To: Prof. Mitter

Theorem 1.1 *If $b \in \mathbb{R}^m$ and $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ has full column rank (so $n \leq m$) , then*

$$\min_x \max_{\|\delta A\|_2 \leq \gamma, \|\delta b\|_2 \leq \lambda} \|(A+\delta A)x - (b+\delta b)\|_2 = \|A\hat{x} - b\|_2 + \gamma\|\hat{x}\|_2 + \lambda,$$

where $\hat{x} = (A'A + \gamma I)^{-1}A'b$.

Proof:
We will first show that: for any $x \in \mathbb{R}^n$

$$\max_{\|\delta A\|_2 \leq \gamma, \|\delta b\|_2 \leq \lambda} \|(A+\delta A)x - (b+\delta b)\|_2 = \|Ax - b\|_2 + \gamma\|x\|_2 + \lambda. \tag{1}$$

Then we will demonstrate that

$$(A'A + \gamma I)^{-1}A'b = arg\min_x(\|Ax - b\|_2 + \gamma\|x\|_2), \tag{2}$$

which will prove the theorem.

For any fixed x and δA, subject to $\|\delta b\|_2 \leq \lambda$, we have

$$\|(A+\delta A)x - (b+\delta b)\|_2 \leq \|(A+\delta A)x - b\|_2 + \|\delta b\|_2 \leq \|(A+\delta A)x - b\|_2 + \lambda$$

with equality achieved for $(\delta b)_0 = \lambda\frac{(A+\delta A)x - b}{\|(A+\delta A)x-b\|_2}$, which satisfies $\|(\delta b)_0\|_2 \leq \lambda$. Therefore,

$$\max_{\|\delta b\|_2 \leq \lambda} \|(A+\delta A)x - (b+\delta b)\|_2 = \|(A+\delta A)x - b\|_2 + \lambda. \tag{3}$$

Now, we claim that: for any fixed x

$$\max_{\|\delta A\|_2 \leq \gamma} \|(A+\delta A)x - b\|_2 = \|Ax + b\|_2 + \gamma\|x\|_2. \tag{4}$$

First observe that subject to the constraint $\|\delta A\|_2 \leq \gamma$, we have

$$\|(A+\delta A)x - b\|_2 \leq \|Ax - b\|_2 + \|\delta Ax\|_2 \leq \|Ax - b\|_2 + \|\delta A\|_2\|x\|_2 \leq \|Ax - b\|_2 + \gamma\|x\|_2. \tag{5}$$

To prove Equation 4 we have to find a $(\delta A)_0$ that satisfies the constraint and achieves equality in 5. Let

$$(\delta A)_0 = U_{m\times m}\begin{bmatrix}\gamma I_{n\times n} \\ 0\end{bmatrix}_{m\times n}, \tag{6}$$

where U is any $m \times m$ unitary matrix that maps the unit vector $\frac{1}{\|x\|_2}\begin{bmatrix}x \\ 0\end{bmatrix}_{m\times 1}$ to the unit vector $\frac{Ax-b}{\|Ax-b\|_2}$, i.e.,

$$U\frac{1}{\|x\|_2}\begin{bmatrix}x \\ 0\end{bmatrix}_{m\times 1} = \frac{Ax-b}{\|Ax-b\|_2}.$$

(Note that one way to construct such a U is as follows: Construct two orthonormal basis of $\mathbb{R}^m$: $v_1, \ldots, v_m$ and $w_1, \ldots, w_m$ with $v_1 = \frac{1}{\|x\|_2}\begin{bmatrix}x \\ 0\end{bmatrix}_{m\times 1}$ and $w_1 = \frac{Ax-b}{\|Ax-b\|_2}$. Then let U be the linear transformation that maps v_i to w_i for $i = 1, \ldots, m$. Because both basis are orthonormal U must be unitary.)

From Equation 6, which is the SVD form of $(\delta A)_0$, we know that all the singular values of $(\delta A)_0$ are equal to γ. So $\|\delta A\|_2 = \gamma$, i.e., $(\delta A)_0$ satisfies the constraint. To see why it achieves equality in 5, note that

$$\|(A + (\delta A)_0)x - b\|_2 = \|Ax - b + \|x\|_2(\delta A)_0\frac{x}{\|x\|_2}\|_2$$

and

$$(\delta A)_0\frac{x}{\|x\|_2} = U\begin{bmatrix}\gamma I_{n\times n} \\ 0\end{bmatrix}\frac{x}{\|x\|_2} = \gamma\left(U\frac{1}{\|x\|_2}\begin{bmatrix}x \\ 0\end{bmatrix}_{m\times 1}\right) = \gamma\frac{Ax-b}{\|Ax-b\|_2}.$$

Therefore

$$\|(A + (\delta A)_0)x - b\|_2 = \|Ax - b + \|x\|_2\gamma\frac{Ax - b}{\|Ax - b\|_2}\|_2 = \|Ax - b\|_2 + \gamma\|x\|_2,$$

which prove the correctness of Equation 4. Using 3 and 4, we get

$$\begin{aligned}
\max_{\|\delta A\|_2 \leq \gamma, \|\delta b\|_2 \leq \lambda} \|(A + \delta A)x - (b + \delta b)\|_2 &= \max_{\|\delta A\|_2 \leq \gamma} \max_{\|\delta b\|_2 \leq \lambda} \|(A + \delta A)x - (b + \delta b)\|_2 \\
&= \max_{\|\delta A\|_2 \leq \gamma} (\|(A + \delta A)x - b\|_2 + \lambda)(\text{ From 3}) \\
&= \|Ax - b\|_2 + \gamma\|x\|_2 + \lambda(\text{ From 4}),
\end{aligned}$$

which proves Equation 1. So we are left with Equation 2. We have

$$\|Ax + b\|_2 + \gamma\|x\|_2 = (Ax - b)'(Ax - b) + x'\sqrt{\gamma}\sqrt{\gamma}x = \begin{bmatrix} Ax - b \\ \sqrt{\gamma}x \end{bmatrix}' \begin{bmatrix} Ax - b \\ \sqrt{\gamma}x \end{bmatrix}.$$

Noting that $\begin{bmatrix} Ax - b \\ \sqrt{\gamma}x \end{bmatrix} = \begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix}$, we obtain

$$\|Ax + b\|_2 + \gamma\|x\|_2 = \left(\begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix}\right)' \left(\begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix}\right). \tag{7}$$

We are assuming that the matrix A has full column rank, so the matrix $\begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix}$ also has a full column rank. Therefore, we have an overconstrained least square estimation problem. So, the value $\hat{x}$ of x that achieves the minimum of expression 7 is

$$\hat{x} = \left(\begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix}' \begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix}\right)^{-1} \begin{bmatrix} A \\ \sqrt{\gamma}I \end{bmatrix}' \begin{bmatrix} b \\ 0 \end{bmatrix} = (A'A + \gamma I)^{-1}A'b,$$

which proves Equation 2. □

Louay