I For this exercise we restrict ourselves to one dimensional functions, \( d = 1 \). Prove the equivalence of the two definitions of convexity shown in class. That is, we defined that \( f : K \mapsto \mathbb{R}^d \) is convex if and only if \( f((1-\alpha)x+\alpha y) \leq (1-\alpha)f(x)+\alpha f(y) \) for all \( x, y \in K \) and \( \alpha \in [0, 1] \). Show that \( f \) (assuming it is differentiable) is convex if and only if

\[
f(x) \geq f(y) + f'(y)(x-y)
\]

**Proof:** We wish to prove that convexity \( \iff \forall x, y f(x) \geq f(y) + f'(y)(x-y) \)

\[ \iff \text{Consider points } x, y \in K, x > y \text{ and } \alpha \in (0, 1), \text{ we know } f((1-\alpha)y+\alpha x) \leq (1-\alpha)f(y)+\alpha f(x). \text{ Rearranging terms, } f(y+\alpha(x-y)) \leq f(y) + \alpha(f(x) - f(y)) \]

Rearranging again and using \( \alpha > 0 \) and \( x > y \),

\[
\frac{f(y + \alpha(x-y)) - f(y)}{\alpha(x-y)} \leq \frac{f(x) - f(y)}{x-y}
\]

Note that \( \lim_{h \to 0} \frac{f(y+h)-f(y)}{h} = f'(y) \), so \( \lim_{\alpha \to 0} \frac{f(y+\alpha(x-y))-f(y)}{\alpha(x-y)} = f'(y) \). Setting \( \alpha \to 0 \) in the above equation and using the fact that if every element of a sequence is lower bounded by a quantity, then the limit of the sequence is also lower bounded by that quantity, we get that \( \frac{f(x)-f(y)}{x-y} \geq f'(y) \). A similar proof works for the \( x < y \) case. Since \( x \) and \( y \) were arbitrarily chosen, this true for \( x, y \in K \)

\[ \iff \text{Consider } x, y \in K, \alpha \in [0, 1]. \text{ Define } z_\alpha = \alpha x + (1-\alpha)y. \text{ We know } f(a) \geq f(b) + f'(b)(a-b), \forall a, b \in K. \text{ We use this inequality for } a = x, b = z_\alpha \text{ and } a = y, b = z_\alpha. \text{ We get the following two inequalities}
\]

\[
f(x) \geq f(z_\alpha) + f'(z_\alpha)(x-z_\alpha)
\]

\[
f(y) \geq f(z_\alpha) + f'(z_\alpha)(y-z_\alpha)
\]
Multiplying the first equation by $\alpha$ and second by $(1 - \alpha)$ and adding the two, we get

$$\alpha f(x) + (1 - \alpha) f(y) \geq \alpha f(z_\alpha) + (1 - \alpha) f(z_\alpha) + f'(z_\alpha)(\alpha x + (1 - \alpha) y - z_\alpha)$$

which gives us $f(z_\alpha) = f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y)$ as desired.

II Prove:

(a) The sum of convex functions is convex.

**Proof:** Let $f$ and $g$ be two convex functions and let $h = f + g$. We know that

$$\forall x, y, \alpha \in [0, 1], f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y) \text{ and } g(\alpha x + (1 - \alpha) y) \leq \alpha g(x) + (1 - \alpha) g(y).$$

Adding these equations we get

$$f(\alpha x + (1 - \alpha) y) + g(\alpha x + (1 - \alpha) y) \leq \alpha(f(x) + g(x)) + (1 - \alpha)(f(y) + g(y))$$

$$h(\alpha x + (1 - \alpha) y) \leq ah(x) + (1 - a)h(y)$$

So $h = f + g$ is also convex

(b) Let $f$ be $\alpha_1$-strongly convex and $g$ be $\alpha_2$-strongly convex. Then $f + g$ is $(\alpha_1 + \alpha_2)$-strongly convex.

**Proof:** We first prove a simple statement about psd matrices which we will use in parts (b) and (c).

**Lemma:** Sum of psd matrices is a psd matrix. **Proof:** If $A \succeq 0$ and $B \succeq 0$, $x^T Ax \geq 0$ and $x^T Bx \geq 0$, $\forall x$. If $C = A + B$, the $x^T Cx = x^T (A + B)x = x^T Ax + x^T Bx \geq 0$. So $C$ is also psd.

Since, $f$ is $\alpha_1$-strongly convex, $\nabla^2 f(x) \succeq \alpha_1 I \implies \nabla^2 f(x) - \alpha_1 I \succeq 0, \forall x.$

Similarly $\nabla^2 g(x) - \alpha_2 I \succeq 0.$ Using the above lemma, we conclude that

$$\nabla^2 f(x) - \alpha_1 I + \nabla^2 g(x) - \alpha_2 I \succeq 0$$

So

$$\nabla^2(f + g)(x) - (\alpha_1 + \alpha_2) I \succeq 0, \forall x$$
So \( f + g \) is \( \alpha_1 + \alpha_2 \)-strongly convex.

(c) Let \( f \) be \( \beta_1 \)-smooth and \( g \) be \( \beta_2 \)-smooth. Then \( f + g \) is \( (\beta_1 + \beta_2) \)-smooth.

**Proof:** The proof is very similar to part (b). Using the definition of smooth functions, we know \( \beta_1 I - \nabla^2 f(x) \succeq 0 \) and \( \beta_2 I - \nabla^2 g(x) \succeq 0 \) for every \( x \). Adding these equations, we get \( (\beta_1 + \beta_2) I - \nabla^2 (f + g)(x) \succeq 0, \forall x \). So \( f + g \) is \( \beta_1 + \beta_2 \)-smooth function.