Math 164 Homework 4

Nick Monozon

Problem 6.2

We are given the function

$$f(x_1, x_2) = \frac{1}{3}x_1^3 - 4x_1 + \frac{1}{3}x_2^3 - 16x_2.$$

Since we have no constraints, this is an unconstrained optimization problem. To find critical points, we want to find points (x_1, x_2) that satisfy the FONC

$$\nabla f(x_1, x_2) = \mathbf{0}.$$

The gradient is therefore

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_1^2 - 4 \\ x_2^2 - 16 \end{bmatrix}.$$

From this, there are four points (x_1, x_2) that satisfy the FONC. There are

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} -2 \\ -4 \end{bmatrix}, \ \boldsymbol{x}^{(2)} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}, \ \boldsymbol{x}^{(3)} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}, \text{ and } \boldsymbol{x}^{(4)} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

We then compute the Hessian matrix of f to be

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 2x_1 & 0 \\ 0 & 2x_2 \end{bmatrix}.$$

Hence, we see that $\boldsymbol{F}(\boldsymbol{x}^{(1)}) < 0$, so $\boldsymbol{x}^{(1)}$ is a strict local maximizer. We also have that $\boldsymbol{F}(\boldsymbol{x}^{(4)}) > 0$, so $\boldsymbol{x}^{(4)}$ is a strict local minimizer. In addition, the Hessian matrix is indefinite for $\boldsymbol{x}^{(3)}$ and $\boldsymbol{x}^{(4)}$, so these points are not extremizers.

Problem 6.3

Proof. Assume that $\boldsymbol{x}^*$ is a global minimizer of f over Ω and $\boldsymbol{x}^* \in \Omega' \subset \Omega$. Let $\boldsymbol{x} \in \Omega'$. Since $\Omega' \subset \Omega$, $\boldsymbol{x} \in \Omega$ and hence $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega'$. Therefore, $\boldsymbol{x}^*$ is a global minimizer of f over Ω'. □

Problem 6.8

(a) We are given the function $f : \mathbb{R}^2 \to \mathbb{R}$ such that

$$f(\boldsymbol{x}) = \boldsymbol{x}^\top \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \boldsymbol{x} + \boldsymbol{x}^\top \begin{bmatrix} 3 \\ 5 \end{bmatrix} + 6.$$

Thus the gradient is given by

$$\nabla f(\boldsymbol{x}) = 2 \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 3 \\ 5 \end{bmatrix},$$

which means that

$$\begin{aligned} \nabla f([1,1]^\top) &= 2 \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} \\ &= 2 \begin{bmatrix} 3 \\ 11 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 6 \\ 22 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 9 \\ 27 \end{bmatrix}. \end{aligned}$$

The Hessian matrix is given by

$$F(x) = 2\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix}.$$

And hence $F([1,1]^\top)$ is simply the Hessian above.

(b) The direction of the maximum rate of increase at $x^* = [1,1]^\top$ is $\nabla f(x^*)$. The directional derivative with respect to a unit vector in this direction is therefore given by

$$\left(\frac{\nabla f(x^*)}{\|\nabla f(x^*)\|}\right)^\top \nabla f(x^*) = \frac{\|\nabla f(x^*)\|^2}{\|\nabla f(x^*)\|} = \|\nabla f(x^*)\|.$$

So, at $x^* = [1,1]^\top$, the rate of increase of f is

$$\|f([1,1]^\top)\| = \sqrt{9^2 + 27^2} = \sqrt{810} = 28.46.$$

(c) The FONC for the interior case is satisfied when

$$\nabla f(x^*) = \mathbf{0}.$$

We solve to find that

$$x^* = \begin{bmatrix} 11/2 \\ -7/2 \end{bmatrix}.$$

We now need to check if this satisfies the second-order necessary condition (SONC). We find that

$$F(x^*) = \begin{bmatrix} 2 & 4 \\ 8 & 14 \end{bmatrix}$$

which has eigenvalues $\lambda_1 = 8 - 2\sqrt{17} < 0$ and $\lambda_2 = 8 + 2\sqrt{17} > 0$. Hence, the Hessian at x^* is not positive semidefinite, so the SONC is not satisfied.

Problem 6.9

(a) Given the function

$$f(x_1, x_2) = x_1^2 x_2 + x_2^3 x_1,$$

the gradient is

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 x_2 + x_2^3 \\ x_1^2 + 3x_1 x_2^2 \end{bmatrix}.$$

At the point $x^{(0)} = [2,1]^\top$, f decreases most rapidly in the direction of the negative gradient. That is, the direction

$$-\nabla f(x^{(0)}) = -\begin{bmatrix} 5 \\ 10 \end{bmatrix}.$$

(b) The rate of increase of f at $x^{(0)}$ in the direction $-\nabla f(x^{(0)})$ is given by

$$\nabla f\left(x^{(0)}\right)^\top \frac{-\nabla f(x^{(0)})}{\|\nabla f(x^{(0)})\|} = -\|\nabla f(x^{(0)})\| = -\sqrt{125} = -5\sqrt{5}.$$

(c) Hence, the rate of increase of f at the point $x^{(0)}$ in the direction $d = [3,4]^\top$ is

$$\left\langle \nabla f([2,1]^\top), \frac{d}{\|d\|} \right\rangle = [5, 10]\begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = 5 \times 3/5 + 10 \times 4/5 = 3 + 8 = 11.$$

Problem 6.12

(a) We have that $f(\boldsymbol{x}) = 5x_2$ and $\Omega = \{\boldsymbol{x} : x_1^2 + x_2 \geq 1\}$. For $\boldsymbol{x}^* = [0,1]^\top$ satisfy the first-order necessary condition (FONC), we require that for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$,

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0.$$

We must first find the set of feasible directions at $\boldsymbol{x}^*$, denoted $\mathcal{D}_{\boldsymbol{x}^*}$. If $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$, then there exists $\bar{\alpha} > 0$ such that for any $\eta \in [0, \bar{\alpha}]$, we have that $\boldsymbol{x}^* + \eta \boldsymbol{d} \in \Omega$. We have that

$$\begin{aligned}
[0,1]^\top + \eta[d_1, d_2]^\top = [\eta d_1, 1 + \eta d_2]^\top \in \Omega \quad &\Leftrightarrow \quad (\eta d_1)^2 + (1 + \eta d_2) \geq 1 \\
&\Leftrightarrow \quad \eta^2 d_1^2 + \eta d_2 \geq 0 \\
&\Leftrightarrow \quad \eta d_1^2 + d_2 \geq 0 \\
&\Leftrightarrow \quad \eta \geq -\frac{d_2}{d_1^2} \qquad (d_1 \neq 0).
\end{aligned}$$

Since $\eta \geq 0$ and $d_1^2 > 0$, we require that $d_2 \geq 0$ and $d_1 \in \mathbb{R}^*$. However, we now consider the case when $d_1 = 0$. We have that

$$[0,1]^\top + \eta[0, d_2]^\top = [0, 1 + \eta d_2]^\top \quad \Leftrightarrow \quad 0^2 + (1 + \eta d_2) \geq 1.$$

So we have that $d_1 = 0$ and $d_2 \geq 0$. This means that

$$\begin{aligned}
\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*} \quad &\Rightarrow \quad \boldsymbol{d} \in \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \in \mathbb{R}, d_2 \geq 0\} \\
&\Rightarrow \quad \mathcal{D}_{\boldsymbol{x}^*} \subset \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \in \mathbb{R}, d_2 \geq 0\}.
\end{aligned}$$

Let's now show the other containment. Let $\mathcal{A} = \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \in \mathbb{R}, d_2 \geq 0\}$. If $\boldsymbol{d} \in \mathcal{A}$, this implies that $d_1 \in \mathbb{R}, d_2 \geq 0$. We have that for all $\eta \geq 0$,

$$[0,1]^\top + \eta[d_1, d_2]^\top = [\eta d_1, 1 + \eta d_2]^\top$$

and $(\eta d_1)^2 + (1 + \eta d_2) \geq 1$, meaning that $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$ and hence $\mathcal{A} \subset \mathcal{D}_{\boldsymbol{x}^*}$. By showing both directions of containment, we have shown that the set of feasible directions is given by

$$\mathcal{D}_{\boldsymbol{x}^*} = \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \in \mathbb{R}, d_2 \geq 0\}.$$

We can now check if the FONC is satisfied. We have that

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) = \boldsymbol{d}^\top \begin{bmatrix} 0 \\ 5 \end{bmatrix} = 5d_2 \geq 0$$

because $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$ where $d_2 \geq 0$. Hence, the FONC is satisfied at $\boldsymbol{x}^* = [1,0]^\top$.

(b) The SONC is satisfied if for any feasible direction $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$,

$$\boldsymbol{d}^\top \frac{\partial^2 f}{\partial x^2} \boldsymbol{d} \geq 0.$$

We find the Hessian matrix to be

$$\frac{\partial^2 f}{\partial x^2} = \mathbf{J}(\nabla f(\boldsymbol{x})) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus, for any $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$,

$$\boldsymbol{d}^\top \frac{\partial^2 f}{\partial x^2} \boldsymbol{d} = 0 \geq 0,$$

so the SONC is satisfied at $\boldsymbol{x}^* = [1,0]^\top$.

(c) Consider points of the form $\boldsymbol{x} = \left[x_1, 1 - x_1^2\right]^\top$ where $x_1 \in \mathbb{R}$. All such points are in Ω, but we see that for $\boldsymbol{x} \neq \boldsymbol{x}^*$,

$$f(\boldsymbol{x}) = -5(-x_1^2 + 1) = 5x_1^2 - 5 < 5 = f(\boldsymbol{x}^*).$$

So $\boldsymbol{x}^* = \left[0, 1\right]^\top$ is not a local minimizer.

Problem 6.16

We have that $f(x) = 4x_1^2 - x_2^2$ and $\Omega = \left\{\boldsymbol{x} : x_1^2 + 2x_1 - x_2 \geq 0, x_1 \geq 0, x_2 \geq 0\right\}$. For $\boldsymbol{x}^* = \left[0, 0\right]^\top$ satisfy the first-order necessary condition (FONC), we require that for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$,

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0.$$

We must first find the set of feasible directions at $\boldsymbol{x}^*$, denoted $\mathcal{D}_{\boldsymbol{x}^*}$. If $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$, then there exists $\bar{\alpha} > 0$ such that for any $\eta \in [0, \bar{\alpha}]$, we have that $\boldsymbol{x}^* + \eta \boldsymbol{d} \in \Omega$. We have that

$$\begin{aligned}
\left[0,0\right]^\top + \eta\left[d_1, d_2\right]^\top = \left[\eta d_1, 1 + \eta d_2\right]^\top \in \Omega \quad &\Leftrightarrow \quad (\eta d_1)^2 + 2(\eta d_1) - \eta d_2 \geq 0 \\
&\Leftrightarrow \quad \eta^2 d_1^2 + 2\eta d_1 - \eta d_2 \geq 0 \\
&\Leftrightarrow \quad \eta d_1^2 + 2d_1 - d_2 \geq 0 \\
&\Leftrightarrow \quad \eta \geq \frac{d_2 - 2d_1}{d_1^2} \qquad (d_1 \neq 0).
\end{aligned}$$

Since $\eta \geq 0$ and $d_1^2 > 0$, we require that $d_2 - 2d_1 \leq 0$, or that $d_2 \leq 2d_1$. However, we must also consider the case when $d_1 = 0$. We have that

$$\begin{aligned}
\left[0,0\right]^\top + \eta\left[0, d_2\right]^\top = \left[0, \eta d_2\right]^\top \quad &\Leftrightarrow \quad 0^2 + 2(0) - \eta d_2 \geq 0 \\
&\Leftrightarrow \quad 0 \geq \eta d_2.
\end{aligned}$$

Since $\eta \geq 0$ and we have the constraint $d_2 \geq 0$, we must have here that $d_2 = 0$. So we see that

$$\begin{aligned}
\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*} \quad &\Rightarrow \quad \boldsymbol{d} \in \left\{\boldsymbol{d} \in \mathbb{R}^2 : 2d_1 \geq d_2, d_1 \geq 0, d_2 \geq 0\right\} \\
&\Rightarrow \quad \mathcal{D}_{\boldsymbol{x}^*} \subset \left\{\boldsymbol{d} \in \mathbb{R}^2 : 2d_1 \geq d_2, d_1 \geq 0, d_2 \geq 0\right\}.
\end{aligned}$$

Let's show the other containment now. Let $\mathcal{A} = \left\{\boldsymbol{d} \in \mathbb{R}^2 : 2d_1 \geq d_2, d_1 \geq 0, d_2 \geq 0\right\}$. If $\boldsymbol{d} \in \mathcal{A}$, this implies that $2d_1 \geq d_2, d_1 \geq 0, d_2 \geq 0$. Hence, for all $\eta \geq 0$,

$$\left[0,0\right]^\top + \eta\left[d_1, d_2\right]^\top = \left[\eta d_1, \eta d_2\right]^\top$$

and therefore $(\eta d_1)^2 + 2(\eta d_1) - \eta d_2 \geq 0$, meaning that $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$ and hence $\mathcal{A} \subset \mathcal{D}_{\boldsymbol{x}^*}$. By showing both directions of containment, we have established that the set of feasible directions is

$$\mathcal{D}_{\boldsymbol{x}^*} = \left\{\boldsymbol{d} \in \mathbb{R}^2 : 2d_1 \geq d_2, d_1 \geq 0, d_2 \geq 0\right\}.$$

We now check whether the FONC is satisfied. We have that

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) = \boldsymbol{d}^\top \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \geq 0.$$

So $\boldsymbol{x}^*$ satisfies the FONC.

(b) The SONC is satisfied if for any feasible direction $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$,

$$\boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} \geq 0.$$

We find the Hessian matrix for $\boldsymbol{x}^*$ to be

$$\boldsymbol{F}(\boldsymbol{x}^*) = \mathbf{J}(\nabla f(\boldsymbol{x}^*)) = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}.$$

So, for any $\boldsymbol{d} \in \mathcal{D}_{\boldsymbol{x}^*}$,

$$\boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} = 8d_1^2 - 2d_2^2 \geq 8d_1^2 - 2(2d_1)^2 = 0 \geq 0.$$

The SONC is satisfied at $\boldsymbol{x}^*$.

(c) We are told that $f(\boldsymbol{x}^*) = 0$. Consider points of the form $\boldsymbol{x} = \left[x_1, x_1^2 + 2x_1\right]^\top$ for $x_1 > 0$. Clearly all of these points are in Ω and hence feasible. We see that

$$f(\boldsymbol{x}) = 4x_1^2 - (x_1^2 + 2x_1)^2 = -(4x_1^3 + x_1^4) < 0 = f(\boldsymbol{x}^*).$$

So although $\boldsymbol{x}^*$ satisfies the FONC and SONC, it is not a local minimizer.

Problem 6.18

Given $x \in \mathbb{R}$, define

$$f(x) = \sum_{i=1}^{n} (x - x_i)^2.$$

Let $\bar{x}$ be the minimizer of f. By the FONC, this means that

$$f'(\bar{x}) = 0,$$

so $\bar{x}$ is the minimizer of f when

$$\begin{aligned}
\sum_{i=1}^{n} 2(\bar{x} - x_i) = 0 \quad &\Leftrightarrow \quad 2\sum_{i=1}^{n} \bar{x} = 2\sum_{i=1}^{n} x_i \\
&\Leftrightarrow \quad \sum_{i=1}^{n} \bar{x} = \sum_{i=1}^{n} x_i \\
&\Leftrightarrow \quad n\bar{x} = \sum_{i=1}^{n} x_i \\
&\Leftrightarrow \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.
\end{aligned}$$

Problem 6.23

The FONC for the interior case is satisfied for all points $\boldsymbol{x}^*$ such that

$$\nabla f(\boldsymbol{x}) = \boldsymbol{0}.$$

Hence, we compute

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 4(x_1 - x_2)^3 + 2x_1 - 2 \\ -4(x_1 - x_2)^3 - 2x_2 + 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which yields the system of equations

$$\begin{cases} 4(x_1 - x_2)^3 + 2x_1 - 2 & = 0, \\ -4(x_1 - x_2)^3 - 2x_2 + 2 & = 0. \end{cases}$$

Adding the two equations, we find that $x_1 = x_2$, which we can substitute back in to find that $x_1 = x_2 = 1$. So, $\boldsymbol{x}^* = \left[1, 1\right]^\top$ is the only point that satisfies the FONC. The Hessian matrix at $\boldsymbol{x}^*$ is

$$\boldsymbol{F}(\boldsymbol{x}^*) = \mathbf{J}(\nabla f(\boldsymbol{x}^*)) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Since the matrix $\boldsymbol{F}(\boldsymbol{x}^*)$ has eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 2$, it is not positive semidefinite. Therefore, the SONC is not satisfied at $\boldsymbol{x}^* = \left[1, 1\right]^\top$.

Problem 6.33

Proof. ($\Rightarrow$) We know that if $\boldsymbol{x}^*$ minimizes f, then $\boldsymbol{x}^*$ satisfies the FONC by necessity, so this direction is complete. We just need to show the other direction.

($\Leftarrow$) We can express $f(\boldsymbol{x})$ as

$$f(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^*)^\top \boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{x}^*) - \frac{1}{2}\boldsymbol{x}^{*\top}\boldsymbol{Q}\boldsymbol{x}^*$$

where we define $\boldsymbol{x}^* = \boldsymbol{Q}^{-1}\boldsymbol{b}$ to be the unique vector that satisfies the FONC. Since $\frac{1}{2}\boldsymbol{x}^{*\top}\boldsymbol{Q}\boldsymbol{x}^*$ is a constant and $\boldsymbol{Q} > 0$, it follows that

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) = -\frac{1}{2}\boldsymbol{x}^{*\top}\boldsymbol{Q}\boldsymbol{x}^*$$

where $f(\boldsymbol{x}) = f(\boldsymbol{x}^*)$ if and only if $\boldsymbol{x} = \boldsymbol{x}^*$. This completes the proof of the second implication, so we are done. $\square$