

Inverse problems: formulation, ill-posedness, and first-order theory

B1 · Graduate course

Hermann Agossou^{1,2}

PhD candidate, UTER MID@S, ECC / UH2C

¹École Centrale Casablanca · ²Université Hassan II de Casablanca

PART 01

Formulation

01

FORMULATION

Linear algebra you already have

Two questions before we start

01

Condition number

What exactly does it bound: the error in the solution, or something weaker?

02

No zero singular value

If none of them vanishes, is the inverse safe to use?

THE SECOND QUESTION IS THE ONE THAT MATTERS TODAY

The running matrix is invertible but sensitive. Later, exact zero singular values will provide a contrasting rank-deficient case.

FORMULATION

Add one part in a thousand of noise

And the reconstruction is unusable

$$\mathcal{K}_{1,\{1,10,32\}} = \text{diag}(1, 10^{-2}, 2^{-10}), \quad \delta\sigma^* = (0.2, 0.02, 0.00625)^\top$$

Three selected modes of one reference derivative. Every diagonal entry is nonzero, and that is the trap.

NOW ADD 2E-4 TO THE MODE-32 DATUM

The exact data on modes 1, 10, and 32 are about 0.2, 0.0002, and 0.0000061. The perturbation is about 0.1% of their norm, but direct inversion recovers a mode-32 increment of about 0.211 instead of 0.00625. Nothing failed and nothing warned.

FORMULATION

The setting: states, parameters, observations

Every symbol used today, defined here

- ▶ $\mathcal{P} = \mathcal{U} = \mathcal{V} = \mathcal{Y} = \ell^2$: parameter, state, residual, and data spaces.
- ▶ $\mathcal{P}_\rho = \{\sigma : \|\sigma\|_\infty \leq \rho < 1\}$ and $\mathcal{Y}_{\text{ad}} = \mathcal{G}(\mathcal{P}_\rho)$: admissible parameter and data sets.
- ▶ $[\mathcal{F}(u, \sigma)]_i = (1 - [\sigma]_i)u_i - q_i$, with known $q_i = 1/i$: the state residual.
- ▶ $\mathcal{M}(u)_i = u_i/i$: the smoothing observation map.

$$[S(\sigma)]_i = \frac{1}{i(1 - [\sigma]_i)}, \quad [\mathcal{G}(\sigma)]_i = \frac{1}{i^2(1 - [\sigma]_i)}$$

The residual, solution map, observation map, and exact parameter-to-data map now belong to one problem.

Operators, boundedness, and compactness

The three properties the rest of the module uses

DEFINITION: BOUNDED OPERATOR

prerequisite: norms

An operator is bounded when it cannot stretch any vector by more than one fixed factor, the same factor for all of them. That factor is called the operator norm. For a matrix it always exists; for a map between function spaces it need not.

$\|T\|$ = the smallest such factor · for linear maps, bounded and continuous mean the same thing

$$\|Tq\| \leq C \|q\| \quad \text{for all } q, \quad \|T\| := \inf\{C \text{ that works}\}$$

The symbolic version of the sentence above. Nothing more is meant by it.

WHERE THE DIFFICULTY LIES

T is always bounded in a finite-dimensional instance. The question is whether its inverse remains bounded across the intended spaces. If T shrinks some directions arbitrarily much, no single factor bounds the inverse.

FORMULATION

The Hadamard criterion: existence, uniqueness, stability

Assess it on named spaces

01 Existence

For every admissible datum y there is a σ with $\mathcal{G}(\sigma) = y$.

02 Uniqueness

There is at most one such σ .

03 Stability

The map $y \mapsto \sigma$ is continuous: small changes in the data give small changes in the parameter.

ILL-POSED MEANS ANY ONE OF THE THREE FAILS

This is a statement about a map between specified spaces. It cannot be diagnosed from a solver failure or from one finite condition number.

See: [1] Clason (2020)

FORMULATION

Assess the running map, condition by condition

The spaces and admissible set decide the answer

Condition	Running modal map	What changes it
Existence	Yes on $\mathcal{Y}_{\text{ad}}$ by definition; no on all of ℓ^2	Admissible data or least-squares target
Uniqueness	$[\sigma]_i = 1 - 1/(i^2 y_i)$ gives at most one admissible coefficient	Observation geometry or constraints
Stability	The inverse derivative multiplies mode i by i^2	Regularize a nearby family and accept bias

THE THREE-MODE DERIVATIVE IS HADAMARD WELL POSED

Its inverse is continuous because the selected space is finite dimensional and every factor is nonzero. The finite condition number explains the sensitivity. The limiting reduced derivative has factors $1/i^2$; its inverse is unbounded.

PART 02

Diagnosis

02

The singular value decomposition, stated

Theorem, notation, and what each factor holds

THEOREM: SINGULAR VALUE DECOMPOSITION

always exists for a finite matrix

Every real matrix is: rotate, stretch each axis by its own factor, rotate again. Those factors are the singular values, and this factorisation always exists.

orthogonal = the rotation preserves lengths and angles, so it never amplifies anything

$$\mathcal{K}_1 = U D_s V^T = \sum_{i=1}^r s_i u_i v_i^T, \quad r = \min(m, n), \quad s_1 \geq \dots \geq s_r \geq 0$$

For a finite $\mathcal{K}_1 \in \mathbb{R}^{m \times n}$, this is the thin SVD: $U \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{n \times r}$ have orthonormal columns; $D_s \in \mathbb{R}^{r \times r}$ is diagonal.

- ▶ $u_i \in \mathbb{R}^m$: left singular vectors, an orthonormal family in the data space.
- ▶ $v_i \in \mathbb{R}^n$: right singular vectors, an orthonormal family in the parameter space.
- ▶ $s_i \geq 0$: ordered singular values. $\mathcal{K}_1$ maps v_i to $s_i u_i$.

An SVD you can compute by hand

Three selected modes of the running derivative

$$\mathcal{K}_{1,\{1,10,32\}} = \text{diag}(1, 10^{-2}, 2^{-10}) \implies U = V = I$$

For $[\mathcal{K}_1 \delta \sigma]_i = [\delta \sigma]_i / i^2$, the selected singular values are read off directly.

WHAT EACH DIRECTION COSTS

Forward, $\mathcal{K}_1$ shrinks mode 32 by a factor of 1024. Backward, recovering that mode multiplies by 1024, along with whatever noise sits there.

VERIFICATION

Multiply the three factors back together and confirm the product reproduces the original operator entrywise. A factorisation that does not reconstruct what you started from is not an SVD.

DIAGNOSIS

Why the generalized inverse is unbounded

A property of the operator, not of your solver

START WITH THE FINITE INTUITION

Mode 32 is shrunk by 1024, so inversion stretches its noise by 1024. Across the full family, mode i is shrunk by i^2 with no largest i : no single number bounds the inverse. That is what unbounded means.

THEOREM: THE GENERALIZED INVERSE IS UNBOUNDED

Clason, Corollary 3.8

For a compact operator between Hilbert spaces with infinite-dimensional range, the generalized inverse is unbounded. In a singular system, the nonzero singular values tend to zero, producing shrinking with no positive lower bound.

$\mathcal{K}_1$ compact $\Rightarrow s_i \rightarrow 0 \Rightarrow 1/s_i \rightarrow \infty \Rightarrow$ no bound on the inverse

WHAT THIS REQUIRES OF A CONVERGENT REGULARIZATION METHOD

Replace the exact inverse by one bounded operator per parameter value. As the noise shrinks, let the parameter shrink too. A parameter fixed once and for all cannot do this.

DIAGNOSIS

The pseudo-inverse, defined

What it computes, and where it breaks

$$\widehat{\delta\sigma}_{\text{mn}} = \mathcal{K}_1^+ \delta y^{\text{obs}} = \sum_i \frac{u_i^\top \delta y^{\text{obs}}}{s_i} v_i, \quad s_i > 0$$

MINIMUM NORM DOES NOT MEAN STABLE

On modes 1, 10, and 32, the perturbed data are about $(0.2, 2 \cdot 10^{-4}, 2.06 \cdot 10^{-4})$. Direct inversion returns increments $(0.2, 0.02, 0.211)$: mode 32 amplifies the added noise by 1024.

DIAGNOSIS

Singular values that decay without a gap

Ill-posed is not the same as rank-deficient

RANK-DEFICIENT

Exact zeros create a null space.

- ▶ $s = (1, 0.8, 0.5, 0, 0)$
- ▶ The unrecoverable null-space directions are exact
- ▶ A minimum-norm solution is one explicit convention

DISCRETIZED INVERSE PROBLEM

Gradual decay, with no clean numerical-rank gap.

- ▶ $s = (1, 0.4, 0.15, 0.06, 0.02, \dots)$
- ▶ Every mode is partly signal, partly noise
- ▶ Where to stop is a decision, not a reading

The second pattern is common when discretizing smoothing operators. Each finite full-rank matrix remains Hadamard well posed, although it may be severely ill conditioned.

DIAGNOSIS

Refine the same modal operator

One case, four discretization levels

WORKFLOW 01

Discretize consistently

Build $K_{1,n} = \text{diag}(1/i^2)_{i=1}^n$ and retain n .

WORKFLOW 02

Compute the spectrum

Use `'np.linalg.svd(K_1_n, compute_uv=False)'`; do not reconstruct vectors you will not inspect.

WORKFLOW 03

Normalize and compare

Plot s_i/s_1 on the same axes for at least three refined grids.

VERIFICATION

Repeat on three consistent grids. Stable normalized decay is evidence about the limiting operator, not proof; a moving curve can expose scaling or discretization artifacts.

DIAGNOSIS

Hansen: the discrete Picard condition

A diagnostic you can plot

CRITERION: DISCRETE PICARD CONDITION

read it from a plot

For exact data, useful regularized approximation requires the generalized Fourier coefficients to decay faster on average than the singular values. With noisy observations, a coefficient plateau marks where division amplifies noise.

$u_i^T \delta y^{\text{obs}} = \text{observed increment along } u_i \cdot s_i = \text{how much } \mathcal{K}_1 \text{ shrinks } v_i$

$$|u_i^T \delta y| \longrightarrow 0 \text{ faster than } s_i \longrightarrow 0$$

Underlying exact-data pattern. The finite DPC concerns average decay over numerically nonzero singular values, not a per-index test.

WHAT IT TELLS YOU

The comparison locates noise amplification under the stated scaling. A short finite plot is a sensitivity diagnosis, not a theorem that no method can recover more.

DIAGNOSIS

A finite coefficient table, symbol by symbol

Sensitivity evidence, not an asymptotic theorem

mode i	s_i	$ u_i^T \delta y^{\text{obs}} $	recovered increment	verdict
1	1	0.2	0.2	supported
10	10^{-2}	$2 \cdot 10^{-4}$	0.02	supported
32	2^{-10}	$2.06 \cdot 10^{-4}$	0.211	noise amplified

READ THE LAST COLUMN BEFORE RECONSTRUCTING

At mode 32, direct inversion multiplies the perturbed coefficient by 1024. Under the stated noise model and without additional information, this datum does not support an accurate mode-32 increment. This is a finite sensitivity diagnosis, not an asymptotic theorem.

DIAGNOSIS

Finish the coefficient-ratio table

Four rows are given, four are yours

i	s_i	$ u_i^\top \delta y^{\text{obs}} $	ratio	ratio trend
1	1.0000	2.008e-1	2.008e-1	falling
2	2.500e-1	2.580e-2	1.032e-1	falling
3	1.111e-1	8.207e-3	7.386e-2	falling
4	6.250e-2	3.925e-3	6.280e-2	falling
5	4.000e-2	2.400e-3	?	?
6	2.778e-2	1.726e-3	?	?
7	2.041e-2	1.383e-3	?	?
8	1.562e-2	1.191e-3	?	?

8-MINUTE EVIDENCE CHECK

Artifact: one completed table per pair, with the first non-falling ratio circled. Check: another pair recomputes that row before either pair interprets it.

DIAGNOSIS

Five-minute break

Resume with one question

WHEN YOU RETURN

Which modes did the data support, and what additional information would license keeping more?

PART 03

Regularization

03

Tikhonov, derived from its own problem

From the penalised least squares to a linear system

$$\widehat{\delta\sigma}_\lambda = \arg \min_h \{ \|\mathcal{K}_1 h - \delta y^{\text{obs}}\|^2 + \lambda^2 \|h\|^2 \} = R_\lambda \delta y^{\text{obs}}$$

Step 1: fit the linearized data increment, but not at unlimited cost in parameter-increment size.

$$2\mathcal{K}_1^\text{T}(\mathcal{K}_1 h - \delta y^{\text{obs}}) + 2\lambda^2 h = 0 \quad \implies \quad (\mathcal{K}_1^\text{T}\mathcal{K}_1 + \lambda^2 I)\widehat{\delta\sigma}_\lambda = \mathcal{K}_1^\text{T}\delta y^{\text{obs}}$$

Step 2: for this penalty and fixed $\lambda > 0$, the matrix is positive definite and the regularized problem is stable.

VERIFICATION

Substitute the solution back into the gradient and confirm it returns zero, then check that the system solved in code is the one on the second line.

REGULARIZATION

Filter factors: one picture for all methods

Truncation and Tikhonov are two filters

$$\widehat{\delta\sigma}_\lambda = \sum_i f_i \frac{u_i^\top \delta y^{\text{obs}}}{s_i} v_i, \quad f_i^{\text{Tik}} = \frac{s_i^2}{s_i^2 + \lambda^2}, \quad f_i^{\text{TSVD}} = 1 \text{ if } i \leq k, \text{ else } 0$$

Obtained by substituting $\mathcal{K}_1 = U D_s V^\top$ into the normal equations of the previous slide.

HOW TO READ A FILTER FACTOR

A filter factor near 1 keeps the component; a factor near 0 discards it. Tikhonov rolls off smoothly around the regularization scale, while truncation cuts hard at one index. The bias is visible before any experiment.

Tikhonov on the running example, by hand

Three filter factors, three outcomes

$$f_1 \approx 1, \quad f_{10} = 0.5, \quad f_{32} \approx 0.0094$$

With $\lambda = 10^{-2}$: retain mode 1, halve mode 10, and suppress mode 32.

THE RESULTING SOLUTION, AND THE HONEST READING

The Tikhonov estimate on modes 1, 10, and 32 is about 0.20, 0.010, and 0.0020, against the true increments 0.20, 0.020, and 0.00625. The weak modes are biased toward zero. Regularization stabilizes the calculation; it does not create missing information.

VERIFICATION

Build this solution twice, once from the filtered expansion and once by solving the regularized normal equations, and confirm they agree. Disagreement means the filter factors were written down wrongly.

REGULARIZATION

Choosing lambda: three rules, three costs

And what each one assumes

01

Discrepancy principle

Match the residual norm to a defensible bound on $\|\varepsilon\|$.

02

Generalized cross-validation

Minimise predicted error under a stated error model.

03

L-curve

Balance solution and residual norms at a visible corner.

THERE IS NO PARAMETER-FREE CHOICE

A defensible reconstruction names the rule that set λ and the assumption that licensed it.

See: [3] Hansen (2010)

REGULARIZATION

Your turn: choose lambda and defend it

Then meet a colleague who chose differently

TASK

12 minutes. Artifact: a one-row decision record naming the λ rule, the value obtained, the licensing assumption, and one piece of supporting or contradicting evidence.

- ▶ Then pair with someone who used a different rule on the same data
- ▶ Adjudicate: whose assumption is better supported here?
- ▶ Check: swap records and reproduce the other student's value before debating the assumption
- ▶ A defence that cites only a number, without its rule, does not count

PART 04

First order

04

The Fréchet derivative, defined

Differentiating a map between spaces

YOU ALREADY KNOW THIS DEFINITION

In calculus you wrote $f(x+h) = f(x) + f'(x)h + o(h)$. The Fréchet derivative is that same sentence with absolute values replaced by norms.

DEFINITION: FRÉCHET DERIVATIVE

generalizes $f'(x)$

$D\Phi(q)$ is the unique bounded linear operator whose remainder vanishes faster than $\|h\|$. In finite dimensions it is the Jacobian.

$o(\|h\|)$ means: divide the remainder by $\|h\|$ and it still tends to 0

$$\|\Phi(q+h) - \Phi(q) - D\Phi(q)h\| = o(\|h\|)$$

This is what 'first order' means, and it is tested numerically later.

The nonlinear map behind the running case

Return to the residual, then differentiate it

$$[\mathcal{F}(u, \sigma)]_i = (1 - [\sigma]_i)u_i - \frac{1}{i} = 0 \quad \Rightarrow \quad [S(\sigma)]_i = \frac{1}{i(1 - [\sigma]_i)}$$

Step 1: solve the state residual exactly on the admissible set.

$$[\mathcal{M}(u)]_i = \frac{u_i}{i} \quad \Rightarrow \quad [\mathcal{G}(\sigma)]_i = \frac{1}{i^2(1 - [\sigma]_i)}$$

Step 2: compose with the same observation map used from the start.

VERIFICATION BEFORE DIFFERENTIATION

Substitute $S(\sigma)$ into $\mathcal{F}$ and $\mathcal{M}$. If either identity fails componentwise, the reduced map is wrong and its derivative is irrelevant.

The coupled first-order problem

State and parameter unknown together

$$\mathcal{F}(u_0 + \delta u, \sigma_0 + \delta \sigma) = L_0 \delta u + B_0 \delta \sigma + o(\|\delta u\| + \|\delta \sigma\|)$$

Expanded at a reference pair, where $\mathcal{F}(u_0, \sigma_0) = 0$ so the leading term drops.

$$L_0 = D_u \mathcal{F}(u_0, \sigma_0), \quad B_0 = D_\sigma \mathcal{F}(u_0, \sigma_0) \quad \implies \quad L_0 \delta u + B_0 \delta \sigma = 0$$

The two frozen tangents. Both increments are still unknown.

VERIFICATION

At $\sigma_0 = 0$, $L_0 = I$ and $[B_0 \delta \sigma]_i = -[\delta \sigma]_i / i$. Evaluate both tangents by finite differences before eliminating the state.

Eliminate the state, get the reduced data map

And the invertibility you just assumed

$$\begin{aligned}\delta u &= -L_0^{-1} B_0 \delta \sigma, & \delta y &= \mathcal{K}_1 \delta \sigma, \\ [\mathcal{K}_1 \delta \sigma]_i &= \frac{[\delta \sigma]_i}{i^2}, & \mathcal{K}_1 &= -D\mathcal{M}(u_0) L_0^{-1} B_0.\end{aligned}$$

The bounded inverse of L_0 is the assumption. The second line closes the loop with the operator used since slide 4.

VERIFICATION

Compare the nonlinear increment with $\mathcal{K}_1 \delta \sigma$ over shrinking steps. The remainder divided by $\|\delta \sigma\|$ must tend to zero. Test second-order decay only after stating a stronger smoothness assumption.

What truncation loses and reduction hides

Higher-order terms, invertibility, and explicit coupling

LINEARIZATION: TRUNCATION

The $o(\cdot)$ term was dropped.

- ▶ Everything beyond first order is gone
- ▶ The guarantee is local in $\delta\sigma$
- ▶ B2 addresses this one

STATE ELIMINATION: REPRESENTATION

The coupling is retained inside $\mathcal{K}_1$ but is no longer explicit.

- ▶ Requires L_0 to have a bounded inverse
- ▶ One composed operator hides the state increment
- ▶ B3 addresses this one

CHECK: NAME THE OPERATION

3 minutes. Label each claim: a remainder was dropped; L_0^{-1} was assumed; δu is hidden but its coupling remains in $\mathcal{K}_1$. Check one label with a neighbor.

What first order cannot reach

The question B2 and B3 will answer

01

Established today

Named spaces define well-posedness; spectra bound what data support; every λ has a rule; truncation and reduction have distinct costs.

02

Left open

What is available beyond first order, and what it costs to get it.

TWO DIRECTIONS, NEITHER ANSWERED TODAY

- ▶ Keep eliminating, but go past first order: B2.
- ▶ Refuse to eliminate, and follow the coupled problem: B3.

The module stops at its evidence boundary; B2 and B3 start beyond it.

PART 05

References

