**1** Note that for any point $x^*$ such that $\nabla f(x^*) = 0$

$$(x(t+1) - x^*) \Lambda^{-1} (x(t+1) - x^*)$$

$$= (x(t) - x^*) \Lambda^{-1} (x(t) - x^*) + \gamma^2 \nabla f(x(t))^T \Lambda \nabla f(x(t))$$

$$- 2\gamma \nabla f(x(t))^T [x(t) - x^*]$$

$$\leq (x(t) - x^*) \Lambda^{-1} (x(t) - x^*) + \gamma^2 \nabla f(x(t))^T \Lambda \nabla f(x(t))$$

$$- \frac{2\gamma}{L} \|\nabla f(x(t))\|^2$$

Since $\lambda_{max}$ is the largest eigenvalue of $\Lambda$,

$$\nabla f(x(t))^T \Lambda \nabla f(x(t)) \leq \lambda_{max} \|\nabla f(x(t))\|^2$$

$$\therefore \quad (x(t+1) - x^*) \Lambda^{-1} (x(t+1) - x^*)$$

$$\leq (x(t) - x^*) \Lambda^{-1} (x(t) - x^*)$$

$$- \left(\frac{2\gamma}{L} - \gamma^2 \lambda_{max}\right) \|\nabla f(x(t))\|^2$$

If $0 < \gamma < \frac{2}{L\lambda_{max}}$, $\frac{2\gamma}{L} - \gamma^2 \lambda_{max} > 0$, we then have

$$(x(t+1) - x^*) \Lambda^{-1} (x(t+1) - x^*)$$

$$\leq (x(0) - x^*) \Lambda^{-1} (x(0) - x^*)$$

$$- \sum_{k=0}^{t} \left(\frac{2\gamma}{L} - \gamma^2 \lambda_{max}\right) \cdot \|\nabla f(x(k))\|^2 \qquad (*)$$

Since $(x(t+1) - x^*) \Lambda^{-1} (x(t+1) - x^*)$ is bounded from below, we have $\lim_{t\to+\infty} \nabla f(x(t)) = 0$. It remains to show that $x(t)$ converges. Let $\{X_{t_n}\}$ be a subsequence of $\{X_t\}$ that converges to a limit point $X_0$. Then $\nabla f(X_0) = 0$, i.e. $X_0$ is also a minimizer of $f$. Replace $x^*$ by $x_0$ in $(*)$, we have

$$\lim_{t\to+\infty} (x(t) - X_0) \Lambda^{-1} (x(t) - X_0) = \lim_{n\to+\infty} (X(t_n) - X_0) \Lambda^{-1} (X(t_n - X_0)) = 0.$$

Problem 2

Let $f(w) = \|Aw - y\|_2^2 + \lambda \|w\|_2$

$= w^T A^T A w + y^T y - 2y^T A w + \lambda w^T w.$

Then,

$$\nabla f(w) = 2A^T A w - 2A^T y + 2\lambda w$$

(a) For smoothness, we have

$$\|\nabla f(w) - \nabla f(w')\|_2$$

$$= \|(2A^T A + 2\lambda)(w - w')\|_2$$

$$\leq \|2A^T A (w - w')\|_2 + 2\lambda \|w - w'\|_2$$

$$\leq (2\lambda_{max} + 2\lambda) \|w - w'\|_2$$

Hence, $f$ is always smooth with

$$L = 2\lambda_{max} + 2\lambda$$

(b) If $\lambda = 0$, then

$$\|\nabla f(w) - \nabla f(w')\|$$

$$= \|2A^T A (w - w')\|_2$$

If $\lambda_{min}(A^T A) = 0$, then there exists

$w - w'$ such that $\|\nabla f(w) - \nabla f(w')\| = 0$

but $\|w - w'\|_2 > 0$. Thus, the func. $f$ may not be strongly convex.

However, if $\lambda > 0$, then

$$\|\nabla f(w) - \nabla f(w')\|_2$$

$$= \| 2(A^T A + \lambda)(w - w')\|_2$$

$$\geq 2(\lambda_{min} + \lambda) \|w - w'\|_2$$

Hence, $f$ is strongly convex with $\alpha = \lambda_{min} + \lambda$.