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 $x_n + b_{n-1}x^{n-1} + \cdots + b_2x^2 + b_1x + b_0 = 0$ we construct a matrix called companion matrix of form

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

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:

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

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

QR decomposition on Hessenberg matrix:

It is a Numerical method for finding eigenvalues of a given matrix

We say a matrix A is in hessenberg form if it is in form shown below

$$H = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \times \end{pmatrix} \tag{0.16}$$

We will use householder method to reduce any matrix into hessenberg form.

It reduces an $n \times n$ matrix to hessenberg form by $n - 2$ orthogonal trasformations. Each transformations annihilates the required part of a whole column at a time rather than element wise elimination. The basic ingredient for a house holder matrix is P which is in the form

$$P = I - 2\mathbf{w}\mathbf{w}^\top \tag{0.17}$$

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

$$P^2 = \left(I - 2\mathbf{w}\mathbf{w}^\top\right) \cdot \left(I - 2\mathbf{w}\mathbf{w}^\top\right) \tag{0.18}$$

$$= I - 4\mathbf{w}\mathbf{w}^\top + 4\mathbf{w} \cdot \left(\mathbf{w}^\top \mathbf{w}^\top\right) \cdot \mathbf{w}^\top \tag{0.19}$$

$$= I \tag{0.20}$$

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

We can rewrite P as

$$P = I - \frac{\mathbf{u}\mathbf{u}^\top}{H} \tag{0.21}$$

where the scalar H is

$$H = \frac{1}{2}\left|\mathbf{u}\right|^2 \tag{0.22}$$

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

$$\mathbf{u} = \mathbf{x} \mp \left|\mathbf{x}\right| \mathbf{e}_1 \tag{0.23}$$

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

$$P \cdot \mathbf{x} = \mathbf{x} - \frac{\mathbf{u}}{H} \cdot \left(\mathbf{u} \mp \left|\mathbf{x}\right| \mathbf{e}_1\right)^\top \cdot \mathbf{x} \tag{0.24}$$

$$= \mathbf{x} - \frac{2\mathbf{u}\left(\left|x\right|^2 \mp \left|x\right| x_1\right)}{2\left|x\right|^2 \mp \left|x\right| x_1} \tag{0.25}$$

$$= \mathbf{x} - \mathbf{u} \tag{0.26}$$

$$= \mp \left|\mathbf{x}\right| \mathbf{e}_1 \tag{0.27}$$

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.

$$P_1 \cdot A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & p_{11} & p_{12} & \cdots & p_{1n} \\ 0 & p_{21} & p_{22} & \cdots & p_{2n} \\ 0 & p_{31} & p_{32} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} a_{00} & \times & \times & \cdots & \times \\ a_{10} & \times & \times & \cdots & \times \\ a_{20} & \times & \times & \cdots & \times \\ a_{30} & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & \times & \times & \cdots & \times \end{pmatrix} \tag{0.28}$$

$$= \begin{pmatrix} a'_{00} & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \times & \times & \cdots & \times \end{pmatrix} \tag{0.29}$$

Now if we multiply the matrix $P_1 A$ with P_1, the eigenvalues will be conserved as it is a similarity transformation.
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

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & p_{22} & \cdots & p_{2n} \\ 0 & 0 & p_{32} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & p_{n2} & \cdots & p_{nn} \end{pmatrix} \tag{0.30}$$

Now if do similarity transform PAP, we will zero out the $n-3$ elements in second column. If we continue this pattern we will get the hessenberg form of a the matrix A. In this algorithm, we decompose matrix given in Hessenberg form to two matrices Q and R such that Q is an orthogonal matrix and R is an upper triangular matrix. Then we assign the new matrix A' to be $A' = RQ$, and we do this iteratively. Theoritically, as the number of iterations go to infinite, the matrix A' will converge to an upper triangular matrix whose diagonal elements are the eigenvalues of A. There will be a minor problem in this method when the entries are real while the eigenvalues are complex, we will solve this issue shortly. The eigenvalues of the matrix A will not change because of the following

$$A = QR \tag{0.31}$$

$$R = Q^\top A \tag{0.32}$$

$$A' = RQ \tag{0.33}$$

$$A' = Q^\top AQ \tag{0.34}$$

As the matrix A is undergoing similarity transformation, the eigenvalues will not change. The rate of covergence of A depends on the ratio of absolute values of the eigenvalues. That is, if the eigenvalues are $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_n|$ then, the elements of A_k below the diagonal to converge to zero like

$$\left|a_{ij}^{(k)}\right| = O\left(\left|\frac{\lambda_i}{\lambda_j}\right|^k\right) \quad i > j \tag{0.35}$$

The QR decomposition is implemented using the Givens rotation technique. This approach is robust and numerically stable, making it ideal for QR decomposition, especially in iterative methods like eigenvalue computations. It is every similar to Jacobian Transformation. We define a rotation matrix G, to zero out the elements which are non-diagonal, since the matrix which we are dealing is a Hessenberg matrix, we need to zero out the elements which are just below the diagonal elements.

$$G = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \iddots & \vdots \\ 0 & \cdots & c & s & \cdots & 0 \\ 0 & \cdots & -s & c & \cdots & 0 \\ \vdots & \iddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{0.36}$$

Where the value of c and s are

$$c = \frac{\overline{x_{i,i}}}{\sqrt{\left|x_{i,i}\right|^2 + \left|x_{i,i+1}\right|^2}} \tag{0.37}$$

$$s = \frac{\overline{x_{i,i+1}}}{\sqrt{\left|x_{i,i}\right|^2 + \left|x_{i,i+1}\right|^2}} \tag{0.38}$$

If we multiply G and A, we can see easily that it nulls out the element in $(i+1)^{\text{th}}$ row and i^{th} column. The following matrix multiplication eliminates all the elements below the diagonal of A

$$A \implies G_n G_{n-1} \cdots G_2 G_1 A \tag{0.39}$$

Now, we store $G_n G_{n-1} \cdots G_2 G_1$ in Q and then

$$A' \implies QAQ^{\top} \tag{0.40}$$

$$\tag{0.41}$$

If we carry out these transformation infinite times, the A will be an upper triangular matrix with diagonal elements as eigenvalues. If all the entries in the matrix are real but the eigenvalues are complex, the matrix A will converge to a Quasi-triangular form, that is

$$A = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & 0 \\ 0 & 0 & 0 & B_n \end{pmatrix} \tag{0.42}$$

Where B_i is a 2×2 block matrix. These matrices are called jordan blocks. In this case, the eigenvalues are calculated by solving the characteristic equation of the 2×2 matrix. Since it will be a quadratic equation, it can be easily solved and the solutions of that characteristic equation will be the eigenvalues.