Math 164 Homework 4

Sava Spasojevic

October 30, 2018

Exercise 1. Let V be a vector space and consider $f : V \longrightarrow \mathbb{R}$. We define the epigraph of f as

$$\text{epi } f = \{(x, \xi) \in V \times \mathbb{R} : f(x) \leq \xi\}$$

Note that epi $f \subseteq V \times \mathbb{R}$. Show that f is conves if and only if epi f is convex.

Proof. To show that epi f is a convex set, we need to show that $\forall x, y \in$ epi f, $(1-\lambda)x + \lambda y \in$ epi f, where $\lambda \in [0,1]$. Let $x, y \in$ epi f. Let $\lambda \in [0,1]$. Then, $f(x) \leq \xi$ and $f(y) \leq \xi$, or

$$(1-\lambda)f(x) \leq (1-\lambda)\xi \tag{1}$$
$$\lambda f(y) \leq \lambda \xi \tag{2}$$

Adding (1) snd (2) gives

$$(1-\lambda)f(x) + \lambda f(y) \leq \xi$$

But, since f is convex we have that,

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \leq \xi$$

Thus, $(1-\lambda)x + \lambda y \in$ epi f. □

Exercise 2. A map $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is said to be monotone if

$$\langle A(y) - A(x), y - x \rangle \geq 0$$

$\forall x, y \in \mathbb{R}^n$. Here $\langle \cdot, \cdot \rangle$ denotes the canonical inner product on $\mathbb{R}^n : \langle u, v \rangle = \sum_{i=1}^n u_i v_i$.

1. Show that if a convex function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable then its gradient ∇f is monotone.
2. Assume that $n = 1$. Show that the notion of monotonicity is a actually a concept that you already know.

Proof. 1. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a convex, differentiable function. Let $x, y \in \mathbb{R}^n$. Then,

$$f(y) \geq f(x) + \nabla f(x)(y - x) \tag{3}$$
$$f(x) \geq f(y) + \nabla f(y)(x - y) \tag{4}$$

Adding (3) and (4) gives

$$f(y) + f(x) \geq f(x) + \nabla f(x)(y-x) + f(y) + \nabla f(y)(x-y)$$
$$\nabla f(y)(y-x) - \nabla f(x)(y-x) \geq 0$$

But, $\nabla f(y)(y-x) - \nabla f(x)(y-x) = \langle \nabla f(y) - \nabla f(x), y - x \rangle$. So,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$$

□

Proof. 2. Let $n = 1$. Then the gradient of f is just a single derivative. We have that

$$\langle f'(y) - f'(x), y - x \rangle \geq 0$$
$$(f'(y) - f'(x))(y - x) \geq 0$$

This implies two cases;

<u>case I:</u> $f'(y) - f'(x) \geq 0$ and $y - x \geq 0$. In this case, the function is the familiar monotone increasing function.

<u>case II:</u> $f'(y) - f'(x) \leq 0$ and $y - x \leq 0$. In this case, the function is the familiar monotone decreasing function. Thus, monotonicity is something we have already seen before. □

Exercise 3. Use Newton's method to devise a method of approximating $\ln(2)$. Use the initial point $x^{(0)} = 1$, and perform two iterations. Hint: Consider the function $x \mapsto e^x - 2$.

Let $f(x) = e^x - 2$. Performing two iterations,

$$
\begin{aligned}
x^{(k+1)} &= x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})} \\
x^{(1)} &= x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})} = 1 - \frac{e^1 - 2}{e^1} \approx 0.735789 \\
x^{(2)} &= x^{(1)} - \frac{f(x^{(1)})}{f'(x^{(1)})} = 0.735789 - \frac{e^{0.735789} - 2}{e^{0.735789}} \approx 0.694044
\end{aligned}
$$

Note $\ln(2) \approx 0.693147$. So this method of approximation is working well.

Exercise 4. Let $[x_1, y_1]^T, \ldots, [x_n, y_n]^T, n \geq 2$, be points on the $\mathbb{R}^2$ place. We wish to find

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

(a) Show that $f(a,b)$ can be written in the form $z^TQz - 2c^Tz + d$ where $z = [a,b]^T, Q = Q^T, c \in \mathbb{R}^2$ and $d \in \mathbb{R}$, and find expressions for $Q, c,$ and d in terms of $\overline{X}, \overline{Y}, \overline{X^2}, \overline{Y^2}$, and $\overline{XY}$.

$$
\begin{aligned}
f(a,b) &= \frac{1}{n}\sum_{i=1}^{n}(ax_i + b - y_i)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} a^2x_i^2 + 2abx_i - 2ax_iy_i + b^2 - 2by_i + y_i^2 \\
&= a^2\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right) + b^2 + 2ab\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right) - 2\left(\frac{1}{n}\sum_{i=1}^{n}x_iy_i\right) - 2b\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right) + \left(\frac{1}{n}\sum_{i=1}^{n}y_i^2\right) \\
&= [a\ b]\begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}x_i^2 & \frac{1}{n}\sum_{i=1}^{n}x_i \\ \frac{1}{n}\sum_{i=1}^{n}x_i & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} - 2\left[\frac{1}{n}\sum_{i=1}^{n}x_iy_i \ \ \frac{1}{n}\sum_{i=1}^{n}y_i\right]\begin{bmatrix} a \\ b \end{bmatrix} + \frac{1}{n}\sum_{i=1}^{n}y_i^2 \\
&= z^TQz - 2c^Tz + d
\end{aligned}
$$

(b) Assume that $x_i, i \in [1,n]$ are not all equal. Find the parameters a^* and b^* for the line of best fit in terms of in terms of $\overline{X}, \overline{Y}, \overline{X^2}, \overline{Y^2}$, and $\overline{XY}$. Show that the point $[a^*\ b^*]^T$ is the only local minimizer of f. Hint: $\overline{X^2} - (\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2$.

Suppose that $[a^*\ b^*]^T$ is a solution. Then, the FONC says that $\nabla f([a^*\ b^*]^T) = 2Q[a^*\ b^*]^T - 2c = 0$ or that $Q[a^*\ b^*]^T = c$. So $\det Q = \overline{X^2} - (\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2$. But, not all x_i are zero, so $\det Q \neq 0$. Thus,

$$[a^*\ b^*]^T = Q^{-1}c = \frac{1}{\overline{X^2} - (\overline{X})^2}\begin{bmatrix} 1 & -\overline{X} \\ -\overline{X} & \overline{X^2} \end{bmatrix}\begin{bmatrix} \overline{XY} \\ \overline{Y} \end{bmatrix}$$

$Q > 0$ so the SOSC guarantees that $[a^*\ b^*]^T$ is a local minimizer. Since it is also the only point satisfying the FONC, it is the only local minimizer of f.

(c) Show that if a^* and b^* are the parameters of the line of best fit, then $\overline{Y} = a^*\overline{X} + b^*$.

$$a^*\overline{X} + b^* = \overline{X}\frac{\overline{XY} - (\overline{X})(\overline{Y})}{\overline{X^2} - (\overline{X})^2} + \frac{(\overline{X^2})(\overline{Y}) - (\overline{X})(\overline{XY})}{\overline{X^2} - (\overline{X})^2} = \overline{Y}$$