Homework 7

1. (a)

Since f is coercive function, $\forall f_0 \in R$, $\exists r > 0$, if $\|x\| > r$, then $f(x) > f_0$

This means all x with $\|x\| > r$ are not in $\mathcal{L}_{f_0}^- = \{x \in R^n : f(x) \le f_0\}$, because $f(x) > f_0$.

Thus, $\forall x \in \mathcal{L}_{f_0}^-$ must satisfy $\|x\| \le r$. This implies $\mathcal{L}_{f_0}^- \subset B_r(0)$

(b)

According to part (a), for a coercive function f, for any given value f_0,

$\mathcal{L}_{f_0}^- \subset B_r(0)$ which means that $\mathcal{L}_{f_0}^-$ is bounded.

Since coercive function f is also continuous, this means $\mathcal{L}_f^-$ contains all

boundary points, $\mathcal{L}_{f_0}^-$ is closed

Because $\mathcal{L}_{f_0}^-$ is closed and bounded, $\mathcal{L}_{f_0}^-$ is compact set

According to Weierstrass extreme value theorem, any continuous function f

defined on compact set D has a global minimizer in D

Thus, coercive function always has global minimizer in $\mathcal{L}_{f_0}^-$

(C)

Since $\|x-y\| \ge |\|x\| - \|y\||$, $f(x) = \|x-y\|^2 \ge (\|x\| - \|y\|)^2 = \|x\|^2 - 2\|x\|\|y\| + \|y\|^2$

As $\|x\| \to +\infty$, $\|x\|^2 - 2\|x\|\|y\| + \|y\|^2 = \|x\|(\|x\| - 2\|y\|) + \|y\|^2 \longrightarrow +\infty$

Thus, $\lim_{\|x\| \to +\infty} f(x) = +\infty$, f is coercive.

(d)

$f(x) = e^{-x}$ is strictly convex function and is not coercive

$f'(x) = -e^{-x}$, $f''(x) = e^{-x} > 0$ which means $f(x)$ is strictly convex

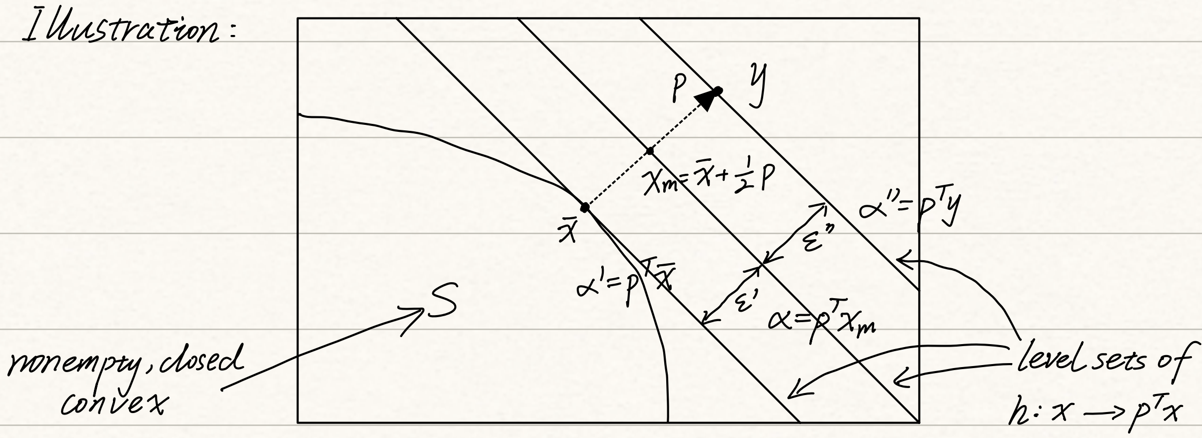
$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} e^{-x} = 0 \ne +\infty$, which means $f(x)$ is not coercive.

2.

Separating Hyperplane Theorem:

Separating Hyperplane between S and $\{y\}$. Let S be a nonempty, closed, convex set in R^n and $y \notin S$. There exists a hyperplane $H = \{x \in R^n : p^T x = \alpha\}$ defined by $p \neq 0$ ($p \in R^n$) and $\alpha \in R$ strongly separates the sets S from the set $\{y\}$. i.e. $p^T x \leq \alpha - \varepsilon$ for $\forall x \in S$ and $p^T y \geq \alpha + \varepsilon$ for some $\varepsilon > 0$

Illustration:



Since S is convex and closed, the distance between y and S is well-defined

Let the Euclidean distance be $d = \inf_{x \in S} \|x - y\|$

The convexity and closedness of S ensure the existence of a unique $\bar{x} \in S$ that achieves this minimum distance. $\bar{x} = \underset{x \in S}{\operatorname{argmin}} \|x - y\| = \underset{x \in S}{\operatorname{argmin}} \|x - y\|^2$

Define the optimization problem

$$\min_{x \in S} f(x) = \min_{x \in S} \|x - y\|^2 = \min_{x \in S} (x-y)^T (x-y)$$

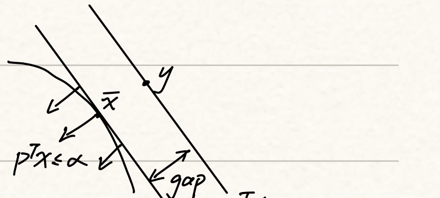
$$\nabla f(x) = 2(x - y)$$

At the minimizer $\bar{x}$, we have $\nabla f(\bar{x})^T (x - \bar{x}) \geq 0$, $\forall x \in S$

$$\Longrightarrow 2(\bar{x} - y)^T (x - \bar{x}) \geq 0 \Longrightarrow (\bar{x} - y)^T (x - \bar{x}) \geq 0 \Longrightarrow (y - \bar{x})^T (x - \bar{x}) \leq 0 \quad \forall x \in S$$

Define $p = y - \bar{x}$, we have $p^T (x - \bar{x}) \leq 0 \quad \forall x \in S$

$$\Longleftrightarrow p^T x \leq p^T \bar{x}, \quad \forall x \in S$$



let $\alpha = p^T \bar{x}$, the above inequality implies $p^T x \leq \alpha$. $\forall x \in S$

$p^T\bar{x}=\alpha$ p^Ty

Since $\bar{x}$ is the closet point in S to y, $y \notin S$, we also have $p^Ty > \alpha$

These inequalities follows from the distance $d = \|\bar{x} - y\| > 0$, ensuring that $\exists\, \varepsilon > 0$

such that $p^Tx \le \alpha - \varepsilon$, $p^Ty \ge \alpha + \varepsilon$

Therefore, there exists the hyperplane $H = \{x \in R^n : p^Tx = \alpha\}$ strongly separates S and $\{y\}$.

3.

Key Lemma:

Given matrices $\bar{A}$, B and H of appropriate dimensions, exactly one of the following two systems has a solution

(1) $\bar{A}x < 0$, $Bx \le 0$, $Hx = 0$

(2) $\bar{A}^Tu + B^Tw + H^Tv = 0$, $u \ge 0$, $w \ge 0$, v arbitrary, $e^Tu = 1$, $e = \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$

According to (1) $\bar{A}x < 0$, $Bx \le 0$, $Hx = 0$, we introduce $\theta > 0$, and reformulate the problem as

$$\bar{A}x + e\theta \le 0$$

$$Bx \le 0$$

$$Hx \le 0$$

$$-Hx \le 0$$

Assuming that system (1) has no solution, the above extened system also has no solution.

The extended system is equivalent to

$$\begin{pmatrix}\bar{A} & e\\ B & 0\\ H & 0\\ -H & 0\end{pmatrix}\begin{pmatrix}x\\ \theta\end{pmatrix} \le \begin{pmatrix}0\\ 0\\ 0\\ 0\end{pmatrix} \quad \text{with } \theta > 0 \Rightarrow (0, 0, \cdots, 0, 1)\begin{pmatrix}x\\ \theta\end{pmatrix} > 0$$

(Farka's Lemma: Let $A \in R^{m\times n}$, $c \in R^n$. Exactly one of the following two systems has a solution (1) $Ax \le 0, c^T x > 0$ (2) $A^T y = c, y \ge 0$)

According to Farka's Lemma, the above system has no solution is equivalent to the existence of a solution to its system (2).

$\exists\, u \ge 0, w \ge 0, v_1 \ge 0, v_2 \ge 0$

$$\begin{pmatrix} \bar{A}^T & B^T & H^T & -H^T \\ e^T & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ w \\ v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$\Longrightarrow \bar{A}^T u + B^T w + H^T v_1 - H^T v_2 = 0, \quad e^T u = 1$

$\Longrightarrow$ let $v = v_1 - v_2$, $\bar{A}^T u + B^T w + H^T v = 0, \quad e^T u = 1$

Thus, system (1) has no solution, system (2) must have a solution.

Therefore, the two alternatives in the key lemma cannot hold at the same time.

4. minimize $(x_1-12)^2 + (x_2+6)^2$

subject to

$$x_1^2 + x_2^2 + 3x_1 - 4.5x_2 - 6.5 \le 0$$

$$(x_1-9)^2 + x_2^2 - 64 \le 0$$

$$8x_1 + 4x_2 - 20 = 0$$

Fritz John Necessary Conditions: Let $\bar{x}$ be a feasible solution of (P) $\begin{cases} \min f(x) \\ \text{s.t. } g_i(x) \le 0, \ i=1,\cdots,m \\ \quad h_i(x) = 0, \ i = 1, \cdots, l \end{cases}$

If $\bar{x}$ is a local minimum of (P), then there exists $(u_0, u, v) \ne 0$ such that

$$u_0 (\nabla f)(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) + \sum_{i=1}^{l} v_i h_i(\bar{x}) = 0$$

$$u_0, u \ge 0,$$

$$u_i g_i(\bar{x}) = 0, \ i = 1, \cdots, m$$

According to the question, $f(x_1, x_2) = (x_1-12)^2 + (x_2+6)^2$,

$$g_1(x_1, x_2) = x_1^2 + x_2^2 + 3x_1 - 4.5x_2 - 6.5$$

$$g_2(x_1, x_2) = (x_1-9)^2 + x_2^2 - 64 \le 0$$

$$h(x_1, x_2) = 8x_1 + 4x_2 - 20$$

$$\Longrightarrow \nabla f(x) = \begin{pmatrix} 2(x_1-12) \\ 2(x_2+6) \end{pmatrix} \quad \nabla g_1(x) = \begin{pmatrix} 2x_1+3 \\ 2x_2 - 4.5 \end{pmatrix} \quad \nabla g_2(x) = \begin{pmatrix} 2(x_1-9) \\ 2x_2 \end{pmatrix} \quad \nabla h(x) = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$$

Thus, we have $u_0 \nabla f(x) + u_1 \nabla g_1(x) + u_2 \nabla g_2(x) + v \nabla h(x) = 0$

$$u_0\begin{pmatrix}2(x_1-12)\\2(x_2+6)\end{pmatrix}+u_1\begin{pmatrix}2x_1+3\\2x_2-4.5\end{pmatrix}+u_2\begin{pmatrix}2(x_1-9)\\2x_2\end{pmatrix}+v\begin{pmatrix}8\\4\end{pmatrix}=0$$

$$\Longrightarrow\begin{cases}2u_0(x_1-12)+u_1(2x_1+3)+2u_2(x_1-9)+8v=0\\2u_0(x_2+6)+u_1(2x_2-4.5)+2u_2x_2+4v=0\end{cases}$$

$u_0, u_1, u_2 \geq 0$

$u_1 g_1(x) = u_1(x_1^2+x_2^2+3x_1-4.5x_2-6.5)=0$

$u_2 g_2(x) = u_2((x_1-9)^2+x_2^2-64)=0$

5. minimize $e^{-(x_1+x_2)}$

subject to $e^{x_1}+e^{x_2}\leq 10$

$-x_1 \leq 0$

$f(x_1,x_2)=e^{-(x_1+x_2)}$ $\quad g_1(x_1,x_2)=e^{x_1}+e^{x_2}-10$, $\quad g(x_1,x_2)=-x_1$

$$\nabla f=\begin{pmatrix}-e^{-(x_1+x_2)}\\-e^{-(x_1+x_2)}\end{pmatrix}\quad \nabla g_1=\begin{pmatrix}e^{x_1}\\e^{x_2}\end{pmatrix}\quad \nabla g_2=\begin{pmatrix}-1\\0\end{pmatrix}$$

$$Hf=\begin{pmatrix}e^{-(x_1+x_2)} & e^{-(x_1+x_2)}\\e^{-(x_1+x_2)} & e^{-(x_1+x_2)}\end{pmatrix}>0\quad Hg_1=\begin{pmatrix}e^{x_1} & 0\\0 & e^{x_2}\end{pmatrix}>0\quad Hg_2=\begin{pmatrix}0&0\\0&0\end{pmatrix}=0$$

Thus, f, g_1, g_2 are convex function. Then, we have

$$\nabla f(x)+u_1\nabla g_1(x)+u_2\nabla g_2(x)=0$$

$$\begin{pmatrix}-e^{-(x_1+x_2)}\\-e^{-(x_1+x_2)}\end{pmatrix}+u_1\begin{pmatrix}e^{x_1}\\e^{x_2}\end{pmatrix}+u_2\begin{pmatrix}-1\\0\end{pmatrix}=0$$

$$\Longrightarrow\begin{cases}-e^{-(x_1+x_1)}+u_1e^{x_1}-u_2=0\\-e^{-(x_1+x_2)}+u_1e^{x_2}=0\end{cases}$$

$u_1, u_2 \geq 0$

$u_1 g_1(x) = u_1(e^{x_1}+e^{x_2}-10)=0$

$u_2 g_2(x) = -u_2 x_1 = 0$

Case 1: $u_1=u_2=0$ which implies that

$-e^{-(x_1+x_2)}=0$ So, there is no solution

Case 2: $u_1>0$, $u_2=0$, which implies that

$$\begin{cases}e^{x_1}+e^{x_2}-10=0\\-e^{-(x_1+x_2)}+u_1e^{x_1}=0\\-e^{-(x_1+x_2)}+u_1e^{x_2}=0\end{cases}\Longrightarrow x_1=x_2\Longrightarrow\begin{cases}x_1=\ln 5\\x_2=\ln 5\end{cases}$$

Case 3: $u_1=0$, $u_2>0$, which implies that

$$\begin{cases}-e^{(x_1+x_2)}-u_2=0\\-e^{(x_1+x_1)}=0\\x_1=0\end{cases}$$ So, there is no solution.

Case 4: $u_1 > 0$, $u_2 > 0$, which implies that

$$\begin{cases} e^{x_1} + e^{x_2} - 10 = 0 \\ x_1 = 0 \\ -e^{-(x_1+x_2)} + u_1 e^{x_1} - u_2 = 0 \\ -e^{-(x_1+x_2)} + u_1 e^{x_2} = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = 0 \\ x_2 = \ln 9 \end{cases} \quad \begin{cases} u_1 = \frac{1}{81} \\ u_2 = -\frac{8}{81} \end{cases}$$

does not satisfy $u_2 > 0$

So, the solution is $x_1 = \ln 5$, $x_2 = \ln 5$, $f(x_1, x_2) = e^{-2\ln 5} = \frac{1}{25}$

6. minimize $x^T Q x + c^T x$
subject to $x^T R x \leq 1$
$e^T x = 1$

where $Q \in R^{n\times n}$ is invertible but not necessarily positive definite, $R > 0$, $e^T = (1, 1, \cdots, 1)$

According to the KKT condition, we can get

$f(x) = x^T Q x + c^T x$, $g(x) = x^T R x - 1$, $h(x) = e^T x - 1$

$\nabla f(x) = 2Qx + c$, $\nabla g(x) = 2Rx$, $\nabla h(x) = e$

$$\nabla f(x) + u \nabla g(x) + v \nabla h(x) = 0$$

$$\Rightarrow \begin{cases} 2Qx + c + 2uRx + ve = 0 \Rightarrow (2Q + 2uR)x + c + ve = 0 \\ u \geq 0 \\ ug(x) = u(x^T R x - 1) = 0 \Longrightarrow u = 0 \text{ or } x^T R x = 1 \end{cases}$$

$$\Rightarrow x = -(2Q + 2uR)^{-1}(c + ve) = -\frac{1}{2}(Q + uR)^{-1}(c + ve)$$

If $u = 0$, $x = -\frac{1}{2} Q^{-1}(c + ve)$

$$e^T x = 1 \Rightarrow e^T\left(-\frac{1}{2}Q^{-1}(c + ve)\right) = -\frac{1}{2} e^T Q^{-1} c - \frac{v}{2} e^T Q^{-1} e = 1$$

$$\Rightarrow v = \frac{-2 - e^T Q^{-1} c}{e^T Q^{-1} e}$$

Thus $x = -\frac{1}{2} Q^{-1}\left(c + \frac{-2 - e^T Q c}{e^T Q^{-1} e} e\right) = -\frac{1}{2} Q^{-1}\left(c + \frac{-2 - e^T Q c}{e^T Q^{-1}}\right)$

If $u > 0$, $x^T R x = 1$,

$$x = -\frac{1}{2}(Q + uR)^{-1}(c + ve)$$

$$e^T x = 1 \Rightarrow e^T\left(-\frac{1}{2}(Q + uR)^{-1}(c + ve)\right) = 1 \Rightarrow -\frac{1}{2} e^T (Q + uR)^{-1} c - \frac{v}{2} e^T (Q + uR)^{-1} e = 1$$

$$\Rightarrow v = \frac{-2 - e^T (Q + uR)^{-1} c}{e^T (Q + uR)^{-1} e}$$

$$x^T R x = \left(-\frac{1}{2}(Q + uR)^{-1}\left(c + \frac{-2 - e^T (Q + uR)^{-1} c}{e^T (Q + uR)^{-1} e}\right)\right)^T R \left(-\frac{1}{2}(Q + uR)^{-1}\left(c + \frac{-2 - e^T (Q + uR)^{-1} c}{e^T (Q + uR)^{-1} e}\right)\right) = 1$$

7. minimize $(x_1-16)^2 + (x_2-14)^2$

subject to $(x_1-8)^2 + (x_2-9)^2 - 49 \le 0$

$$-x_1+2 \le 0$$
$$x_1 - 13 \le 0$$
$$x_1+x_2-24 = 0$$

Apply the Fritz John Necessary Condition to this problem.

$f(x_1,x_2) = (x_1-16)^2+(x_2-14)^2$, $g_1(x_1,x_2) = (x_1-8)^2+(x_2-9)^2-49$. $g_2(x_1,x_2) = -x_1+2$,

$g_3(x_1,x_1) = x_1-13$, $h(x_1,x_2) = x_1+x_2-24$

$$\nabla f(x) = \begin{pmatrix} 2(x_1-16) \\ 2(x_2-14) \end{pmatrix},\ \nabla g_1(x) = \begin{pmatrix} 2(x_1-8) \\ 2(x_2-9) \end{pmatrix}\ \nabla g_2(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}\ \nabla g_3(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\quad \nabla h(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$u_0 \nabla f(x) + u_1 \nabla g_1(x) + u_2 \nabla g_2(x) + u_3 \nabla g_3(x) + v \nabla h(x) = 0$$

$$\Rightarrow\ u_0 \begin{pmatrix} 2(x_1-16) \\ 2(x_2-14) \end{pmatrix} + u_1 \begin{pmatrix} 2(x_1-8) \\ 2(x_2-9) \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + u_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

$$\Rightarrow \begin{cases} 2u_0(x_1-16) + 2u_1(x_1-8) - u_2 + u_3 + v = 0 \\ 2u_0(x_2-14) + 2u_1(x_2-9) + v = 0 \end{cases}$$

$u_0, u_1, u_2, u_3 \ge 0$

$u_1 g_1(x) = u_1[(x_1-8)^2+(x_2-9)^2-49] = 0$

$u_2 g_2(x) = u_2(-x_1+2) = 0$

$u_3 g_3(x) = u_3(x_1-13) = 0$

$h(x_1,x_2) = 0 \Rightarrow x_1+x_2-24=0 \Rightarrow x_1+x_2 = 24$

If $u_1 > 0$, $(x_1-8)^2+(x_2-9)^2-49=0 \Rightarrow x_1=15, x_2=9$ or $x_1=8, x_2=16$

If $u_2 > 0$, $-x_1+2=0 \Rightarrow x_1=2$, $x_2=22$

If $u_3 > 0$, $x_1-13=0 \Rightarrow x_1=13$, $x_2=11$

Only $x_1=8, x_2=16$; $x_1=13, x_2=11$ satisfy all constraints

$f(8,16) = (8-16)^2+(16-14)^2 = 64+4 = 68$

$f(13,11) = (13-16)^2+(11-14)^2 = 9+9 = 18$

Thus, the solution is $x_1=13, x_2=11$, $f(x_1,x_2) = 18$