PROBLEM 1. (optional: not for marking) For the following NLP, write down the KKT conditions and (making sure to check the constraint qualifications) identify all stationary points of the NLP together with their corresponding Lagrange multipliers. At each stationary point, describe the critical cone and check that one of the second-order conditions holds. Find all local minima. Sketch the feasible region, lines of constant objective function and the stationary points you found.

\[
\min x_1^3 + 2x_1^2 - 10x_1 + x_2^2 - 8x_2
\]
\[
\begin{align*}
\text{s.t.} & \quad x_1 + x_2 \leq 2 \\
& \quad x_1, x_2 \geq 0 \\
& \quad x \in \mathbb{R}^2
\end{align*}
\]

PROBLEM 2. Use the KKT conditions to determine whether \( x = (1, 1)^T \) can be optimal for the NLP below.

\[
\min x_1^4 + 2x_2^2
\]
\[
\begin{align*}
\text{s.t.} & \quad x_1^2 + x_2^2 \geq 2 \\
& \quad x \in \mathbb{R}^2
\end{align*}
\]

SOLUTION:

\[
L(x, \lambda) = x_1^4 + 2x_2^2 + \lambda(x_1^2 + x_2^2 - 2)
\]
\[
\nabla_x L(x, \lambda) = \begin{pmatrix} 4x_1^3 - 2\lambda_1 x_1 \\ 4x_2^3 - 2\lambda_1 x_2 \end{pmatrix}
\]
\[
\nabla^2_x L(x, \lambda) = \begin{pmatrix} 12x_1^2 - 2\lambda_1 & 0 \\ 0 & 4 - 2\lambda_1 \end{pmatrix}
\]
\[
\nabla g(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}
\]

KKT(a): Evaluate \( \nabla L \) at the point \( (1, 1)^T \) and set to zero:

\[
\nabla^2_L(1, 1, \lambda) = \begin{pmatrix} 12 - 2\lambda_1 & 0 \\ 0 & 4 - 2\lambda_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]

implies \( 4 - 2\lambda_1^* = 0, \lambda_1^* = 2 \).

KKT(b): \( g(x^*) = 2 - 1 - 1 = 0 \), and \( \lambda^* = 2 > 0 \), so \( \lambda^* g(x^*) = 0 \) as required. KKT(c): Not applicable.

The point is stationary for this problem (CQ hold because \( g \) is affine). To verify whether it is a local minimum, we evaluate the Hessian at this point to check the second order condition.

\[
\nabla^2_L(1, 1, 2) = \begin{pmatrix} 12 - 4 & 0 \\ 0 & 4 - 4 \end{pmatrix}
\]

1
The critical cone at \((1, 1, 2)\) is given by:

\[
C(x^*, \lambda^*) = \left\{ d \in \mathbb{R}^2: (-2, -2) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0 \right\}
\]

so \(d_2 = -d_1\). Verification of second order condition:

\[
(d_1 - d_1) \begin{pmatrix} 8 \\ 0 \\ 4 - 4 \end{pmatrix} \begin{pmatrix} d_1 \\ -d_1 \end{pmatrix} = 8d_1^2 > 0
\]

so the point \((1, 1)^T\) is indeed a local minimum.

***

**PROBLEM 3.** Consider minimising \(f(x), x \in \mathbb{R}^n\) subject to \(x \in F, F = \{ x: g(x) \leq 0, h(x) = 0 \}\), where \(f, g, h \in C^2\), and \(g\) and \(h\) are affine. Let \(B_\epsilon(x^*) = \{ x \in F: ||x^* - x||^2 \leq \epsilon \}\) be a feasible ball around a point \(x^*\).

(a) Show that \(B_\epsilon(x^*)\) contains points of the form \((x^* + td), d \in C(x^*)\), (for sufficiently small \(t\)), where:

\[
C(x^*) = \{ d \in \mathbb{R}^n: \nabla h(x^*)^T d = 0, \nabla g_i(x^*)^T d \leq 0, i \in I(x^*) \}\.
\]

(b) Use Taylor approximation to show that if for all directions \(d \in C(x^*)\), the Hessian of the Lagrangian satisfies the second order condition \(d^T \nabla^2_{xx} L(x^*, \lambda^*, \eta^*) d > 0\), then \(x^*\) is a local minimum.

**SOLUTION:**

(a) Because the constraint functions are affine, we can write:

\[
h_j(x) = a_h(j)^T x + c_h(j), \quad g_i(x) = a_g(i)^T x + c_g(i)
\]

for vectors \(a_h(j) \in \mathbb{R}^n, j = 1, \ldots, q\) and \(a_g(i) \in \mathbb{R}^n, i = 1, \ldots, p\) and constants \(c_h(j), j = 1, \ldots, q\) and \(c_g(i), i = 1, \ldots, p\). Let \(x \in B_\epsilon(x^*)\), then there exist \(t > 0, t \leq \epsilon\) and \(d \in \mathbb{R}^n\) such that \(x = x^* + td\). Because \(x\) is feasible, it follows that \(h(x) = 0, g(x) \leq 0\).

\[
h_j(x) = 0 \Rightarrow a_h(j)^T (x^* + td) = 0 \Rightarrow a_h(j)^T d = 0,
\]

but this is equivalent to \(\nabla h(x^*)^T d = 0\). Suppose now that \(i \in I(x^*)\), then \(g(x^*) = 0\) and for feasible \(x \in B_\epsilon(x^*)\) we have:

\[
g_i(x^*) \leq 0 \Rightarrow a_g(i)^T (x^* + td) + c_g(i) \leq 0 \Rightarrow a_g(i)^T d \leq 0,
\]

which is equivalent to \(nablag_i(x^*)^T d \leq 0, i \in I(x^*)\), so that \(g(x^*) < 0\). For \(i \notin I(x^*)\), by continuity of \(g\) there is a sufficiently small \(t\) such that for all possible directions \(d, g_i(x^* + td) = g_i(x^*) + ta_g(i)^T d < 0\). In other words, no condition is further required for satisfying the inactive constraints in a small neighbourhood of the stationary point \(x^*\).

(b) Consider \(x = x^* + td \in B_\epsilon(x^*)\). Then using Taylor,

\[
f(x) - f(x^*) = \nabla f(x^*)^T d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + o(t^2)
\]

\[
- \sum \lambda_i \nabla g_i(x^*)^T d - \sum \eta_j \nabla h(x^*)^T d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + o(t^2)
\]

\[
\geq \frac{t^2}{2} d^T \nabla^2 L(x^*, \lambda^*, \eta^*) d + o(t^2)
\]

2
where we have used that $x^*$ is a stationary point, so that:

$$\nabla f(x^*) = -\sum_i \lambda_i \nabla g_i(x^*) - \sum_j \eta_j \nabla h(x^*),$$

ad also either $g_i(x^*) = 0$, in which case $\lambda_i \nabla g_i(x^*)^T d \leq 0$, or $g_i(x^*) < 0$, in which case $\lambda_i = 0$. Also, we used that $\nabla^2 L(x^*, \lambda^*, \eta^*) = \nabla^2 f(x^*)$, because $\nabla^2 g_i(x) = \nabla^2 h(x) = 0$ (affine functions). Thus for sufficiently small $\epsilon$, $f(x) > f(x^*), x \in B_\epsilon(x^*)$, establishing that $x^*$ is indeed a local minimum.

***