10.4.1.1.8

EE24BTECH11024 - G. Abhimanyu Koushik

Question:
Find the roots of the equation $x^3 - 4x^2 - x + 1 = (x-2)^3$
Solution:
Theoritical solution:
The equation can be simplified to

$$x^3 - 4x^2 - x + 1 = (x-2)^3 \tag{0.1}$$

$$x^3 - 4x^2 - x + 1 = x^3 - 6x^2 + 12x - 8 \tag{0.2}$$

$$2x^2 - 13x + 9 = 0 \tag{0.3}$$

Applying quadratic formula gives solution as

$$x_1 = \frac{13 - \sqrt{97}}{4} \tag{0.4}$$

$$x_2 = \frac{13 + \sqrt{97}}{4} \tag{0.5}$$

Computational solution:
Two methods to find solution of a quadratic equation are:
Matrix-Based Method:
For a polynomial equation of form $x_n + b_{n-1}x^{n-1} + \dots + b_2x^2 + b_1x + b_0 = 0$ we construct a matrix called companion matrix of form

$$\Lambda = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \vdots & 1 \\ -b_0 & -b_1 & -b_2 & \dots & -b_{n-1} \end{pmatrix} \tag{0.6}$$

The eigenvalues of this matrix are the roots of the given polynomial equation.
The solution given by the code is

$$x_1 = 0.7878 \tag{0.7}$$

$$x_2 = 5.7122 \tag{0.8}$$

Newton-Raphson Method:
Start with an initial guess x_0, and then run the following logical loop,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{0.9}$$

where,

$$f(x) = 2x^2 - 13x + 9 \tag{0.10}$$

$$f'(x) = 4x - 13 \tag{0.11}$$

The update equation will be

$$x_{n+1} = x_n - \frac{2{x_n}^2 - 13x_n + 9}{4x_n - 13} \tag{0.12}$$

$$\tag{0.13}$$

The problem with this method is if the roots are complex but the coeffcients are real, x_n either converges to an extrema or grows continuously without any bound. To get the complex solutions, however , we can just take the initial guess point to be a random complex number.
The output of a program written to find roots is shown below:

$$r_1 = 0.7878 \tag{0.14}$$

$$r_2 = 5.7122 \tag{0.15}$$

Companion matrix

$$\Lambda_0 = \begin{pmatrix} 0 & 1 \\ -\frac{9}{2} & \frac{13}{2} \end{pmatrix} \tag{0.16}$$

The update equation will be

$$\Lambda_{n+1} = Q(\Lambda_n - \sigma I)Q^\top + \sigma I \tag{0.17}$$

$$\tag{0.18}$$

As n tends to infinite, Λ converges to Upper triangular matrix

$$\sigma = \frac{13}{2} \tag{0.19}$$

$$\Lambda_{\text{shifted}} = \begin{pmatrix} 0 & 1 \\ -\frac{9}{2} & \frac{13}{2} \end{pmatrix} - \sigma I \tag{0.20}$$

$$\Lambda_{\text{shifted}} = \begin{pmatrix} \frac{-13}{2} & 1 \\ \frac{-9}{2} & 0 \end{pmatrix} \tag{0.21}$$

Where σ is the last diagonal element.

$$Q = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{0.22}$$

$$\tag{0.23}$$

Where

$$c = \cos\phi \tag{0.24}$$

$$s = \sin\phi \tag{0.25}$$

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \times \begin{pmatrix} \frac{-13}{2} & 1 \\ \frac{-9}{2} & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \tag{0.26}$$

$$c\left(\frac{9}{2}\right) + s\left(\frac{13}{2}\right) = 0 \tag{0.27}$$

$$c^2 + s^2 = 1 \tag{0.28}$$

Solving for c and s gives

$$c = \frac{\frac{-13}{2}}{\sqrt{\left(\frac{13}{2}\right)^2 + \left(\frac{9}{2}\right)^2}} \tag{0.29}$$

$$s = \frac{\frac{-9}{2}}{\sqrt{\left(\frac{13}{2}\right)^2 + \left(\frac{9}{2}\right)^2}} \tag{0.30}$$

Now, as we got the Q matrix we will do the following

$$\Lambda_{\text{new}} = Q\Lambda_{\text{shifted}}Q^\top + \sigma I \tag{0.31}$$

$$\Lambda_{\text{new}} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} \frac{-13}{2} & 1 \\ \frac{-9}{2} & 0 \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix} + \sigma I \tag{0.32}$$

$$\Lambda_{\text{new}} \approx \begin{pmatrix} 0.468 & 5.176 \\ -0.324 & 6.032 \end{pmatrix} \tag{0.33}$$

Run the same sequence of steps for 20 iterations after which you end up with the following matrix

$$\Lambda_{\text{new}} = \begin{pmatrix} 0.78778555 & 5.5 \\ 0 & 5.71221445 \end{pmatrix} \tag{0.34}$$

The eigenvalues are same as its diagonal elements
Hence the roots of given equation are 0.78778555 and 5.71221445

To reduce a matrix into QR form we do the following
For a general matrix, we should bring to hessenberg form H, for that let

$$P = I - 2\mathbf{w}\mathbf{w}^\top \tag{0.35}$$

where w is a vector with $|w|^2 = 1$. The matrix P is orthogonal as

$$P^2 = \left(I - 2\mathbf{w}\mathbf{w}^\top\right) \cdot \left(I - 2\mathbf{w}\mathbf{w}^\top\right) \tag{0.36}$$

$$= I - 4\mathbf{w}\mathbf{w}^\top + 4\mathbf{w} \cdot \left(\mathbf{w}^\top\mathbf{w}^\top\right) \cdot \mathbf{w}^\top \tag{0.37}$$

$$= I \tag{0.38}$$

Therefore, $P = P^{-1}$ but $P = P^\top$, so $P = P^\top$

We can rewrite P as

$$P = I - \frac{\mathbf{u}\mathbf{u}^\top}{H} \tag{0.39}$$

where the scalar H is

$$H = \frac{1}{2}\,|\mathbf{u}|^2 \tag{0.40}$$

Where $\mathbf{u}$ can be any vector. Suppose $\mathbf{x}$ is the vector composed of the first column of A. Take

$$\mathbf{u} = \mathbf{x} \mp |\mathbf{x}|\,\mathbf{e}_1 \tag{0.41}$$

Where $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & \dots \end{pmatrix}^\top$, we will take the choice of sign later. Then

$$P \cdot \mathbf{x} = \mathbf{x} - \frac{\mathbf{u}}{H} \cdot (\mathbf{u} \mp |\mathbf{x}|\,\mathbf{e}_1)^\top \cdot \mathbf{x} \tag{0.42}$$

$$= \mathbf{x} - \frac{2\mathbf{u}\left(|x|^2 \mp |x|\,x_1\right)}{2\,|x|^2 \mp |x|\,x_1} \tag{0.43}$$

$$= \mathbf{x} - \mathbf{u} \tag{0.44}$$

$$= \mp |\mathbf{x}|\,\mathbf{e}_1 \tag{0.45}$$

To reduce a matrix A into Hessenberg form, we choose vector $\mathbf{x}$ for the first householder matrix to be lower $n-1$ elements of the first column, then the lower $n-2$ elements will be zeroed.

$$P_1 \cdot A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & p_{11} & p_{12} & \cdots & p_{1n} \\ 0 & p_{21} & p_{22} & \cdots & p_{2n} \\ 0 & p_{31} & p_{32} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} a_{00} & \times & \times & \cdots & \times \\ a_{10} & \times & \times & \cdots & \times \\ a_{20} & \times & \times & \cdots & \times \\ a_{30} & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & \times & \times & \cdots & \times \end{pmatrix} \tag{0.46}$$

$$= \begin{pmatrix} a'_{00} & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \times & \times & \cdots & \times \end{pmatrix} \tag{0.47}$$

Now we choose the vector $\mathbf{x}$ for the householder matrix to be the bottom $n-2$ elements of the second column, and from it construct the P_2

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & p_{22} & \cdots & p_{2n} \\ 0 & 0 & p_{32} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & p_{n2} & \cdots & p_{nn} \end{pmatrix} \tag{0.48}$$

Continue the pattern
Let H be the Hessenberg matrix. To find Q, we do the following, let

$$G = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \iddots & \vdots \\ 0 & \cdots & c & s & \cdots & 0 \\ 0 & \cdots & -s & c & \cdots & 0 \\ \vdots & \iddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{0.49}$$

Where the value of c and s are

$$c = \frac{\overline{x_{i,i}}}{\sqrt{\left|x_{i,i}\right|^2 + \left|x_{i,i+1}\right|^2}} \tag{0.50}$$

$$s = \frac{\overline{x_{i,i+1}}}{\sqrt{\left|x_{i,i}\right|^2 + \left|x_{i,i+1}\right|^2}} \tag{0.51}$$

Multiplying G and A nulls out the element in $(i+1)^{\text{th}}$ row and i^{th} column. Now

$$Q = G_{n-1}G_{n-2}\cdots G_2G_1 \tag{0.52}$$

QR decomposition:
Given H_0

$$H_n = Q_nR_n \tag{0.53}$$

$$Q^\top H_n = R_n \tag{0.54}$$

$$H_{n+1} = R_nQ_n \tag{0.55}$$

The update equation will be

$$H_{n+1} = Q^\top H_n Q \tag{0.56}$$

As n tends to infinite, H will converge to upper triangular matrix, whose eigenvalues are the roots of the equation
QR decomposition with Shift:

Given H_0

$$H_n - \sigma I = Q_n R_n \tag{0.57}$$

$$Q^{\top} (H_n - \sigma I) = R_n \tag{0.58}$$

$$H_{n+1} = R_n Q_n + \sigma I \tag{0.59}$$

The update equation will be

$$H_{n+1} = Q^{\top} (H_n - \sigma I) Q_n + \sigma I \tag{0.60}$$

Where σ is the last diagonal element As n tends to infinite, H will converge to upper triangular matrix faster than QR decomposition