Suggested Solutions to Assignment 3

1. (a) f^{**} is $\frac{1}{\sigma}$-smooth, and $f^{**} = f$ since f is proper closed convex.

   (b) Notice that $f^* = (f^* - \frac{\sigma}{2}\|\cdot\|^2) + \frac{\sigma}{2}\|\cdot\|^2$ and $f^* - \frac{\sigma}{2}\|\cdot\|^2$ is proper convex, then by Theorem 4.17,

$$\begin{aligned} f = f^{**} &= \left((f^* - \frac{\sigma}{2}\|\cdot\|^2) + \frac{\sigma}{2}\|\cdot\|^2\right)^* \\ &= \left(f^* - \frac{\sigma}{2}\|\cdot\|^2\right)^* \Box \left(\frac{\sigma}{2}\|\cdot\|^2\right)^* \\ &= \left(f^* - \frac{\sigma}{2}\|\cdot\|^2\right)^* \Box \left(\frac{1}{2\sigma}\|\cdot\|^2\right) \end{aligned}$$

   Finally, $\left(f^* - \frac{\sigma}{2}\|\cdot\|^2\right)^*$ is convex since $f^* - \frac{\sigma}{2}\|\cdot\|^2$ is proper convex.

   (c) Let $h_1(x) = x^3 - x^2$ if $x \geq 0$, while $h(x) = \infty$ otherwise. Let $h_2(x) = x^2$. Define $f = h_1 \Box h_2$.

2. (a) By the second prox theorem, for any x,

$$\begin{aligned} & x - \mathrm{prox}_f(x) \in \partial f\left(\mathrm{prox}_f(x)\right) \\ \Leftrightarrow & x \in \mathrm{prox}_f(x) + \partial f\left(\mathrm{prox}_f(x)\right) = (I + \partial f)\left(\mathrm{prox}_f(x)\right) \\ \Leftrightarrow & \mathrm{prox}_f(x) = (I + \partial f)^{-1}(x). \end{aligned}$$

   (b)

$$\begin{aligned} \nabla M_f^\mu(x) &= \frac{1}{\mu}\left(x - \mathrm{prox}_{\mu f}(x)\right) = \frac{1}{\mu}\left(I - \mathrm{prox}_{\mu f}\right)(x) \\ &= \frac{1}{\mu}\left(I - (I + \partial(\mu f)^{-1})\right)(x) = \frac{1}{\mu}\left(I - (I + \mu\partial f)^{-1}\right)(x). \end{aligned}$$

3.

4. Hint:

   (b) Characterize the positive definiteness of $\lambda A + I$. Hence we need $\lambda \geq 0$ if A is positive semidefinite and $0 \leq \lambda < \frac{-1}{\lambda_{\min}(A)}$ otherwise. In this case, $\mathrm{prox}_{\lambda f} = (\lambda A + I)^{-1}(x + \lambda b)$.

5. Recall Theorem 6.13 for

$$f(x) = \left(g - \frac{\sigma}{2}\|\cdot\|^2\right)(x) + \frac{\sigma - c}{2}\|x\|^2 + \langle e, x\rangle,$$

   Then

$$\mathrm{prox}_f(x) = \mathrm{prox}_{\frac{1}{\sigma - c + 1}\left(g - \frac{\sigma}{2}\|\cdot\|^2\right)}\left(\frac{x - e}{\sigma - c + 1}\right).$$

6. (a) For any x, $M_f\mu(x) \geq \inf_{u\in\mathbb{R}^n} f(u) > -\infty$ and choose $u_x \in \mathrm{dom}(f)$, then $M_f^\mu \leq f(u_x) + \frac{1}{2\mu}\|x - u_x\|^2 < \infty$.

   (b) $M_f^\mu(x) \leq f(x) + \frac{1}{2\mu}\|x - x\|^2 = f(x)$.

7. (i). $(a) \Rightarrow (b)$.

   (ii). $(b) \Rightarrow (c)$. We know $\exists\, \alpha, \beta, \gamma \in \mathbb{R}$, $a \in \mathbb{R}^n$ such that $q(u) = \alpha \|u\|^2 + \beta a^T u + \gamma$. Then choose $\mu = \frac{1}{4|\alpha|}$.

$$\begin{aligned} f(u) + \frac{1}{2\mu}\|u\|^2 \geq& \alpha \|u\|^2 + \beta a^T u + \gamma + \frac{1}{2\mu}\|u\|^2 \\ \geq& |\alpha| \|u\|^2 + \beta a^T u + \gamma. \end{aligned}$$

   (iii). $(c) \Rightarrow (d)$. $\exists c$ constant such that $f(u) + \frac{1}{2\mu}\|u\|^2 > c$. Then

$$\liminf_{\|x\|\to\infty} \frac{f(x)}{\|x\|^2} + \frac{1}{2\mu} \geq \liminf_{\|x\|\to\infty} \frac{c}{\|x\|^2} = 0.$$

   (iv). $(d) \Rightarrow (a)$. Define

$$\liminf_{\|x\|\to\infty} \frac{f(x)}{\|x\|^2} = \alpha$$

   and $\mu = 1/6|\alpha|$. Notice that there exists $M > 0$ such that $f(u)/\|u\|^2 \geq -2|\alpha|$ for which $u \notin \mathbb{B}_M(0)$. Then

$$\begin{aligned} M_f^\mu(x) =& \inf_{u\in\mathbb{R}^n}\left\{f(u) + \frac{1}{2\mu}\|u-x\|^2\right\} \\ =& \min\left\{\inf_{u\in B_M(0)}\left\{f(u) + \frac{1}{2\mu}\|u-x\|^2\right\}, \inf_{u\notin B_M(0)}\left\{f(u) + \frac{1}{2\mu}\|u-x\|^2\right\}\right\} \\ >& -\infty, \end{aligned}$$

   since

$$\inf_{u\in B_M(0)}\left\{f(u) + \frac{1}{2\mu}\|u-x\|^2\right\} > -\infty \text{ by the Weierstrass theorem,}$$

   and

$$\begin{aligned} &\inf_{u\notin B_M(0)}\left\{f(u) + \frac{1}{2\mu}\|u-x\|^2\right\} \\ \geq& \inf_{u\notin B_M(0)}\left\{-2|\alpha|\|u\|^2 + 3|\alpha|\|u-x\|^2\right\} > -\infty. \end{aligned}$$

8. Fix $\bar{\mu} > 0, \bar{x} \in \mathbb{R}^n$. Let $\bar{y} = \varphi(\bar{\mu}, \bar{x})$.

   For any $\{x^k\} \to \bar{x}$, $\{\mu_k\} \to \bar{\mu}$, define $y^k = \varphi(\mu_k, x^k)$. Just show $y^k \to \bar{y}$.

   Notice that

$$f(y^k) + \frac{1}{2\mu_k}\|y^k - x^k\|^2 \leq f(\bar{y}) + \frac{1}{2\mu_k}\|\bar{y} - x^k\|^2.$$

   Take the upper limit,

$$\limsup_{k\to\infty} f(y^k) + \frac{1}{2\bar{\mu}}\|y^k - \bar{x}\|^2 \leq f(\bar{y}) + \frac{1}{2\bar{\mu}}\|\bar{y} - \bar{\mu}\|^2 = M_f^{\bar{\mu}}(\bar{x}) < \infty. \qquad (\star)$$

   If $\|y^k\| \to \infty$, then

$$\limsup_{k\to\infty} f(y^k) + \frac{1}{4\bar{\mu}}\|y^k - \bar{x}\|^2 \leq \limsup_{k\to\infty} f(y^k) + \frac{1}{2\bar{\mu}}\|y^k - \bar{x}\|^2 - \lim_{k\to\infty}\frac{1}{4\bar{\mu}}\|y^k - \bar{x}\|^2 = -\infty.$$

   While

$$f(y^k) + \frac{1}{4\bar{\mu}}\|y^k - \bar{x}\|^2 > M_f^{2\bar{\mu}}(\bar{x}),$$

i.e.,

$$\left\{f\left(y^k\right)+\frac{1}{4\bar{\mu}}\left\|y^k-\bar{x}\right\|^2\right\}$$

has a finite lower bound. It's a contradiction with ($\star$). Hence $\left\{y^k\right\}$ is bounded. Let y^* be a limit point of $\left\{y^k\right\}$, then by ($\star$) and the lower-semicontinuity of f, we have

$$M_f^{\bar{\mu}}(\bar{x}) \geq \limsup_{k\to\infty} f\left(y^k\right)+\frac{1}{2\bar{\mu}}\left\|y^k-\bar{x}\right\|^2 \geq \liminf_{k\to\infty} f\left(y^k\right)+\frac{1}{2\bar{\mu}}\left\|y^*-\bar{x}\right\|^2 \geq f\left(y^*\right)+\frac{1}{2\bar{\mu}}\left\|y^*-\bar{x}\right\|^2 \geq M_f^{\bar{\mu}}(\bar{x}),$$

Then $y^*=\bar{y}$.

9.

10. (a) Apply the second projection theorem and the definition of normal cone.
    (b) Notice that

$$\operatorname{dist}\left(\bar{x}+\alpha v, C\right)=\alpha\|v\|=\|(\bar{x}+\alpha v)-\bar{x}\|$$

implies that $P_C\left(\bar{x}+\alpha v\right)=\bar{x}$. Hence

$$N_C\left(\bar{x}\right)=N_C^{\text{prox}}(\bar{x}).$$

11. (a) fundamental prox-grad inequality.
    (b) Plugging $y=\operatorname{proj}(x;X^*)$ into the fundamental prox-grad inequality, we have

$$\begin{aligned}\langle G_L(x), x-\operatorname{proj}(x;X^*)\rangle &\geq F(T_L(x))-F(\operatorname{proj}(x;X^*))+\frac{1}{2L}\|G_L(x)\|^2\\ &\geq F(T_L(x))-F_{\text{opt}}+\frac{\alpha}{2L}\operatorname{dist}^2(x,X^*)\\ &\geq \frac{\alpha}{2L}\operatorname{dist}^2(x,X^*).\end{aligned}$$

12. Define $F_\mu=\mu h+(1-\mu)f$, it's easy to see that F_μ is convex and L-smooth. The fundamental prox-grad inequality implies that

$$F_\mu(x)-F_\mu\left(y-\frac{1}{L}\nabla F_\mu(y)\right)\geq\frac{L}{2}\left\|x-\left(y-\frac{1}{L}\nabla F_\mu(y)\right)\right\|^2-\frac{L}{2}\|x-y\|^2.$$

For any $n\geq 0$, substituting $\mu=\mu_k$, $x=x^*$ and $y=x^n$ in the above inequality, we obtain

$$F_{\mu_n}(x^*)-F_{\mu_n}(x^{n+1})\geq\frac{L}{2}\left\|x^*-x^{n+1}\right\|^2-\frac{L}{2}\left\|x^*-x^n\right\|^2.$$

Summing the above inequality over $n=0,1,\cdots,k-1$ we obtain

$$\begin{aligned}&\frac{L}{2}\left\|x^*-x^k\right\|^2-\frac{L}{2}\left\|x^*-x^0\right\|^2\\ \leq&\sum_{n=0}^{k-1}\left(F_{\mu_n}(x^*)-F_{\mu_n}(x^{n+1})\right)\\ =&\sum_{n=0}^{k-1}\left(\mu_n\left(h(x^*)-h(x^{n+1})\right)+(1-\mu_n)\left(f(x^*)-f(x^{n+1})\right)\right)\\ \leq&\left[h(x^*)-\min_{n=0,1,\cdots,k-1}h(x^n)\right]\left(\sum_{n=0}^{k-1}\mu_n\right)+\left[f(x^*)-\min_{n=0,1,\cdots,k-1}f(x^n)\right]\left(k-\sum_{n=0}^{k-1}\mu_n\right).\end{aligned}$$

Thus

$$\left[\min_{n=0,1,\cdots,k-1}h(x^n)-h(x^*)\right]\left(\sum_{n=0}^{k-1}\mu_n\right)+\left[\min_{n=0,1,\cdots,k-1}f(x^n)-f(x^*)\right]\left(k-\sum_{n=0}^{k-1}\mu_n\right)\leq\frac{L}{2}\left\|x^*-x^0\right\|^2.$$

Consequently,

$$\min_{n=0,1,\cdots,k-1} f(x^n) - f(x^*) \leq \frac{L\left\|x^* - x^0\right\|^2}{2\left(k - \sum_{n=0}^{k-1}\mu_n\right)} \leq \frac{L\left\|x^* - x^0\right\|^2}{2\left(k - \frac{(k-1)^\alpha}{\alpha} - \frac{1}{\alpha}\right)} = O\left(\frac{1}{k}\right),$$

and

$$\min_{n=0,1,\cdots,k-1} h(x^n) - h(x^*) \leq \frac{L\left\|x^* - x^0\right\|^2}{2\sum_{n=0}^{k-1}\mu_n} \leq \frac{\alpha L\left\|x^* - x^0\right\|^2}{2(k-1)^\alpha} = O\left(\frac{1}{k^\alpha}\right),$$