1. *Consider applying the* ***basic Newton method*** *to minimise the function*

$$h(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 + 3x_2 - 1$$

on $\mathbb{R}^2$*, starting from the point* $\boldsymbol{x}^{(1)} = (1,\ 1)^\top$.

(i) *Calculate the Newton direction* $\boldsymbol{s}^{(1)}$ *for* h *at* $\boldsymbol{x}^{(1)}$.

Solution. The gradient and Hessian are given by

$$\nabla h(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - x_2 \\ 2x_2 - x_1 + 3 \end{pmatrix}, \ \nabla^2 h(\boldsymbol{x}) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

The Newton direction $\boldsymbol{s}^{(1)}$ satisfies

$$\nabla^2 h(\boldsymbol{x}^{(1)})\boldsymbol{s}^{(1)} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \boldsymbol{s}^{(1)} = -\nabla h(\boldsymbol{x}^{(1)}) = \begin{pmatrix} -1 \\ -4 \end{pmatrix}.$$

By row reduction:

$$\begin{pmatrix} 2 & -1 & | & -1 \\ -1 & 2 & | & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -\frac{1}{2} & | & -\frac{1}{2} \\ 1 & -2 & | & 4 \end{pmatrix} \to \begin{pmatrix} 1 & -\frac{1}{2} & | & -\frac{1}{2} \\ 0 & -\frac{3}{2} & | & \frac{9}{2} \end{pmatrix},$$

giving $\boldsymbol{s}^{(1)} = (-2, -3)^\top$.

(ii) *Is the Newton direction* $\boldsymbol{s}^{(1)}$ *a descent direction for* h *at* $\boldsymbol{x}^{(1)}$*. Give reasons for your answer.*

Solution. It is a descent direction as

$$\nabla h(\boldsymbol{x}^{(1)})^\top \boldsymbol{s}^{(1)} = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ -3 \end{pmatrix} = -14 < 0.$$

(iii) *Find the next Newton iterate* $\boldsymbol{x}^{(2)}$.

Solution. $\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + \boldsymbol{s}^{(1)} = (1, 1)^\top + (-2, -3)^\top = (-1, -2)^\top$.

(iv) *Stating reasons, show that* $\boldsymbol{x}^{(2)}$ *is the unique global minimiser for the function* h.

Solution. The point $\boldsymbol{x}^{(2)}$ is a stationary point because

$$\nabla h(\boldsymbol{x}^{(2)}) = \begin{pmatrix} (2)(-1) - (-2) \\ (2)(-2) - (-1) + 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Moreover, $\nabla^2 h(\boldsymbol{x}) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ with $\mathrm{tr}(\nabla^2 h(\boldsymbol{x})) = 4 > 0$ and $\det(\nabla^2 h(\boldsymbol{x})) = 3 > 0$ has positive eigenvalues, so $\nabla^2 h(\boldsymbol{x})$ is positive definite for all $\boldsymbol{x} \in \mathbb{R}^2$. Hence, h is a strictly convex function over $\mathbb{R}^2$. Hence, the stationary point $\boldsymbol{x}^{(2)}$ is the unique global minimiser.

2. *Consider the* **equality constrained optimization** *problem*

$$\begin{array}{lll} (\mathrm{P}_1) & \min\limits_{\boldsymbol{x} \in \mathbb{R}^3} & -x_1x_2x_3 \\ & s.t. & x_1 + x_2 + x_3 - 40 = 0 \\ & & x_1 + x_2 - x_3 = 0, \end{array}$$

where $f(\boldsymbol{x}) = -x_1x_2x_3$, $c_1(\boldsymbol{x}) = x_1 + x_2 + x_3 - 40$ and $c_2(\boldsymbol{x}) = x_1 + x_2 - x_3$. It is **given** *that $\boldsymbol{x}^* = (10,\ 10,\ 20)^\top$ is a regular constrained stationary point for (P1) with the Lagrange multipliers $\lambda_1^* = 150$ and $\lambda_2^* = 50$. Using the second-order sufficient optimality conditions, determine whether or not $\boldsymbol{x}^*$ is a (strict) local minimiser for (P1).*

Solution. The Lagrangian function is $L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \lambda_1 c_1(\boldsymbol{x}) + \lambda_2 c_2(\boldsymbol{x})$ with gradient and Hessian

$$\nabla L(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{pmatrix} -x_2x_3 + \lambda_1 + \lambda_2 \\ -x_1x_3 + \lambda_1 + \lambda_2 \\ -x_1x_2 + \lambda_2 - \lambda_2 \end{pmatrix}, \ \nabla^2 L(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{pmatrix} 0 & -x_3 & -x_2 \\ -x_3 & 0 & -x_1 \\ -x_2 & -x_1 & 0 \end{pmatrix}, \ \nabla^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \begin{pmatrix} 0 & -20 & -10 \\ -20 & 0 & -10 \\ -10 & -10 & 0 \end{pmatrix}.$$

The matrix $\nabla^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ is not positive definite. So, to calculate the reduced Hessian matrix, we want to find $Z^* = (\alpha, \beta, \gamma)^\top \in \mathbb{R}^{3\times 1}$ with full rank such that

$$(Z^*)^\top \begin{pmatrix} \nabla c_1(\boldsymbol{x}^*) & \nabla c_2(\boldsymbol{x}^*) \end{pmatrix} = (Z^*)^\top \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \alpha + \beta + \gamma & \alpha + \beta - \gamma \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

This means

$$\alpha + \beta + \gamma = 0, \text{ and } \alpha + \beta - \gamma = 0.$$

This gives $\gamma = 0$ and $\alpha = -\beta$. So Z^* is of the form $Z^* = (-\beta, \beta, 0)$, $\beta \neq 0$. Choose $Z^* = (-1, 1, 0)^\top$. The reduced Hessian is

$$W^* = (Z^*)^\top \nabla L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) Z^* = \begin{pmatrix} -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -20 & -10 \\ -20 & 0 & -10 \\ -10 & -10 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 40,$$

which is positive definite (1×1) matrix. By the second-order sufficient optimality condition, $\boldsymbol{x}^*$ is a strict local minimiser.

3. *Consider the* ***convex optimization*** *problem*

$$(P2) \qquad \min_{\boldsymbol{x} \in \mathbb{R}^2} \quad e^{-2(x_1+x_2)}$$
$$\text{subject to} \quad e^{x_1} + e^{x_2} - 4 \leq 0.$$

(i) *Find a constrained stationary point $\boldsymbol{z}^*$ of the problem $(P2)$.*

Solution. Let $c_1(x_1, x_2) := e^{x_1} + e^{x_2} - 4$. Suppose that the constraint is active at $\boldsymbol{z}^*$, i.e., $c_1(z_1^*, z_2^*) = e^{z_1^*} + e^{z_2^*} - 4 = 0$. The Lagrangian function is $L(\boldsymbol{x}, \lambda_1) = e^{-2(x_1+x_2)} + \lambda_1(e^{x_1} + e^{x_2} - 4)$. For $\boldsymbol{z}^*$ to be a constrained stationary point, we require $\nabla_{\boldsymbol{z}} L(\boldsymbol{z}^*, \lambda_1^*) = \mathbf{0}$ and $c_1(\boldsymbol{z}^*) = 0$ for some λ_1^*, i.e.,

$$-2e^{-2(z_1^*+z_2^*)} + \lambda_1^* e^{z_1^*} = 0, \tag{1}$$
$$-2e^{-2(z_1^*+z_2^*)} + \lambda_1^* e^{z_2^*} = 0. \tag{2}$$
$$e^{z_1^*} + e^{z_2^*} - 4 = 0. \tag{3}$$

Subtracting Eqn. (1) from Eqn. (2) yields $\lambda_1^*(e^{z_2^*} - e^{z_1^*}) = 0$. Then either $\lambda_1^* = 0$ or $e^{z_1^*} = e^{z_2^*}$. If $\lambda_1^* = 0$, then Eqn. (1) implies $-2e^{-2(z_1^*+z_2^*)} = 0$ which is not possible. So, $e^{z_1^*} = e^{z_2^*}$, and $z_1^* = z_2^*$. Now, Eqn. (3) gives $z_1^* = z_2^* = \log 2$. Finally, from Eqn. (1), $\lambda_1^* = 2^{-4}$. Hence, $\boldsymbol{z}^* = (\log 2, \log 2)$ is a constrained stationary point with the Lagrange multiplier $\lambda_1^* = 2^{-4}$.

(ii) Is $\boldsymbol{z}^*$ of part (i) the only constrained stationary point for the problem (P2)? Give reasons for your answer.

Solution. Suppose that the constraint is not active at $\boldsymbol{z}^*$, i.e., $e^{z_1^*} + e^{z_2^*} - 4 < 0$. Then, the first-order necessary optimality condition requires

$$\nabla_{\boldsymbol{z}}(e^{-2(z_1^*+z_2^*)}) = \begin{pmatrix} -2e^{-2(z_1^*+z_2^*)} \\ -2e^{-2(z_1^*+z_2^*)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which is not possible. Hence, $\boldsymbol{z}^* = (\log 2, \log 2)$ is the **only** constrained stationary point.

(iii) Explain why the constrained stationary point $\boldsymbol{z}^*$ of part (i) is a global minimizer for the problem (P2).

Solution. It is given that (P_2) is a convex problem, and we have found $\lambda_1^* > 0$ for the active inequality constraint. By the KKT sufficient optimality conditions, the constrained stationary point $\boldsymbol{z}^*$ is a global minimiser.

(iv) Stating clear reasons, show that $\boldsymbol{z}^*$ of part (i) is the unique global minimizer for the problem (P2).

Solution. Suppose that $\boldsymbol{y}^* = (y_1^*, y_2^*)$ is another global minimiser for (P_2). Then, it is a regular local minimizer because $\nabla c_1(\boldsymbol{y}^*) = (e^{y_1^*}, e^{y_2^*})^\top \neq 0$. So, by the first-order necessary optimality conditions, it must be a constrained stationary point for (P_2). By part *(ii)*, $\boldsymbol{z}^*$ is the only constrained stationary point for (P_2), so $\boldsymbol{z}^* = \boldsymbol{y}^*$. Note that the constraint cannot be active at $\boldsymbol{y}^*$. Otherwise,the first-order necessary optimality condition gives

$$\nabla_{\boldsymbol{y}^*}(e^{-2(y_1^*+y_2^*)}) = \begin{pmatrix} -2e^{-2(z_1^*+z_2^*)} \\ -2e^{-2(z_1^*+z_2^*)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which is not possible. Hence, $\boldsymbol{z}^*$ is the unique global minimiser for (P_2).