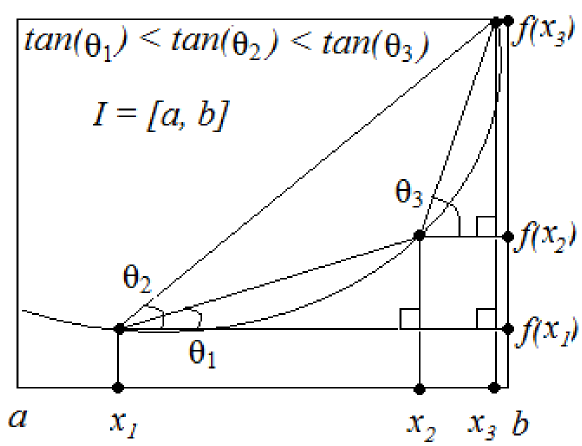
Math 650, Foundations of Optimization, Spring 2010

Sandipan Dey,
Homework Assignment - 3

April 4, 2010

Problem 9 Solution

(a)



$x_1 < x_2 < x_3$
$\Rightarrow \exists\lambda \in (0,1) \mid x_2 = (1-\lambda)x_1 + \lambda x_3$.
$\Rightarrow x_2 - x_1 = \lambda(x_3 - x_1)$ and $x_3 - x_2 = (1-\lambda)(x_3 - x_1) \ldots\ldots (1)$.

Also, f is convex
$\Rightarrow f(x_2) = f((1-\lambda)x_1 + \lambda x_3) \leq (1-\lambda)f(x_1) + \lambda f(x_3)$, by Jensen's inequality.
$\Rightarrow f(x_2) - f(x_1) \leq \lambda\,(f(x_3) - f(x_1))$ and $f(x_3) - f(x_2) \geq (1-\lambda)\,(f(x_3) - f(x_1))$.
$\Rightarrow \frac{f(x_2)-f(x_1)}{\lambda} \leq f(x_3) - f(x_1) \leq \frac{f(x_3)-f(x_2)}{1-\lambda}$, since $\lambda \neq 0 \wedge \lambda \neq 1$.

$\Rightarrow \frac{f(x_2)-f(x_1)}{\lambda(x_3-x_1)} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{(1-\lambda)(x_3-x_1)}$, since $x_3 \neq x_1$, dividing by $x_3 - x_1$.
$\Rightarrow \frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$, from (1).

When f is strictly convex, $\leq$ in Jensen's inequality can be replaced by strict inequaltiy, hence the above inequalities (that come from Jensen's) will be strict.

(b)

Let $g(t) = \frac{f(x+t)-f(x)}{t}$, with $x \in (a,b),\ t > 0$.
Let $t_1 > t_2 > 0,\ x_1 = x,\ x_2 = x + t_2,\ x_3 = x + t_1 \Rightarrow x_1 < x_2 < x_3$
$\Rightarrow \frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \Rightarrow \frac{f(x+t_2)-f(x)}{t_2} \leq \frac{f(x+t_1)-f(x)}{t_1}$, from (a)
Hence $t_1 > t_2 \Rightarrow g(t_1) \geq g(t_2) \Rightarrow g$ is increasing.

Now, let's consider the decreasing sequence $\{t_k\}_{k\geq 1}$, i.e., $t_1 > t_2 > \ldots \geq 0$. From the above result, $\{g(t_k)\}_{k\geq 1}$ i.e., $g(t_1) \geq g(t_2) \geq \ldots$, is a monotonically decreasing sequence too. Now, any decreasing sequence must have a limit and the limit is the greatest-lower bound (infimum) of the numbers. Hence, the sequence must have a limit, which is $\underset{k}{inf}\{g(t_k)\}_{k\geq 1}$.

Hence, $f\prime_+(x) = \underset{t\downarrow 0}{lim}\frac{f(x+t)-f(x)}{t} = \underset{t\downarrow 0}{inf}\frac{f(x+t)-f(x)}{t}$. Now, if x is an interior point, $f(x)$ is finite and when t si arbitrarily small, we have $f(x+t)$ approaching $f(x)$, which is finite, hence the limit is finite too.

Similarly consider the increasing sequence $t_1 < t_2 < \ldots \leq 0$, with $g(t_1) \leq g(t_2)\ldots$, another increasing sequence (by (a)), hence have the following limit $f\prime_-(x) = \underset{t\uparrow 0}{lim}\frac{f(x+t)-f(x)}{t} = \underset{t\uparrow 0}{sup}\frac{f(x+t)-f(x)}{t}$.

(c)

Let $t > 0,\ x_1 = x - t,\ x_2 = x,\ x_3 = x + t \Rightarrow x_1 < x_2 < x_3$
$\Rightarrow \frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \Rightarrow \frac{f(x)-f(x-t)}{t} \leq \frac{f(x+t)-f(x)}{t}$, from (a)
$\Rightarrow \underset{t\to 0}{lim}\frac{f(x)-f(x-t)}{t} \leq \underset{t\to 0}{lim}\frac{f(x+t)-f(x)}{t}$
$\Rightarrow f\prime_-(x) \leq f\prime_+(x)$

(d)

When x is not an interior point in I, still the limits exists, but $f(x)$ can be $\pm\infty$, hence the limits can be infinite.

For instance, consider the function $f(x) = -ln(x)$ which is convex and defined in the interval $(0,\infty)$, with $g(t) = \frac{-ln(x+t)+ln(x)}{t} \Rightarrow \underset{t\to 0}{lim} - \frac{ln(1+\frac{t}{x})}{\frac{t}{x}}.\frac{1}{x} = \frac{1}{x}$.
At $x = 0$, the limits exist, but equals ∞.

(e)

Consider a monotonically decreasing sequence t_n of positive numbers tending to 0. Since f is convex, from the part $(a)-(c)$, the sequence $\frac{f(x+t_n d)-f(x)}{t_n}$ is monotonically decreasing, so has a limit $\Rightarrow f\prime(x;d) = \underset{t\geq 0}{inf} \frac{f(x+td)-f(x)}{t}$ exists and it's finite if $f(x)$ is finite, i.e., x is an interior point.

Problem 11 Solution

(i) Proof $(b) \Rightarrow (c)$:

$$\begin{aligned} f(y) &\geq f(x) + \langle \nabla f(x), y-x \rangle \,, \ \forall x,y \in C & (1)\\ \Rightarrow f(x) &\geq f(y) + \langle \nabla f(y), x-y \rangle \,, \ \forall x,y \in C & (2)\\ (1)+(2) \Rightarrow 0 &\geq \langle \nabla f(x), y-x \rangle + \langle \nabla f(y), x-y \rangle \,, \ \forall x,y \in C & \\ \Rightarrow \langle \nabla f(y), y-x \rangle &- \langle \nabla f(x), y-x \rangle \geq 0, \ \forall x,y \in C & \\ \Rightarrow \langle \nabla f(y) &- \nabla f(x), y-x \rangle \geq 0, \ \forall x,y \in C & (3) \end{aligned}$$

(ii) Proof $(c) \Rightarrow (b)$:

Given

$$\langle \nabla f(y) - \nabla f(x), y-x \rangle \geq 0, \forall x,y \in C \tag{4}$$

Let's define the function $g(t) = f(x+t(y-x)) - t\langle \nabla f(x), y-x \rangle$, for any $x,y \in C, \ t \in \Re$.

$$\begin{aligned} &\Rightarrow g\prime(t) = Df(x+t(y-x)).(y-x) - \langle \nabla f(x), y-x \rangle \,, \text{ by chain rule}\\ &\Rightarrow g\prime(t) = \langle \nabla f(x+t(y-x)), y-x \rangle - \langle \nabla f(x), y-x \rangle\\ &\Rightarrow g\prime(t) = \langle \nabla f(x+t(y-x)) - \nabla f(x), y-x \rangle\\ \text{Let } z = x+t(y-x) \Rightarrow g\prime(t) &= \begin{cases} 0, & t=0\\ \frac{1}{t}\langle \nabla f(z) - \nabla f(x), z-x \rangle \geq 0, \text{ by (4)}, & t>0 \end{cases}\\ &\Rightarrow g\prime(t) \geq 0, \ \forall t \geq 0\\ &\Rightarrow \text{g(t) is nondecreasing}, \forall t \geq 0 \Rightarrow g(1) \geq g(0)\\ &\text{Also } g(0) = f(x), \ g(1) = f(y) - \langle \nabla f(x), y-x \rangle\\ &\Rightarrow f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle \,, \ \forall x,y \in C \end{aligned}$$

(iii) Proof $(c) \Rightarrow (d)$:

$$\langle \nabla f(z) - \nabla f(x), z - x \rangle \geq 0 \; \forall x, z \in C$$

$$\text{Let } z = x + t(y-x), 0 \neq t \in \Re \Rightarrow \langle \nabla f(x + t(y-x)) - \nabla f(x), t(y-x) \rangle \geq 0$$
$$\Rightarrow \frac{\langle \nabla f(x + t(y-x)) - \nabla f(x), t(y-x) \rangle}{t^2} \geq 0$$
$$\Rightarrow \frac{\langle \nabla f(x + t(y-x)) - \nabla f(x), (y-x) \rangle}{t} \geq 0$$
$$\Rightarrow \left\langle \frac{\nabla f(x + t(y-x)) - \nabla f(x)}{t}, (y-x) \right\rangle \geq 0$$

But f is twice Frechet (hence Gateaux) differentiable, hence taking limit $t \downarrow 0$ from both sides of the above inequality,

$$\Rightarrow \left\langle \lim_{t \downarrow 0} \frac{\nabla f(x + t(y-x)) - \nabla f(x)}{t}, (y-x) \right\rangle \geq 0$$
$$\Rightarrow \left\langle \left\langle \nabla^2 f(x), y - x \right\rangle, y - x \right\rangle \geq 0$$
$$\Rightarrow \langle H(f(x)).z, z \rangle \geq 0, \; z = y - x \in C$$
$$\Rightarrow \langle H(f(x)).z, z \rangle \geq 0, \; \forall z \in C, \text{ z being arbitrary, since y is arbitrary}$$
$$\Rightarrow Hf(x) \text{ is positive semidefinite, at } x \in C$$
$$\text{but } x \text{ is arbitrary} \Rightarrow Hf(x) \text{ is positive semidefinite, } \forall x \in C$$

(iv) Proof $(d) \Rightarrow (c)$:

Given $H(f(x))$ is positive semidefinite, $\forall x \in C$.

We again use the function $g(t) = f(x + t(y-x)) - t \langle \nabla f(x), y - x \rangle$, for any $x, y \in C, \; t \in \Re$.

$$\Rightarrow g\prime(t) = Df(x + t(y-x)).(y-x) - \langle \nabla f(x), y - x \rangle, \text{ by chain rule}$$
$$\Rightarrow g\prime(t) = \langle \nabla f(x + t(y-x)) - \nabla f(x), y - x \rangle, \; g\prime(0) = 0$$
$$\Rightarrow g\prime\prime(t) = D^2 f(x + t(y-x)).(y-x)^2 = \left\langle \nabla^2 f(x + t(y-x)).(y-x), y - x \right\rangle \geq 0,$$
$$\text{since f(z) is p.s.d. } \forall z \in C$$
$$\text{By FTC, } g\prime(1) - g\prime(0) = \int_0^1 g\prime\prime(t) dt \geq \int_0^1 0 dt = 0$$
$$\Rightarrow \langle \nabla f(y)) - \nabla f(x), y - x \rangle \geq 0, \forall x, y \in C$$

Problem 14 Solution

Choose any $\lambda \mid 0 \leq \lambda \leq 1$ and any $x, y \in \Re^n$.

(a)

1. $x, y \in C \Rightarrow \lambda x + (1-\lambda)y \in C,\ \delta_C(x) = \delta_C(y) = 0.$
   $\Rightarrow \delta_C(\lambda x + (1-\lambda)y) = 0 \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) = 0.$

2. $x \in C, y \notin C \Rightarrow \delta_C(x) = 0, \delta_C(y) = +\infty.$

   (a) $\lambda x + (1-\lambda)y \in C \Rightarrow \delta_C(\lambda x + (1-\lambda)y) = 0 \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) = \infty.$
   (b) $\lambda x + (1-\lambda)y \notin C \Rightarrow \delta_C(\lambda x + (1-\lambda)y) = \infty \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) = \infty.$

   Similarly, Jensen's inequality holds for $x \in C, y \notin C$.

3. $x, y \notin C \Rightarrow \delta_C(x) = \delta_C(y) = \infty.$

   (a) $\lambda x + (1-\lambda)y \in C \Rightarrow \delta_C(\lambda x + (1-\lambda)y) = 0 \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) = \infty.$
   (b) $\lambda x + (1-\lambda)y \notin C \Rightarrow \delta_C(\lambda x + (1-\lambda)y) = \infty \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) = \infty.$

Hence, $\forall x, y \in \Re^n,\ 0 \leq \lambda \leq 1$, δ_C satisfies the Jensen's inequality
$\delta_C(\lambda x + (1-\lambda)y) \leq \lambda\delta_C(x) + (1-\lambda)\delta_C(y) \Rightarrow \delta_C$ is convex.

(b)

Let $d_C(x) = inf\{||z - x|| : z \in C\} = ||z_1 - x||$, inf is achieved at $z_1 \in C$
$\Rightarrow ||z_1 - x|| \leq ||z - x||,\ \forall z \in C \quad \ldots \ldots (1).$

Let $d_C(y) = inf\{||z - y|| : z \in C\} = ||z_2 - y||$, inf is achieved at $z_2 \in C$
$\Rightarrow ||z_2 - y|| \leq ||z - y||,\ \forall z \in C \quad \ldots \ldots (2).$

Let $d_C(\lambda x + (1-\lambda)y) = inf\{||z - \lambda x - (1-\lambda)y|| : z \in C\} = ||z_3 - \lambda x - (1-\lambda)y||$,
inf is achieved at $z_3 \in C \Rightarrow ||z_3 - \lambda x - (1-\lambda)y|| \leq ||z - \lambda x - (1-\lambda)y||,\ \forall z \in C.$

$d_C(\lambda x + (1-\lambda)y) = ||z_3 - \lambda x - (1-\lambda)y|| \leq ||z - \lambda x - (1-\lambda)y||,\ \forall z \in C$
$\Rightarrow \delta_C(\lambda x + (1-\lambda)y) \leq ||z - \lambda x - (1-\lambda)y|| = ||\lambda(z - x) + (1-\lambda)(z - y)||$
$\leq ||\lambda(z - x)|| + ||(1-\lambda)(z - y)||$ (by the triangle inequality)
$= \lambda||z - x|| + (1-\lambda)||z - y||,\ \forall z \in C$
$\Rightarrow d_C(\lambda x + (1-\lambda)y) \leq \lambda||z - x|| + (1-\lambda)||z - y||,\ \forall z \in C$
$\Rightarrow d_C(\lambda x + (1-\lambda)y) \leq \lambda.\underset{z \in C}{inf}||z - x|| + (1-\lambda).\underset{z \in C}{inf}||z - y||$
$= \lambda||z_1 - x|| + (1-\lambda)||z_2 - y|| = \lambda d_C(x) + (1-\lambda)d_C(y)$, from (1) and (2).
$\Rightarrow d_C(\lambda x + (1-\lambda)y) \leq \lambda d_C(x) + (1-\lambda)d_C(y),\ \forall x, y \in C.$
$\Rightarrow d_C$ is convex.

(c)

$$
\begin{aligned}
epi\,(\sigma_C) &= \{(x,\alpha) \mid \underset{z\in C}{sup}\,\langle z,x\rangle \le \alpha\} \\
&= \{(x,\alpha) \mid\ \langle z,x\rangle \le \alpha, \forall z \in C\} \\
&= \bigcap_{z\in C} \{(x,\alpha) \mid\ \langle z,x\rangle \le \alpha\} \\
&\Rightarrow \text{intersection of convex sets} \\
&\Rightarrow epi(\sigma_C) \text{ is convex} \\
&\Rightarrow \sigma_C \text{ is convex.}
\end{aligned}
$$