Math 164 Homework 5

Nick Monozon

Problem 6.7

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x}_0 \in \mathbb{R}^n$, $\Omega \subset \mathbb{R}^n$, and $\Omega' = \{\boldsymbol{y} : \boldsymbol{y} - \boldsymbol{x}_0 \in \Omega\}$. For notation ease, let $\boldsymbol{w}_0 = \boldsymbol{x}_0 + \arg\min_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x})$. Therefore, we need to show that $\boldsymbol{w}_0 = \arg\min_{\boldsymbol{y}\in\Omega'} f(\boldsymbol{y} - \boldsymbol{x}_0)$. Hence, we must show that for all $\boldsymbol{z} \in \Omega'$,

$$f(\boldsymbol{z} - \boldsymbol{x}_0) \geq f(\boldsymbol{w}_0 - \boldsymbol{x}_0).$$

Choose some arbitrary $\boldsymbol{z} \in \Omega'$. By definition, $\boldsymbol{z} - \boldsymbol{x}_0 \in \Omega$. It follows that

$$\begin{aligned} f(\boldsymbol{z} - \boldsymbol{x}_0) &\geq \min_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x}) \\ &= f\left(\operatorname*{argmin}_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x})\right) \\ &= f(\boldsymbol{w}_0 - \boldsymbol{x}_0). \end{aligned}$$

And we're done. □

Problem 6.11

(a) Since $d_1 \geq 0$ and $|d_2| \leq d_1^2$ are conditions, the set of feasible directions is hence

$$\mathcal{D} = \{\boldsymbol{d} : d_1 \geq 0, d_2 = 0\}$$

as $d_1 \geq 0$, so $|d_2| \leq d_1^2$ for any $d_1 \geq 0$ only when $d_2 = 0$. This means that any feasible direction is of the form $[d_1, 0]^\top$ for $d_1 \geq 0$. Since $f(\boldsymbol{x}) = -x_2^2$, the gradient of f is

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 0 \\ -2x_2 \end{bmatrix}.$$

For the FONC to be satisfied at $\boldsymbol{x}^*$, we must have that

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

So if we consider the point $[x_1, x_2]^\top = \boldsymbol{0}$, then for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{0}$,

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{0}) = [d_1, 0] \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \geq 0.$$

So it is true.

(b) We have that $[x_1, x_2]^\top = \boldsymbol{0}$ is a local maximizer because $f(\boldsymbol{0}) = 0$ while any feasible point $\boldsymbol{x}$ satisfies just $f(\boldsymbol{x}) \leq 0$. However $\boldsymbol{0}$ is not a strict local maximizer because for any point $\boldsymbol{y}$ of the form $\boldsymbol{y} = [y_1, 0]^\top$, we can see that $f(\boldsymbol{y}) = 0 = f(\boldsymbol{0})$, which exist in any neighborhood of $\boldsymbol{0}$. Additionally, we know that $\boldsymbol{0}$ is not a local minimizer because for any point $\boldsymbol{y}$ of the form $\boldsymbol{y} = [y_1, y_1^2]^\top$ where $y_1 > 0$, $f(\boldsymbol{y}) = -y_1^4 < 0$. Such $\boldsymbol{y}$ exist in any neighborhood of $\boldsymbol{0}$. So $\boldsymbol{y}$ is neither a strict local minimizer not a local minimizer.

Problem 6.13

(a) The point $\boldsymbol{x}^* = [3,0]^\top$ satisfies the FONC if for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$, $\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0$ is true. We therefore must find the set of feasible directions at $\boldsymbol{x}^*$, which we will denote $\mathcal{D}^*$. If we consider $\boldsymbol{d} \in \mathcal{D}^*$, this implies that there exists $\bar{\alpha} > 0$ such that for any $\eta \in [0, \bar{\alpha}]$, $\boldsymbol{x}^* + \eta \boldsymbol{d} \in \Omega$. We have that

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

Since $\eta \geq 0$ and $d_2^2 > 0$, we require that $d_2 \in \mathbb{R} \backslash \{0\}$ and $d_1 \leq 0$. If we consider the case where $d_2 = 0$ (not considered in the above), we have that

$$\begin{aligned}
[2,0]^\top + \eta[d_1, 0]^\top = [2+\eta d_1, 0]^\top \in \Omega \quad &\Leftrightarrow \quad (2+\eta d_1) + 0^2 \leq 2 \\
&\Leftrightarrow \quad \eta d_1 \leq 0 \\
&\Leftrightarrow \quad d_1 \leq 0.
\end{aligned}$$

So $\boldsymbol{d} = [d_1 = 0, d_2 \in \mathbb{R}]^\top$ is feasible. Hence, the set of feasible directions is

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

Clearly, $\boldsymbol{d} \in \mathcal{D}^*$ implies that $\boldsymbol{d} \in \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \leq 0, d_2 \in \mathbb{R}\}$, so $\mathcal{D}^* \subset \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \leq 0, d_2 \in \mathbb{R}\}$. Let's show the other containment now. Let $\mathcal{A} = \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \leq 0, d_2 \in \mathbb{R}\}$. We have that $\boldsymbol{d} \in \mathcal{A}$ implies that $d_1 \leq 0$, $d_2 \in \mathbb{R}$. So it follows that for any $\eta \geq 0$,

$$[2,0]^\top + \eta[d_1, d_2]^\top = [2+\eta d_1, \eta d_2]^\top$$

and $(2+\eta d_1) + (\eta d_2)^2 \leq 2$, so $\boldsymbol{d} \in \mathcal{D}^*$ and hence $\mathcal{A} \subset \mathcal{D}^*$. Our two containments mean that $\mathcal{A} \subset \mathcal{D}^* \subset \mathcal{A}$ and thus that $\mathcal{D}^* = \mathcal{A}$, so $\mathcal{D}^* = \{\boldsymbol{d} \in \mathbb{R}^2 : d_1 \leq 0, d_2 \in \mathbb{R}\}$ is the set of feasible directions at $\boldsymbol{x}^*$. Now that we have established the set of feasible directions, we can check if the FONC is satisfied at $\boldsymbol{x}^*$. Consider $\boldsymbol{d} \in \mathcal{D}^*$. We have that

$$\nabla f(\boldsymbol{x}^*) = \begin{bmatrix} -3 \\ 0 \end{bmatrix},$$

which means that

$$[d_1, d_2] \begin{bmatrix} -3 \\ 0 \end{bmatrix} = -3d_1 \geq 0$$

as $d_1 \leq 0$. So the FONC is satisfied at $\boldsymbol{x}^*$.

(b) The SONC is satisfied at $\boldsymbol{x}^*$ if for any $\boldsymbol{d} \in \mathcal{D}^*$, $\boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}^*) \boldsymbol{d} \geq 0$. We compute the Hessian matrix to be

$$\boldsymbol{F}(\boldsymbol{x}) = \mathcal{J}(\nabla f(\boldsymbol{x})) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

so $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{F}(\boldsymbol{x})$. Hence, for any $\boldsymbol{d} \in \mathcal{D}^*$,

$$\boldsymbol{d}^\top \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{d} \geq 0,$$

so the SONC is satisfied at $\boldsymbol{x}^* = [2,0]^\top$.

(c) Yes, $\boldsymbol{x}^*$ is a local minimizer. This is because for any feasible point $\boldsymbol{x} = [x_1, x_2]^\top \neq \boldsymbol{x}^*$, we have that $x_1 < 2$. Thus, for each $\boldsymbol{x} \neq \boldsymbol{x}^*$,

$$f(\boldsymbol{x}) = -3x_1 > -6 = f(\boldsymbol{x}^*).$$

To be more precise, $\boldsymbol{x}^*$ is a strict local minimizer.

Problem 6.17

(a) We have that $f(\boldsymbol{x}) = \ln x_1 + \ln x_2$. Taking the gradient, we find that

$$f(\boldsymbol{x}) = \begin{bmatrix} 1/x_1 \\ 1/x_2 \end{bmatrix}$$

for $x_1, x_2 > 0$. Suppose that $\boldsymbol{x}^*$ is an optimal solution for this maximization problem. For $\boldsymbol{x}^*$ to be an interior point of Ω, we require that $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$. However, $1/x_1$ and $1/x_2$ cannot be 0, so it's not possible for $\boldsymbol{x}^*$ to be an interior point of Ω.

(b) We compute the Hessian matrix to be

$$\boldsymbol{F}(\boldsymbol{x}) = \mathcal{J}(f(\boldsymbol{x})) = \begin{bmatrix} -\frac{1}{x_1^2} & 0 \\ 0 & -\frac{1}{x_2^2} \end{bmatrix}$$

which is negative definite for every $\boldsymbol{x} = [x_1, x_2]^\top \in \Omega$. Since this is a maximization problem, the negative definiteness of the Hessian is analogous to the positive definiteness of the Hessian for a minimization problem. As such, by Corollary 6.2, since $\boldsymbol{F}(\boldsymbol{x}^*)$ is negative definite for every feasible $\boldsymbol{x} = [x_1, x_2]^\top \in \Omega$, the SONC is satisfied at every point $\boldsymbol{x} \in \Omega$.

Problem 6.26

Since $x_1, x_2 \geq 0$, $\boldsymbol{d} = [1, 1]^\top$ is clearly a feasible direction at $\mathbf{0}$. We see that

$$\boldsymbol{d}^\top \nabla f(\mathbf{0}) = \frac{\partial f}{\partial x_1}(\mathbf{0}) + \frac{\partial f}{\partial x_2}(\mathbf{0}).$$

Since $\nabla f(\mathbf{0}) \leq \mathbf{0}$ and $\nabla f(\mathbf{0}) \neq 0$, this means that

$$\boldsymbol{d}^\top \nabla f(\mathbf{0}) < 0.$$

This does not satisfy the FONC, so $\mathbf{0}$ cannot be a local minimizer.

Problem 6.28

For this linear programming question, the objective function is $f(\boldsymbol{x}) = -c_1 x_1 - c_2 x_2$. Hence, it follows that

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} -c_1 \\ -c_2 \end{bmatrix} \neq \mathbf{0}$$

for every $\boldsymbol{x}$ as $c_1 > c_2 \geq 0$. By the FONC, the optimal feasible solution $\boldsymbol{x}^*$ therefore cannot be an interior point of the feasible set. For each $\boldsymbol{x} \in L_1 \cup L_2$, $\boldsymbol{d} = [1, 1]^\top$ is a feasible direction, so we have that

$$\boldsymbol{d}^\top f(\boldsymbol{x}) = -c_1 - c_2 < 0.$$

So by the FONC, the optimal feasible solution is not contained in $L_1 \cup L_2$. Next, for each $\boldsymbol{x} \in L_3$, $\boldsymbol{d} = [1, -1]^\top$ is a feasible direction, so we have that

$$\boldsymbol{d}^\top f(\boldsymbol{x}) = c_2 - c_1 > 0.$$

Again, the FONC for the interior case is not satisfied, so the optimal feasible solution is not in L_3. By the process of elimination, the unique optimal feasible solution must be $[1, 0]^\top$.

Problem 6.32

Proof. Let Ω be a convex subset of $\mathbb{R}^n$, $f \in C^2$ be a real-valued function on Ω, and $\boldsymbol{x}^*$ be a point in Ω. Since $f \in C^2$, we know that $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{F}^\top(\boldsymbol{x}^*)$. Suppose that $\boldsymbol{d} \neq \mathbf{0}$ is a feasible direction at $\boldsymbol{x}^*$. Then, by Taylor's Theorem,

$$f(\boldsymbol{x}^* + \boldsymbol{d}) - f(\boldsymbol{x}^*) = \frac{1}{2}\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) + \boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} + o(\|\boldsymbol{d}\|^2).$$

By conditions 1 and 2, we have that

$$f(\boldsymbol{x}^* - \boldsymbol{d}) - f(\boldsymbol{x}^*) \geq c\|\boldsymbol{d}\|^2 + o(\|\boldsymbol{d}\|^2).$$

Hence, for any $\boldsymbol{d}$ that that $\|\boldsymbol{d}\|$ is sufficiently small, we have that

$$f(\boldsymbol{x}^* + \boldsymbol{d}) > f(\boldsymbol{x}^*),$$

and so $\boldsymbol{x}^*$ is a strict local minimizer for f. This completes the proof. $\square$