Homework 2 (draft)

Problem 1

(a)

$\because A, B \in S_+^n$

$\therefore \exists U \in R^{k\times n}, k = rank(B), B = U^TU$

$\therefore tr(AB) = tr(AU^TU) = tr(UAU^T) \geq 0$ with assumption $tr(AB) = tr(BA)$

let us proof $tr(AB) = tr(BA), \forall A, B \in S^n$:

let $A = (a_{ij})_{n\times n}, B = (b_{ij})_{n\times n}$

$tr(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ji}$

$tr(BA) = \sum_{i=1}^n \sum_{j=1}^n b_{ij}a_{ji}$

we can rearrange the equation above and get $tr(AB) = tr(BA)$

(b)

First we proof $\mathcal{S}_+^n \subseteq \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

$\because$ the result in (a), we have $\mathcal{S}_+^n \subseteq \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

$\therefore \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\} \supseteq \bigcap_{A\in\mathcal{S}_+^n} S_+^n = S_+^n$

then we proof $\mathcal{S}_+^n \supseteq \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

we only have to proof : $\forall X \notin S_+^n, \exists A \in S_+^n, s.t. A \bullet X < 0$

$\because X \notin S_+^n$

$\therefore \exists \mu \in R^n, s.t. \mu^TX\mu < 0$

let $A = \mu\mu^T$, we can get $tr(AX) = tr(\mu\mu^TX) = tr(\mu^TX\mu) < 0$

note that $A = \mu\mu^T$ because $\forall z \in R^n, z^TAz = z^T\mu\mu^Tz = (\sum z_iu_i)^2 >= 0$

$\therefore \forall X \notin S_+^n, X \notin \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

$\therefore \mathcal{S}_+^n \supseteq \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

In summary, $\mathcal{S}_+^n = \bigcap_{A\in\mathcal{S}_+^n} \{X \in \mathcal{S}^n : A \bullet X \geq 0\}$

Problem 2

(a)

First let us proof f possesses Property C $\Leftrightarrow L_t = \{x \in \mathbb{R}^n : f(x) \leq t\}$ is compact for any $t \in \mathbb{R}$

① proof of $\Rightarrow$:

first let us show L_t is bounded, which means $\forall x \in L_t, \exists T \in \mathbb{R}, s.t. ||x||_2 < T$

Proof by contradiction:

assume L_t is not bounded, so there exists a sequence $\{x^k\}_{k\geq 0} \subseteq L_t, ||x^k||_2 \to +\infty$

$\because f$ possesses Property C

$\therefore f(x^k) \to +\infty$

However, with the limitation of $f(x) \leq t$, we can find $\{x^k\}_{k\geq 0} \subsetneq L_t$

Contradicting the assumption, thus L_t is bounded

then let us show L_t is close, which means $\forall \{x^k\}_{k\geq 0} \subseteq L_t, \{x^k\} \to x, x \in L_t$

$\because f$ is a continuous function

$\therefore \{f(x^k)\} \to f(x)$

$\because \{x^k\}_{k\geq 0} \subseteq L_t$

$\therefore \forall x^i \in \{x^k\}, f(x^i) \leq t$

$\therefore \{f(x^k)\} \to f(x) \leq t$

$\therefore x \in L_t, L_t$ is close

In summary, L_t is compact

② proof of $\Leftarrow$:

Proof by contradiction:

assume f not possesses Property C, which means $\exists \{x^k\}_{k\geq 0} \subseteq \mathbb{R}, ||x^k||_2 \to +\infty, f(x^k) \leq T, T \in \mathbb{R}$

let $t > T$, then $\{x^k\}_{k\geq 0} \subseteq L_t$. However $||x^k||_2 \to +\infty$, thus L_t is not bounded.

Contradicting the assumption, thus f possesses Property C

In summary, f possesses Property C $\Leftrightarrow L_t = \{x \in \mathbb{R}^n : f(x) \leq t\}$ is compact for any $t \in \mathbb{R}$

Then let us proof f is continuous and possesses Property C, $\inf_{x\in\mathbb{R}^n} f(x)$ always has an optimal solution.

$\inf_{x\in\mathbb{R}^n} f(x)$ always has an optimal solution, which means we can find a $x_0 \in \mathbb{R}^n s.t. \forall x \in R^n, f(x) \geq f(x_0)$

$\because f$ is continuous and possesses Property C, thus we can find a infimum c, which is the max c_i satisfy $f(x) \geq c_i$

then we only need to show that exist $x_0 \in \mathbb{R}^n, s.t. f(x_0) = c$

let $t = c + 1$, L_t is compact and c is also infimum of L_t

with the definition of infimum and continuous, $\exists \{x^k\}_{k\geq 0} \subseteq L_t, \{f(x^k)\} \to c, \{x^k\} \to x_0$

$\because L_t$ is compact

$\therefore x_0 \in L_t \subseteq R^n$

Thus we can always find an optimal solution x_0

(b)

No.

assume $f(0) < +\infty$, let $A = 0^{m\times n}$, thus $\forall x \in \mathbb{R}^n, g(x) \equiv f(0) < \infty, g$ not possesses Property C.

Problem 3

(a)

$\forall x, y \in K^\circ, \alpha, \beta \in R^+$, let $z = \alpha x + \beta y$

$\forall u \in K, z^T u = (\alpha x + \beta y)^T x = \alpha x^T u + \beta y^T u \leq 0$

$\therefore z \in K^\circ, K^\circ$ is a convex cone

(b)

First we know that, $z^* = \prod_K(x) \Leftrightarrow z^* \in K, (z - z^*)^T(x - z^*) \leq 0, \forall z \in K$

① proof of $\Rightarrow$:

With the property above, we have $z^* \in K$, then let us show that $(x - z^*)^T z^* = 0$

assume $(x - z^*)^T z^* \neq 0$, let $t = x - z^*, z^* = x - t$

$\because t^T z^* \neq 0$

$\therefore t = \alpha z^* + z$, where $\alpha \neq 0, z^* \neq 0, z^T z^* = 0$ and $t^T z^* = \alpha(z^*)^T z^*$

$\therefore ||x - z^*||_2^2 = ||x - (x - t)||_2^2 = ||t||_2^2 = ||\alpha z^*||_2^2 + ||z||_2^2 = |\alpha|||z^*||_2^2 + ||z||_2^2$

by the definition of closed convex cone and $z^* \neq 0, \alpha \neq 0$, we can find a
$z^o = (1 - sign(\alpha)min(\alpha, 1))z^* \in K, ||x - z^o|| = (1 - sign(\alpha)min(\alpha, 1))|\alpha|||z^*||_2^2 + ||z||_2^2 < ||x - z^*||_2^2$

so that $z^* \neq \prod_K(x)$, contradicting the assumption, thus $(x - z^*)^T z^* = 0$

last let us shat that $x - z^* \in K^o$

$\because (x - z^*)^T z^* = 0$ and $(z - z^*)^T(x - z^*) \leq 0, \forall z \in K$

$\therefore z^T(x-z^*)=(x-z^*)^Tz\le 0, \forall z\in K$

by the definition of K^o, we get $x-z^*\in K^o$

② **proof of** $\Leftarrow$:

$\because x-z^*\in K^o$ and $(x-z^*)^Tz^*=0$

$\therefore (z-z^*)^T(x-z^*)\le 0, \forall z\in K$

so that, we get $z^*\in K, (z-z^*)^T(x-z^*)\le 0, \forall z\in K \Rightarrow z^*=\prod_K(x)$

In summary, $z^*=\prod_K(x) \Leftrightarrow z^*\in K, x-z^*\in K^o, (x-z^*)^Tz^*=0$

(c)

Using the result in (b), we only need to show that $z^*=\prod_K(x), x-z^*=\prod_{K^o}(x)$

what's more, we only need to show that $K=\{\omega\in\mathbb{R}^n:\omega^Tx\le 0, \forall x\in K^o\}, K^o=\{\omega\in\mathbb{R}^n:\omega^Tx\le 0,\forall x\in K\}, K\subseteq\mathbb{R}^n$ be a closed convex cone

$\because x^T\omega=\omega^Tx$, so that $K=\{\omega\in\mathbb{R}^n:\omega^Tx\le 0,\forall x\in K^o\}$ is true obviously by definition

Using the result in (b):

$$x-z^*\in K^o, x-(x-z^*)=x\in K, (x-(x-z^*))^T(x-z^*)=x^T(x-z^*)=(x-z^*)^Tx=0\Rightarrow x-z^*=\prod_{K^o}(x)$$

In summary, $x=\prod_K(x)+\prod_{K^o}(x)$

Problem 4

① 1 $\Leftrightarrow$ 2

let $g(x)=f(x)+\frac{\rho}{2}||x||_2^2,\ g(x)$ is convex which means $g(\alpha x+(1-\alpha)y)\le\alpha g(x)+(1-\alpha)g(y), \forall x,y\in\mathbb{R}^n, \alpha\in[0,1]$

which is

$$f(\alpha x+(1-\alpha)y)+\frac{\rho}{2}||\alpha x+(1-\alpha)y||_2^2\le\alpha(f(x)+\frac{\rho}{2}||x||_2^2)+(1-\alpha)(f(y)+\frac{\rho}{2}||y||_2^2)$$

$$\Leftrightarrow f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)+\alpha\frac{\rho}{2}||x||_2^2+(1-\alpha)\frac{\rho}{2}||y||_2^2-\frac{\rho}{2}||\alpha x+(1-\alpha)y||_2^2$$

$$\Leftrightarrow f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)+\frac{\rho}{2}(\alpha||x||_2^2+(1-\alpha)||y||_2^2-||\alpha x+(1-\alpha)y||_2^2)$$

$$\Leftrightarrow f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)+\frac{\rho}{2}(\alpha||x||_2^2+(1-\alpha)||y||_2^2-(\alpha^2||x||_2^2+(1-\alpha)^2||y||_2^2+2\alpha(1-\alpha)\sum_{i=1}^n x_iy_i)$$

$$\Leftrightarrow f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)+\frac{\rho}{2}\alpha(1-\alpha)(||x||_2^2+||y||_2^2-2\sum_{i=1}^n x_iy_i)$$

$$\Leftrightarrow f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)+\frac{\rho\alpha(1-\alpha)}{2}||x-y||_2^2$$

$\because f$ is ρ-convex, so the inequality is true by the definition

$\therefore f$ is ρ-convex $\Leftrightarrow\ g(x)$ is convex

② 2 $\Leftrightarrow$ 3

$g(x)=f(x)+\frac{\rho}{2}||x||_2^2, \nabla g(x)=\nabla f(x)+\rho x$

$\because g$ is convex

$\therefore \forall x,y\in\mathbb{R}^n, g(y)\ge g(x)+\nabla g(x)^T(y-x)$

which is

$$f(y)+\frac{\rho}{2}||y||_2^2\ge f(x)+\frac{\rho}{2}||x||_2^2+(\nabla f(x)+\rho x)^T(y-x)$$

$$\Leftrightarrow f(y)+\frac{\rho}{2}||y||_2^2\ge f(x)+\frac{\rho}{2}||x||_2^2+\nabla f(x)^T(y-x)+\rho x^T(y-x)$$

$$\Leftrightarrow f(y)\ge f(x)+\nabla f(x)^T(y-x)+\frac{\rho}{2}||x||_2^2-\frac{\rho}{2}||y||_2^2+\rho x^T(y-x)$$

$$\Leftrightarrow f(y)\ge f(x)+\nabla f(x)^T(y-x)+\frac{\rho}{2}(||x||_2^2-||y||_2^2+2x^T(y-x))$$

$$\Leftrightarrow f(y)\ge f(x)+\nabla f(x)^T(y-x)-\frac{\rho}{2}(||y||_2^2+||x||_2^2-2x^Ty)$$

$$\Leftrightarrow f(y)\ge f(x)+\nabla f(x)^T(y-x)-\frac{\rho}{2}||y-x||_2^2$$

$\therefore g$ is convex $\Leftrightarrow \forall x,y\in\mathbb{R}^n, f(y)\ge f(x)+\nabla f(x)^T(y-x)-\frac{\rho}{2}||y-x||_2^2$

In conclusion, $1 \Leftrightarrow 2 \Leftrightarrow 3$, the statements are equivalent.