PROBLEM 10-1

(a)

we have
$$\begin{cases} x^{LS} = A^{+}y & A^T A x^{LS} = A^T y \\ A^{+} = V\Sigma' U^T & \Sigma'_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ \begin{cases} 1/\Sigma_{ii} & \text{if } \Sigma_{ii} \neq 0 \\ 0 & \text{if } \Sigma_{ii} = 0 \end{cases} \end{cases} \\ A = U\Sigma V^T \end{cases}$$

Recall from HW8: if $\text{rank}(A) = r$,
a basis of $\text{Ker}(A)$ is $(v_{r+1}, \dots v_m)$. Take $k \in \{r+1, \dots, m\}$

$$v_k^T x^{LS} = v_k A^{+} y = v_k^T V \Sigma' U^T y$$

$$v_k^T V = (\langle v_k, v_1 \rangle, \langle v_k, v_2 \rangle, \dots \langle v_k, v_m \rangle) \quad \text{row vector}$$
$$= (0, \dots, \underset{k^{th}\text{ position}}{1}, 0 \dots 0) \in \mathbb{R}^{1\times m}$$

$$\underbrace{(0, \dots, 1, \dots 0)}_{\mathbb{R}^{1\times m}} \underbrace{\begin{pmatrix} 1/\Sigma_{11} & & & & \\ & \ddots & & & \\ & & 1/\Sigma_{r} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}}_{\mathbb{R}^{m\times n}} U^T y$$

$$= \underbrace{(0, \dots 0)}_{0 \text{ vector in } \mathbb{R}^n} U^T y = 0$$

So any basis vector of $\text{Ker}\, A$ is orthogonal to x^{LS}

by using linear combinations: x^{LS} is orthogonal to $\text{Ker}(A)$.

(b) All the solutions are of the form

$$x = x^{LS} + v \quad \text{for some } v \in \ker(A) \quad \text{(lecture fact)}$$

So $\|x\|^2 = \|x^{LS}\|^2 + \|v\|^2$ by Pythagorean theorem

$\Rightarrow \|x\|^2 \geq \|x^{LS}\|^2 \Rightarrow$ x^{LS} is the solution with smallest norm.

PROBLEM 10.2

$f(x) = \|Ax - y\|^2 + \lambda\|x\|^2$ is a λ-strongly convex function since $g(x) = \|Ax-y\|^2$ is convex (see HW9).

- HW9 gives us that $f(x)$ is then strictly convex (9.3)a.
- Then the minizer is unique (9.1)(b).

The gradient of $f(x)$ is given by:

$$\nabla f(x) = \nabla\|Ax-y\|^2 + \lambda\nabla\|x\|^2 = 2(A^TAx - A^Ty) + \lambda 2x \quad \text{(HW9)}$$

$$\Rightarrow \nabla f(x) = 0 \iff A^TAx + \lambda Id\, x = A^Ty$$

$$\iff \underbrace{(A^TA + \lambda Id)}_{} x = A^Ty$$

invertible as soon as $\lambda > 0$ as spectrum as only strictly possitive eigenvalue

$$\iff x = (A^TA + \lambda Id)^{-1}A^Ty.$$

PROBLEM 10.4

$$\|M\|_{sp} = \max_{\|x\|=1} \|Ax\|$$

(a) if $x \neq 0$

$$\|Ax\| = \|A \underbrace{\frac{x}{\|x\|}}_{\text{norm 1 vector}}\| \|x\| \quad \text{and} \quad \|A \frac{x}{\|x\|}\| \leq \|A\|_{sp} \text{ by def.}$$

$$\Rightarrow \|Ax\| \leq \|A\|_{sp} \|x\|.$$

if $x = 0 \rightarrow \|Ax\| = 0$ so ok.

(b) $$\|AB\|_{sp} = \max_{\|x\|=1} \|ABx\|$$

$$\|ABx\| \leq \|A\|_{sp} \|Bx\| \quad \text{for any } x$$

$$x^* = \arg\max_{\|x\|=1} \|ABx\|$$

$$\underbrace{\|ABx^*\|}_{=\|AB\|_{sp}} \leq \|A\|_{sp} \underbrace{\|Bx^*\|}_{\|Bx^*\| \leq \|B\|_{sp}}$$

$\|Bx^*\| \leq \|B\|_{sp} \leftarrow$ since this is the max

$$\Rightarrow \boxed{\|AB\|_{sp} \leq \|A\|_{sp} \|B\|_{sp}}.$$

PROBLEM 10.4

TRUE: $\|AB\|_F^2 = \sum_{i=1}^{n} (AB)_{i,j}^2 = \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\sum_{\ell=1}^{m} A_{i,\ell} B_{\ell,j}\right)^2$

Now for any pair (i,j), by Cauchy Schwartz:

$$\left(\sum_{\ell=1}^{m} A_{i,\ell} B_{\ell,j}\right)^2 \leq \left(\sum_{\ell=1}^{m} A_{i,\ell}^2\right)\left(\sum_{\ell=1}^{m} B_{\ell,j}^2\right)$$

Hence $\|AB\|_F^2 \leq \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{k} \left(\sum_{\ell=1}^{m} A_{i\ell}^2\right)\left(\sum_{\ell=1}^{m} B_{\ell j}^2\right)}$

$$= \left(\sum_{i=1}^{n} \sum_{\ell=1}^{m} A_{i\ell}^2\right)\left(\sum_{j=1}^{k} \sum_{\ell=1}^{m} B_{\ell j}^2\right)$$

$$\leq \|A\|_F^2 \, \|B\|_F^2 .$$