Homework 3

Math 467

Minyoung Heo

6.1

a)

We have to specify a feasible direction $d = \begin{pmatrix} a \\ b \end{pmatrix}$ such that $a > 0$ since x^* is on the boundary. When we plug this into $d^T \nabla f = a + b$, which may yield a negative value when $b < -a$. Therefore, this is **(ii) definitely not a local minimizer**.

b)

We specify a feasible direction $d = \begin{pmatrix} a \\ b \end{pmatrix}$ such that $a > 0$ or $b > 0$ since x^* is on the boundary. Pluggin this into $d^T \nabla f = a \geq 0$ where equality if achieved with $b > 0$ and $a = 0$. This seems like **(iii) possibly a local minimizer** because x^* is on the boundary and $d^T \nabla f$ is greater than 0 (or equal to zero on the boundary) for any feasible direction.

c)

Since x^* is an interior point, we can use any feasible direction $d = \begin{pmatrix} a \\ b \end{pmatrix}$. It's easy to see $d^T \nabla f = 0$ and $F(x^*) > 0$, so this is **(i) definitely a local minimizer**.

d)

Since x^* is on the boundary, $d = \begin{pmatrix} a \\ b \end{pmatrix}$ where $a > 0$ or $b > 0$. This is similar to part (b) but with an addition of the Hessian $F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. We can see that F is not posittive definite, so this is **(ii) definitely not a local minimizer**.

6.4

Since x^* is an interior point, we can make an $\epsilon > 0$ such that a ball around x^* $B_\epsilon(x^*) \subset \Omega$. We can then say that x^* is a local minimizer of $B_\epsilon(x^*)$. But since $\Omega \subset \Omega'$, then $B_\epsilon(x^*) \subset \Omega'$, meaning x^* is a local minimizer to Ω'.

We know this is not true if x^* is not an interior point. We can take a 1-dimensional example. For a function $f(x) = -x$, if we let $\Omega = [a, b]$, we can say that a is the minimum point. But for $\Omega' = [a-1, b]$, the minimum is at $a-1$.

6.9

a)

We can find this by calculating the gradient. The gradient finds the direction of the greatest increase, meaning we can negate it to find the greatest decrease.

$$\nabla f([2,1]^T) = \begin{bmatrix} 2x_1x_2 + x_2^3 \\ x_1^2 + 3x_2^2x_1 \end{bmatrix} = \begin{bmatrix} 4+1 \\ 4+6 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

Thus the direction must be $d = -\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

b)

The rate of increase can be found with $d^T \nabla f = -\frac{1}{\sqrt{5}} [1 \quad 2] \begin{bmatrix} 5 \\ 10 \end{bmatrix} = -5\sqrt{5}$.

c)

First we can normalize d, $d = \frac{1}{5} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Then we do $d^T \nabla f = \frac{1}{5} [3 \quad 4] \begin{bmatrix} 5 \\ 10 \end{bmatrix} = \frac{1}{5}(15 + 40) = 11$.

6.10

Before we proceed, for the sake of convenience, we can first do this:

$$\begin{aligned} f(x) = (x_1 \quad x_2) \begin{bmatrix} 2 & 5 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (x_1 \quad x_2) \begin{pmatrix} 3 \\ 4 \end{pmatrix} + 7 \\ = 2x_1^2 - x_2x_1 + 5x_1x_2 + x_2^2 + 3x_1 + 4x_2 + 7 \\ = 2x_1^2 + x_2^2 + 4x_1x_2 + 3x_1 + 4x_2 + 7 \end{aligned}$$

a)

$$\nabla f([0,1]^T) = \begin{bmatrix} 4x_1 + 4x_2 + 3 \\ 2x_2 + 4x_1 + 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$$

$$d^T \nabla f = (1 \quad 0) \begin{bmatrix} 7 \\ 6 \end{bmatrix} = 7$$

b)

First let's find all the critical points. Let $4x_1 + 4x_2 + 3 = 0$ and $2x_2 + 4x_1 + 4 = 0$. By negating the second and adding it to the first, we get $2x_2 - 1 = 0$, or $x_2 = \frac{1}{2}$ which gives us $x_1 = -\frac{5}{4}$. To see if any of these are minimizers, we look at the Hessian

$$F = \begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix} = 2 \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}$$

When computing for the eigenvalues, we get $p(\lambda) = \lambda^2 - 3\lambda - 2 = 0$. By completing the square

$$\begin{aligned} \lambda^2 - 3\lambda - 2 &= 0 \\ \to \lambda^2 - 3\lambda + \left(\frac{3}{2}\right)^2 &= 2 + \left(\frac{3}{2}\right)^2 \\ \to \lambda_{1,2} &= \frac{3}{2} \pm \frac{\sqrt{17}}{2} \end{aligned}$$

which means F is not definite. Therefore, there is no minimizer.

6.11

a)

$$\nabla f = \begin{pmatrix} 0 \\ -2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Yes, because $\nabla f = 0$.

b)

It is a local maximizer. The graph of $-x_2$ is a parabola upside down. Any other x_2 value will be less than 0. Any x_1 values will just be the same.

6.13

a)

$$\nabla f = \begin{pmatrix} -3 \\ 0 \end{pmatrix}$$

Since x^* is on the boundary, $d = \begin{pmatrix} a \\ b \end{pmatrix}$ such that $a < 0$ and $|b| < \sqrt{|a|}$ (i.e. we can only decrease a and we can increase b as long as we don't do it as much).

$$d^T \nabla f = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} -3 \\ 0 \end{pmatrix} = -3a > 0$$

Yes, it does satisfy the first order necessary condition.

b)

The Hessian:

$$F = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

Yes, it does satisfy the second order necessary condition.

c)

Yes. If we consider the graph $-3x_1$, the *boundary point* $x^* = [2, 0]^T$ can only go uphill. Those are the only feasible directions.

6.29

a)

First I show that

$$f(a, b) = \frac{1}{n} \sum_{i=1}^{n} (ax_i + b - y_i)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} a^2 x_i^2 + ax_i b - ax_i y_i + bax_i + b^2 - by_i - ax_i y_i - by_i + y_i^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} a^2 x_i^2 + 2abx_i + b^2 - 2ax_i y_i - 2by_i + y_i^2$$

Keep note of this until the very end. Now we let

$$
\begin{aligned}
Q &= \begin{pmatrix} \overline{X^2} & \overline{X} \\ \overline{X} & 1 \end{pmatrix} \\
c &= \begin{pmatrix} \overline{XY} \\ \overline{Y} \end{pmatrix} \\
d &= \overline{Y^2}
\end{aligned}
$$

This means,

$$
\begin{aligned}
f(a,b) &= z^T Q z - 2c^T z + d \\
&= \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} \overline{X^2} & \overline{X} \\ \overline{X} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} - 2 \begin{pmatrix} \overline{XY} \\ \overline{Y} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \overline{Y^2} \\
&= a^2\overline{X^2} + 2ab\overline{X} + b^2 - 2a\overline{XY} - 2b\overline{Y} + \overline{Y^2} \\
&= a^2\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) + 2ab\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + b^2 - 2a\left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right) - 2b\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) + \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right)
\end{aligned}
$$

We can bring out the $\left(\frac{1}{n}\sum_{i=1}^{n}\right)$, and get

$$
\rightarrow \frac{1}{n}\sum_{i=1}^{n} a^2 x_i^2 + 2abx_i + b^2 - 2ax_i y_i - 2by_i + y_i^2
$$

which is the equation at the top.

b)

$$
f(a,b) = a^2\overline{X^2} + 2ab\overline{X} + b^2 - 2a\overline{XY} - 2b\overline{Y} + \overline{Y^2}
$$

$$
\nabla f = \begin{pmatrix} 2a\overline{X^2} + 2b\overline{X} - 2\overline{XY} \\ 2b + 2a\overline{X} - 2\overline{Y} \end{pmatrix} = 0
$$

This yields

$$
\begin{aligned}
2a\overline{X^2} + 2b\overline{X} - 2\overline{XY} &= 0 \\
2b + 2a\overline{X} - 2\overline{Y} &= 0
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
a\overline{X^2} + b\overline{X} - \overline{XY} &= 0 \\
b + a\overline{X} - \overline{Y} &= 0
\end{aligned}
$$

which we can derive from the second equation:

$$b = \overline{Y} - a\overline{X}$$

(*which we'll hold onto for part (c)*).

plugging this into the other equation,

$$a\overline{X^2} + \overline{YX} - a\overline{X}^2 - \overline{XY} = 0$$
$$\rightarrow a(\overline{X^2} - \overline{X}^2) = \overline{XY} - \overline{XY}$$
$$\rightarrow a = \frac{\overline{XY} - \overline{XY}}{\overline{X^2} - \overline{X}^2}$$

Using this for b, and we get

$$b = \overline{Y} - \overline{X}\frac{\overline{XY} - \overline{XY}}{\overline{X^2} - \overline{X}^2}$$

We know that this critical point is the local minimima because $f(a,b)$ has a shape like a paraboloid going upward. There is only one critical point for a graph this shape, and it's at this point (a,b) shown above.

c)

Turns out we already derived this in part a (*Look for the note about part (c) in part (b)*)