Exercise 06

$\cancel{A}$. $E_\gamma f(x) = \min_z \{ f(z) + \frac{1}{2\gamma} \|z-x\|_2^2 \} \le f(x)$ *Moreau envelope*

$\text{prox}_\gamma f(x) = \arg\min_z \{ f(z) + \frac{1}{2\gamma} \|z-x\|_2^2 \}$.

(A). $\hat{f}(x;x_0) = f(x_0) + (x-x_0)' \nabla f(x_0)$.

$E_\gamma \hat{f}(x) = \min_z \{ f(x_0) + (z-x_0)' \nabla f(x_0) + \frac{1}{2\gamma} \|z-x\|_2^2 \}$

Objective function $\tilde{f} = f(x_0) + (z-x_0)' \nabla f(x_0) + \frac{1}{2\gamma} \sum_{i=1}^{p} (z_i - x_i)^2$

$\tilde{f}'_z = \nabla f(x_0) + \frac{1}{\gamma}(z-x) = 0$. $z = x - \gamma \nabla f(x_0)$. $z|_{x=x_0} = x_0 - \gamma \nabla f(x_0)$

$\therefore \text{prox}_\gamma \hat{f}(x;x_0) = x_0 - \gamma \nabla f(x_0)$ is identical to a gradient-descent algorithm with step size γ, and direction $\nabla f(x_0)$, starting from x_0.

B) ① Objective function $\tilde{l}(z) = l(x) + \frac{\gamma}{2} \|z-x\|_2^2 = \frac{1}{2} z' P z - q' z + r + \frac{\gamma}{2} \sum_{i=1}^{p} (z_i - x_i)^2$

$\frac{d\tilde{l}(x)}{dz} = Pz^* - q^* + \gamma(z-x) = 0 \Rightarrow (P+\gamma I) z^* = \gamma x + q \Rightarrow z^* = (P+\gamma I)^{-1} (\gamma x + q)$

$\therefore \text{prox}_{\frac{1}{\gamma}} l(x) = (P+\gamma I)^{-1} (\gamma x + q)$

② $L(y|x) = (2\pi)^{-\frac{p}{2}} |\Omega^{-1}|^{-\frac{p}{2}} e^{-\frac{1}{2}(y-Ax)' \Omega (y-Ax)}$.

$-l(y|x) \propto \frac{1}{2}(y-Ax)'\Omega(y-Ax) = \frac{1}{2} y'\Omega y - \frac{1}{2} x'A'\Omega y - \frac{1}{2} y'\Omega Ax + \frac{1}{2} x'A'\Omega Ax$

$= \frac{1}{2} y'\Omega y - x'A'\Omega y + \frac{1}{2} x'A'\Omega Ax = \frac{1}{2} y' P y - q' y + r$

Here $P = \Omega$, $q = \Omega A x$, $r = \frac{1}{2}(x'A'\Omega Ax - n\log(2\pi) - \log|\Omega|)$.

Symmetrix

(c). $\phi(x) = \tau\|x\|_1$

Objective function $\tilde{\Phi}(x) = \tau\|z\|_1 + \frac{1}{2\gamma}\|z - x\|_2^2$

$\arg\min_z \tilde{\Phi}(x) = \arg\min_z \{\gamma\tau\|z\|_1 + \frac{1}{2}\|z - x\|_2^2\}$.

$= \arg\min_z S_{\gamma\tau}(x)$ with $\theta = z$ as in Exercise 05

$\therefore \operatorname{prox}_\gamma \phi(x) = \operatorname{sgn}(x)(|x| - \gamma\tau)$.

The proximal gradient method

(A). $\operatorname{prox}_\gamma \phi(u) = \operatorname{prox}_\gamma \phi(x_0 - \gamma \nabla l(x_0))$

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}\|x - [x_0 - \gamma\nabla l(x_0)]\|_2^2\}$.

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}\|(x - x_0) + \gamma\nabla l(x_0)\|_2^2\}$

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}[(x - x_0) + \gamma\nabla l(x_0)]'[(x - x_0) + \gamma\nabla l(x_0)]\}$

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}[(x - x_0)'(x - x_0) + 2\gamma(x - x_0)'\nabla l(x_0) + \gamma^2 \nabla l^2(x_0)]\}$

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}\|x - x_0\|_2^2 + (x - x_0)'\nabla l(x_0) + \frac{1}{2}\nabla l^2(x_0)\}$.

Since $\frac{1}{2}\nabla l^2(x_0)$ is indep. of x, and $l(x_0)$ is also indep. of x.

$= \arg\min_x \{\phi(x) + \frac{1}{2\gamma}\|x - x_0\|_2^2 + (x - x_0)'\nabla l(x_0) + l(x_0)\}$

$= \arg\min_x \{\phi(x) + \tilde{l}(x; x_0)\}$.