Problem 10.1.

a) Let $A = U\Sigma V^T$ be the SVD of A. We write $r = \operatorname{rank} A$ and $v_1 \dots v_m$ the columns of A.

<u>Claim</u>: $\operatorname{Ker}(A) = \operatorname{Span}(v_{r+1}, \dots v_m)$.

Indeed, for $i \in \{r+1, \dots m\}$, $\Sigma V^T v_i = \Sigma e_i = 0$ because V is orthogonal and only the r first diagonal elements of Σ are non-zero.

This proves that for all $i \geq r+1$, $v_i \in \operatorname{Ker}(A)$. Since $\operatorname{Ker}(A)$ is a subspace, we get $\operatorname{Span}(v_{r+1}, \dots, v_m) \subset \operatorname{Ker}(A)$.

Now, by the rank-nullity theorem:

$$\dim \operatorname{Ker}(A) = m - \operatorname{rank}(A) = m - r = \dim \operatorname{Span}(v_{r+1}, \dots v_m)$$

because $(v_{r+1}, \dots v_m)$ is orthonormal, hence lin. independent.

This proves the claim.

It suffices now to show that for all $i \geq r+1$, $\langle v_i, x^{LS} \rangle = 0$.

Compute $$\langle x^{LS}, v_i \rangle = y^T (A^+)^T v_i = y^T \underbrace{U (\Sigma^+)^T V^T v_i}_{=0 \text{ for the same reason as above}} = 0.$$

b) Let x be a solution of (1).
From the lecture we know that $x - x^{LS} \in \mathrm{Ker}(A)$
hence $x - x^{LS} \perp x^{LS}$: one can apply
Pythagoras theorem:

$$\|x\|^2 = \|x - x^{LS} + x^{LS}\|^2 = \|x^{LS}\|^2 + \|x - x^{LS}\|^2$$

$$\geqslant \|x^{LS}\|^2$$

with a strict inequality whenever $x \neq x^{LS}$.

Problem 10.2.

Define $f(x) = \|Ax - y\|^2 + \lambda\|x\|^2$. f is convex,
as a sum of two convex functions.
Hence

x is a minimizer of f.

$$\iff \nabla f(x) = 0$$

$$\iff 2A^TAx - 2A^Ty + 2\lambda x = 0$$

$$\iff (A^TA + \lambda \mathrm{Id}_d)\, x = A^Ty.$$

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

The matrix $A^TA + \lambda \mathrm{Id}$ is indeed invertible:
it is symmetric with eigenvalues all $\geqslant \lambda > 0$.
(A^TA is positive semi-definite and hence has
non-negative eigenvalues, and the eigenvalues
of $A^TA + \lambda \mathrm{Id}$ are the eigenvalues of A^TA plus λ)

Problem 10.3

a) Let $x \in \mathbb{R}^n$.

- if $x = 0$, then $\|Ax\| = 0 = \|A\|_{sp} \|x\|$ → ok!
- if $x \neq 0$, then $\|\frac{x}{\|x\|}\| = 1$. Using the definition

of $\|A\|_{sp}$:

$$\|A \frac{x}{\|x\|}\| \leq \max_{\|u\|=1} \|Au\| = \|A\|_{sp}.$$

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

b) Let $u \in \mathbb{R}^k$ such that $\|u\| = 1$. Using (a) twice

$\|ABu\| \leq \|A\|_{sp} \|Bu\| \leq \|A\|_{sp} \|B\|_{sp} \|u\| = \|A\|_{sp} \|B\|_{sp}$

This is true for all unit u, therefore:

$$\|AB\|_{sp} = \max_{\|u\|=1} \|ABu\| \leq \|A\|_{sp} \|B\|_{sp}.$$

c) This is true.

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

Now, for $i \in \{1 \dots n\}$ and $j \in \{1, \dots k\}$, Cauchy-Schwartz gives:

$$\left| \sum_{l=1}^m A_{i,l} B_{l,j} \right| \leq \sqrt{\sum_{l=1}^m A_{i,l}^2} \sqrt{\sum_{p=1}^m B_{p,j}^2}$$

Hence $\|AB\|_F^2 \leq \sum_{i=1}^n \sum_{j=1}^k \sum_{l=1}^m A_{i,l}^2 \sum_{p=1}^m B_{p,j}^2$

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

$$= \|A\|_F^2 \|B\|_F^2$$

Problem 10.4

a) $x^* \simeq (14,1 ;\ 3,61 ;\ -7,87 ;\ -1,75)$

b) $v \simeq (0,0352 ;\ -0,266 ;\ 0,782 ;\ 0,518 ;\ 0,219)$

c) ~~(0,0352; -0,266; 0,782; 0,518; 0,219)~~

$w = (0.177,\ 0.179,\ -0.521,\ 0,801)$

~~(0,0352; -0,266; 0,782; 0,518; 0,219)~~

Error $\simeq 1,77$.

Problem 10.5.

Let $v = A^+ y = A^T y$ be the least-square solution (A is invertible hence $A^+ = A^{-1} = A^T$ because A orthogonal). obt:

$$f(x) = \frac{1}{2}\|Ax-y\|^2 + \frac{\alpha}{2}\|x\|^2 + \lambda\|x\|_1$$

$$= \frac{x^T \underbrace{A^T A}_{= Id} x}{2} - \langle v, x\rangle + \frac{1}{2}\|y\|^2 + \frac{\alpha}{2}\|x\|^2 + \lambda\|x\|_1$$

because A orthogonal.

$$= \frac{1}{2}(1+\alpha)\|x\|^2 - \langle v, x\rangle + \lambda\|x\|_1 + \frac{1}{2}\|y\|^2.$$

$$= (1+\alpha)\sum_{i=1}^{n}\left(\frac{x_i^2}{2} - \frac{v_i}{1+\alpha}x_i + \frac{\lambda}{1+\alpha}|x_i|\right) - \frac{\|y\|^2}{2}$$

We see that x minimizes f, if and only if for all i x_i minimizes the function

$$g_{a,b}(t) = \frac{t^2}{2} - at + b|t|$$

for $a = \frac{v_i}{\alpha+1}$ and $b = \frac{\lambda}{1+\alpha}$

Lemma: For all $a \in \mathbb{R}$ and $b \geq 0$ the fonction $g_{a,b}(t) = \frac{t^2}{2} - at + b|t|$ admits a unique minimizer on $\mathbb{R}$, given by:

$$t^* = \eta(a,b) \overset{\text{def}}{=} \begin{cases} a-b & \text{if } a \geq b \\ 0 & \text{if } -b \leq a \leq b \\ a+b & \text{if } a \leq -b. \end{cases}$$

Proof: $g_{a,b}$ is continuous on $\mathbb{R}$, and differentiable on $\mathbb{R} \setminus \{0\}$.

For $t \neq 0$, $g'_{a,b}(t) = t - a + b \underbrace{\text{sign}(t)}_{\substack{+1 \text{ if } t>0 \\ -1 \text{ if } t<0}}$.

Case 1: $a > b$.

In that case $t^* = a - b > 0$ and we see that

$g'_{a,b}(t^*) = 0$

$\forall t > t^*$, $g'_{a,b}(t) > 0$

$\forall t < t^*$, s.t $t \neq 0$, $g'_{a,b}(t) < 0$.

$g_{a,b}$ is continous, hence $t^* = \eta(a,b)$ is the unique minimizer of $g_{a,b}$.

Case 2: $a < -b$: We prove that $t^* = \eta(a,b)$ is the unique minimizer using the same arguments.

Case 3: $-b \leq a \leq b$.

In that case, $\begin{cases} g'_{a,b}(t) > 0 & \text{for all } t > 0 \\ g'_{a,b}(t) < 0 & \text{for all } t < 0 \end{cases}$

We get that $t^* = \eta(a,b)$ is the unique minimizer of $g_{a,b}$.

This proves the lemma ∎

Using the lemma we get that the unique minimizer of f is given by:

$$x_i = \eta\left(\frac{v_i}{1+\alpha}, \frac{\lambda}{1+\alpha}\right)$$

$$= \eta\left(\frac{(A^T y)_i}{1+\alpha}, \frac{\lambda}{1+\alpha}\right).$$