Solutions (midterm)

1. T F T T T

2. (a) F.

   Let Ω be an open set. Then it is locally closed around any point $x \in \Omega$.

   (b) F

   Notice that the unique minimizer is $x^* = Q^{-1}b$. For any starting point x^0, the first iteration gives

$$x^1 = x^0 - \left[\nabla^2 f\left(x^0\right)\right]^{-1} \nabla f\left(x^0\right) = x^0 - Q^{-1}\left(Qx^0 - b\right) = Q^{-1}b = x^*.$$

5. (a) Suppose $\exists d \in T(\bar{x})$ with $d^T \nabla f(\bar{x}) < 0$. Then $\exists d^k \to d, t^k \downarrow 0$ with $\left\{\bar{x} + t_k d^k\right\}_{k \geq 0} \subset \Omega$. Then $\left(d^k\right)^T \nabla f(\bar{x}) < 0$ and

$$f\left(\bar{x} + t_k d^k\right) = f(\bar{x}) + t_k \left(d^k\right)^T \nabla f(\bar{x}) + o(t_k) < f(\bar{x})$$

   for k sufficiently large. Then $\bar{x}$ is not locally optimal.

   (b) Notice that $\bar{d} = 0 \in T(\bar{x})$. Then

$$0 = \bar{d}^T \nabla f(\bar{x}) \geq \min_{d \in T(\bar{x})} d^T \nabla f(\bar{x}).$$

   On the other hand, by (a)

$$\min_{d \in T(\bar{x})} d^T \nabla f(\bar{x}) \geq 0 = \bar{d}^T \nabla f(\bar{x}).$$

   That is to say $\bar{d} = 0$ is a minimizer.

   (c) Recall

$$f(x) - f(x^*) = (x - x^*)^T \nabla f(x^*) + o(\|x - x^*\|).$$

   Notice $\exists \delta > 0$ such that for any $x \in \mathbb{B}_\delta(x^*)$, we have

$$o(\|x - x^*\|) \geq -\frac{1}{2}\left\|(x - x^*)^T \nabla f(x^*)\right\|.$$

Then $\forall x^* \neq x \in \mathbb{B}_\delta(x^*) \cap \Omega$, we have $x - x^* \in T(x^*)$ (think of why) and

$$f(x) - f(x^*) = (x - x^*)^T \nabla f(x^*) - \frac{1}{2}\left\|(x - x^*)^T \nabla f(x^*)\right\| \geq \frac{1}{2}\eta \|x - x^*\| > 0.$$

That is to say, x^* is a strict local minimizer.

6. (a) By the strong convexity of f, we have

$$f(x) \leq f\left(y^k\right) + \left\langle \nabla f\left(y^k\right), x - y^k \right\rangle + \frac{\mu}{2}\left\|x - y^k\right\|^2.$$

Then

$$\varphi_{k+1}(x) \leq (1-\theta)\,\varphi_k(x) + \theta f(x).$$

(b) Notice that

$$\varphi_0(x) \leq \left[1 - (1-\theta)^0\right] f(x) + (1-\theta)^0 \varphi_0(x).$$

Suppose for k, we have

$$\varphi_k(x) \leq \left[1 - (1-\theta)^k\right] f(x) + (1-\theta)^k \varphi_0(x).$$

Then for $k+1$,

$$\begin{aligned}
\varphi_{k+1}(x) &\leq (1-\theta)\,\varphi_k(x) + \theta f(x) \\
&\leq (1-\theta)\left[1 - (1-\theta)^k\right] f(x) + (1-\theta)^k \varphi_0(x) + \theta f(x) \\
&= \left[1 - (1-\theta)^{k+1}\right] f(x) + (1-\theta)^{k+1} \varphi_0(x).
\end{aligned}$$

(c) By hint, suppose φ_k has the form $\varphi_k(x) = \varphi_k^* + \frac{\mu}{2}\left\|x - z^k\right\|^2$ (think of why). Then

$$\begin{aligned}
\varphi_{k+1}^* =& \varphi_{k+1}\left(z^{k+1}\right) \\
=& (1-\theta)\left(\varphi_k^* + \frac{\mu}{2}\left\|z^{k+1} - z^k\right\|^2\right) + \theta\left[f\left(y^k\right) + \left\langle \nabla f\left(y^k\right), z^{k+1} - y^k\right\rangle + \frac{\mu}{2}\left\|z^{k+1} - y^k\right\|^2\right] \\
=& (1-\theta)\,\varphi_k^* + \theta f\left(y^k\right) - \frac{1}{2L}\left\|\nabla f\left(y^k\right)\right\|^2 + \theta(1-\theta)\left[\left\langle \nabla f\left(y^k\right), z^k - y^k\right\rangle + \frac{\mu}{2}\left\|z^k - y^k\right\|^2\right],
\end{aligned}$$

where the last equality holds by

$$z^{k+1} = (1-\theta) z^k + \theta\left(y^k - \frac{1}{\mu}\nabla f\left(y^k\right)\right).$$

Observe that φ_k has the quadratic form. Let $w^k = \arg\min \varphi_k(x)$ and so

$$\varphi_k(x) = \varphi_k^* + \frac{\mu_k}{2}\left\|x - z^k\right\|^2.$$

Then

$$
\begin{aligned}
\varphi_{k+1}(x) =& \varphi_{k+1}^* + \frac{\mu_{k+1}}{2} \left\|x - w^{k+1}\right\|^2. \\
\varphi_{k+1}(x) =& (1-\theta)\left(\varphi_k^* + \frac{\mu_k}{2}\left\|x - w^k\right\|^2\right) + \theta\left[f\left(y^k\right) + \left\langle \nabla f\left(y^k\right), x - y^k\right\rangle + \frac{\mu}{2}\left\|x - y^k\right\|^2\right].
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mu_0 = \mu, \\
&\mu_{k+1} = (1-\theta)\,\mu_k + \theta\mu, \\
&0 = \nabla\varphi_{k+1}\left(w^{k+1}\right) = (1-\theta)\,\mu_k\left(w^{k+1} - w^k\right) + \theta\nabla f\left(y^k\right) + \theta\mu\left(w^{k+1} - y^k\right),
\end{aligned}
$$

i.e.,

$$
\begin{aligned}
\mu_k \equiv& \mu, w^0 = x^0 \\
w^{k+1} =& \frac{(1-\theta)\,\mu_k}{[(1-\theta)\,\mu_k + \theta\mu]} w^k + \frac{\theta\mu_k}{[(1-\theta)\,\mu_k + \theta\mu]}\left(y^k - \frac{1}{\mu}\nabla f\left(y^k\right)\right) \\
=& (1-\theta)\, w^k + \theta\left(y^k - \frac{1}{\mu}\nabla f\left(y^k\right)\right)
\end{aligned}
$$

(d) sufficient descent.

(e) Notice that

$$
z^0 = x^0 = \frac{1+\theta}{\theta}x^0 - \frac{1}{\theta}x^0 = \frac{1+\theta}{\theta}y^0 - \frac{1}{\theta}x^0.
$$

Suppose for k, we have

$$
z^k = \frac{1+\theta}{\theta}y^k - \frac{1}{\theta}x^k.
$$

Then for $k+1$,

$$
\begin{aligned}
z^{k+1} =& (1-\theta)\, z^k + \theta\left(y^k - \frac{1}{\mu}\nabla f\left(y^k\right)\right) \\
=& (1-\theta)\left(\frac{1+\theta}{\theta}y^k - \frac{1}{\theta}x^k\right) + \theta\left(y^k - \frac{1}{\mu}\nabla f\left(y^k\right)\right) \\
=& \frac{1}{\theta}y^k - \frac{1-\theta}{\theta}x^k - \frac{\theta}{\mu}\nabla f\left(y^k\right),
\end{aligned}
$$

and

$$\begin{aligned}\frac{1+\theta}{\theta}y^{k+1}-\frac{1}{\theta}x^{k+1}=&\frac{1+\theta}{\theta}\left(x^{k+1}+\frac{1-\theta}{1+\theta}\left(x^{k+1}-x^{k}\right)\right)-\frac{1}{\theta}x^{k+1}\\ =&\frac{1}{\theta}x^{k+1}-\frac{1-\theta}{\theta}x^{k}\\ =&\frac{1}{\theta}y^{k}-\frac{1}{\theta L}\nabla f\left(y^{k}\right)-\frac{1-\theta}{\theta}x^{k}.\end{aligned}$$

Then by $\mu=\theta^2 L$, we have

$$z^{k+1}=\frac{1+\theta}{\theta}y^{k+1}-\frac{1}{\theta}x^{k+1}.$$

(f) Notice that

$$\varphi_0^*\geq f\left(x^0\right).$$

Suppose for k, we have

$$\varphi_k^*\geq f\left(x^k\right).$$

Then for $k+1$,

$$\begin{aligned}\varphi_{k+1}^*\geq&\left(1-\theta\right)f\left(x^k\right)+\theta f\left(y^k\right)-\frac{1}{2L}\left\|\nabla f\left(y^k\right)\right\|^2+\theta\left(1-\theta\right)\left\langle\nabla f\left(y^k\right),\frac{y^k-x^k}{\theta}\right\rangle\\ =&\left(1-\theta\right)\left[f\left(x^k\right)+\left\langle\nabla f\left(y^k\right),y^k-x^k\right\rangle\right]+\theta f\left(y^k\right)-\frac{1}{2L}\left\|\nabla f\left(y^k\right)\right\|^2\\ \geq&\left(1-\theta\right)f\left(y^k\right)+\theta f\left(y^k\right)-\frac{1}{2L}\left\|\nabla f\left(y^k\right)\right\|^2\\ \geq&f\left(x^{k+1}\right).\end{aligned}$$

(g)

$$\begin{aligned}f\left(x^k\right)-f\left(x^*\right)\leq&\varphi_k^*-f\left(x^*\right)\leq\varphi\left(x^*\right)-f\left(x^*\right)\\ \leq&\left[1-\left(1-\theta\right)^k\right]f(x^*)+\left(1-\theta\right)^k\varphi_0(x^*)-f\left(x^*\right)\\ \leq&\left(1-\theta\right)^k\left[\varphi\left(x^*\right)-f\left(x^*\right)\right].\end{aligned}$$