

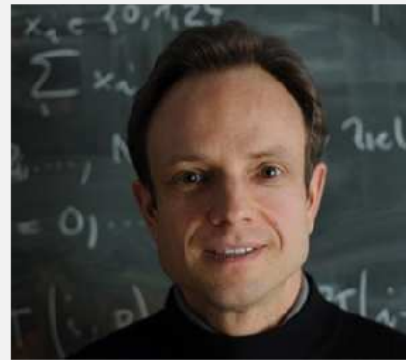
Course Video

Section 2.2.2: Sensitivity of Linear Systems

Prof. R. Hiptmair, SAM, ETH Zurich

Date: September 18, 2020

(C) Seminar für Angewandte Mathematik, ETH Zürich



Prerequisites.

- Ability to compute with matrix and vector norms

Dependencies. Section 1.5.5

Duration: 30 minutes

II. Direct Methods for Square Linear Systems of Equations

2.2. Theory : Linear Systems of Equations

2.2.2. Sensitivity/Conditioning of LSE

↓
quantifies how small (relative) perturbation of data lead to changes of the output

Needed : vector norms & (induced) matrix norms $\|\cdot\|$

$$\|M\| = \sup_{x \neq 0} \frac{\|Mx\|}{\|x\|} \quad \triangleright \quad \|Mx\| \leq \|M\| \|x\|$$

$$\|x + y\| \leq \|x\| + \|y\|$$

$$\Rightarrow \|x - y\| \geq \|x\| - \|y\|$$

Ex.: 2.2.1 (Sensitivity of linear mappings)

$A \in \mathbb{R}^{n,n}$ Fixed : input $x \mapsto y := Ax \hat{=}$ output
& regular

②

$$Ax = y \Leftrightarrow x = A^{-1}y \Rightarrow \|x\| \leq \|A^{-1}\| \|y\|$$

$$A(\underline{x} + \Delta x) = \underline{y} + \Delta y \Rightarrow \|\Delta y\| \leq \|A\| \|\Delta x\|$$

rel. pert. $\frac{\|\Delta y\|}{\|\underline{y}\|} \leq \frac{\|A\| \|\Delta x\|}{\|A^{-1}\|^{-1} \|\underline{x}\|} = \underbrace{\|A\| \|A^{-1}\|}_{\text{sensitivity}} \frac{\|\Delta x\|}{\|\underline{x}\|}$

! Sensitivity heavily depends on choice of norms

LSE: $A \in \mathbb{R}^{n,n}$ regular: $Ax = \underline{b}$
 $\underline{(A + \Delta A)(\underline{x} + \Delta x) = \underline{b} + \Delta b}$
 $\Rightarrow (A + \Delta A) \Delta x = \Delta b - \Delta A \underline{x}$

Tool:

Lemma 2.2.2.5. Perturbation lemma $\rightarrow$
 $B \in \mathbb{R}^{n,n}, \|B\| < 1 \Rightarrow I + B \text{ regular} \wedge \|(I + B)^{-1}\| \leq \frac{1}{1 - \|B\|}$

Proof: Auxiliary estimate, $x \in \mathbb{R}^n$

$$\|(I + B)x\| \geq \|x\| - \|Bx\| \geq \overbrace{(1 - \|B\|)}^{>0} \|x\|$$

$\underline{x} \neq 0 \Rightarrow (I + B)x \neq 0 \Leftrightarrow \mathcal{N}(I + B) = \{0\}$

$\cdot \| (I + B)^{-1} \| = \sup_{\underline{x} \neq 0} \frac{\| (I + B)^{-1} \underline{x} \|}{\|\underline{x}\|} = \sup_{\substack{\underline{y} \in \mathbb{R}^n \\ \underline{y} \neq 0}} \frac{\|\underline{y}\|}{\| (I + B) \underline{y} \|}$
 $\leq \sup_{\underline{y} \neq 0} \frac{\|\underline{y}\|}{(1 - \|B\|) \|\underline{y}\|}$ □

Apply to $A + \Delta A = A(I + A^{-1} \Delta A)$

$\| (A + \Delta A)^{-1} \| = \| (I + A^{-1} \Delta A)^{-1} A^{-1} \|$
 $\stackrel{[\text{Lemma 2.2.2.5}]}{\leq} \|A^{-1}\| \cdot \frac{1}{1 - \|A^{-1}\| \|\Delta A\|}$

$\Rightarrow \|\Delta x\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|\Delta A\|} \cdot (\|\Delta b\| + \|\Delta A\| \|\underline{x}\|)$

$\|\Delta x\| \leq \frac{\|A^{-1}\| \|A\|}{1 - \|A^{-1}\| \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|A\| \|\underline{x}\|} + \frac{\|\Delta A\|}{\|A\|} \right) \|\underline{x}\|$

Theorem 2.2.2.4. Conditioning of LSEs $\rightarrow$,

If A regular, $\|\Delta A\| < \|A^{-1}\|^{-1}$ and (2.2.2.3), then

- (i) $A + \Delta A$ is regular/invertible,
- (ii)

$\frac{\|\underline{x} \Delta x\|}{\|\underline{x}\|} \leq \frac{\|A^{-1}\| \|A\|}{1 - \|A^{-1}\| \|\Delta A\| / \|A\|} \left(\frac{\|\Delta b\|}{\|\underline{b}\|} + \frac{\|\Delta A\|}{\|A\|} \right)$
relative error of data result
relative perturbations

Definition 2.2.2.7. Condition (number) of a matrix

Condition (number) of a matrix $\mathbf{A} \in \mathbb{R}^{n,n}$: $\text{cond}(\mathbf{A}) := \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} \\ (\mathbf{A} + \Delta\mathbf{A})\tilde{\mathbf{x}} = \mathbf{b} + \Delta\mathbf{b} \end{aligned} \Rightarrow \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \frac{\text{cond}(\mathbf{A})}{1 - \text{cond}(\mathbf{A})\|\Delta\mathbf{A}\|/\|\mathbf{A}\|} \left(\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \right). \quad (2.2.2.9)$$

$\text{cond}(\mathbf{A}) \approx 1$: LSE *well-conditioned*

$\text{cond}(\mathbf{A}) \gg 1$: LSE *ill-conditioned*

Ex 2.2.2.10 : Visualizing sensitivity:
Intersection of lines in 2D

Here normal norms



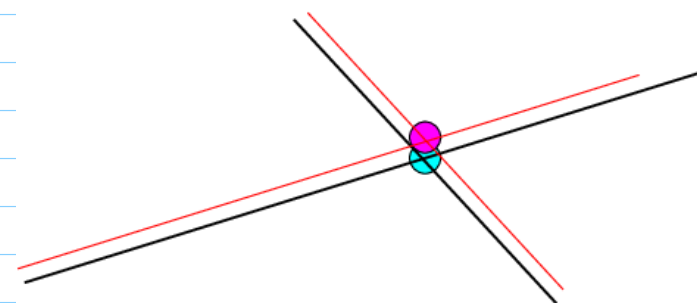
$$L_i = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}^T \mathbf{n}_i = d_i\}, \quad \mathbf{n}_i \in \mathbb{R}^2, d_i \in \mathbb{R}, \quad i = 1, 2.$$



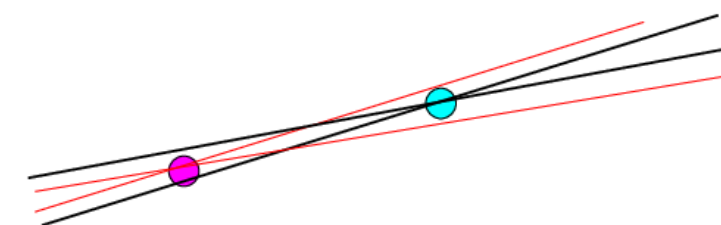
LSE for finding intersection:

$$\underbrace{\begin{bmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \end{bmatrix}}_{=: \mathbf{A}} \mathbf{x} = \underbrace{\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}}_{=: \mathbf{b}},$$

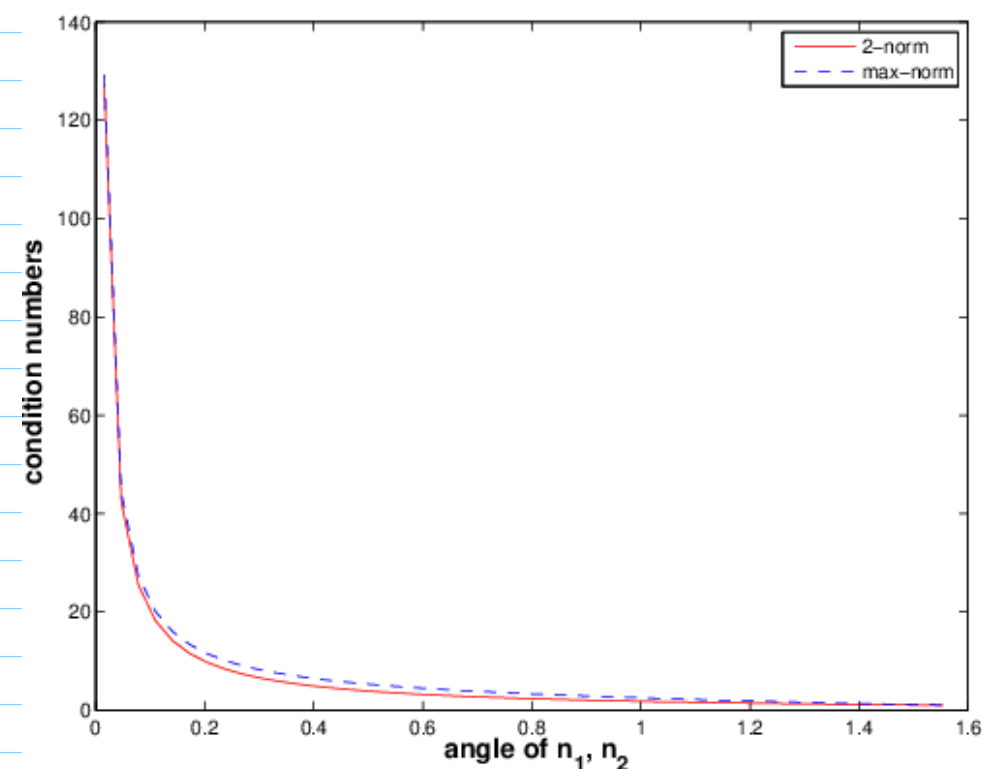
↑
coordinates of $L_1 \cap L_2$



nearly orthogonal intersection: well-conditioned



glancing intersection: ill-conditioned



$\mathbf{n}_1$ "almost linearly dependent" of $\mathbf{n}_2$
 $\rightarrow \text{cond}(\mathbf{A}) \rightarrow \infty$

Heuristics for predicting large $\text{cond}(\mathbf{A})$

$\text{cond}(\mathbf{A}) \gg 1 \leftrightarrow$ columns/rows of $\mathbf{A}$ "almost linearly dependent"

4

5

7