AMS230 – Homework #5

Zayd Hammoudeh

June 3, 2018

Name: Zayd Hammoudeh
Course Name: AMS230
Assignment Name: Homework #5
Due Date: June 13, 2018
Student Discussions: I discussed the problems with the following students. All write-ups were prepared separately and independently.

- Bernardo Torres

Problem #1
Exercise 11.2 in Nocedal and Wright.

Consider the function $r : \mathbb{R} \to \mathbb{R}$ defined by $r(x) = x^q$, where q is an integer greater than 2. Note that $x^* = 0$ is the sole root of this function and that it is degenerate. Show that Newton's method converges Q-linearly, and find the value of the convergence ratio r in (A.34)

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq r, \qquad \textbf{for all } k \textbf{ sufficiently large.}$$

The Jacobian for a system of non-linear equations equals:

$$J(x_k) = \left[\frac{\partial \mathbf{f}}{\partial x_1}, \dots, \frac{\partial \mathbf{f}}{\partial x_n},\right].$$

For this function, the Jacobian is simply $J = qx^{q-1}$. Therefore,

$$\begin{aligned} p_k &= -\frac{r(x_k)}{J(x_k)} \\ &= -\frac{x^q}{qx^{q-1}} \\ &= -\frac{1}{q}x \end{aligned}$$

Since $x^* = 0$ and $x \in \mathbb{R}$, this makes the ratio:

$$\begin{aligned} \frac{\|x_k + p_k\|}{\|x_k\|} &= \\ \frac{x_k + p_k}{x_k} &= \\ \frac{x_k - (1/q)x_k}{x_k} &= \\ \boxed{\frac{q-1}{q}} &\leq r. \end{aligned}$$

Problem #2

Exercise 12.5 in Nocedal and Wright.

Let $v : \mathbb{R}^n \to \mathbb{R}^m$ **be a smooth vector function, and consider the unconstrained optimization problems of minimizing** $f(x)$ **where**

$$f(x) = \|v(x)\|_\infty \qquad f(x) = \max_{i=1,2,\ldots,m} v_i(x).$$

Reformulate these (generally nonsmooth) problems as smooth constrained optimization problems.

(a) $f(x) = \|v(x)\|_\infty \coloneqq \max\{|v_1(x)|, \ldots, |v_n(x)|\}$

This problem is very similar to the example in Nocedal and Wright for (12.5) and (12.7). The constrained, smooth optimization version of the problem is:

$$\min f(x) = \min t \text{ s.t. } \exists x \forall i \big(t \geq v_i(x),\ t \geq -v_i(x)\big).$$

(b) $f(x) = \max_{i=1,2,\ldots,m} v_i(x)$

This is a simplified version of constrained optimization problem in part (a). The constrained, smooth optimization version of the problem is:

$$\min f(x) = \min t \text{ s.t. } \exists x \forall i \big(t \geq v_i(x)\big).$$

Problem #3

Exercise 12.14 in Nocedal and Wright (assume that a is not a zero vector).

Consider the half space defined by $H = \{x \in \mathbb{R}^n | a^{\mathrm{T}}x + \alpha \geq 0\}$ where $a \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ are given. Formulate and solve the optimization problem for finding the point x in H that has the smallest Euclidian norm.

This can be framed as the constrained optimization problem below.

$$\min f(x) = \sum_i x_i^2 \text{ s.t. } a^{\mathrm{T}}x + \alpha \geq 0.$$

If $\alpha \geq 0$, then $f(x)$ is minimized at $x^* = \vec{0}$ since $f(\vec{0}) = 0$. For all other $x \neq \vec{0}$, $f(x)$ is strictly positive. This also means $\lambda^* = 0$.

Consider the more interesting case when α is less than 0. First, the gradient of $f(x)$ equals:

$$\nabla f(x) = (2x_1, \ldots, 2x_n).$$

Similarly, the gradient of the constraint $c(x)$ is:

$$\nabla c(x) = a.$$

By the first KKT condition,

$$\begin{aligned}
\mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda^* \nabla c(x^*) &= 0 \\
2x^* - \lambda^* a &= 0 \\
x^* &= \frac{\lambda^*}{2} a.
\end{aligned}$$

By the fifth KKT condition,

$$\begin{aligned}
\lambda^* c(x^*) &= 0 \\
\lambda^* (a^{\mathrm{T}} x^* + \alpha) &= 0
\end{aligned}$$

$\lambda^* = 0$ is not a valid value if $\alpha < 0$. Instead continuing by the fourth KKT condition where $\lambda^* > 0$,

$$\begin{aligned}
a^{\mathrm{T}} x^* + \alpha &= 0 \\
a^{\mathrm{T}} \left(\frac{\lambda^*}{2} a \right) &= -\alpha \\
\lambda^* &= -\frac{2\alpha}{\|a\|^2}
\end{aligned}$$

This makes

$$\boxed{x^* = -\frac{\alpha}{\|a\|^2}a}.$$

Since α is negative, λ^* is positive satisfying the fourth KKT condition. It is also trivial to see the third and fifth KKT conditions hold for these values of x^* and λ^*. The second KKT condition does not apply since there are no equality constraints. LICQ also holds since $\nabla c = a \neq \vec{0}$.

Problem #4
Exercise 12.17 in Nocedal and Wright.

Prove that when the KKT conditions (12.34) and the LICQ are satisfied at a given point x^*, the Lagrange multiplier λ^* in (12.34) is unique.

By the first KKT condition, it holds that:

$$\nabla \mathcal{L}(x^* \lambda^*) = \nabla f - \sum_{I} \lambda_i^* \nabla c_i(x^*) = 0.$$

This can be rewritten as a matrix equation as shown below:

$$\begin{aligned} \sum_{i} \lambda_i^* \nabla c_i(x^*) &= \nabla f \\ [\nabla c_1(x^*) \ \cdots \ \nabla c_k(x^*)]\,[\lambda_1, \cdots, \lambda_k]^{\mathrm{T}} &= \nabla f. \end{aligned}$$

Since the constraint gradients are linearly independent by LICQ property, then matrix $[\nabla c_1(x^*) \ \cdots \ \nabla c_k(x^*)]$ is invertible. By the Invertible Matrix Theorem, the solution for $[\lambda_1, \cdots, \lambda_k]^{\mathrm{T}}$ is unique making all of the Lagrange multipliers unique.

Problem #5
Exercise 12.18 in Nocedal and Wright.

Consider the problem of finding the point on the parabola $y = \frac{1}{5}(x-1)^2$ **that is closest to** $(x, y) = (1, 2)$ **in the Euclidian norm sense. We can formulate this problem as:**

$$\min f(x, y) = (x-1)^2 + (y-2)^2 \qquad \textbf{subject to } (x-1)^2 = 5y.$$

The gradient of f is:

$$\nabla f = \begin{bmatrix} 2(x-1) \\ 2(y-2) \end{bmatrix}.$$

Rewrite $c_1(x, y) = (x-1)^2 - 5y = 0$. Its gradient is:

$$\nabla c_1 = \begin{bmatrix} 2(x-1) \\ -5 \end{bmatrix}.$$

(a) Find all the KKT points for this problem. Is the LICQ satisfied?

For a single constraint, Lagrangian multiplier λ_1 only holds when $\nabla f = \lambda_1 \nabla c_1$. In the case of the first dimension, $\nabla f = \lambda_1 \nabla c_1$. There are two possible cases:

Case #1: $\lambda = 1$. This means $2(y-2) = -5$ making $y = -1/2$. Clearly, this not valid so we can drop this case.

Case #2: $x = 1$. Trivially, this makes $y = 0$ and $\lambda_1 = 4/5$. Therefore, $(x^*, y^*) = (1, 0)$. Clearly then

$$\nabla \mathcal{L}\left(0, 1, \frac{4}{5}\right) = \nabla f(1, 0) - \frac{4}{5} \nabla c_1(1, 0) = \vec{0}$$

and $c_1(1, 0) = 0$ satisfying Eq. (12.34a) and Eq. (12.34b). Since there is only an equality constraint, we can ignore KKT inequality constraints Eq. (12.34c) and Eq. (12.34d). In addition, Eq. (12.34e) follows from proving $c_1(x^*) = 0$.

Since there is only a single constraint, LICQ is established by showing only that $\nabla c_1(x^*) = (0, -5) \neq \vec{0}$.

(b) Which of these points are solutions?

The point $(x^*, y^*) = (1, 0)$ is a solution.

(c) By directly substituting the constraint into the objective function and eliminating the variable x**, we obtain an unconstrained optimization problem. Show that the solutions of this problem cannot be solutions to the original problem.**

Substituting the constraint into the objective function yields:

$$f(y) = 5y + (y - 2)^2.$$

Taking the derivative and solving for the root yields:

$$\begin{aligned}
\frac{df}{dy} = 5 + 2(y-2) &= 0 \\
2(y-2) &= -5 \\
y &= 2 - 5/2 \\
y &= -1/2.
\end{aligned}$$

Clearly this is an invalid root since the parabola is strictly positive.