The Chinese University of Hong Kong, Shenzhen
SDS · School of Data Science

Andre Milzarek
Fall Semester 2024-25

DDA 3005 — Numerical Methods

Solutions — Final Sample

- The exam time is 120 minutes.
- There are five exercises.
- The total number of achievable points is 100 points.
- You are allowed to bring one self-made sheet of A4 paper (with arbitrary notes on both sides of it) for your personal use in this exam. Other tools are not allowed.
- Please abide by the honor codes of CUHK-SZ.

Exercise 1 (QR Algorithm): (*20* points)
We consider the 2×2 matrix

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

a) Does the matrix $\boldsymbol{A}$ have an eigendecomposition? Explain your answer!

b) Perform two steps of the QR algorithm to generate the iterates $\boldsymbol{X}^1$ and $\boldsymbol{X}^2$.

c) Based on your observations and computations, does the QR algorithm converge for this example (i.e., can it recover a Schur factorization of $\boldsymbol{A}$)?

Givens Rotations – Recalled: For a given vector $\boldsymbol{a} = [a_1, a_2]^\top$ with $\|\boldsymbol{a}\| \neq 0$, the associated Givens rotation is defined via:

$$\boldsymbol{G} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \frac{a_1}{\|\boldsymbol{a}\|}, \quad s = \frac{a_2}{\|\boldsymbol{a}\|}, \quad \boldsymbol{G}\boldsymbol{a} = \begin{bmatrix} \|\boldsymbol{a}\| \\ 0 \end{bmatrix}.$$

Solution :

a) The matrix $\boldsymbol{A}$ is a lower triangular matrix and thus, its eigenvalues can be found on the diagonal: $\lambda_1 = -1$ and $\lambda_2 = 1$. As λ_1 and λ_2 are distinct, $\boldsymbol{A}$ must have an eigendecomposition.

b) We first compute a QR of $\boldsymbol{X}^0 = \boldsymbol{A}$ using Givens rotations (Householder transformation and Gram-Schmidt orthogonalization can also be used):

- Step 1: QR factorization of $\boldsymbol{X}^0$. $\boldsymbol{X}^0 = \boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{R}_1$:

$$\boldsymbol{Q}_1^\top \boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} = \boldsymbol{R}_1.$$

- Step 2: Compute $\boldsymbol{X}^1$:

$$\boldsymbol{X}^1 = \boldsymbol{R}_1\boldsymbol{Q}_1 = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

- Step 3: Factorizing $\boldsymbol{X}^1$. We can again use Givens rotations to compute $\boldsymbol{X}^1 = \boldsymbol{Q}_2\boldsymbol{R}_2$:

$$\boldsymbol{Q}_2^\top \boldsymbol{X}^1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & -\sqrt{2} \end{bmatrix} = \boldsymbol{R}_2.$$

- Step 4: Compute $\boldsymbol{X}^2$:

$$\boldsymbol{X}^2 = \boldsymbol{R}_2\boldsymbol{Q}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \boldsymbol{X}^0.$$

To use a Householder transformation, let us denote the first column of $\boldsymbol{X}^0$ as $\boldsymbol{a}$. Then:

$$\boldsymbol{v}_1 = \boldsymbol{a} - \|\boldsymbol{a}\|\boldsymbol{e}_1 = \begin{bmatrix} 1 - \sqrt{2} \\ 1 \end{bmatrix}, \quad \|\boldsymbol{v}_1\|^2 = 4 - 2\sqrt{2},$$

and

$$
\boldsymbol{H}_{v_1} = \boldsymbol{I} - 2\frac{\boldsymbol{v}_1\boldsymbol{v}_1^\top}{\|\boldsymbol{v}_1\|^2} = \begin{bmatrix} 1 - \frac{(1-\sqrt{2})^2}{2-\sqrt{2}} & -\frac{1-\sqrt{2}}{2-\sqrt{2}} \\ -\frac{1-\sqrt{2}}{2-\sqrt{2}} & 1 - \frac{1}{2-\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1-\sqrt{2}}{2-\sqrt{2}} & -\frac{1-\sqrt{2}}{2-\sqrt{2}} \\ -\frac{1-\sqrt{2}}{2-\sqrt{2}} & \frac{1-\sqrt{2}}{2-\sqrt{2}} \end{bmatrix}
$$

$$
= -\frac{1-\sqrt{2}}{2-\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \cdots
$$

c) Due to $\boldsymbol{X}^2 = \boldsymbol{X}^0$, the QR algorithm does not converge in this case. Indeed, since the two eigenvalues have the same modulus, the QR iteration is not guaranteed to "converge".

Exercise 2 (Power Iteration with Shift): (*21* points)

Let the matrix $\boldsymbol{A} \in \mathbb{R}^{4\times4}$ and the initial point $\boldsymbol{x}^0 \in \mathbb{R}^4$ be given via

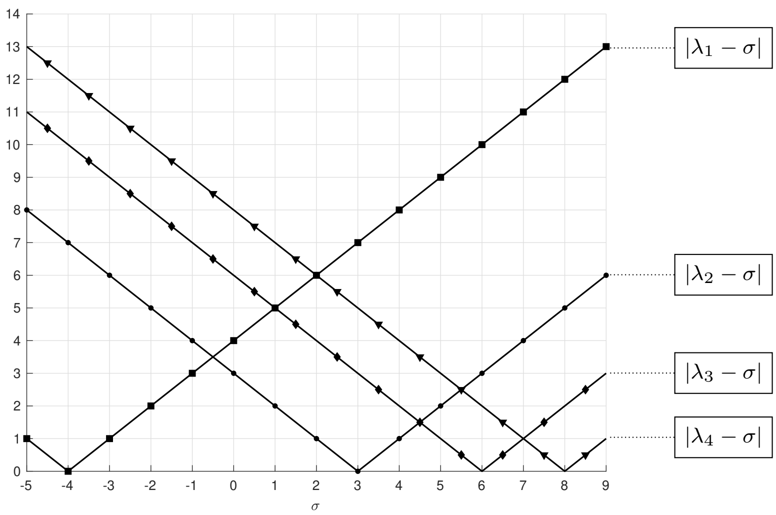
$$
\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 5 & 0 \\ 0 & 3 & 0 & 0 \\ 5 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \qquad \tilde{\boldsymbol{x}}^0 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \boldsymbol{x}^0 = \frac{\tilde{\boldsymbol{x}}^0}{\|\tilde{\boldsymbol{x}}^0\|}. \tag{1}
$$

In this problem, we want to apply the (normalized) power iteration with shift $\sigma \in \mathbb{R}$ to $\boldsymbol{A}$:

$$
\tilde{\boldsymbol{x}}^k = (\boldsymbol{A} - \sigma\boldsymbol{I})\boldsymbol{x}^k, \quad \boldsymbol{x}^k = \tilde{\boldsymbol{x}}^k/\|\tilde{\boldsymbol{x}}^k\|, \quad \sigma_k = (\boldsymbol{x}^k)^\top\boldsymbol{A}\boldsymbol{x}^k, \quad k = 1, 2, \ldots
$$

Let $(\lambda_i, \boldsymbol{v}_i)$, $i = 1, \ldots, 4$, further denote the corresponding eigenpairs of $\boldsymbol{A}$. It can be shown that $(\boldsymbol{x}^0)^\top\boldsymbol{v}_i \neq 0$ for all $i = 1, \ldots, 4$.

The plot below depicts the mappings $\sigma \mapsto |\lambda_i - \sigma|$ for the different eigenvalues of $\boldsymbol{A}$ and choices of σ.



a) State the eigenvalues $\lambda_1, \ldots, \lambda_4$ of $\boldsymbol{A}$.

b) Discuss which eigenpairs $(\lambda_i, \boldsymbol{v}_i)$ of $\boldsymbol{A}$ can be recovered by the (normalized) power iteration using suitable choices of the shift σ. Can all eigenpairs of $\boldsymbol{A}$ be recovered? Provide detailed explanations!

c) Derive the optimal choice of the shift σ for which the power iteration (1) converges with the fastest possible (optimal) convergence rate.

Solution :

a) We can directly read the eigenvalues of $\boldsymbol{A}$ from the plot: $\lambda_1 = -4$, $\lambda_2 = 3$, $\lambda_3 = 6$, and $\lambda_4 = 8$.

b) The power iteration with shift σ "converges" to the eigenpair of $\boldsymbol{A} - \sigma\boldsymbol{I}$ corresponding to the largest eigenvalue of $\boldsymbol{A} - \sigma\boldsymbol{I}$ (in magnitude). As the eigenpairs of $\boldsymbol{A} - \sigma\boldsymbol{I}$ are given by $(\lambda_i - \sigma, \boldsymbol{v}_i)$, we can use the depicted plot to determine which eigenvalue has the largest magnitude. In particular, we are interested in finding

$$\max_{i=1,2,3,4} |\lambda_i - \sigma|$$

Clearly, for $\sigma < 2$, $\lambda_4 - \sigma = 8 - \sigma$ has the largest magnitude; for $\sigma > 2$, $\lambda_1 - \sigma = -4 - \sigma$ has the largest magnitude. For $\sigma = 2$, it holds that $|\lambda_4 - 2| = |\lambda_1 - 2|$, i.e., there is no gap between the first and second largest eigenvalue and the power method does not converge.

As $(\boldsymbol{x}^0)^\top \boldsymbol{v}_1 \neq 0$ and $(\boldsymbol{x}^0)^\top \boldsymbol{v}_4 \neq 0$, we can conclude that the power method with shift only allows to recover the eigenpairs $(-4, \boldsymbol{v}_1)$ and $(8, \boldsymbol{v}_4)$.

c) The rate of convergence depends on the ratio between the second and first largest eigenvalue of $\boldsymbol{A} - \sigma\boldsymbol{I}$ (in magnitude). Let us denote those two eigenvalues by $\tilde{\lambda}_1 - \sigma$ and $\tilde{\lambda}_2 - \sigma$ ($\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ can change with the choice of σ). Then, the rate is given by:

$$\frac{|\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|} = 1 - \frac{|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|}$$

We see that the best rate can be obtained when $|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|$ is as large as possible and $|\tilde{\lambda}_1 - \sigma|$ is a small as possible. Using the plot, we can easily find $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$. For $\sigma \leq 1$, we have

$$\frac{|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|} = \frac{2}{8 - \sigma} \qquad \Longrightarrow \qquad \text{best value at } \sigma = 1\text{: } \frac{2}{7}.$$

For $\sigma \geq 5.5$, we have

$$\frac{|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|} = \frac{7}{4 + \sigma} \qquad \Longrightarrow \qquad \text{best value at } \sigma = 5.5\text{: } \frac{14}{19}.$$

For $\sigma \in (1, 2]$, it holds that $\frac{|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|} = \frac{8-\sigma-4-\sigma}{8-\sigma} = \frac{4-2\sigma}{8-\sigma}$ – this is largest for $\sigma = 1$ (again yielding $\frac{2}{7}$).

For $\sigma \in [2, 5.5)$, it holds that $\frac{|\tilde{\lambda}_1 - \sigma| - |\tilde{\lambda}_2 - \sigma|}{|\tilde{\lambda}_1 - \sigma|} = \frac{4+\sigma-8+\sigma}{4+\sigma} = \frac{2\sigma-4}{4+\sigma}$ – which is largest for $\sigma = 5.5$.

Overall, we can conclude that the optimal rate is attained for the shift $\sigma = 5.5$.

Exercise 3 (Miscellaneous): (*16* points)

State whether each of the following statements is *True* or *False*. For each statement provide a short explanation or proof. Only answers with full explanations will be graded. (Short answers of the form "true" or "false" will not be accepted).

a) The problem "*find x such that $f(x) = 0$*" for $f(x) := x^2 - 1$ is well-conditioned at its roots.

b) The eigenvalues of the matrix

$$\boldsymbol{A} = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

are all nonnegative and lie in region $[-1, 0]$. (* not covered this year).

c) Let $\lambda \in \mathbb{R}$ be an eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Then, there is $\boldsymbol{v} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ such that $\boldsymbol{v}^\top \boldsymbol{A} = \lambda \boldsymbol{v}^\top$.

d) Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ be given with $m > n$ and suppose that $\boldsymbol{A}$ has full column rank. Then, the CG method can find a solution to the least-squares problem $\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2$ in at most n steps.

Solution :

a) *True*. We have $\frac{1}{|f'(x)|} = \frac{1}{2|x|}$. This term is $\frac{1}{2}$ for $x = \pm 1$.

b) *False*. According to Gershgorin, all eigenvalues of $\boldsymbol{A}$ need to lie in the union of the disks

$$D_1 := \{z \in \mathbb{C} : |z + 2| \leq 1\} \quad \text{and} \quad D_2 := \{z \in \mathbb{C} : |z + 2| \leq 2\}.$$

Since $\boldsymbol{A}$ is symmetric, all eigenvalues are real and hence must lie in $[-4, 0]$. Moreover, due to $\operatorname{tr}(\boldsymbol{A}) = -8 = \sum_{i=1}^{4} \lambda_i$, there must exists some eigenvalue $\lambda < -1$. (Can be also argued by computing its characteristic polynomial, while less efficient) (* not covered this year).

c) *True*. Since $0 = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \det(A^\top - \lambda \boldsymbol{I})$ (where we used $\det(\boldsymbol{B}) = \det(\boldsymbol{B}^\top)$ for any matrix $\boldsymbol{B} \in \mathbb{R}^{n \times n}$), λ is also an eigenvalue of $\boldsymbol{A}^\top$. The stated property then holds for the associated eigenvector of $\boldsymbol{A}^\top$.

d) *True*. We can apply the CG-method to the normal equation $\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{b}$. Since $\boldsymbol{A}$ has full column rank, the matrix $\boldsymbol{A}^\top \boldsymbol{A}$ is positive definite (which is a requirement for the convergence of CG).

Exercise 4 (Singular Value Decomposition): (*22* points)
This exercise concerns singular value decompositions of matrices and their properties.

a) Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be a given matrix with $m > n$ with singular values σ_i, $i = 1, \ldots, n$. Prove that $\|\boldsymbol{A}\|_F^2 = \sum_{i=1}^{n} \sigma_i^2$.

Hint: The Frobenius norm of $\boldsymbol{A}$ is given by $\|\boldsymbol{A}\|_F^2 = \operatorname{tr}(\boldsymbol{A}^\top \boldsymbol{A})$. The identity $\operatorname{tr}(\boldsymbol{B}\boldsymbol{C}) = \operatorname{tr}(\boldsymbol{B}\boldsymbol{C})$ (for matrices $\boldsymbol{B}$, $\boldsymbol{C}$ with matching dimensions) might be helpful.

b) We now consider the specific matrix:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & \sqrt{2} \\ -1 & 1 & 1 & 0 \\ -1 & -1 & -1 & 0 \end{bmatrix}.$$

– Determine the rank of $\boldsymbol{A}$ and compute $\|\boldsymbol{A}\|_F^2$.

– Let $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$ be a SVD of $\boldsymbol{A}$ with $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_4)$ and $\sigma_1 \geq \cdots \geq \sigma_4$. After some first calculations, the following partial information about $\boldsymbol{U}$ and $\boldsymbol{\Sigma}$ is available:

$$\boldsymbol{U} = \begin{bmatrix} \star & 0 & 0 & \frac{\sqrt{6}}{3} \\ 0 & -\frac{\sqrt{6}}{3} & \star & 0 \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \star & \star & \star & \star \\ 0 & 2 & 0 & \star \\ 0 & 0 & 1 & \star \\ 0 & 0 & 0 & \star \end{bmatrix}.$$

Determine the missing entries “$\star$” in the matrices $\boldsymbol{U}$ and $\boldsymbol{\Sigma}$.

– Does the vector $\boldsymbol{b} = [1,\ 2023,\ 1,\ -1]^\top$ belong to the range of $\boldsymbol{A}$? Explain your answer!

Solution :

a) Let $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$ be an associated SVD of $\boldsymbol{A}$. Following the hint, it holds that

$$\|\boldsymbol{A}\|_F^2 = \mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{V}\boldsymbol{\Sigma}^\top\boldsymbol{U}^\top\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top) = \mathrm{tr}(\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}\boldsymbol{V}^\top\boldsymbol{V}) = \mathrm{tr}(\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}),$$

where we used the orthogonality of the matrices $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ and $\boldsymbol{V} \in \mathbb{R}^{n\times n}$. Due to $m > n$, we have $\boldsymbol{\Sigma}^\top = [\tilde{\boldsymbol{\Sigma}}\ \boldsymbol{0}] \in \mathbb{R}^{n\times m}$ where $\tilde{\boldsymbol{\Sigma}} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$. Hence, the result follows from $\mathrm{tr}(\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}) = \mathrm{tr}(\tilde{\boldsymbol{\Sigma}}^2) = \sum_{i=1}^n \sigma_i^2$.

b) Clearly, we have $\mathrm{rank}(\boldsymbol{A}) = 3$ and $\|\boldsymbol{A}\|_F^2 = 11$.

Let $\boldsymbol{u}_1$ be the first column of $\boldsymbol{U}$. As each of the columns of $\boldsymbol{U}$ is normalized, we can conclude $\star^2 + \frac{2}{3} = 1$, i.e., $\star = \pm\frac{\sqrt{3}}{3}$. As $\boldsymbol{u}_1$ and $\boldsymbol{u}_4$ need to be orthogonal, we can immediately conclude $\star = -\frac{\sqrt{3}}{3}$. Similarly, the missing entry in $\boldsymbol{u}_3$ needs to be $\star = -\frac{\sqrt{3}}{3}$.

$\boldsymbol{\Sigma}$ is diagonal, so we only need to find the missing diagonal elements. Due to $\mathrm{rank}(\boldsymbol{A}) = 3$, we can infer $\sigma_4 = \boldsymbol{\Sigma}_{44} = 0$. Moreover, applying part a), it follows $11 = \|\boldsymbol{A}\|_F^2 = \sigma_1^2 + 4 + 1 + 0 = \sigma_1^2 + 5$. This implies $\sigma_1 = \boldsymbol{\Sigma}_{11} = \sqrt{6}$. (Singular values are nonnegative).

The condition $\boldsymbol{b} \in \mathrm{range}(\boldsymbol{A})$ is equivalent to $\boldsymbol{u}_4^\top\boldsymbol{b} = 0$. Indeed, it holds that $\boldsymbol{u}_4^\top\boldsymbol{b} = \frac{\sqrt{6}}{3} - 2\cdot\frac{\sqrt{6}}{6} = 0$. Hence, $\boldsymbol{b}$ belongs to the range of $\boldsymbol{A}$!

Exercise 5 (Newton's Method): (*21* points)

We consider the following nonlinear equation:

$$F(\boldsymbol{x}) = \boldsymbol{0} \quad \text{where} \quad F : \mathbb{R}^3 \to \mathbb{R}^3, \quad F(\boldsymbol{x}) := \begin{bmatrix} 4(x_1 - 1)^3 \\ x_2 + x_3 \\ 2x_3 + x_2 \end{bmatrix}. \tag{2}$$

The goal of this exercise is to study the behavior of Newton's method applied to (2).

a) Show that $\boldsymbol{x}^* = [1,\ 0,\ 0]^\top$ is the unique solution to the equation (2).

b) Compute the Jacobian $DF(\boldsymbol{x})$ and its inverse $DF(\boldsymbol{x})^{-1}$ for all $\boldsymbol{x}$ with $x_1 \neq 1$.

c) Let us set $\boldsymbol{x}^0 = [2,\ 2,\ 2]^\top$. Using this initial point, perform one step of Newton's method to solve the equation (2).

d) Find all initial points for which Newton's method converges (to $\boldsymbol{x}^*$). When starting from those initial points, will Newton's method converge quadratically?

Hint: Consider the general update formula of Newton's method. What can you say about $\boldsymbol{x}^{k+1}$?

Notice that the sequence $\{\boldsymbol{x}^k\}_k$ converges quadratically to $\boldsymbol{x}^*$ if $\boldsymbol{x}^k \to \boldsymbol{x}^*$ and if there is a constant C such that $\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^*\| \le C\|\boldsymbol{x}^k - \boldsymbol{x}^*\|^2$ as $k \to \infty$.

Solution :

a) Clearly, we have $F(\boldsymbol{x}) = \mathbf{0}$ if and only if $x_1 = 0$. Furthermore, it holds that $x_2 = -x_3$ which yields $x_3 = 0$ (in the third equation). In summary, $\boldsymbol{x}^* = [1,\, 0,\, 0]^\top$ is the only root of F.

b) It holds that

$$DF(\boldsymbol{x}) = \begin{bmatrix} 12(x_1-1)^2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

For $x_1 \neq 1$, the inverse of $DF(\boldsymbol{x})$ is given by

$$[DF(\boldsymbol{x})]^{-1} = \begin{bmatrix} \frac{1}{12}(x_1-1)^{-2} & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

c) Setting $\boldsymbol{x}^0 = [2,\, 2,\, 2]^\top$, we have

$$\boldsymbol{x}^1 = \boldsymbol{x}^0 - [DF(\boldsymbol{x}^0)]^{-1}F(\boldsymbol{x}^0) = \boldsymbol{x}^0 - \begin{bmatrix} \frac{1}{12} & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ 0 \\ 0 \end{bmatrix}.$$

d) Following the hint, we first consider a general update of Newton's method. Let $\boldsymbol{x}^k \in \mathbb{R}^3$ be given. Then, we have

$$x_1^{k+1} = x_1^k - \frac{1}{12}(x_1^k - 1)^{-2} \cdot 4(x_1^k - 1)^3 = 1 + \frac{2}{3}(x_1^k - 1) \tag{3}$$

and

$$\begin{bmatrix} x_2^{k+1} \\ x_3^{k+1} \end{bmatrix} = \begin{bmatrix} x_2^k \\ x_3^k \end{bmatrix} - \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_2^k + x_3^k \\ 2x_3^k + x_2^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{4}$$

Hence, from (4), we know $x_2^k = x_3^k = 0$ for all $k \geq 1$. Combining this with (3), it follows

$$\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^*\| = \frac{2}{3}\|\boldsymbol{x}^k - \boldsymbol{x}^*\| \quad \forall\, k \geq 1.$$

Thus, $\{\boldsymbol{x}^k\}_k$ converges to $\boldsymbol{x}^*$ (with linear rate) for every choice of $\boldsymbol{x}^0 \in \mathbb{R}^n$. Moreover, $\{\boldsymbol{x}^k\}_k$ cannot converge quadratically since

$$\frac{\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}^k - \boldsymbol{x}^*\|^2} = \frac{2}{3\|\boldsymbol{x}^k - \boldsymbol{x}^*\|} \to +\infty \quad \text{as } k \to \infty.$$

(This is because $DF(\boldsymbol{x}^*)$ is singular).