ERG 2012B

Advanced Engineering Mathematics II

Part III

Introduction to Numerical Methods

Lecture #20

Numerical Integrations, Differentiation & LU Factorization

Simpson's Rule

Rectangular rule - a piecewise constant approximation of f

Trapezoidal rule - a piecewise linear approximation of f

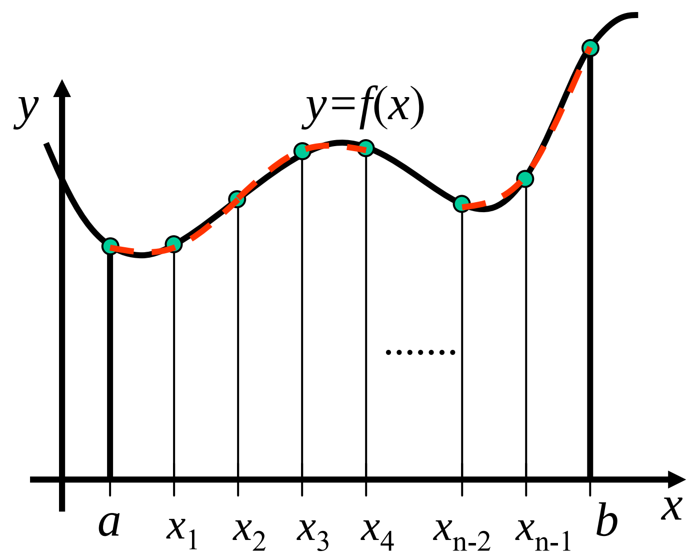
Simpson's rule - a piecewise quadratic approximation of f

Great practical importance - sufficiently accurate, but still simple.

Divide the interval into an **even** number ($n = 2m$) of equal subintervals of length $h=(b-a)/2m$

Take two subintervals at a time and approximate $f(x)$ in the interval by the Lagrange polynomial $p_2(x)$ for the first two from x_0 to x_2 we get:

$$p_2(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f_2$$



Simpson's Rule

for the first two subintervals from x_0 to x_2 we get:

$$p_2(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f_2$$

The denominators are $2h^2$, $-h^2$ and $2h^2$ respectively

Setting $s=(x-x_1)/h$, we have $x-x_0=(s+1)/h$, $x-x_1=sh$, $x-x_2=(s-1)h$

$$p_2(x) = \tfrac{1}{2} s(s-1) f_0 - (s+1)(s-1) f_1 + \tfrac{1}{2}(s+1) s f_2$$

Now integrate wrt x from x_0 to x_2 This corresponds to integrating wrt s *from -1 to 1. Since* dx = $h\, ds$, the result is:

$$\int_{x_0}^{x_2} f(x)dx \approx \int_{x_0}^{x_2} p_2(x)dx = h(\tfrac{1}{3} f_0 + \tfrac{4}{3} f_1 + \tfrac{1}{3} f_2)$$

We can generalize this for all pairs of subintervals and sum them

$$\int_{x_0}^{x_2} f(x)dx \approx \frac{h}{3}(f_0 + 4 f_1 + 2 f_2 + 4 f_3 + \ldots + 2 f_{2m-2} + 4 f_{2m})$$

Simpson's Rule

Simpson's rule is easy to construct as a program - see text book

Bounds for the error: ε_s can be obtained in a similar way to that in the case of the trapezoidal rule.

Assuming that the fourth derivative of *f* exists and is continuous in the region of integration then the results is:

$$CM_4 \leq \varepsilon_s \leq CM_4^* \quad \text{where} \quad C = -\frac{(\mathrm{b}-\mathrm{a})^5}{180(2m)^4}$$

and M_4 and M_4^* are the largest and smallest value of the fourth derivative of *f* in the interval of integration.

Example 3a

Evaluate $J = \int_0^1 e^{-x^2}\,dx$ by simpson's rule with $2m = 10$

Computational Table					
j	***x j***	***x j*^2**	**exp(-*x* j^2)**		
0	0.0	0.0	1.000000		
1	0.1	0.0		0.990050	
2	0.2	0.0			0.960789
3	0.3	0.1		0.913931	
4	0.4	0.2			0.852144
5	0.5	0.3		0.778801	
6	0.6	0.4			0.697676
7	0.7	0.5		0.612626	
8	0.8	0.6			0.527292
9	0.9	0.8		0.444858	
10	1.0	1.0	0.367879		
Sums			1.367879	3.740266	3.037902

$$\int_{x_0}^{x_2} f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 2f_{2m-2} + 4f_{2m})$$

So $J \approx 0.3333(1.367879 + 4 \bullet 3.740266 + 2 \bullet 3.037902) = 0.746826$

Example 3b

Estimate the error in Example 3a.

Solution: From

$$CM_4 \le \varepsilon_s \le CM_4^* \quad \text{where} \quad C = -\frac{(b-a)^5}{180(2m)^4}$$

where M_4 and M_4^* are the largest and smallest values of $f^4(x)$ in the region of integration

By differentiation $f^4(x) = 4(4x^4\text{-}12x^2 + 3)\exp(-x^2)$

Also $f^5(x)$ shows max of f^4 is at x=0 and min at $x^*=2.5+0.5\sqrt{10}$

Therefore $M_4 = f^4(0)=12$ and $M_4^* = f^4(x^*) = -7.359$

and C=-1/1800000 so that

$$-0.000007 \le \varepsilon \le 0.000005$$

and exact value of J lies between 0.746818 and 0.746830

far better than was obtained from the trapezoid rule.

Example 4

Determine *n* in previous example such that we have 6D accuracy

Solution: As M_4 =12 (the biggest in absolute value of the two boundaries) we find that

$$\varepsilon = \left|CM_4\right| = -\frac{12(\mathrm{b}-\mathrm{a})^5}{180(2m)^4} = -\frac{12}{180(2m)^4} = \frac{1}{2}10^{-6} \quad \text{(required accuracy)}$$

$$\text{or} \quad m = \left[\frac{2 \cdot 10^6 \cdot 12}{180 \cdot 2^4}\right]^{\frac{1}{4}} = 9.55$$

Hence we should choose $n = 2m = 20$ for the required accuracy.

Numerical Differentiation

Numerical differentiation should be avoided whenever possible, because, whereas **integration** is a **smoothing** process and not affected much by small inaccuracies in values, **differentiation** tends to **roughen** and gives values of f' much less accurate then those of f

We use the notation $f_j' = f'(x_j)$, $f_j'' = f''(x_j)$, etc.

Rough approximation formulas can be found from

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

$$\boxed{f_1'' \approx \frac{\dfrac{(f_2 - f_1)}{h} - \dfrac{(f_1 - f_0)}{h}}{h}}$$

Which suggests

$$f_{\frac{1}{2}}' \approx \frac{\delta f_{\frac{1}{2}}}{h} = \frac{f_1 - f_0}{h} \quad \text{and} \quad f_1'' \approx \frac{\delta^2 f_1}{h^2} = \frac{f_2 - 2f_1 + f_0}{h^2}$$

Numerical Differentiation

More accurate approximations are obtained by differentiating suitable Lagrange polynomials.

$$f'(x) = p_2'(x) = \frac{2x - x_1 - x_2}{2h^2} f_0 - \frac{2x - x_0 - x_2}{h^2} f_1 + \frac{2x - x_0 - x_1}{2h^2} f_2$$

Evaluating this at x_0, x_1, x_2 we obtain the *three point formulas*

(a) $\quad f_0' \approx \dfrac{1}{2h}(-3f_0 + 4f_1 - f_2)$

(b) $\quad f_1' \approx \dfrac{1}{2h}(-f_0 + f_2)$

(c) $\quad f_2' \approx \dfrac{1}{2h}(f_0 - 4f_1 + 3f_2)$

Applying the same idea to $p_4(x)$ we get similar formula, particularly

$$f_2' \approx \frac{1}{12h}(f_0 - 8f_1 + 8f_3 - f_4)$$

LU Factorization

To solve a linear system **Ax** = **b**
where **A** is nonsingular, we can make use of **LU factorization of A** that find **L** and **U** such that **A** = **LU**
where **L** is lower triangular, and **U** is upper triangular

Example:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 8 & 5 \end{bmatrix} = \mathbf{LU} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & -7 \end{bmatrix}$$

L is the matrix of multipliers m_{jk} from the Gauss elimination and has main diagonals = 1.

U is the matrix at the end of the Gauss elimination

L and **U** can be computed directly without using Gauss elimination - which requires $2n^3/3$ operations – in $n^3/3$ operations
And once we have **L** and **U** we can use them to solve **Ax**=**b** in two steps, involving only n^2 operations, by letting **y**=**Ux** so that **Ly**=**b** as **Ax**=**LUx**=**b**

Doolittle's Method

We use **Ly=b** to solve for **y** first

Then use **Ux=y** to solve for **x**

This is known as **Doolittle's Method**

A similar method, **Crout's Method** is obtained if **U** (instead of **L**) is required to have main diagonal =1.

Example: solve the system

$$\begin{aligned} 3x_1 + 5x_2 + 2x_3 &= 8 \\ 8x_2 + 2x_3 &= -7 \\ 6x_1 + 2x_2 + 8x_3 &= 26 \end{aligned}$$

The **LU** decomposition is obtained from:

$$A = \begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} \mathrm{u}_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Doolittle's Method

m_{jk} and u_{jk} are determined using matrix multiplication:

$a_{11} = 3 = u_{11}$	$a_{12} = 5 = u_{12}$	$a_{13} = 2 = u_{13}$
$a_{21} = 0 = m_{21}u_{11}$ $m_{21} = 0$	$a_{22} = 8 = m_{21}u_{21} + u_{22}$ $u_{22} = 8$	$a_{23} = 2 = m_{21}u_{13} + u_{23}$ $u_{23} = 2$
$a_{31} = 6 = m_{31}u_{11}$ $m_{31} = 2$	$a_{32} = 2 = m31u_{12} + m_{32}u_{22}$ $m_{32} = -1$	$a_{33} = 8 = m_{31}u_{13} + m_{32}u_{23} + u_{33}$ $u_{33} = 6$

so that

$$A = \begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{bmatrix}$$

first solve **Ly=b**

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y1 \\ y2 \\ y3 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \\ 26 \end{bmatrix} \Rightarrow \mathbf{y} = \begin{bmatrix} 8 \\ -7 \\ 3 \end{bmatrix}$$

Doolittle's Method

so that

$$A=\begin{bmatrix}3&5&2\\0&8&2\\6&2&8\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\2&-1&1\end{bmatrix}\begin{bmatrix}3&5&2\\0&8&2\\0&0&6\end{bmatrix}$$

first solve **Ly=b**

$$\begin{bmatrix}1&0&0\\0&1&0\\2&-1&1\end{bmatrix}\begin{bmatrix}y1\\y2\\y3\end{bmatrix}=\begin{bmatrix}8\\-7\\26\end{bmatrix}\Rightarrow \mathbf{y}=\begin{bmatrix}8\\-7\\3\end{bmatrix}$$

Then solve **Ux=y**

$$\begin{bmatrix}3&5&2\\0&8&2\\0&0&6\end{bmatrix}\begin{bmatrix}x1\\x2\\x3\end{bmatrix}=\begin{bmatrix}8\\-7\\3\end{bmatrix}\Rightarrow \mathbf{x}=\begin{bmatrix}4\\-1\\\frac{1}{2}\end{bmatrix}$$

Doolittle's Method

The formulas obtained in the example suggest that for general n the elements of the matrices $\mathbf{L}$ =[m_{jk}] and $\mathbf{U}$ = [u_{jk}] in the Doolittle Method are computed from:

$$u_{ik} = a_{1k} \qquad k = 1, \cdots, n$$

$$u_{jk} = a_{jk} - \sum_{s=1}^{j-1} m_{js} u_{sk} \qquad k = j, \cdots, n; \ \ j \geq 2$$

$$m_{ji} = \frac{a_{ji}}{u_{11}} \qquad j = 2, \cdots, n$$

$$m_{jk} = \frac{1}{u_{kk}} \left(a_{jk} - \sum_{s=1}^{k-1} m_{js} u_{sk} \right) \qquad j = k+1, \cdots, n; \ \ k \geq 2$$

Inclusion of Eigenvalues

By **inclusion** we mean the determination of approximate values of eigenvalues and corresponding error bounds.

The next, important, theorem gives a region consisting of closed circular disks in the complex plane which include the eigenvalues of a given matrix

For each $j = 1,\ldots\ldots,n$ the inequality in the theorem determines a closed circular disk in the complex plane with center a_{jj} and radius given by the right hand side

The theorem states that each of the eigenvalues lies inside one of these n disks

Gerchgorin's Theorem

Theorem 1: Let λ be an eigenvalue of an arbitrary nxn matrix **A**. Then for some integer j ($1 \le j \le n$) we have:

(1) $|a_{jj} - \lambda| \le |a_{j1}| + |a_{j2}| + \ldots + |a_{jj-1}| + |a_{jj+1}| + \ldots + |a_{jn}|$

Proof: Let **x** be an eigenvector corresponding to λ. Then

(2) $\mathbf{Ax} = \lambda\mathbf{x}$ or $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$

Let x_j be the component of **x** that has the largest absolute value. Then we have $|x_m/x_j| \le 1$ for $m = 1,\ldots\ldots,n$

The vector equation (2) is equivalent to a system of n equations for the n components of the vectors on both sides and the j^{th} of these n equations is:

$$a_{j1}x_1 + \ldots\ldots + a_{jj-1}x_{j-1} + (a_{jj} - \lambda)x_j + a_{jj+1}x_{j+1} + \ldots + a_{jn}x_n = 0$$

Divide by x_j and rearrange gives:

$$(a_{jj} - \lambda) = -a_{j1}x_1/x_j - \ldots\ldots - a_{jj-1}x_{j-1}/x_j - a_{jj+1}x_{j+1}/x_j - \ldots - a_{jn}x_n/x_j$$

Taking the absolute values on both sides, recalling $|a+b| \le |a|+|b|$ and because of our choice of j $|x_m/x_j| \le 1$ we get (1)

Example

For the eigenvalues of the matrix

$$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 5 & 1 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}$$

$$|a_{11} - \lambda| \leq |a_{12}| + |a_{13}| \Rightarrow |\lambda| \leq ½ + ½ \Rightarrow |\lambda| \leq 1$$

$$|a_{22} - \lambda| \leq |a_{21}| + |a_{23}| \Rightarrow |5-\lambda| \leq ½ + 1 \Rightarrow |5-\lambda| \leq 1.5$$

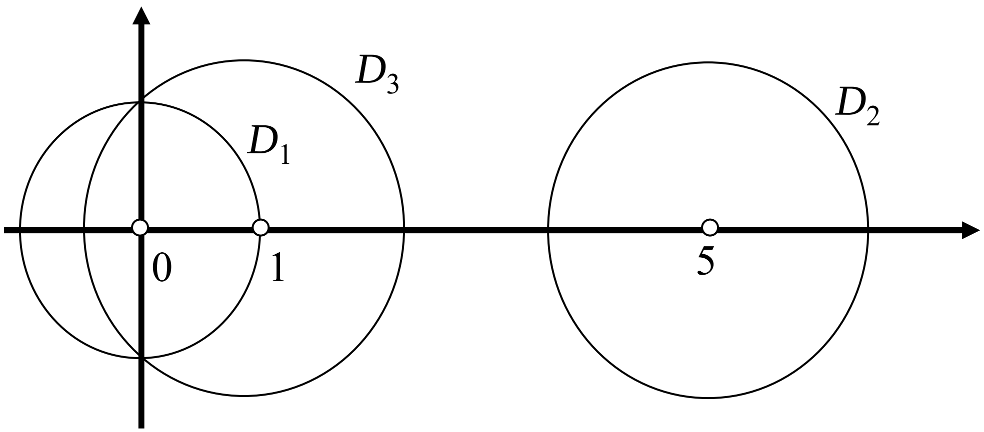
$$|a_{33} - \lambda| \leq |a_{31}| + |a_{32}| \Rightarrow |1-\lambda| \leq ½ + 1 \Rightarrow |1-\lambda| \leq 1.5$$

we get the Greschgorin disks:

D_1: Center 0, radius 1

D_2: Center 5, radius 1.5

D_3: Center 1, radius 1.5



Since **A** is symmetric it follows that the spectrum of **A** must lie in the intervals [-1, 2.5] and [3.5, 6.5] on the real axis

Note how the Gerschgorin disks form two disjoint sets

Extension to Gerchgorin's Theorem

Theorem 2: If p Gerschgorin disks form a set S that is disjoint from the n-p other disks of a given matrix $\mathbf{A}$ then S contains precisely p eigenvalues of $\mathbf{A}$ (each counted with its algebraic multiplicity)

Proof: This is a *continuity proof*. Let $S = D_1 \cup D_2 \cup \dots \cup D_p$ where D_j is the Gerschgorin disk with center a_{jj}.

Consider $\mathbf{A} = \mathbf{B} + \mathbf{C}$, where $\mathbf{B} = \text{diag}(a_{jj})$ is the diagonal matrix with main diagonal of $\mathbf{A}$ as its diagonal. Next consider

$$\mathbf{A}t = \mathbf{B} + t\,\mathbf{C} \qquad \text{for } 0 \le t \le 1$$

Then if $\mathbf{A}_0 = \mathbf{B}$ and $\mathbf{A}_1 = \mathbf{A}$. The eigenvalues of $\mathbf{A}_t$ change continuously from $a_{11},\dots,a_{nn}$ (t=0) to those of $\mathbf{A}$ (t=1) if we change t continuously from 0 to 1. Thus the radii of the disks change continuously from 0 (t=0) to those for $\mathbf{A}$ at the same time Since at t=1, S is disjoint from the other disks there is no way for the p values to move to the other set.

Schur's Theorem

Theorem 3: Let $\mathbf{A} = [a_{jk}]$ be an nxn matrix. Let $\lambda_1, \ldots \lambda_n$ be its eigenvalues. Then:

$$\sum_{j=1}^{n} |\lambda_i|^2 \le \sum_{j=1}^{n}\sum_{k=1}^{n} |a_{jk}|^2 \qquad \textbf{Scur's inequality}$$

The equality holds iff **A** is such that $\overline{\mathbf{A}}^{\mathbf{T}}\mathbf{A} = \mathbf{A}\overline{\mathbf{A}}^{\mathbf{T}}$

Matrices that satisfy this are called **normal matrices**. It can be shown that Hermitian, skew-Hermitian and unitary matrices are normal and hence their real equivalents.

Let λ_m be any eigenvalue of the matrix **A**. Then $|\lambda_m|^2$ is also less than or equal to the sum on the right hand side so that

$$|\lambda_m| \le \sqrt{\sum_{j=1}^{n}\sum_{k=1}^{n} |a_{jk}|^2}$$

Example

Bounds for eigenvalues from Schlur's Theorem

For the matrix:

$$\mathbf{A} = \begin{bmatrix} 26 & -2 & 2 \\ 2 & 21 & 4 \\ 4 & 2 & 28 \end{bmatrix}$$

we get from Schlur's inequality:

$$|\lambda| \leq \sqrt{1949} < 44.2$$

The eigenvalues of **A** are 30, 25 and 20 all < 44.2;

and $30^2 + 25^2 + 20^2 = 1925 < 1949$

Note: A is not a normal matrix

Perron-Frobenius's Theorem

Let **A** be a real square matrix whose elements are all positive. Then **A** has at least one real positive eigenvalue λ, and the corresponding eigenvector can be chosen real and such that all its components are positive.

Collatz's Theorem

Let **A** =[a_{jk}] be a real *n*x*n* matrix whose elements are all positive. Let **x** be any real vector whose components $x_1,\ldots,x_n$ are positive, and let $y_1,\ldots.,y_n$ be the components of the vector $\mathbf{y} = \mathbf{A}\mathbf{x}$. Then the closed interval on the real axis bounded by the smallest and the largest of the *n* quotients $q_j = y_j/x_j$ contains at least one eigenvalue of **A**

Eigenvalues by Iteration

Power Method: a simple procedure for computing approximate values of the eigenvalues of an nxn matrix $\mathbf{A}=[a_{jk}]$. In this method we start from any vector $\mathbf{x}_0$ ($\neq\mathbf{0}$) with n components and compute successively:

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0, \quad \mathbf{x}_2=\mathbf{A}\mathbf{x}_1, \; \ldots.., \quad \mathbf{x}_s = \mathbf{A}\mathbf{x}_{s\text{-}1}$$

To simplify the notation we denote $\mathbf{x}_{s\text{-}1}$ by $\mathbf{x}$ and $\mathbf{x}_s$ by $\mathbf{y}$ so that $\mathbf{y} = \mathbf{A}\mathbf{x}$. If $\mathbf{A}$ is real symmetric, the following theorem gives an approximation and error bounds:

Theorem: Let $\mathbf{A}$ be an nxn real symmetric matrix. Let $\mathbf{x}$ ($\neq \mathbf{0}$) be any real vector with n components and let:

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad m_0=\mathbf{x}^\mathbf{T}\mathbf{x}, \quad m_1 = \mathbf{x}^\mathbf{T}\mathbf{y}, \; m_2= \mathbf{y}^\mathbf{T}\mathbf{y}$$

Then the quotient $q = m_1/m_0$ (**Rayleigh quotient**) is an approximation for an eigenvalue of $\mathbf{A}$ and the error is given by:

$$|\varepsilon| \leq \sqrt{\frac{m_2}{m_0} - q^2}$$

see text book for proof

Example

Consider the real symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 6 & -4 \\ 2 & -4 & 6 \end{bmatrix} \text{ and choose } \mathbf{x_0} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Then:

$$\mathbf{x}_1 = \begin{bmatrix} 8 \\ 0 \\ 4 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 72 \\ -32 \\ 40 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 720 \\ -496 \\ 512 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} 7776 \\ -6464 \\ 6496 \end{bmatrix},$$

Taking $\mathbf{x} = \mathbf{x}_3$ and $\mathbf{y} = \mathbf{x}_4$, we have

$m_0 = \mathbf{x}^\mathbf{T}\mathbf{x} = 1026560$, $m_1 = \mathbf{x}^\mathbf{T}\mathbf{y} = 12130816$, $m_2 = \mathbf{y}^\mathbf{T}\mathbf{y} = 14447488$

And from this we calculate:

$q = m_1/m_0 = 11.817$, $|\varepsilon| \leq \sqrt{}(m_2/m_1 - q^2) = 1.034$

showing that q =11.817 is an approximation for an eigenvalue that must lie between 10.783 and 12.851. (In fact λ=12 is one)