1. (1) $\frac{1}{2}x_1^2+|x_2|$: $\partial f = \begin{cases} \{x_1\}\times[-1,1], & x_2=0, x_1\in\mathbb{R}. \\ \{x_1\}\times\{\mathrm{sign}(x_2)\}, & x_2\neq 0, x_1\in\mathbb{R}. \end{cases}$

~~$\partial f = \{(x_1, t), t\in[-1,1]\}$~~

(2) $\|X\|_2$: $\partial f = \{y \mid \langle y, x\rangle = \|x\|$ 且 $\|y\|_2 \le 1\}$ as $\|\cdot\|_2$ 的 dual norm 为 $\|\cdot\|_2$.

$\partial f = \begin{cases} \{y \mid \|y\|_2\le 1\}, & x=0 \\ \{y \mid y=\frac{x}{\|x\|_2}\}, & x\neq 0 \end{cases}$ 且 $\|y\|\le 1$.

(3) $\|x\|_\infty$: $\partial f = \{y \mid \langle y, x\rangle = \max(|x_t|)$ 且 $\|y\|_1\le 1\}$

$\therefore \partial f = \{y \mid y_i x_i \ge 0$ 且 $\sum|y_i|\le 1$, 在 $y_t\neq 0$ 时取到 $|x_t| = \|\cdot\|_\infty\}$.

(4) $\|X\|_{2,1} = \left(\sum \|X_i\|_2^1\right)^{1/1} = \sum\|X_i\|_2$

$\partial f = \partial\left(\sum\|X_i\|_2\right) = \sum\left(\partial\|X_i\|_2\right) = \{\sum M_i \mid M_i\in\partial\|X_i\|_2\}$

$\therefore \partial f = \{\sum M_i \mid M_i \in \partial\|X_i\|_2\}$ 而 $\partial\|X\|_2 = \begin{cases} \{y \mid \|y\|_2\le 1\} \\ \{y \mid y = \frac{x}{\|x\|_2}\} \end{cases}$

2. (1) ~~先证 $\langle x-y, \partial f(x)-\partial f(y)\rangle \ge 0$~~

~~由定义. $\partial f(x) = \{g \mid f(y)\ge f(x)+\langle g, y-x\rangle\}$~~

~~即 $f(y)\ge f(x)+\langle g, y$~~

由定义: $\partial f(x) = \{g \mid f(y)\ge f(x)+\langle g, y-x\rangle\}$ (*)

取 $x=x_1$, $y=x_2$.

得 $f(x_2)\ge f(x_1)+\langle g_1, x_2-x_1\rangle$

同理 $f(x_1)\ge f(x_2)+\langle g_2, x_1-x_2\rangle$

$\Rightarrow \langle g_1, x_2-x_1\rangle \ge -\langle g_2, x_1-x_2\rangle$

$\therefore \langle g_1-g_2, x_1-x_2\rangle \ge 0$.

(2) 由(*)式.

取 $x = \cancel{\theta x_1+(1-\theta)x_2}\ x_1$, $y=\theta x_1+(1-\theta)x_2$.

$\therefore f(\theta x_1+(1-\theta)x_2)\ge f(x_1)+\langle g_1, (1-\theta)(x_2-x_1)\rangle$ ①

又 $\because f$ is μ-convex

$\therefore f(\theta x_1+(1-\theta)x_2)\le \theta f(x_1)+(1-\theta)f(x_2)-\frac{\mu}{2}\theta(1-\theta)\|x_1-x_2\|^2$ ②

化简 ①,② $\Rightarrow \langle g_1, x_1-x_2\rangle \ge f(x_1)-f(x_2)+\frac{\mu}{2}\theta\|x_1-x_2\|^2$.

同理 $\langle g_2, x_2-x_1\rangle \ge f(x_2)-f(x_1)+\frac{\mu}{2}\theta\|x_1-x_2\|^2$

$\therefore \langle g_1-g_2, x_1-x_2\rangle \ge \mu\theta\|x_1-x_2\|^2$. ③

③式取 $\theta\to 1$ 则得 $\langle g_1-g_2, x_1-x_2\rangle \ge \mu\|x_1-x_2\|^2$

3. $f(x)=\sum_{i=1}^{r}\alpha_i x_i$

从 x 中任取 r 项 $\{x_{a_1}\cdots x_{a_r}\}$, $g(x,a)=\sum_{i=1}^{r}\alpha_i x_{a_i}$

当 $g_i=\alpha_i x_{a_i}$ 时, 此为线性 $\Rightarrow g_i$ is convex.

$\therefore g(x,a)$ is combine of convex $\Rightarrow$ convex

而 $f(x)=\max_{\{a\}} g(x,a)$, 由 C-thm

$\therefore f(x)$ is convex.

4. $f(x) = -\log(-\log(\sum e^{a_i^T x+b_i}))$, $\sum e^{\sim} < 1$.

显有 $A_i = a_i^T x + b_i$ is convex

$g(A_i) = e^{A_i}$ is convex.

$h(g) = \sum g_i$ is convex.

$m(h) = \log h = \log(\sum g_i)$ is convex

$\therefore$ $p(m) = \log(-m)$ is concave

$\therefore$ $f(x) = -p(m)$ is convex #.

5. $f_k(x) = \sum^{k} \sigma_i(x)$

设 $A = \sum \sigma_i u_i v_i^T$

~~$\langle A, B\rangle = \langle \Sigma, U^T B V\rangle \overset{设}{=} \langle \Sigma, G\rangle = \sum \sigma_i g_i$.~~

~~由正设 $G = U^T B V$ 的对角阵且 $g_i \le 1$~~

$k=1$ 时 $f_1(x) = \sigma_1(x)$ is convex

$k>1$ 时 $f_{k+1}(x) = f_k(x) + \sigma_{k+1}$

$\sigma_{k+1} = \sup\{\|XV\| ; |V|=1$ 且 $V \perp V_1 \cdots V_k\}$

显 σ_{k+1} is convex.

由数学归纳法 $f_k(x)$ is convex #.

6. $M_C(x) = \inf\{t>0 \mid t^{-1}x \in C\}$ C is convex.

① 设 $M_C(x_1) = t_1$ $M_C(x_2) = t_2$, $x_1, x_2 \in C$.

$\therefore$ x_1/t_1 , $x_2/t_2 \in C$.

取 $\theta \in [0,1]$

$M_C(\theta x_1 + (1-\theta)x_2) \le t_3$, $\frac{\theta x_1 + (1-\theta)x_2}{t_3} \in C$.

不妨取 t_3 s.t. $t_3 \le t_1, t_2$.

$\therefore M_C(\theta x_1 + (1-\theta)x_2) \le \theta \cdot \frac{x_1}{t_1} + (1-\theta)\frac{x_2}{t_2} = \theta M_C(x_1) - (\theta-1) M_C(x_2)$

$\therefore$ is convex.

② $M_C^*(y) = \sup_x(\langle y, x\rangle - M_C(x))$

$= \sup_x \inf_t (yx - t \mid \frac{x}{t} \in C)$

7. (a) $f(x) = \sum_1^k x_{[i]}$ $f^*(y) = \sup_x \{\langle y, x\rangle - f(x)\}$

设 $g(x,y) = y^T x - f(x)$.

~~① $y_i < 0$ 时 $g(x,y) = x_{[i]} \cdot y_{[i]} - x_{[i]} \to \infty$ 当 $x_{[i]} \to \infty$~~

① $y_i > 1$ 时 $g(x,y) = x_{[i]} \cdot y_i - x_{[i]} \to \infty$. 当 $x_{[i]} \to \infty$.

② $y_i < 0$ 时 $g(x,y) = x_{[i]} \cdot y_i \to \infty$ 当 $x_{[i]} \to \infty$ 时

③ $0 < y_i < 1$ 时. $g(x,y) = x^T y - f(x) = (1^T y - k) \cdot x_{[i]} \le 0$.

取 sup 得 $g = 0$.

$\Rightarrow \boxed{f^*} = \begin{cases} 0 & , x^T y \le f(x), \|y\|_2 \le 1 \\ \infty & , \text{else}. \end{cases}$

(b) $f(x) = \max_i(a_i x + b_i) \to$ 间断点为 $\{\frac{b_{i+1}-b_i}{a_{i+1}-a_i}\}$.

① $a_i \le y \le a_{i+1}$ 时 $f^* = \sup_x\{xy - \underbrace{\max(a_i x + b_i)}_{= a_i x + b_i}\} = x(y-a_i) - b_i$ 且 $x = a_{i+1}$ 时刚好最大.

即 $\boxed{f^*} = -b_i - \frac{b_{i+1}-b_i}{a_{i+1}-a_i}(y-a_i)$.

② $y \notin [a_1, a_m]$ 时, x 也不再存在.

$\therefore \boxed{f^*(y)} = \begin{cases} -b_i - \frac{b_{i+1}-b_i}{a_{i+1}-a_i}(y-a_i) & , y \in [a_i, a_{i+1}] 时. \\ \text{none} & \text{else}. \end{cases}$

(3) $f(x)=x^p \begin{cases} p>1 \\ p<0. \end{cases}$

① $p>1$ 时 取 $\frac{1}{p}+\frac{1}{q}=1 \ (q>1)$

$\boxed{f^*} = \underbrace{\sup(yx-x^p)}_{} = \begin{cases} 0, & y\le 0 \\ (p-1)(y/p)^q, & y>0. \end{cases}$

$(yx-x^p)' = y-px^{p-1}=0 \Rightarrow x=(y/p)^{1/p-1}$

此时 $q=\frac{p}{p-1} \quad \Rightarrow f^* = y\left(\frac{y}{p}\right)^{\frac{1}{1-p}} - \left(\frac{y}{p}\right)^{\frac{p}{p-1}} = p(y/p)^q-(y/p)^q = (p-1)(y/p)^q$

② $p<0$ 时

同理 $f^* = \begin{cases} \\ -\frac{p}{q}\left(-\frac{y}{p}\right)^q. \end{cases}$

(4) $f(x) = -(\prod x_i)^{1/n}$

① 设 $y_k>0$ 取 $x_k=t$ 且其余 $=1$.

则 $x^Ty-f(x)=ty_k+\sum y_i - t^{1/n} \Rightarrow \infty$.

② 设 $(\prod(-y_i))^{1/n}<\frac{1}{n}$. 取 $x_i=t/y_i$

则 $x^Ty-f(x)=-tn-t(\prod-\frac{1}{y_i})^{1/n} \to \infty$.

③ 反之,

$\frac{x^Ty}{n} \ge \left(\prod(-y_ix_i)\right)^{-\frac{1}{n}} \ge \frac{1}{n}(\prod x_i)^{\frac{1}{n}}$.

即 $x^Ty \ge f(x)$, $x_i=\frac{1}{y_i}$

$\therefore f^*(y)=0$.

$\therefore \boxed{f^*} = \begin{cases} 0, & y\le 0 \ (\prod(-y_i))^{\frac{1}{n}} \ge \frac{1}{n} \\ \infty, & \text{else} \end{cases}$

(5) $f(x,t) = -\log(t^2-x^Tx)$, $\|x\|_2<t$.

① 设 $\|y\|_2 \ge u$, 取 $x=s\cdot y$, $t=s(\|x\|_2+1)>s\|y\|$

有 $y^Tx+tu>s\|y\|_2^2-su^2\ge 0$.

$\therefore y^Tx+tu\to\infty$ when $\log(t^2\|x\|_2^2) \to -\infty$.

② $\|y\|_2<u$. 考虑 $g=y^Tx+tu+\log(t^2-x^Tx)$.

$\frac{\partial g}{\partial x}=0; \ \frac{\partial g}{\partial t}=0 \Rightarrow \begin{cases} x=\frac{2y}{u^2-\|y\|^2} \\ t=\frac{-2u}{u^2-\|y\|^2} \end{cases} \Rightarrow f^*=-2+\log 4-\log(\|y\|^2-u^2)$.

$\therefore \boxed{f^*} = -2+2\log 2-\log(\|y\|^2-u^2)$, $\|y\|_2<-u$ 时.

8 (a) $\|x\|_2$: $\text{Prox}_c f = x-c\cdot P_B(\frac{x}{c}) \quad B=\{\|x\|_2\le c\}$.

$\therefore \text{Prox}_c f = \begin{cases} 0, & \|x\|_2\le c \\ x-c\cdot\frac{x}{\|x\|_2}, & \|x\|_2 \ge c. \end{cases}$

(b) $\|x\|_\infty$: $\text{Prox}_c f = x-c\cdot P_B(\frac{x}{c}) \quad B=\{x \mid \|x\|_1\le c\}$.

$\text{Prox}_c f = \begin{cases} [x_i], & |x_i|\le r \\ [\text{sgn}(x_i)\cdot r], & |x_i|>r \end{cases} \quad r=\arg\max_x\{\sum \max(x_i-k,0)\}$

(c) $\|X\|_2$:

$B=\{X \mid \|X\|_* \le c\}$

$\text{Prox}_c f = \begin{cases} 0 & \|X\|_*\le c. \\ X-c\cdot\frac{X}{\|X\|_*} & \|X\|_*>c \end{cases}$

9. 设 C 为 2-order cone.

若 $(x,t)\in C$, $\mathrm{Prox}(x,t)=(x,t)$

else: ~~考虑如此~~: $\min \frac{1}{2}\|x-y\|_2^2+\frac{1}{2}(s-t)^2$, s.t. $\|x\|_2^2\le s$.

由KKT条件: $\begin{cases} x=\frac{y}{1-2\lambda} \\ s=\frac{t}{1+2\lambda} \end{cases}$

∴ ~~原式满足KKT条件~~

$\mathrm{Prox}(x,t)=\langle \frac{(x,\|x\|_2)}{\sqrt{2}\|x\|_2},(x,t)\rangle\cdot\frac{(x,\|x\|_2)}{\sqrt{2}\|x\|_2}=\frac{\|x\|_2+t}{2\|x\|_2}(x,\|x\|_2)$.

10. (1) $\mathrm{Prox}\, d(x)$:

① $x\in C$, $\mathrm{Prox}\, d(x)=x$

② $x\notin C$, 设 $p(x)=\arg\inf\|x-y\|_2$.

而 $\inf\{d(w)+\frac{1}{2c}\|w-x\|^2\}=\min g(y)=y+\frac{1}{2c}(d(x)-y)^2$.

∴ 当 $d(x)>c$ 或 $d(x)=0$ 时 $g(y)$ 取得 $d(x)-c$.

$\mathrm{Prox}=\begin{cases} x & x\in C \\ d(x)-c & x\notin C \end{cases}$.

(2) $\mathrm{Prox}\,\frac{1}{2}d^2(x)$.

类似 $\mathrm{Prox}=\begin{cases} x & x\in C \\ \frac{1}{c+1}x+\frac{c}{c+1}p(x) & x\notin C. \end{cases}$ 其中 c,p 如(1)所设.

11. (1) $f(x)=g(x)+a^Tx$

$\mathrm{Prox}\,f=\arg\min\{\frac{1}{2c}\|z-x\|^2+g(z)+a^Tz\}$.

$=\arg\min\{\frac{1}{2c}\|z\|_2^2-\langle x-ca,z\rangle+g(z)\}$

$=\arg\min\{\frac{1}{2c}\|z-(x-ca)\|^2$.

∴ $=\mathrm{Prox}\,g(x-ca)$ #.

(2) $f(x)=g(x)+\frac{1}{2\mu}\|x-a\|^2$

$\mathrm{Prox}_c f=\arg\min\{\frac{1}{2}\|z-cx\|^2+g(z)+\frac{1}{2\mu}\|z-a\|^2\}$.

$=\arg\min\{\frac{1}{2}\|z\|^2-\lambda\langle z,\frac{x}{c}+\frac{a}{\mu}\rangle+\lambda g(z)\}$.

$=\arg\min\{\frac{1}{2}\|z-\lambda(\frac{x}{c}+\frac{a}{\mu})\|^2+\lambda g z\}$.

$=\mathrm{Prox}_\lambda g(\frac{\lambda}{c}\cdot x+\frac{\lambda}{\mu}\cdot a)$