Problem Set 5

Natasha Watkins

Exercise 6.6

$$Df(x, y) = \begin{bmatrix} 6xy + 4y^2 + y \\ 3x^2 + 8xy + x \end{bmatrix}$$

The critical points, where $Df(x, y) = \mathbf{0}$, are $(0, 0)$, $(-\frac{1}{9}, -\frac{1}{12})$, $(0, -\frac{1}{4})$ and $(-\frac{1}{3}, 0)$.

$$D^2f(x, y) = \begin{bmatrix} 6y & 6x + 8y + 1 \\ 6x + 8y + 1 & 8x \end{bmatrix}$$

Evaluating the second derivative at $(0, 0)$, we see

$$D^2f(x, y)\big|_{(0,0)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$\lambda_1 = -1$ and $\lambda_2 = 1$, so $(0, 0)$ is a saddle point.

Evaluating the second derivative at $(-\frac{1}{9}, -\frac{1}{12})$, we see

$$D^2f(x, y)\big|_{(\frac{1}{9}, -\frac{1}{12})} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{8}{9} \end{bmatrix}$$

This is a negative definite matrix, so $(-\frac{1}{9}, -\frac{1}{12})$ is a maximizer.

Evaluating the second derivative at $(0, -\frac{1}{4})$, we see

$$D^2f(x, y)\big|_{(0, -\frac{1}{4})} = \begin{bmatrix} -\frac{6}{4} & -1 \\ -1 & 0 \end{bmatrix}$$

This matrix has mixed sign eigenvalues, so $(0, -\frac{1}{4})$ is a saddle point.

Evaluating the second derivative at $(-\frac{1}{3}, 0)$, we see

$$D^2f(x, y)\big|_{(-\frac{1}{3}, 0)} = \begin{bmatrix} 0 & -1 \\ -1 & -\frac{8}{3} \end{bmatrix}$$

This matrix has mixed sign eigenvalues, so $(-\frac{1}{3}, 0)$ is a saddle point.

Exercise 6.11

Using Newton's method,

$$\begin{aligned}
x_1 &= x_0 - \frac{2ax_0 + b}{2a} \\
&= \frac{2ax_0 - 2ax_0 - b}{2a} \\
&= -\frac{b}{2a}
\end{aligned}$$

Substituting this value into $f'(x)$, we find $f'(-\frac{b}{2a}) = 0$, confirming that this is a critical point of f. $f''(-\frac{b}{2a}) > 0$, so $-\frac{b}{2a}$ is a minimizer of f. We know this is the unique minimizer of f as $f'(x) = 0$ where $x = -\frac{b}{2a}$.

Exercise 7.1

Take two points, $\mathbf{x}, \mathbf{y} \in conv(S)$ such that $\mathbf{x} = \delta_1\mathbf{x}_1 + \cdots \delta_k\mathbf{x}_k$ and $\mathbf{y} = \gamma_1\mathbf{x}_1 + \cdots \gamma_k\mathbf{x}_k$, with $\sum_{i=1}^k \delta_i = 1$ and $\sum_{i=1}^k \gamma_i = 1$. We can write their convex combination as $\lambda\mathbf{x} + (1+\lambda)\mathbf{y}$.

$$\begin{aligned}
\lambda\mathbf{x} + (1+\lambda)\mathbf{y} &= \lambda(\delta_1\mathbf{x}_1 + \cdots \delta_k\mathbf{x}_k) + (1-\lambda)(\gamma_1\mathbf{x}_1 + \cdots \gamma_k\mathbf{x}_k) \\
&= (\lambda\delta_1 + (1-\lambda)\gamma_1)\mathbf{x}_1 + \cdots + (\lambda\delta_k + (1-\lambda)\gamma_k)\mathbf{x}_k
\end{aligned}$$

We know that the convex combination of $\mathbf{x}$ and $\mathbf{y}$ are contained in $conv(S)$ because

$$\begin{aligned}
\sum_{i=1}^k(\lambda\delta_i + (1-\lambda)\gamma_i) &= \lambda\sum_{i=1}^k \delta_i + (1-\lambda)\sum_{i=1}^k \gamma_i \\
&= \lambda + (1-\lambda) = 1
\end{aligned}$$

which satisfies the definition of a convex set.

Exercise 7.2

<u>Part i)</u>

Take two points, $\mathbf{x}, \mathbf{y}$, on the hyperplane described by $\langle \mathbf{a}, \mathbf{x}\rangle = b$, such that $a_1x_1 + \cdots + a_nx_n =$

b and $a_1y_1 + \cdots + a_ny_n = b$. Taking their convex combination,

$$\begin{aligned}\lambda\mathbf{x} + (1-\lambda)\mathbf{y} &= \lambda(a_1x_1 + \cdots + a_nx_n) + (1-\lambda)a_1y_1 + \cdots + a_ny_n \\ &= \lambda b + (1-\lambda)b = b\end{aligned}$$

Therefore, the convex combination also lies on the hyperplane $\implies$ the hyperplane is a convex set.

<u>Part ii)</u>

Take two points, $\mathbf{x}, \mathbf{y}$, in the half space described by $\langle \mathbf{a}, \mathbf{x} \rangle \leq b$, such that $a_1x_1 + \cdots + a_nx_n \leq b$ and $a_1y_1 + \cdots + a_ny_n \leq b$. Taking their convex combination,

$$\begin{aligned}\lambda\mathbf{x} + (1-\lambda)\mathbf{y} &= \lambda(a_1x_1 + \cdots + a_nx_n) + (1-\lambda)a_1y_1 + \cdots + a_ny_n \\ &\leq \lambda b + (1-\lambda)b = b\end{aligned}$$

Therefore, the convex combination also lies in the half space $\implies$ the half space is convex.

Exercise 7.4

<u>Part i)</u>

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|^2 &= \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle \\ &= \langle \mathbf{x} - \mathbf{p} + \mathbf{p} - \mathbf{y}, \mathbf{x} - \mathbf{p} + \mathbf{p} - \mathbf{y} \rangle\end{aligned}$$

Let $\mathbf{z} = \mathbf{x} - \mathbf{p}$ and $\mathbf{k} = \mathbf{p} - \mathbf{y}$.

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|^2 &= \langle \mathbf{z} + \mathbf{k}, \mathbf{z} + \mathbf{k} \rangle \\ &= \langle \mathbf{z}, \mathbf{z} \rangle + 2\langle \mathbf{z}, \mathbf{k} \rangle + \langle \mathbf{k}, \mathbf{k} \rangle \\ &= \langle \mathbf{x} - \mathbf{p}, \mathbf{x} - \mathbf{p} \rangle + \langle \mathbf{p} - \mathbf{y}, \mathbf{p} - \mathbf{y} \rangle + 2\langle \mathbf{x} - \mathbf{p}, \mathbf{p} - \mathbf{y} \rangle \\ &= \|\mathbf{x} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{y}\|^2 + 2\langle \mathbf{x} - \mathbf{p}, \mathbf{p} - \mathbf{y} \rangle\end{aligned}$$

<u>Part ii)</u>

If 7.14 holds, we know

$$
\begin{aligned}
&\|\mathbf{x}-\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{p}\|^2 - \|\mathbf{p}-\mathbf{y}\|^2 \geq 0 \\
\implies &\|\mathbf{x}-\mathbf{y}\|^2 - \|\mathbf{p}-\mathbf{y}\|^2 \geq \|\mathbf{x}-\mathbf{p}\|^2 \\
\implies &\|\mathbf{x}-\mathbf{y}\| > \|\mathbf{x}-\mathbf{p}\| \qquad \text{as } \mathbf{y} \neq \mathbf{p}
\end{aligned}
$$

Part iii)

$$
\begin{aligned}
\|\mathbf{x}-\mathbf{z}\|^2 &= \langle \mathbf{x}-\mathbf{p}+\lambda\mathbf{p}-\lambda\mathbf{y}, \mathbf{x}-\mathbf{p}+\lambda\mathbf{p}-\lambda\mathbf{y} \rangle \\
&= \|\mathbf{x}-\mathbf{p}\|^2 + \lambda^2\|\mathbf{p}-\mathbf{y}\|^2 + 2\lambda\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle
\end{aligned}
$$

Exercise 7.13

Suppose f is not constant, ie. $f(\mathbf{x}) > f(\mathbf{y})$ for some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then,

$$
\begin{aligned}
f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) &\leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \\
f(\mathbf{z}) &\leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \\
\frac{f(\mathbf{z}) - (1-\lambda)f(\mathbf{y})}{\lambda} &\leq f(\mathbf{x})
\end{aligned}
$$

Exercise 7.20

f is convex so $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})$. $-f$ is also convex so

$$
\begin{aligned}
&-f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \leq -\lambda f(\mathbf{x}) - (1-\lambda)f(\mathbf{y}) \\
\implies &f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \geq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \\
\implies &f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) = \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})
\end{aligned}
$$

Therefore f is an affine function (this is the definition of an affine function with $L(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$).

Exercise 7.21

$\mathbf{x}^*$ is a minimizer for the function $f(\mathbf{x})$. Suppose $\mathbf{x}^*$ is not the minimizer for $\phi \circ f(\mathbf{x})$, ie. $\phi \circ f(\mathbf{x}) \leq \phi \circ f(\mathbf{x}^*)$ for some $\mathbf{x}$. As ϕ is strictly increasing, this implies $f(\mathbf{x}) \leq f(\mathbf{x}^*)$, which contradicts the fact that $\mathbf{x}^*$ is the minimizer of f.

$\mathbf{x}^*$ is a minimizer for the function $\phi \circ f(\mathbf{x})$. Suppose $\mathbf{x}^*$ is not the minimizer for $f(\mathbf{x})$, ie. $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ for some $\mathbf{x}$. Taking $\phi \circ f(\mathbf{x})$, as ϕ is strictly increasing, the previous inequality implies $\phi \circ f(\mathbf{x}) \leq \phi \circ f(\mathbf{x}^*)$, which contradict the fact that $\mathbf{x}^*$ is the minimizer of $\phi \circ f(\mathbf{x})$.