EE381K : Convex Optimization Fall 2016 Po-Cheng Pan UT EID: pp22828

Written Problems:

1. Over-determined

① $f(\beta) = \|X\beta - y\|_2^2 = \beta^T X^T X \beta - \beta^T X^T y - y^T X \beta + y^T y$

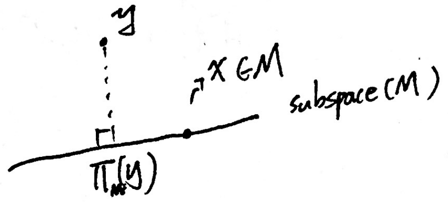
$\nabla^2 f(\beta) = 2X^T X \quad \because z^T \cdot \nabla^2 f(\beta) z = 2 \cdot z^T X^T X z = 2\|Xz\|^2 \geq 0, \forall z$

$\Rightarrow f(\beta)$ is a convex function of β #

If $\nabla f(\beta_{LS}) = 0. \Rightarrow 2X^T X \beta_{LS} - X^T y - (y^T X)^T = 0$

$\Rightarrow \beta_{LS} = (X^T X)^{-1} X^T y$ #

② Pic:



As the description of the question says,

$\langle y - \Pi_M(y), x \rangle = 0$ for all $x \in M$

We can say that $M = \text{Range}(X)$

$\Pi_M(y) = X\beta_{LS}$

Hence, $\langle y - \Pi_M(y), x \rangle = \langle y - X\beta_{LS}, X\beta \rangle = 0, \forall \beta$

Because $\forall \beta$, it all satisfies

$\Rightarrow \langle y - X\beta_{LS}, X \rangle = 0$

$(y - X\beta_{LS})^T \cdot X = 0$

$y^T X - \beta_{LS}^T X^T X = 0 \Rightarrow \beta_{LS} = (y^T X (X^T X)^{-1})^T$

$= (X^T X)^{-1} X^T y$ #

Under-determined

③ min: $\|\beta\|_2^2$

subject to $X\beta = y$

Consider any other solution $\beta_1 = \beta_0 + z$

$\|\beta_1\|^2 = \|\beta_0 + z\|^2 = \|\beta_0\|^2 + \|z\|^2 \quad (\because \beta_0 \perp z \Rightarrow \beta_0 \cdot z = 0)$

$\because \|z\|^2 \geq 0 \Rightarrow \|\beta_1\|^2 \geq \|\beta_0\|^2$

So, β_0 is the minimum norm solution #

④ $y = X\beta_0$ & $\beta_0 \perp z$ for any $z \in \text{Null}(X)$

$\because$ Null(X) is the set of vectors perpendicular to the rows of X

$\therefore$ The set of vectors perpendicular to Null(X) must be in span of the rows of X

$\Rightarrow \beta_0$ is in the span of the rows of X $\Rightarrow \beta_0 = X^T z = \sum z_i x_i$ (x_i is the ith row of X, z for some vector)

$y = X\beta_0$

$= XX^T z$

$z = (XX^T)^{-1} y$

$\beta_0 = X^T z$

$= X^T (XX^T)^{-1} y$ #

12. $C=\{x\in\mathbb{R}^n: x^TAx+b^Tx+c\le 0\}$ where $A\in S^n$, $b\in\mathbb{R}^n$, $c\in\mathbb{R}^n$

(a) if $A\in S_+^n \Rightarrow x^TAx\ge 0$

Consider $\{x_1+tv \mid t\in\mathbb{R}\}$ an arbitrary line

C's intersection with $\{x_1+tv\mid t\in\mathbb{R}\}$

$\Rightarrow (x_1+tv)^TA(x_1+tv)+b^T(x_1+tv)+c\le 0$

$\Rightarrow v^TAvt^2+(2x_1^TAv+b^Tv)t+(b^Tx_1+c+x_1^TAx_1)\le 0$

It's convex because $A\in S_+^n$ s.t. $v^TAv\ge 0$

So C is convex because its intersection with an arbitrary line is convex.

Q.E.D.

(b) We consider the intersection of $C\cap H=C_1$ with an arbitrary line $\{x_1+tv\mid t\in\mathbb{R}\}$ #

Without loss of generality, we can assume $g^Tx_1+h=0$

Hence, the intersection defined by x_1 & v is

$$\{x_1+tv \mid v^TAvt^2+(2x_1^TAv+b^Tv)t+(b^Tx_1+c+x_1^TAx_1)\le 0,\ g^Ttv=0\}$$

If $g^Tv=0$, the intersection is the singleton $\{x_1\}$

The set can be reduced to $\{x_1+tv \mid v^TAvt^2+(2x_1^TAv+b^Tv)t+(b^Tx_1+c+x_1^TAx_1)\le 0\}$

which is convex if $v^TAv\ge 0$

therefore C_1 is convex if $g^Tv=0\Rightarrow v^TAv\ge 0$

Consider $\lambda gg^T\in\mathbb{R}^{n\times n}$ $\lambda\in\mathbb{R}$

$$v^TAv=v^T(A+\lambda gg^T)v\ge 0 \quad (\because \lambda gg^Tv=0)$$

In conclusion, C_1 is convex if there exists $\lambda\in\mathbb{R}$ s.t. $A+\lambda gg^T\in S_+^n$

Q.E.D. #

13. $S=\{(a,b): a^Tx\le b\ \forall x\in C,\ a^Tx\ge b\ \forall x\in D\}$

It forms a set of homogenous linear inequalities in (a,b)

That means it's the intersection of many half spaces that pass through the origin

$\because a^Tx\le b\Rightarrow a^Tx-b\le 0\Rightarrow [x^T\ -1]\begin{bmatrix}a\\b\end{bmatrix}\le 0$ $\forall x\in C$ & $a^Tx-b\ge 0\Rightarrow [x^T\ -1]\begin{bmatrix}a\\b\end{bmatrix}\ge 0$ $\forall x\in D$

Hence, it will form a convex cone.

So S is convex.

17. We can consider a hyperplane perpendicular to (V_2-V_1) & lying on $\frac{V_1+V_2}{2}$

Suppose this hyperplane as $(V_2-V_1)^T X = a$

$$\Rightarrow a = (V_2-V_1)^T \cdot \frac{V_1+V_2}{2} = \frac{1}{2}\left(\|V_2\|^2 - V_1^T V_2 + V_2^T V_1 + \|V_2\|^2\right)$$
$$= \frac{1}{2}\left(\|V_2\|^2 - \|V_1\|^2\right)$$

$$\{X : \|X-V_1\| \le \|X-V_2\|\} = \{X : \|X-V_1\|^2 \le \|X-V_2\|^2\}$$
$$= \{X : \|X\|^2 - 2V_1^T X + \|V_1\|^2 \le \|X\|^2 - 2V_2^T X + \|V_2\|^2\}$$
$$= \{X : 2V_2^T X - 2V_1^T X \le \|V_2\|^2 - \|V_1\|^2\}$$
$$= \{X : (V_2^T - V_1^T) X \le \frac{\|V_2\|^2 - \|V_1\|^2}{2}\}$$

$c = V_2^T - V_1^T$ # $\quad d = \frac{\|V_2\|^2 - \|V_1\|^2}{2}$ #

18. $A \in \mathbb{R}^{l\times m}$ $\quad B \in \mathbb{R}^{k\times m}$

for every $X \in \mathbb{R}^m$, $AX=0 \Rightarrow BX=0 \Rightarrow Null(A) \subseteq Null(B)$

From 11.(c), we know that $U \subseteq W \Leftrightarrow U^\perp \supseteq W^\perp$

So $Null(A) \subseteq Null(B) \Rightarrow$ Row Space$(A) \supseteq$ Row Space(B)

$\Rightarrow Range(A^T) \supseteq Range(B^T)$

For every $b \in Range(B^T)$, we can find a vector $c \in \mathbb{R}^n$

s.t. $A^T c = b$

Hence, For Each column $b_i \in B^T$, we can find a $c_i \in \mathbb{R}^n$ s.t. $A^T c_i = b_i$

$$\Rightarrow A^T \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_k \end{bmatrix} = B^T = \begin{bmatrix} b_1 & b_2 & \cdots & b_k \end{bmatrix}$$

$$A^T C_1 = B^T \Rightarrow (A^T C_1)^T = (B^T)^T \Rightarrow C_1^T A = B$$

$\because C_1 \in \mathbb{R}^{n\times k}$ $\quad \therefore C_1^T \in \mathbb{R}^{k\times n}$

$\Rightarrow$ There exists a $k\times n$ real matrix C s.t. $CA = B$

Q.E.D.

#