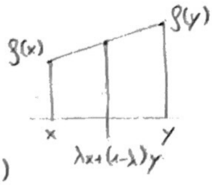
EXERCISES 1.4

1) (OPTIONAL) Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex function such that $-f$ is a convex function. Prove that exist $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $f(x) = a^T x + c$

Proof

$$\left.\begin{array}{l} f:\mathbb{R}^n \to \mathbb{R} \text{ convex} \\ -f:\mathbb{R}^n \to \mathbb{R} \text{ convex} \end{array}\right| \begin{array}{l} \forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n \\ \forall \lambda \in [0,1] \end{array} \quad \begin{array}{l} f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda) f(x) \\ -f(\lambda y + (1-\lambda)x) \leq -\lambda f(y) - (1-\lambda) f(x) \\ \hookrightarrow f(\lambda y + (1-\lambda)x) \geq \lambda f(y) + (1-\lambda) f(x) \end{array}$$

Thus $f(\lambda y + (1-\lambda)x) = \lambda f(y) + (1-\lambda) f(x) \quad (*)$



But we have also proven the following:

$$f \text{ convex}, f \in C^1 \iff f(y) \geq f(x) + [\nabla f(x)]^T (y-x) \quad \forall x,y \in dom(f)$$

Thus, assuming $f \in C^1$ we also have $-f(y) \geq -f(x) + [\nabla -f(x)]^T (y-x) \quad \forall x,y \in dom(f)$

$$\hookrightarrow f(y) \leq f(x) + [\nabla f(x)]^T (y-x)$$

Which gives $f(y) = f(x) + [\nabla f(x)]^T (y-x) \quad \forall x,y \in \mathbb{R}^n$

and fixing $x = x_0 \in \mathbb{R}^n$: $f(y) = \underbrace{[\nabla f(x_0)]^T}_{a^T \in \mathbb{R}^n} y + \underbrace{f(x_0) - [\nabla f(x_0)]^T x_0}_{c \in \mathbb{R}} = a^T y + c \quad \forall y \in \mathbb{R}^n \quad (**)$

$a \in \mathbb{R}^n$, $c \in \mathbb{R}$

In order to show $f \in C^1$, consider $(*)$, rearranging it we obtain the following:

$$f(\lambda y + (1-\lambda)x) = \lambda f(y) + (1-\lambda) f(x)$$

$$f(x + \lambda(y-x)) - f(x) = \lambda [f(y) - f(x)]$$

$$\underbrace{\frac{f(x+\lambda(y-x)) - f(x)}{\lambda}} = f(y) - f(x) \in \mathbb{R}$$

$D_{(y-x)} f(x)$ is the directional derivative for $\lambda \to 0$: $D_v f(x) = \lim_{h \to 0} \frac{f(x+hv) - f(x)}{h}$, $v \in \mathbb{R}^n$

Since the directional derivative $D_{(y-x)} f(x)$ exists real for any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, and thus for any $v = (y-x) \in \mathbb{R}^n$, then $f \in C^1$ and the above mentioned result $(**)$ holds. ■

2) Use the Kuhn-Tucker conditions to solve the following problems

(a) $\begin{cases} \text{Minimize } f(x) = x_1 x_2 \\ \text{subject to} \\ x_1 + x_2 \geq 2 \\ x_2 \geq x_1 \end{cases} \longrightarrow \begin{cases} \text{Minimize } f(x) = x_1 x_2 \\ \text{subject to} \\ -x_1 - x_2 + 2 \leq 0 \\ x_1 - x_2 \leq 0 \end{cases}$

$$L(x,\lambda) = f(x) + \sum_{i=1}^{2} \lambda_i g_i(x) = x_1 x_2 + \lambda_1(-x_1 - x_2 + 2) + \lambda_2(x_1 - x_2)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x_1}(x,\lambda) &= x_2 - \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial x_2}(x,\lambda) &= x_1 - \lambda_1 - \lambda_2 = 0 \end{aligned}\right\} \quad x_2 - x_1 + 2\lambda_2 = 0, \quad \lambda_2 = \frac{x_1 - x_2}{2}, \quad \lambda_1 = \frac{x_1 + x_2}{2}$$

$$H_L(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad tr = 0, \ det = -1$$

$eig_1 = 1$ ✓, $eig_2 = -1$

$$\lambda_1 g_1(x) = \lambda_1(-x_1 - x_2 + 2) = 0 \longrightarrow \frac{1}{2}(x_1 + x_2)(-x_1 - x_2 + 2) = 0$$

$$\lambda_2 g_2(x) = \lambda_2(x_1 - x_2) = 0$$

$x_1 = -x_2 \qquad x_1 = 2 - x_2$

$$\frac{1}{2}(x_1 - x_2)^2 = 0 \longrightarrow x_1 = x_2 \longrightarrow x_2 = 2 - x_2$$

$\hookrightarrow x_2 = -x_2$

$\cancel{\begin{pmatrix} x_2 = 0 \\ x_1 = 0 \end{pmatrix}}$

$g_1(0,0) = 2 \not\leq 0$

$\begin{pmatrix} x_2 = 1 \\ x_1 = 1 \end{pmatrix} \longrightarrow \lambda_1 = 1, \ \lambda_2 = 0$

$g_1(1,1) = 0 \leq 0$

$g_2(1,1) = 0 \leq 0$

but it is not a minimum; it doesn't satisfy $L(\bar{x},\bar{\lambda}) = \min_{x \in S} L(x,\bar{\lambda})$

(b) $\begin{cases} \text{Minimize } f(x) = (x_1 - 1)^2 + x_2 - 2 \\ \text{subject to} \\ x_2 - x_1 = 1 \\ x_1 + x_2 \leq 2 \end{cases} \longrightarrow \begin{cases} \text{Minimize } f(x) = (x_1 - 1)^2 + x_2 - 2 \\ \text{subject to} \\ x_2 - x_1 - 1 = 0 \\ x_1 + x_2 - 2 \leq 0 \end{cases}$

$$L(x,\lambda,\mu) = (x_1 - 1)^2 + x_2 - 2 + \lambda(x_1 + x_2 - 2) + \mu(x_2 - x_1 - 1)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x_1} &= 2x_1 - 2 + \lambda - \mu = 0 \\ \frac{\partial L}{\partial x_2} &= 1 + \lambda + \mu = 0 \end{aligned}\right\} \quad \mu = -\lambda - 1, \ 2x_1 - 2 + \lambda + \lambda + 1 = 0, \quad \mu = x_1 - 3/2, \quad \lambda = 1/2 - x_1$$

$$H_L = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \quad tr = 2, \ det = 0$$

$$\lambda g(x) = \lambda(x_1 + x_2 - 2) = 0 \longrightarrow (1/2 - x_1)(x_1 + x_2 - 2) = 0$$

$x_1 = 1/2 \qquad x_1 = 2 - x_2$

$$h(x) = x_2 - x_1 - 1 = 0 \longrightarrow x_2 = x_1 + 1 \longrightarrow x_2 = 3/2 \qquad x_1 = 1/2, \ x_2 = 3/2$$

$\begin{pmatrix} x_1 = 1/2 \\ x_2 = 3/2 \end{pmatrix} \longrightarrow \lambda = 0, \ \mu = -1$

$g(1/2, 3/2) = 0$

MIN $\longrightarrow L(\bar{x},\bar{\lambda}) = \min_{x \in S} L(x,\bar{\lambda})$ ✓

$\hookrightarrow x_2 = x_1 + 1$

c) $\begin{cases} \text{Minimize } f(x) = x_1^2 + 2x_2^2 + 3x_3^2 \\ \text{subject to} \\ x_1 - x_2 - 2x_3 \le 12 \\ x_1 + 2x_2 - 3x_3 \le 8 \end{cases} \longrightarrow \begin{cases} \text{Minimize } f(x) = x_1^2 + 2x_2^2 + 3x_3^2 \\ \text{subject to} \\ x_1 - x_2 - 2x_3 - 12 \le 0 \\ x_1 + 2x_2 - 3x_3 - 8 \le 0 \end{cases}$

$$L(x,\lambda) = x_1^2 + 2x_2^2 + 3x_3^2 + \lambda_1(x_1 - x_2 - 2x_3 - 12) + \lambda_2(x_1 + 2x_2 - 3x_3 - 8)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x_1}(x,\lambda) &= 2x_1 + \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial x_2}(x,\lambda) &= 4x_2 - \lambda_1 + 2\lambda_2 = 0 \end{aligned}\right\} \quad 2x_1 + 4x_2 + 3\lambda_2 = 0 \qquad \lambda_1 = -2x_1 - \lambda_2$$

$$\lambda_2 = \frac{2x_1 + 4x_2}{3}, \quad \lambda_1 = -\frac{8x_1 + 4x_2}{3}$$

$$\frac{\partial L}{\partial x_3}(x,\lambda) = 6x_3 - 2\lambda_1 - 3\lambda_2 = 0 \longrightarrow 6x_3 + \frac{2}{3}(8x_1 + 4x_2) - 2x_1 - 4x_2 = 0$$

$$6x_3 = -\frac{10}{3}x_1 + \frac{4}{3}x_2 \longrightarrow x_3 = \frac{-5x_1 + 2x_2}{9}$$

$$\lambda_1 g_1(x) = \lambda_1(x_1 - x_2 - 2x_3 - 12) = 0 \longrightarrow -\frac{4}{3}(2x_1 + x_2)(x_1 - x_2 - 2x_3 - 12) = 0$$

$$-\frac{4}{3}(2x_1 + x_2)\left(\frac{19}{9}x_1 - \frac{13}{9}x_2 - 12\right) = 0$$

$$\lambda_2 g_2(x) = \lambda_2(x_1 + 2x_2 - 3x_3 - 8) = 0$$

$$\cancel{-\frac{4}{27}}(2x_1 + x_2)(19x_1 - 13x_2 - 108) = 0$$

$$\frac{2}{3}(x_1 + 2x_2)(x_1 + 2x_2 - 3x_3 - 8) = 0$$

$$\frac{2}{3}(x_1 + 2x_2)\left(\frac{8}{3}x_1 + \frac{4}{3}x_2 - 8\right) = 0$$

$$\cancel{\frac{8}{9}}(x_1 + 2x_2)(2x_1 + x_2 - 12) = 0$$

$x_1 = -2x_2$ ① ② $x_2 = 12 - 2x_1$

① $-3x_2(-51x_2 - 108) = 0$

$\begin{pmatrix} x_2 = 0 \\ x_1 = 0 \\ x_3 = 0 \end{pmatrix}$ $\quad \cancel{\begin{pmatrix} x_2 = -36/17 \\ x_1 = 72/17 \\ x_3 = -48/17 \end{pmatrix}}$

$\lambda_1 = 0 \qquad \lambda_1 = -144/17$

$\lambda_2 = 0 \qquad \lambda_2 = 0$

$g_1(\bar{x}) = -12 \le 0 \qquad g_1(\bar{x}) = 0$

$g_2(\bar{x}) = -8 \le 0 \qquad g_2(\bar{x}) = -8/17 \le 0$

② $12(45x_1 - 264) = 0$

$\cancel{\begin{pmatrix} x_1 = 88/15 \\ x_2 = 4/15 \\ x_3 = -16/15 \end{pmatrix}}$

$\lambda_1 = -16$

$\lambda_2 = 64/15$

$g_1(\bar{x}) = -64/15 \le 0$

$g_2(\bar{x}) = 8/5 \cancel{\le 0}$

$$H_L = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{matrix} \text{tr} > 0 \\ \det > 0 \end{matrix} \longrightarrow \underline{\text{MIN}} \quad L(\bar{x},\bar{\lambda}) = \min_{\lambda \in S} L(x,\bar{\lambda}) \ ✓$$

3) (OPTIONAL) Consider the problem $\begin{cases} \text{Minimize } f(x) \\ \text{subject to} \\ g(x) \leq 0 \\ x \in S \end{cases}$

where $f: \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R}^n \to \mathbb{R}$ are two convex functions and $S \subseteq \mathbb{R}^n$ is a convex set. If x^* is an optimal solution of this problem such that $g(x^*) < 0$, show that x^* is also an optimal solution of the problem $\begin{cases} \text{Minimize } f(x) \\ \text{subject to} \\ x \in S \end{cases}$

<u>Proof</u>

We need to prove that under those assumptions x^* is a global minimum, i.d. not only on $\{x \in S : g(x) \leq 0\}$ but also on $\{x \in S\}$.
Since $g(x^*) < 0$, x^* is not only an optimal solution of the first problem but also a minimum of f. In fact, if x^* is such that $g(x^*) < 0$, then we have a neighborhood $B_\varepsilon(x^*) = \{x \in S : |x - x^*| < \varepsilon, \varepsilon > 0\}$ for ε small enough such that $B_\varepsilon(x^*) \subset \{x \in S : g(x) \leq 0\}$, and $f(x) \leq x^* \ \forall x \in B_\varepsilon(x^*)$.
Since f is convex, we have, as we have previously shown, that x^* is also a global minimum for f. Therefore, it is an optimal solution for the second problem. ■