
2	(0.816, −0.408, 0.408)	5
3	(0.710, −0.497, 0.497)	5.67
4	(0.647, −0.539, 0.539)	5.91
5	(0.613, −0.559, 0.559)	5.98
6	(0.595, −0.568, 0.568)	5.99

Order of convergence

$\mu^{(m)} \to \lambda_1$ in $O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2m}\right)$

Inverse Power method

The inverse power method is a modification of the power method that gives faster convergence. It is used to determine the eigenvalue of A that is close to a specified number q.

Suppose the matrix A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with linearly independent eigenvectors $v^{(1)}, v^{(2)}, \ldots, v^{(n)}$.

The eigenvalues of $(A - qI)^{-1}$ where $q \neq \lambda_i$ for $i = 1, 2, \ldots, n$ are

$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \ldots, \frac{1}{\lambda_n - q}$$

with eigenvectors $v^{(1)}, \ldots, v^{(n)}$

[Proof $(A - qI)v_i = (\lambda_i - q)v_i$

$v_i = (\lambda_i - q)(A - qI)^{-1} v_i$

$(A - qI)^{-1} v_i = \frac{1}{\lambda_i - q} v_i$]

We assume q is close to λ_1

So $\frac{1}{|q-\lambda_1|} > \frac{1}{(q-\lambda_i)}$ $\quad i \geq 2$

We apply power method to $(A-qI)^{-1}$ gives

$$y^{(m)} = (A-qI)^{-1} x^{(m-1)}$$

$$\mu^{(m)} = y^{(m)}_{p_{m-1}}$$

$$x^{(m)} = \frac{y^{(m)}}{y^{(m)}_{p_m}}$$

where at each step p_m represents the smallest integer with $|y^{(m)}_{p_m}| = \|y^{(m)}\|_\infty$

The closer q is to λ_1 the faster will be the convergence of $\mu^{(m)}$ to $\frac{1}{\lambda_1 - q}$

How to set $y^{(m)}$?

$$y^{(m)} = (A-qI)^{-1} x^{(m-1)}$$

$$(A-qI) y^{(m)} = x^{(m-1)}$$

One can get $y^{(m)}$ from Gaus Elimination

One can get PLU factorization of $A-\varepsilon I$.

This allows determination of $y^{(m)}$ in $O(n^2)$ operation.

ε can be determined by Gersgorins circle theorem or some other method

If A is symmetric then so is $(A-\varepsilon I)^{-1}$

Then in this case symmetric power method can be used to speed up convergence