Condition number derivation

1) $Ax = b$ is the orignal problem, but due to errors in 'b', there will be errors in 'x'. We want to know how bad can these errors be, and the condition number helps in this.

2) With error in b: $A(x+\Delta x) = b + \Delta b$.

$\Rightarrow A\Delta x = \Delta b \Rightarrow \Delta x = A^{+}\Delta b$.

Recall: $A = U\Sigma V^T$, $A^{+} = V\Sigma^{+}U^T$

where $\underbrace{(\Sigma^{+})}_{\substack{n\times m \\ \text{diag matrix}}}{}_{ii} = \begin{cases} 1/\sigma_i & \text{if } \sigma_i > 0 \\ 0 & \text{else}. \end{cases}$

3) Define $K = \dfrac{\text{rel error in 'x'}}{\text{rel error in 'b'}}$ soln meas

$$= \frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|}$$

$$K = \underbrace{\left(\frac{\|\Delta x\|}{\|\Delta b\|}\right)}\left(\frac{\|x\|}{\|b\|}\right)^{-1} \quad \longleftarrow \text{Condition no}$$

4) Consider $(\downarrow)^2 = \dfrac{\|A^{+}\Delta b\|^2}{\|\Delta b\|^2} = \dfrac{\|V\Sigma^{+}U^T\Delta b\|^2}{\|\Delta b\|^2}$

Since V is unitary, num $= \|\Sigma^{+}U^T\Delta b\|^2$

Since U^T is unitary, den $= \|U^T\Delta b\|^2$

Let $y = U^T\Delta b$, a vector.

$$\|\Sigma^{+} y\|^2 = y^T\underbrace{(\Sigma^{+})^T(\Sigma^{+})}_{m \times m \text{ diag matrix}}\, y = \sum_{i=1}^{m} y_i^2 \, 1/\sigma_i^2 .$$

$$\therefore \underbrace{\frac{\|\Delta x\|^2}{\|\Delta b\|^2}} = \sum_{i=1}^{m} 1/\sigma_i^2 \, y_i^2 \times \frac{1}{\sum_{i=1}^{m} y_i^2} .$$

5) Since we want to know the max possible error in x, we want an upper bound on this term ( ).

Now, we know $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_m > 0$.

$\Rightarrow \quad 1/\sigma_1 \leq 1/\sigma_2 \cdots \leq 1/\sigma_m$

$$\Rightarrow \|\Delta x\|^2 = \sum_{i=1}^{m} 1/\sigma_i^2 \, y_i^2 \leq \underbrace{\frac{1}{\sigma_m^2} \sum_{i=1}^{m} y_i^2}_{\text{max}} .$$

Also $\|\Delta b\|^2 = \sum_{i=1}^{m} y_i^2$

Together, $\max\left(\frac{\|\Delta x\|^2}{\|\Delta b\|^2}\right) = \frac{1}{\sigma_m^2}$

6) Similarly, looking at the defn of κ, we need to find the min value of $\frac{\|x\|^2}{\|b\|^2}$. By the same logic as in pt ④, and $U^T b = z$, we get:

$$\underbrace{\frac{\|x\|^2}{\|b\|^2}} = \frac{\sum_{i=1}^{m} \frac{1}{\sigma_i^2} z_i^2}{\sum_{i=1}^{m} z_i^2} .$$

7) We need the min possible value of (*) to get the max error in 'x', so, using: $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_m > 0$; we get $1/\sigma_1 \leq 1/\sigma_2 \cdots \leq 1/\sigma_m$

$$\|x\|^2 = \sum_{i=1}^{m} \frac{1}{\sigma_i^2} z_i^2 \geq \frac{1}{\sigma_1^2} \sum_{i=1}^{m} z_i^2 .$$

$\|b\|^2 = \sum_{i=1}^{m} z_i^2$. Together,

$$\min\left(\frac{\|x\|^2}{\|b\|^2}\right) = \frac{1}{\sigma_1^2}$$

8) Thus, condn no: $\kappa^2 = \left(\frac{1}{\sigma_m^2}\right)\left(\frac{1}{\sigma_1^2}\right)^{-1}$

$$\boxed{\Rightarrow \kappa = \frac{\sigma_1}{\sigma_m} = \frac{\sigma_{max}}{\sigma_{min}}}$$

Note: By defn, the condn no of a non-invertible matrix is ∞.