Mathematical methods for inverse problems: Least-squares, conditioning and Singular Value Decomposition

Fallstudien der mathematische Modelbildung, Teil 2

20.10.2023 - 21.11.2023,

paul.catala@tum.de

- We want to solve

$$Ax = y \tag{P}$$

  where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).
- The linear mapping A

- We want to solve

$$Ax = y \tag{P}$$

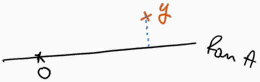
  where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).
- The linear mapping A
  1) might not be surjective

- We want to solve

$$Ax = y \tag{P}$$

  where $x \in E$ $(\mathbb{R}^n)$, $y \in F$ $(\mathbb{R}^m)$ and $A \in \mathcal{L}(E, F)$ $(\mathbb{R}^{m\times n})$.

- The linear mapping A
  1) might not be surjective



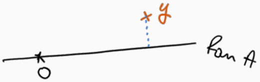
  2) might not be injective

- We want to solve

$$Ax = y \tag{P}$$

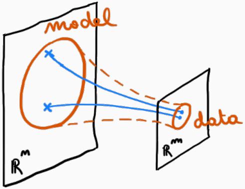
where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).

- The linear mapping A
  1) might not be surjective



  2) might not be injective

  3) might be close to singular, *i.e.* A^{-1} (if it exists) might be "almost" discontinuous.



Linear Systems

- We want to solve

$$Ax = y \tag{P}$$

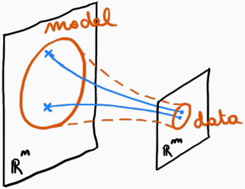
where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).

- The linear mapping A
  1) might not be surjective

     $\longrightarrow$ easily rectifiable through weaker reformulation

  2) might not be injective

  3) might be close to singular, *i.e.* A^{-1} (if it exists) might be "almost" discontinuous.



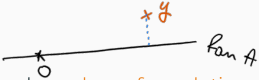
- We want to solve

$$Ax = y \tag{P}$$

where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).

- The linear mapping A

1) might not be surjective

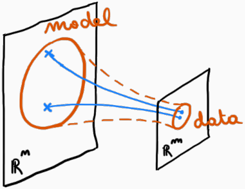


$\longrightarrow$ easily rectifiable through weaker reformulation

2) might not be injective

$\longrightarrow$ needs prior information (structural properties s.a. sparsity, low-rank, etc...)

3) might be close to singular, *i.e.* A^{-1} (if it exists) might be "almost" discontinuous.



- We want to solve

$$Ax = y \tag{P}$$

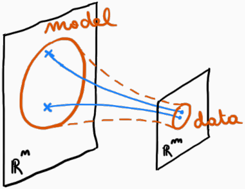
where $x \in E$ ($\mathbb{R}^n$), $y \in F$ ($\mathbb{R}^m$) and $A \in \mathcal{L}(E, F)$ ($\mathbb{R}^{m \times n}$).

- The linear mapping A

  1) might not be surjective
     $\longrightarrow$ easily rectifiable through weaker reformulation

  2) might not be injective
     $\longrightarrow$ needs prior information (structural properties s.a. sparsity, low-rank, etc...)

  3) might be close to singular, *i.e.* A^{-1} (if it exists) might be "almost" discontinuous.
     $\longrightarrow$ numerical instabilities: although A^{-1} is continuous (finite dimension), might be ill-conditoned



- $y \in \operatorname{Ran} A$ too restrictive in practice:
  - over-determined systems ($m > n$) in approximation
  - noise in data
  - model mismatches...
- Least-squares formulation

$$\min_{x \in E} \|y - Ax\|_2^2 \tag{LS}$$

Example (Regression).

Normal equations

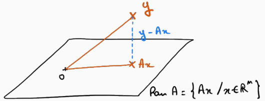
$$\min_{x \in E} \|y - Ax\|_2^2 \tag{LS}$$

- (LS) is equivalent to a linear system, for a different map than A

Proposition (Normal equations). x is a solution of (LS) if and only if

$$A^\top A x = A^\top y, \tag{N}$$

or equivalently $A^\top (y - Ax) = 0$.



Memo (Adjoint).

$$\operatorname{Ran} A^\top = (\operatorname{Ker} A)^\perp$$
$$\operatorname{Ker} A^\top = (\operatorname{Ran} A)^\perp$$

Proof.
Let x be a solution of (N). Hence $y - Ax \in (\operatorname{Ran} A)^\perp$. Let $z \in E$. Then

$$\|y - Az\|_2^2 = \|y - Ax\|_2^2 + \|A(x - z)\|_2^2 \geqslant \|y - Ax\|_2^2$$

Reciprocally, assume that x is not a solution of (N), that is $w := A^\top (Ax - y) \neq 0$. Let $z = x + \varepsilon w$ for $\varepsilon > 0$. Then

$$\|y - Az\|_2^2 = \|y - Ax\|_2^2 - 2\varepsilon \|w\|_2^2 + \varepsilon^2 \|Aw\|_2^2 < \|y - Ax\|_2^2$$

for small ε, hence x is not a solution of (LS). □

$$\min_{x \in E} \|y - Ax\|_2^2 \tag{LS}$$

- ***Proposition (Existence).*** (LS) always admits at least one solution (in finite dimension).

Memo (Ranges, kernels).

$$\operatorname{Ran} A^\top = (\operatorname{Ker} A)^\perp$$
$$\operatorname{Ker} A^\top = (\operatorname{Ran} A)^\perp$$
$$\operatorname{Ker} A^\top A = \operatorname{Ker} A$$

Proof.
We have that $A^\top \in \operatorname{Ran} A^\top$, hence

$$A^\top y \in (\operatorname{Ker} A)^\perp = (\operatorname{Ker}(A^\top A))^\perp = \operatorname{Ran}(A^\top A)$$

hence (N), and therefore (LS), always has a solution. □

- ***Proposition (Unicity).*** The solution is unique if and only if A is injective.

Proof.
(LS) has a unique solution $\iff$ (N) has a unique solution $\iff$ $\operatorname{Ker} A^\top A = \{0\} \iff \operatorname{Ker} A = \{0\} \iff A$ is injective. □

Example (A toy example).

- Consider $Ax = y$ where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 + \frac{1}{n} \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

  Then

$$A^{-1} = n \begin{bmatrix} 1 + \frac{1}{n} & -1 \\ -1 & 1 \end{bmatrix}, \quad \text{and} \quad x = A^{-1}y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- Now, assume we have a small additive noise in our measurement vector, *e.g.* $\tilde{y} = \begin{bmatrix} 1 \\ 1 + \varepsilon \end{bmatrix}$.

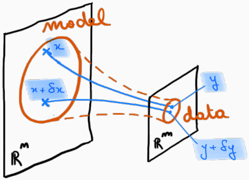
  Then

$$\tilde{x} = A^{-1}\tilde{y} = \begin{bmatrix} 1 - n\varepsilon \\ n\varepsilon \end{bmatrix}$$

- A small perturbation in the data may cause arbitrarily large variations in the solution ($n \gg \varepsilon$).

- Assume for now that A is invertible. Hence (P) and (LS) have a unique solution $x = A^{-1}y$.
- Stability: let $Ax = y$ be perturbed in $A(x + \delta x) = y + \delta y$. We want to compare

$$\frac{\|\delta y\|_2}{\|y\|_2} \quad \text{and} \quad \frac{\|\delta x\|_2}{\|x\|_2}$$



- We have

$$\begin{cases}\delta x = A^{-1}\delta x \\ y = Ax\end{cases} \implies \begin{cases}\|\delta x\|_2 \leqslant \|A^{-1}\|_{2,2}\|\delta x\|_2 \\ \|y\|_2 \leqslant \|A\|_{2,2}\|x\|_2\end{cases}$$

where $\|A\|_{2,2}$ is the operator norm of A

- ***Definition (Operator norm).*** The operator norm of a linear operator $A \in \mathcal{L}(E, F)$ is

$$\|A\|_{E,F} := \sup_{x \neq 0} \frac{\|Ax\|_F}{\|x\|_E} = \sup_{\|x\|_E = 1} \|Ax\|_F = \sup_{\|x\|_E \leqslant 1} \|Ax\|_F$$

In our case, working with the ℓ^2-norm, we write $\|A\|_{2,2}$, or simply $\|A\|$.

- We have
$$\begin{cases}\delta x = A^{-1}\delta x\\ y = Ax\end{cases} \implies \begin{cases}\|\delta x\|_2 \leqslant \|A^{-1}\|_{2,2}\|\delta x\|_2\\ \|y\|_2 \leqslant \|A\|_{2,2}\|x\|_2\end{cases}$$
where $\|A\|_{2,2}$ is the operator norm of A

- ***Definition (Operator norm).*** The operator norm of a linear operator $A \in \mathcal{L}(E, F)$ is
$$\|A\|_{E,F} := \sup_{x\neq 0} \frac{\|Ax\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|Ax\|_F = \sup_{\|x\|_E\leqslant 1} \|Ax\|_F$$
In our case, working with the ℓ^2-norm, we write $\|A\|_{2,2}$, or simply $\|A\|$.

Proof.
(of the two last equalities)

Let $S = \{x\ ;\ \|x\| = 1\}$. The mapping $E \setminus \{0\} \to S, x \mapsto x/\|x\|$ is surjective, hence $\|A\| = \sup_{x\neq 0} \|A(x/\|x\|)\| = \sup_{x\in S} \|Ax\|$.

Let $B = \{x\ ;\ \|x\| \leqslant 1\}$. Since $S \subset B$, necessarily $\sup_{x\in S} \|Ax\| \leqslant \sup_{x\in B} \|Ax\|$. On the other hand, we have by definition that for any $x \in B$,
$$\|Ax\| \leqslant \|A\|\|x\| \leqslant \|A\|$$
and hence $\sup_{x\in B} \|Ax\| \leqslant \|A\|$. Altogether, we reach the desired equalities. □

Condition Number

- Rearranging

$$\begin{cases} \|\delta x\| \leqslant \|A^{-1}\| \|\delta y\| \\ \|y\| \leqslant \|A\| \|x\| \end{cases}$$

  we obtain

$$\frac{\|\delta x\|}{\|x\|} = \|A^{-1}\| \|A\| \frac{\|\delta y\|_2}{\|y\|_2}$$

- ***Definition (Condition number).*** The condition number of an invertible matrix A is

$$\kappa(A) := \|A\| \|A^{-1}\|$$

- Rearranging

$$\begin{cases} \|\delta x\| \leqslant \|A^{-1}\| \|\delta y\| \\ \|y\| \leqslant \|A\| \|x\| \end{cases}$$

we obtain

$$\frac{\|\delta x\|}{\|x\|} = \|A^{-1}\| \|A\| \frac{\|\delta y\|_2}{\|y\|_2}$$

- ***Definition (Condition number).*** The condition number of an invertible matrix A is

$$\kappa(A) := \|A\| \|A^{-1}\|$$

- Note that in general $\kappa(A)$ depends on the norms on E and F. We only consider the ℓ^2-norm, in which case $\kappa(A)$ can actually be given explicitely as we will see.

- Rearranging

$$\begin{cases} \|\delta x\| \leqslant \|A^{-1}\| \|\delta y\| \\ \|y\| \leqslant \|A\| \|x\| \end{cases}$$

we obtain

$$\frac{\|\delta x\|}{\|x\|} = \|A^{-1}\| \|A\| \frac{\|\delta y\|_2}{\|y\|_2}$$

- ***Definition (Condition number).*** The condition number of an invertible matrix A is

$$\kappa(A) := \|A\| \|A^{-1}\|$$

- Note that in general $\kappa(A)$ depends on the norms on E and F. We only consider the ℓ^2-norm, in which case $\kappa(A)$ can actually be given explicitly as we will see.

- ***Proposition (Some properties).*** For $A \in \mathbb{R}^{n \times n}$ invertible
  1. $\kappa(A) \geqslant 1$
  2. $\kappa(A) = \kappa(A^{-1})$
  3. $\kappa(\lambda A) = \kappa(A)$ for any $\lambda \in \mathbb{R} \setminus \{0\}$

Proof.
For all A, B, for all x, $\|ABx\| \leqslant \|A\| \|Bx\| \leqslant \|A\| \|B\| \|x\|$, hence $\|AB\| \leqslant \|A\| \|B\|$. In particular, $1 = \|I_n\| \leqslant \|A\| \|A^{-1}\| = \kappa(A)$. □

- The singular value decomposition of A will help us solve the least-squares problem.
- ***Theorem (Singular Value Decomposition).*** Let $A \in \mathbb{R}^{m \times n}$ of rank r. There exist $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ orthogonal (*i.e.* $U^\top U = UU^\top = I_m$ and $V^\top V = VV^\top = I_n$) and $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ such that

$$A = U\mathbf{\Sigma}V^\top \quad \text{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix}$$

  where $\mathbf{\Sigma}_1 = \mathbf{diag}(\sigma_1, \ldots, \sigma_r)$, and $\sigma_1 \geqslant \ldots \geqslant \sigma_r > 0$.
- Component by component, we have

$$\begin{array}{llll} 1) & Av_i = \sigma_i u_i, & A^\top u_i = \sigma_i v_i & \text{for} \quad i = 1, \ldots, r \\ 2) & Av_i = 0, & A^\top u_i = 0 & \text{for} \quad i \geqslant r + 1 \end{array}$$

  The u_i and v_i are the left and right *singular vectors* of A respectively, associated with the *singular value* σ_j.