Inverse problems: an overview

Marc Bonnet

[1]POEMS, CNRS-INRIA-ENSTA, ENSTA Paris, Palaiseau, France (*mbonnet@ensta.fr*)

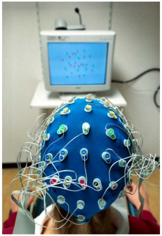
What is inversion?

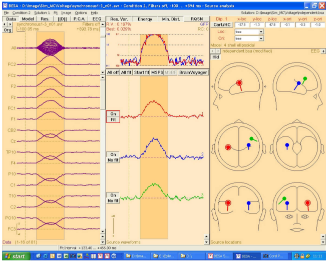
- Quantitative knowledge and understanding of physical systems often rests on obtaining and exploiting experimental data on system responses.
- Sought informations often “hidden”, with measurable quantities as consequences.

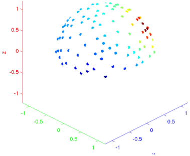
Example (electrostatics / heat equilibrium)

$$-\text{div}\,(\sigma\boldsymbol{\nabla}u) = f \text{ in } \Omega, \qquad \partial_n u = 0 \text{ on } \partial\Omega$$

Identify source $f(\boldsymbol{x})$ or material parameter $\sigma(\boldsymbol{x})$ from measurements of potential u.







Reconstruct brain electrical activity (sources). Measurements = potential (at electrodes)

Baratchard, Clerc, Leblond (2013)

Indirect measurement: quantify hidden $\boldsymbol{p}$ using measurable quantity $\boldsymbol{d}$.
Mathematical model of underlying physics links $\boldsymbol{d}$ and $\boldsymbol{p}$.

(i) Model yields $\boldsymbol{d}$ given $\boldsymbol{p}$ (ODEs, PDEs, variational problem, integral equations...)

$\boxed{\mathbf{G}(\boldsymbol{p},\boldsymbol{d})=0}$ (implicit), or occasionally $\boxed{\boldsymbol{d}-\mathbf{G}(\boldsymbol{p})=0}$ (explicit).

If implicit, finding $\boldsymbol{d}$ (model prediction of data) given $\boldsymbol{p}$ is a well-posed problem.

- Well-posed mathematical problem (Hadamard's definition): solution (a) exists, (b) is unique, (c) depends continuously in the data.

Example $(\boldsymbol{p}=(\Omega,\sigma,f),\ \boldsymbol{d}=u)$

Given Ω,σ,f, find u such that $\quad -\operatorname{div}(\sigma\boldsymbol{\nabla}u)=f$ in Ω, $\qquad \partial_n u=0$ on $\partial\Omega$

(ii) Inversion (find $\boldsymbol{p}$ given data on $\boldsymbol{d}$) is generally an ill-posed problem.

- Ill-posed mathematical problem: at least one of (a), (b), (c) above fails.

Inverse problem: (usually) quantitative exploitation and interpretation of data on configurations involving complex modeling. Needs solving an ill-posed problem.

Relevant situations (far from comprehensive!)

Reconstruction of underground physical characteristics

- Data: seismic, gravimetric...,
- Scientific knowledge, resource prospection.... . .

Non destructve evaluation

- Identification (flaws, cracks,. . .)
- Monitoring (civil engineering, industry)

Data: ultrasound, eddy currents, thermography. . .

Tomography, medical imaging, elastography

- Data: X-ray, MRI, electrostatic, kinematic fields. . .

Reconstruction of temperatures, thermal fluxes. . . on hidden surfaces. . .

Identification using mechanical tests

- mechanical constitutive parameters
- model updating

Deconvolution

- restoration of blurry imges,
- interpretation of dynamical measurements

. . . and many more !

1. **Inverse and ill-posed problems: examples**

2. Short overview of solution approaches

3. Finite-dimensional ill-conditioned linear systems

4. Regularization by promotion of sparsity

5. (a glimpse of) Bayesian approach to inverse problems

(numerical) differentiation

- Antiderivative $f(t) = \mathsf{G}u(t) := \int_0^t u(\tau)\,\mathrm{d}\tau$: $\mathsf{G} : C^0([0,T]) \to C^0([0,T])$ (say) is continuous.
- Inverse problem $u = f'$, i.e. solve $\mathsf{G}u = f$ given $f \in C^0([0,T])$ is ill-posed.
- For example, let $f \in C^1([0,T])$ and set $f^\delta(t) = f(t) + \delta \sin(t/\delta^2) \in C^1([0,T])$ (small-amplitude, highly-oscillatory data perturbation). Then:

$$\sup_{t\in[0,T]} |f^\delta - f| = C\delta \quad \text{but} \quad \sup_{t\in[0,T]} |f^{\delta\prime} - f'| = C\delta^{-1}$$

- Likewise, numerical differentiation with data noise is unstable: consider

$$\hat{f}(x) = f(x) + b(x) \qquad b(x) : \text{ (noise)}$$

$$D_h\hat{f}(x) := \frac{1}{2h}[f(x+h) - f(x-h)] \qquad \text{(central finite difference)}$$

Assume f'' Lipschitz (with modulus L) and $|b(x)| \leq B$, then

$$\boxed{|D_h\hat{f}(x) - f'(x)| \leq Lh^2 + \frac{1}{h}B}$$

In particular, derivative estimation error increases under sampling grid refinement.

Hadamard's classic example: Cauchy ill-posed problem

$$\Delta u = 0 \text{ in } (0,+\infty)\times\mathbb{R}, \qquad u(0,y)=0,\ \partial_x u(0,y) = \frac{\sin ay}{a}\ (a>0)$$

Solution: $u(x,y) = \dfrac{\text{sh}\, ax \, \sin ay}{a^2}$

Unfortunately, $\boxed{\lim_{a\to\infty} \dfrac{\text{sh}\, ax}{a^2} = \infty}$ for any $x>0$: if $a\to\infty$, data magnitude vanishes while solution blows up at any internal point´

Hadamard's views on ill-posed problems

Mais il est remarquable, d'autre part, qu'on trouve un guide sûr dans l'interprétation physique : un problème analytique est toujours correctement posé (au sens précédemment indiqué), quand il est la traduction d'une question mécanique ou physique; et nous avons vu que ceci est le cas du problème de Cauchy dans les exemples notés en premier lieu.

Au contraire, aucun des problèmes physiques en rapport avec $\Delta u = 0$ ne se formule analytiquement sous la forme de Cauchy (2). Chacun d'eux conduit à des énoncés tels que celui de Dirichlet, c'est-à-dire avec une seule donnée numérique en chaque point de la frontière. C'est aussi le cas de l'équation de la chaleur. Tout ceci est d'accord avec le fait que les données de Cauchy, si elles ne sont pas analytiques, ne sont pas de nature à déterminer une solution d'une quelconque de ces deux équations.

... and relevance to physics and engineering of solving *ill-posed* problems not recognized until the 60s (e.g. [Phillips 62; Tikhonov 63; Twomey 65; Tikhonov, Arsenin 75])

J. Hadamard, *Le problème de Cauchy et les équations aux dérivées partielles lineraires hyperboliques* (1932)

Example: scalar equilibrium problem

Forward problem: find u (potential) in Ω given internal source f and conductivity σ

$$-\text{div}\,(\sigma\boldsymbol{\nabla}u) = f \quad \text{in } \Omega, \qquad \partial_n u = 0 \quad \text{on } \partial\Omega$$

- Variational formulation, with $\mathcal{U} = \big\{ w \in H^1(\Omega), \langle w \rangle_\Omega = 0 \big\}$:

$$\text{Find } u \in \mathcal{U}, \quad A(\sigma, u, w) - F(w) = 0 \quad \text{for all } w \in \mathcal{U} \qquad \begin{cases} A(\sigma, u, w) := \displaystyle\int_\Omega \sigma \boldsymbol{\nabla}u \cdot \boldsymbol{\nabla}w \, \mathrm{d}V \\ F(w) := \displaystyle\int_\Omega f\, w,\, \mathrm{d}S \end{cases}$$

- Well-posedness: the potential $u \in \mathcal{U}$ exists, is unique, and is continuous in f with

$$\boxed{\|u\|_{1,\Omega} \leq \frac{1}{\alpha}\|f\|_{0,\Omega}}$$

for any $f \in L^2(\Omega)$ (by Lax-Milgram lemma, $\alpha\|w\|^2_{1,\Omega} \leq A(\sigma, w, w)$, $F(w) \leq \|f\|_{0,\Omega}\,\|w\|_{1,\Omega}$)

Example: scalar equilibrium problem

Inverse source problem: given σ, find f from (possibly noisy) measurements u^{obs} of u (in Ω, say).

- Measured data u^{obs} usually assumed in $L^2(\Omega)$ (rather than $H^1(\Omega)$) due to lack of control on measurement noise
- Source-to-solution mapping: $f \in L^2(\Omega) \to u = \mathcal{S}(f) \in H^1(\Omega)$, continuous
- Source-to-measurement mapping: $f \in L^2(\Omega) \to u = \mathcal{G}(f) \in L^2(\Omega)$, continuous too
- We have $\mathcal{G} = \mathcal{I} \circ \mathcal{S}$ where $\mathcal{I}$ involves a: compact embedding, hence $\mathcal{G}$ is compact.
- Likewise (measurement on $\partial\Omega$), $u \in H^{1/2}(\partial\Omega)$ but $u^{\text{obs}} \in L^2(\partial\Omega)$; again, $\mathcal{G}$ is compact.

1D example: $\boxed{-u'' = f,\ u(0) = u(1) = 0}$. Take $f_n = f + A\cos 2n\pi x$, then

$$u[f_n] - u[f] = \frac{A}{(2n\pi)^2}\big[\cos 2n\pi x - 1\big]$$

Small-amplitude, oscillatory measurement perturbation if n large.
Data noise consistent with **finite-amplitude, oscillatory** conductivity perturbation

Example: scalar equilibrium problem

Inverse conductivity problem: given f, find σ from (possibly noisy) measurement u^{obs} of u in Ω.

- Conductivity-to-solution mapping: $\sigma \in L^\infty(\Omega) \to u = \mathcal{S}(\sigma) \in H^1(\Omega)$, continuous
- Conductivity-to-measurement mapping: $\sigma \in L^\infty(\Omega) \to u = \mathcal{G}(\sigma) \in L^2(\Omega)$, continuous too
- Again, $\mathcal{G} = \mathcal{I} \circ \mathcal{S}$ and $\mathcal{G}$ is compact.

1D example: $\boxed{-(\sigma u')' = 2,\ u(0) = u(1) = 0}$. For $\sigma = 1$, we have $u[\sigma] = x(1-x)$.
Perturbed conductivity: $\sigma_n = 1 - \frac{\cos n\pi x}{2+\cos n\pi x}$. Perturbed solution (known in closed form) verifies

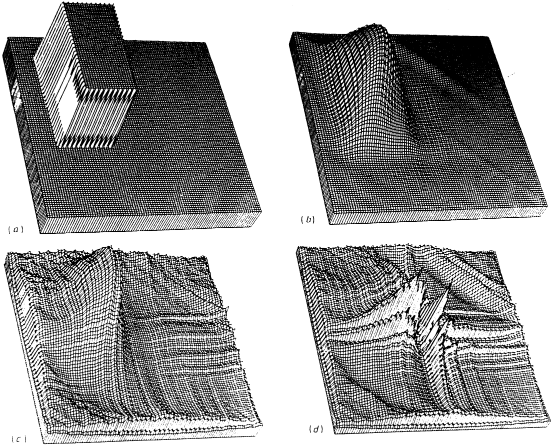
$$u[\sigma_n] - u[\sigma] = \frac{1-2x}{4n\pi} \sin 2n\pi x + \frac{1}{(2\pi n)^2} \big[1 - \cos 2n\pi x \big]$$

Small-amplitude, oscillatory measurement perturbation if n large. Moreover:

$$\|u[\sigma_n] - u[\sigma]\|_{L^2} = O(n^{-1}) \quad \text{but} \quad \|u[\sigma_n] - u[\sigma]\|_{H^1} = O(1)$$

Data noise consistent with **finite-amplitude, oscillatory** conductivity perturbation

Example (scalar equilibrium problem): conductivity reconstruction from (simulated) boundary data on potential



(a) "true" σ; reconstructions (b) no data noise, (c) 3% data noise (11 iterations), (d) 3% data noise (50 iterations).

Kohn R.V., McKenney A., *Inverse Problems* **6**:389–414 (1990).

Example: annulus, reconstruction of internal boundary data

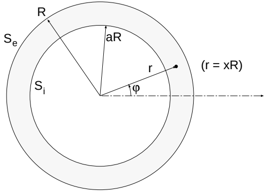
2D heat equilibrium equation: $\Delta\theta = \partial_{rr}\theta + \frac{1}{r}\partial_r\theta + \frac{1}{r^2}\partial_{\varphi\varphi}\theta = 0$

General solution:

$$\theta(r,\varphi) = a_0 + b_0 \operatorname{Log} r + \sum_{n\geq 1}\Big\{(a_n r^n + c_n r^{-n})\cos n\varphi + (b_n r^n + d_n r^{-n})\sin n\varphi\Big\}$$

Data on external boundary $r = R$:

$$\theta(R,\varphi) = f(\varphi) = \alpha_0 + \sum_{n\geq 1} \alpha_n \cos n\varphi + \beta_n \sin n\varphi$$
$$\partial_n\theta(R,\varphi) = g(\varphi) = \gamma_0 + \sum_{n\geq 1} \gamma_n \cos n\varphi + \delta_n \sin n\varphi$$



Unknown temperature and flux at depth $r = xR$:

$$\theta(xR,\varphi) = A_0 + \sum_{n\geq 1} A_n \cos n\varphi + B_n \sin n\varphi$$
$$\theta_{,n}(xR,\varphi) = C_0 + \sum_{n\geq 1} C_n \cos n\varphi + D_n \sin n\varphi$$

Unknowns A_n, B_n, C_n, D_n linked to data $\alpha_n, \beta_n, \gamma_n, \delta_n$ (upon elimination of a_n, c_n, b_n, d_n) by

$$\mathbb{G}_n \begin{Bmatrix} nA_n \\ xRC_n \end{Bmatrix} = \begin{Bmatrix} 2n\alpha_n \\ 2R\gamma_n \end{Bmatrix}, \qquad \mathbb{G}_n \begin{Bmatrix} nB_n \\ xRD_n \end{Bmatrix} = \begin{Bmatrix} 2n\beta_n \\ 2R\delta_n \end{Bmatrix} \qquad n \geq 1$$

with $\mathbb{G}_n = \begin{bmatrix} x^n + x^{-n} & x^n - x^{-n} \\ x^n - x^{-n} & x^n + x^{-n} \end{bmatrix} \qquad (x = r/R \leq 1)$

Singular (actually, eigen-)values of G_n: $\sigma_{n,1} = 2x^{-2n}, \quad \sigma_{n,2} = 2x^{2n}$ (exponentially decreasing).

Example: backward heat conduction equation

$$\begin{aligned} \kappa\partial_{xx}u - \partial_t u &= 0 && (0 \le t \le T,\ 0 \le x \le \ell) && \kappa := k/(\rho c) \\ u(0,t) = u(\ell,t) &= 0 && (0 \le t \le T) \\ u(x,0) &= u_0(x) && (0 \le x \le \ell) \end{aligned}$$

Initial temperature: $u_0(x) = \sum_{n\ge 0} a_n \sin\frac{n\pi x}{\ell}, \quad a_n = \frac{2}{\ell}\int_0^\ell u_0(x)\sin\frac{n\pi x}{\ell}\,\mathrm{dx}$

final temperature: $u(x,T) = \sum_{n\ge 0} b_n \sin\frac{n\pi x}{\ell}, \quad b_n = a_n \underbrace{e^{-(n\pi)^2\kappa T/\ell^2}}_{\sigma_n}$

- Reconstruction of u_0 given $u(x,T)$ (explicit inversion):

$$u_0(x) = \sum_{n\ge 0} \sigma_n^{-1} b_n \sin\frac{n\pi x}{\ell} \qquad \sigma_n^{-1} = O(e^{Cn^2})\,!$$

- Inverse problem: solve compact operator equation $u(.,T) = \mathcal{K}u_0$, where

$$\boxed{\mathcal{K}w := \int_0^\ell K(x,y)w(y)\,\mathrm{dy}}, \qquad K(x,y) := \sum_{n\ge 0} \sigma_n \sin\frac{n\pi x}{\ell}\sin\frac{n\pi y}{\ell}$$

More generally (regularizing effect of heat diffusion): the solution of

$$\partial_t u - \kappa\Delta u = 0 \text{ in } \Omega\times[0,T], \quad u = 0 \text{ on } \partial\Omega, \quad u(\cdot,0) = u_0 \in L^2(\Omega)$$

verifies $u \in C^\infty\big([\varepsilon,T]; H_0^1(\Omega)\big)$ for any $\varepsilon > 0$

Compact operators

Let $\mathcal{X}, \mathcal{Y}$ Hilbert spaces, and let $\mathsf{G} : \mathcal{X} \to \mathcal{Y}$ be linear and bounded (i.e. continuous).

- G is compact iff for any bounded sequence $(x_n)_n \in \mathcal{X}$, $y_n := \mathsf{G}(x_n)$ contains a subsequence converging in $\mathcal{Y}$.
- G is compact iff there exist $(\sigma_n)_n \geq 0$ (singular values, with $\sigma_n \to 0$) and orthonormal sets of functions $(f_n)_n \in \mathcal{X}$, $(g_n)_n \in \mathcal{Y}$ such that

$$\boxed{\mathsf{G} = \sum_{n\geq 0} \sigma_n \big(f_n, \bullet \big)\, g_n}, \qquad \text{i.e.} \qquad \mathsf{G}x = \sum_{n\geq 0} \sigma_n \big(f_n, x \big)\, g_n \text{ for any } x \in \mathcal{X}$$

with the series converging (in operator and $\|\bullet\|_{\mathcal{Y}}$ norms, resp.)

G^{-1} : Range(G) $\subset \mathcal{Y} \to \mathcal{X}$ cannot be continuous: $\mathsf{G}^{-1} g_n = \sigma_n^{-1} f_n$, hence for any $C = 0$ there exists $y \in \mathcal{Y}$ such that $\|\mathsf{G}^{-1}y\|_{\mathcal{X}} / \|y\|_{\mathcal{Y}} \geq C$.

Hence compact operator equations are ill-posed (they routinely occur in inverse problems!)

Prototype of linear ill-posed problem: first-kind integral equation with kernel $K \in L^2(\Omega_1 \times \Omega_2)$ and data $f \in L^2(\Omega_2)$:

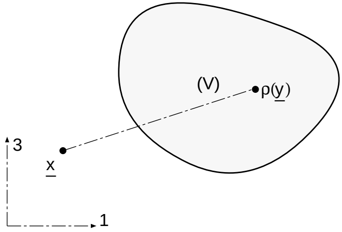
$$\boxed{\mathsf{G}u = f \text{ in } \Omega_2}, \qquad \mathsf{G} : L^2(\Omega_1) \to L^2(\Omega_2), \quad \mathsf{G}u(\bullet) := \int_{\Omega_1} K(\boldsymbol{x}_1, \bullet) u(\boldsymbol{x}_1)\, \mathrm{d}\boldsymbol{x}_1 \quad \text{is compact}$$

Example: gravimety

Physical model (explicit): gravity field induced by mass density ϱ in region V:

$$\boldsymbol{g} = \boldsymbol{A}\varrho, \quad [\boldsymbol{A}\varrho](\boldsymbol{x}) = \mathcal{G}\boldsymbol{\nabla}_x \int_V \frac{\varrho(\boldsymbol{y})}{\|\boldsymbol{x}-\boldsymbol{y}\|}\, \mathrm{d}V(\boldsymbol{y}) =: \mathcal{G}(\rho)(\boldsymbol{x})$$

($\mathbb{G} \approx 6.67408\, 10^{-11}$ m^3 kg^{-1} s^{-2})



Inverse problem: given gravity measurements $\boldsymbol{g}^{\text{obs}}$, solve $\boldsymbol{A}\varrho = \boldsymbol{g}^{\text{obs}}$ (ill-posed 1st-kind linear integral equation)

Solution multiplicity: Let $z(\boldsymbol{y})$ defined in V such that $z = \partial_n z = 0$ on ∂V. Then (3rd Green identity and $\boldsymbol{y} \mapsto 1/\|\boldsymbol{x}-\boldsymbol{y}\|$ harmonic):

$$\boldsymbol{A}(\varrho + \Delta z) = \boldsymbol{A}\varrho \quad \text{in } \mathbb{R}^3 \setminus \overline{V}$$

Example: gravimetry

Source: http://www.csr.utexas.edu/grace/

Example: deconvolution

First-kind integral equation of the form

$$f(\boldsymbol{y}) = \int_{\Omega} K(\boldsymbol{y}-\boldsymbol{x})\, u(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

E.g. model linear response of measurement devices:

- $K(\boldsymbol{z}) = \delta(\boldsymbol{z})$ (perfect device)
- $k(\boldsymbol{z}) = A \exp(-\|\boldsymbol{z}\|^2/2b)$ (image blurring caused by atmospheric turbulence)
- ...
- **Deconvolution** e.g. used for the restoration of blurred images

Linear response of dynamical system:

- $k(\boldsymbol{x}, t)$: response at $(\boldsymbol{x}, t)$ to impulsive point load $\delta(\boldsymbol{x})\delta(t)$ in an infinite medium
- Then, for an arbitrary excitation $\phi(\boldsymbol{x}, \tau)$ $(\tau \geq 0$:

$$u(\boldsymbol{y}, t) = \int_0^T \int_{\mathbb{R}^3} k(\boldsymbol{x}-\boldsymbol{y}, \tau - t)\phi(\boldsymbol{x}, \tau)\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\tau$$

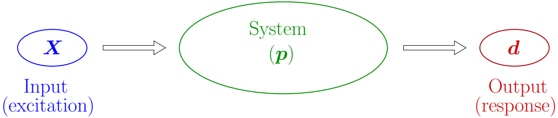
$\Longrightarrow$ identification of dynamical source is a deconvolution problem.

Deconvolution usually an ill-posed problem, due to smoothing character of forward convolution.

Forward and inverse problems

Forward problem (mechanics, acoustics, electromagnetism, heat transfer...):

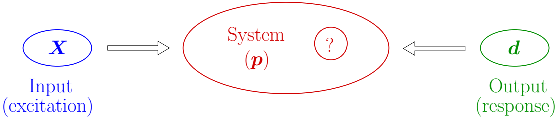
- Compute *response* $\boldsymbol{d}$ (displacement, stress, temperature, potentials. . .) to *excitations* $\boldsymbol{X}$ (sources, applied loads. . .)



- System depends on known parameters (geometry, material, constraints. . .)
- (well-posed) forward problem: find response $\boldsymbol{d}$ given excitation $\boldsymbol{X}$ and parameters $\boldsymbol{p}$.

Inverse problem: System at least partially unknown (sometimes, excitation unknown).

- (ill-posed) inverse problem: find missing information on system, given measurements of responses under given excitations;
  Sometimes, find missing information on excitations, given system and measured responses.



Equation-solving vs. optimization

Equation solving viewpoint, e.g. find $\boldsymbol{p}$ such that $\boldsymbol{d}^{\text{obs}} = \mathbf{G}(\boldsymbol{p};\boldsymbol{X})$, often unsuitable:

- No guarantee on solution existence or uniqueness (model underdetermination, data inconsistent with model due to data noise and modeling assumptions);
- Discretized (or intrinsically finite-dimensional) inverse problems often such that

$$\boxed{\text{Dim}(\mathcal{D}) \neq \text{Dim}(\mathcal{P})} \qquad (\mathcal{D},\ \mathcal{P} : \text{data and parameter spaces})$$

- Overdetermination often desirable, up to

$$\boxed{\text{Dim}(\mathcal{D}) \gg \text{Dim}(\mathcal{P})}$$

- Consequently, inverse problem often set as minimization:

$$\boldsymbol{p}^{\star} \in \arg\min_{\boldsymbol{p}\in\mathcal{P}} J(\boldsymbol{p}), \qquad \text{e.g. } J(\boldsymbol{p}) = \|\boldsymbol{d}^{\text{obs}} - \mathbf{G}(\boldsymbol{p};\boldsymbol{X})\|_{\mathcal{D}}$$

$J(\boldsymbol{p})$ (usually) depends implicitly on $\boldsymbol{p}$: evaluating $\boldsymbol{d} = \mathbf{G}(\boldsymbol{p};\boldsymbol{X})$ requires solving forward problem (ODE, PDE. . .).

$\Longrightarrow$ frequent use of ODE / PDE constrained optimization

Regularization methods for inverse problems

- Let $\mathsf{G} : \mathcal{P} \to \mathcal{D}$. A family $\mathsf{R}_\alpha : \mathcal{D} \to \mathcal{P}$ of operators defines a regularization strategy for G if
  - → R_α is continuous;
  - → $\|\mathsf{R}_\alpha \mathsf{G} p - p\|_{\mathcal{P}} \to 0$ as $\alpha \to 0$ for any $p \in \mathcal{P}$.

  Consequently:
  - → Assume $\mathsf{G}p = d$. Then, $\mathsf{R}_\alpha d = \mathsf{R}_\alpha \mathsf{G} p \to p$ as $\alpha \to 0$;
  - → If G is invertible with continuous inverse, we can set $\mathsf{R}_\alpha = \mathsf{G}^{-1}$ as expected;
- Consider noisy data d^δ such that $\delta = \|d^\delta - d\|$ (data noise level). Then, a regularized solution p^δ_α of the inverse problem is defined as $\boxed{p^\delta_\alpha := \mathsf{R}_\alpha d^\delta}$.
- **Splitting of inversion error** (for linear G and R_α): stability obtained at some cost in accuracy:
  $$\|p^\delta_\alpha - p\| \leq \|\mathsf{R}_\alpha (d^\delta - d)\| + \|\mathsf{R}_\alpha \mathsf{G} p - p\|$$
- **Regularization parameter choice:** aims at tuning $\alpha = \alpha(\delta)$ so that
  $$\|\mathsf{R}_{\alpha(\delta)}(d^\delta - d)\| \to 0 \quad \text{as } \delta \to 0,$$
  so that $p^\delta_{\alpha(\delta)} := \mathsf{R}_{\alpha(\delta)} d^\delta$ achieves $\|p^\delta_{\alpha(\delta)} - p\| \to 0$ as $\delta \to 0$.

Main general approaches to regularization

- **Optimization with penalization:** solve inverse problem as optimization problem of the form

  $$\min_p \Phi\big(\|\mathsf{G}p - d\|\big) + \alpha B(p),$$

  where B is some positive functional which controls p. Archetypal example: Tikhonov regularization

  $$\min_p \|\mathsf{G}p - d\|^2 + \alpha\|p\|^2$$

- **Reduced / sparse discretization of p:** seek p in some finite-dimensional space $\mathcal{U}$, with $\alpha = (\text{dim}(\mathcal{U}))^{-1}$
- **Iterative solver with premature stopping** ("Landweber method"): apply iterative algorithm to $\mathsf{G}p = d$, set $p_{\alpha(\delta)} = p_{1/N(\delta)}$ where some selection rule defines prtemature stopping $N(\delta)$ for given noisy data d^δ.

- **Regularization parameter choice strategies:**
  - ▷ A posteriori rules, rely on availability of data noise estimate δ (e.g. Morozov principle $\|\mathsf{G}p - d\| \leq \delta$), value found in the process of computing p_α;
  - ▷ A priori rules, rely on the former and some prior information on solution smoothness, value fixed without computing p_α;
  - ▷ Data error-free rules: typically aim at balancing data fidelity and regularization. In particular: (i) Generalized cross-validation, (ii) L-curve heuristics.

Ill-conditioned linear systems

- Discretized (or finite-dimensional) inverse problems involve linear systems $Gp = d$ ($G \in \mathbb{K}^{m \times n}$, most often $m \geq n$) with often "unpleasant" characteristics:
  - → G often has no specific properties (e.g. invertibility, symmetry, sign...) and can even be rectangular;
  - → G may have full (theoretical) column rank;
  - → However, G ill-conditioned with very fast decay of singular values, i.e. numerically rank-deficient:

    $$\boxed{\|G - G_r\| \ll \|G\|} \quad \text{for some rank-}r\text{ matrix } G_r,\ r \ll n$$

  - → imperfect data d (e.g. measurement errors)
- Studying them useful (a) in their own right, and (b) in preparation to more-complex settings.
- Overall concern: how to make the most of general linear systems lacking specific structure?

Condition number of matrices and linear systems

Solution of linear system $\boxed{Ay = b}$: sensitivity to data A, b ($A \in \mathbb{K}^{n\times n}$ invertible)

- Perturbation z of solution y satisfies perturbed system $(A+E)(y+z) = b+f$.
- Relative sensitivity of solution w.r.t. data depends on condition number $\kappa(A) = \|A^{-1}\|\|A\|$:

$$\frac{\|z\|/\|y\|}{\|f\|/\|b\| + \|E\|/\|A\|} \leq \kappa(A) + O\big(\|E\|/\|A\|$$

- **Properties** of $\kappa(A)$:
  - ▷ $\kappa(A)$ depends on choice of (matrix) norm. $\boxed{\kappa(A) \geq 1}$ for any induced norm.
  - ▷ For arbitrary $A \in \mathbb{K}^{m\times n}$, $\kappa_2(A)$ given in terms of either singular values or pseudo-inverse of A.
  - ▷ $\kappa_2(Q) = 1$ if Q orthogonal or unitary (since $\|Q\|_2 = 1$ and $\|Q^{-1}\|_2 = 1$).
  - ▷ Discretization of ill-posed linear equations yield ill-conditioned linear systems

A simple numerical example

- Example (exact matrix inverse):

$$A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \implies A^{-1} = \begin{bmatrix} 25 & 41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{bmatrix}$$

  Note: $AA^{\mathsf{T}} = A^{\mathsf{T}}A$ (i.e. A is normal)

- Effect of perturbations of A or b on solution of x of $Ax = b$:

$$b = [\,32 \;\; 23 \;\; 33 \;\; 31\,]^{\mathsf{T}} \implies x = [\,1 \;\; 1 \;\; 1 \;\; 1\,]^{\mathsf{T}}$$

$$\delta b = [\,0.1 \;\; -0.1 \;\; 0.1 \;\; -0.1\,]^{\mathsf{T}} \implies x = [\,9.2 \;\; -12.6 \;\; 4.5 \;\; -1.1\,]^{\mathsf{T}}$$

$$\delta A_{23} = 0.1 \implies x \approx [\,-4.86 \;\; -10.7 \;\; -1.43 \;\; -2.43\,]^{\mathsf{T}}$$

- Eigenvalues of A:

$$\Lambda \approx \text{Diag}[\,30.29 \;\; 3.858 \;\; 0.8431 \;\; 0.01015\,], \qquad \kappa_2(A) \approx 3\,10^3$$

  A is a rather ill-conditioned 4×4 matrix.

Singular value decomposition (SVD)

- Diagonalization of square matrices: $A = X\Lambda X^{-1}$ for some invertible X and diagonal Λ
  - A normal if and only if A diagonalizable with X unitary. Includes all Hermitian matrices.
  - There are (non-normal) non-diagonalizable matrices, e.g. $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$

Singular value decomposition generalizes diagonalization to any (even rectangular) matrix.

- Let $A \in \mathbb{K}^{m\times n}$, then $A^{\mathsf{H}}A \in \mathbb{K}^{n\times n}$ and $AA^{\mathsf{H}} \in \mathbb{K}^{m\times m}$ are square Hermitian.
  Well-defined (symmetric positive) eigenvalue problems: $\boxed{A^{\mathsf{H}}Av = \lambda v}$ and $\boxed{AA^{\mathsf{H}}u = \mu u}$

  (λ, v) eigenpair of $A^{\mathsf{H}}A \implies (\lambda, Av)$ eigenpair of AA^{H}
  (μ, u) eigenpair of $AA^{\mathsf{H}} \implies (\mu, A^{\mathsf{H}}u)$ eigenpair of $A^{\mathsf{H}}A$

  (equal multiplicities if $\lambda = \mu > 0$, unequal multiplicity in general for $\lambda = \mu = 0$).

Singular value decomposition

Any $A \in \mathbb{K}^{m\times n}$ has a SVD $A = USV^{\mathsf{H}}$, where:

- $U = [u_n, \ldots, u_m] \in \mathbb{K}^{m\times m}$, $V = [v_1, \ldots, v_n] \in \mathbb{K}^{n\times n}$: unitary square matrices,
- $S \in \mathbb{R}^{m\times n}$ "diagonal", with $S_{ii} = \sigma_i$, $S_{ij} = 0$ if $i \neq j$.
- Singular values σ_i are real positive; conventionally $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_{\min(m,n)} \geq 0$.

- Available operators: `[U,S,V]=svd(A)` (MATLAB), `F=svdfact(A)` (Julia; `F` contains U, S, V)

Singular value decomposition (SVD): some properties

- $\lambda_i = \mu_i = \sigma_i^2 \; (1 \leq i \leq r)$.
- $A = USV^{\mathsf{H}} = \sum_{i=1}^{\min(m,n)} \sigma_i u_i v_i^{\mathsf{H}}$ (SVD is weighted sum of rank-one matrices).
- **Reduced SVD.** For $A \in \mathbb{K}^{m\times n}$, $\mathrm{rank}(A) = r \leq \min(m,n)$:

  $$A = USV^{\mathsf{H}} = U_r S_r V_r^{\mathsf{H}} \qquad U_r = [u_1, \ldots, u_r], \quad S_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r), \quad V_r = [v_1, \ldots, v_r]$$

  (vectors $u_{r+1}, \ldots, u_m$ and $v_{r+1}, \ldots, v_n$ inactive, generate $\mathcal{N}(A)$ and $\mathcal{R}(A)^{\perp}$).
- SVD is rank-revealing: $\mathrm{rank}(A)$ equal to number of nonzero singular values.
- Matrix 2-norm: we have $\|A\|_2 = \sigma_1$, since

  $$\|Ax\|_2 = \|U_r S_r V_r x\|_2 = \|S_r V_r x\|_2 \leq \|S_r\|_2 \|V_r x\|_2 \leq \|S_r\|_2 \|x\|_2 = \sigma_1 \|x\|_2$$
- 2-norm condition number: we have $\kappa_2(A) = \sigma_1 / \sigma_r$
- Computing a SVD
  - $\rightarrow$ requires solving an eigenvalue problem,
  - $\rightarrow$ takes $O(m^2 n)$ operations for $A \in \mathbb{K}^{m\times n}$.

SVD applied to arbitrary linear systems

Solvability of $Ax = b$ ($A \in \mathbb{K}^{m\times n}$):

$$Ax = b \implies USV^{\mathsf{H}}x = b \implies S(V_r^{\mathsf{H}}x) = U^{\mathsf{H}}b \qquad \text{i.e.} \quad \begin{cases} \sigma_i(v_i^{\mathsf{H}}x) = u_i^{\mathsf{H}}b & (1 \le i \le r) \\ 0 = u_i^{\mathsf{H}}b & (r+1 \le i \le m) \end{cases}$$

- Solvability condition: $u_i^{\mathsf{H}}b = 0 \ (r+1 \le i \le m)$, expresses $b \in \mathcal{R}(A)$.
  Then, $\boxed{\sigma_i(v_i^{\mathsf{H}}x) = u_i^{\mathsf{H}}b}$ $(1 \le i \le r)$ determine these projections $v_i^{\mathsf{H}}x$ uniquely
  Remaining projections $v_i^{\mathsf{H}}x$ $(r+1 \le i \le m)$ arbitrary.
- General solution (if it exists):

$$\boxed{x = \sum_{i=1}^{r} \frac{u_i^{\mathsf{H}}b}{\sigma_i} v_i + \sum_{i=r+1}^{n} x_i v_i} \qquad (x_{r+1}, \ldots, x_n) \in \mathbb{K}^{n-r} \text{ arbitrary}$$

  Setting $x_{r+1} = \ldots = x_n = 0$ gives minimum-norm solution
- Uniqueness condition: $\boxed{r = n}$ (implies $m \ge n$).

SVD applied to least-squares problems

Linear least-squares problem $\min_{x\in\mathbb{K}^n} \|Ax-b\|^2$:

$$\begin{aligned}\|Ax-b\|^2 &= \|USV^{\mathsf{H}}x-b\|^2 = \|U(SV^{\mathsf{H}}x-U^{\mathsf{H}}b)\|^2 = \|S(V^{\mathsf{H}}x)-U^{\mathsf{H}}b\|^2 \\ &= \sum_{i=1}^{r} |\sigma_i(v_i^{\mathsf{H}}x)-(u_i^{\mathsf{H}}b)|^2 + \sum_{i=r+1}^{m} |u_i^{\mathsf{H}}b|^2\end{aligned}$$

Solutions always exist and are given by:

$$\boxed{x = \sum_{i=1}^{r} \frac{u_i^{\mathsf{H}}b}{\sigma_i} v_i + \sum_{i=r+1}^{n} x_i v_i} \qquad (x_{r+1},\ldots,x_n)\in\mathbb{K}^{n-r} \text{ arbitrary}$$

Residual:

$$\min_{x\in\mathbb{K}^n} \|Ax-b\|^2 = \sum_{i=r+1}^{n} |u_i^{\mathsf{H}}b|^2$$

Pseudo-inverse of a matrix

Generalized inverse (i.e. pseudo-inverse) of $A \in \mathbb{K}^{m\times n}$: a matrix $A^\dagger \in \mathbb{K}^{n\times m}$ verifying

(a)	$AA^\dagger A = A,$	(b)	$(AA^\dagger)^{\mathsf{H}} = AA^\dagger,$
(c)	$A^\dagger AA^\dagger = A^\dagger,$	(d)	$(A^\dagger A)^{\mathsf{H}} = A^\dagger A.$

(Moore-Penrose conditions)

Algebraic properties of $A^\dagger$

- $(A^\dagger)^+ = A$.
- (Moore-Penrose) Pseudo-inverse $A^\dagger$ satisfying (a)-(d) exists and is unique
- If A invertible, $A^\dagger = A^{-1}$
- If $\text{rank}(A) = n$ (full column rank, hence $m \geq n$, $A^{\mathsf{H}}A$ invertible): $A^\dagger = (A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}}$.
- If $\text{rank}(A) = m$ (full row rank, hence $m \leq n$, AA^{H} invertible), $A^\dagger = A^{\mathsf{H}}(AA^{\mathsf{H}}A)^{-1}$.
- Explicit formula using reduced SVD: $A^\dagger = V_r S_r^{-1} U_r^{\mathsf{H}}$

General solution of least-squares problem:

$x = A^\dagger b + (I - A^\dagger A)w$ ($A^\dagger b$ minimum-norm solution, $, w \in \mathbb{K}^n$ arbitrary)

- $A^\dagger$ does not depend continuously on A. Example ($\text{rank}(A) = r$, $\text{rank}(A_\varepsilon) = r+1$):

$$A_\varepsilon = U_r S_r V_r^{\mathsf{H}}, \qquad A_\varepsilon = U_r S_r V_r^{\mathsf{H}} + \varepsilon u_{r+1} v_{r+1}^{\mathsf{H}} \implies \boxed{\frac{\|A_\varepsilon^\dagger - A^\dagger\|}{\|A^\dagger\|} = \frac{\sigma_r}{\varepsilon} \gg 1}$$

Regularization of least squares by filtering

- Consider reference situation $p \in \mathbb{K}^n$, $d \in \mathbb{K}^m$ such that (using SVD (u_n, v_n, σ_n) of G)

$$Gp = d, \quad \|p\| \text{ minimum when } r < n, \qquad \text{i.e.} \quad p = \sum_{i=1}^{r} \frac{1}{\sigma_i} (u_i^{\mathsf{H}} d)\, v_i$$

- Filtered inversion for noisy data $d^\delta = d + b$ ($\delta = \|b\|$: noise level): set

$$p_{\alpha,\delta} = \mathsf{R}_\alpha d^\delta = \sum_{i=1}^{n} w_\alpha(\sigma_i) \frac{1}{\sigma_i} (u_i^{\mathsf{H}} d)\, v_i$$

for some filter function w_α depending on a parameter $\alpha > 0$ such that

$$\lim_{\sigma\to 0} w_\alpha(\sigma)/\sigma = 0, \qquad \lim_{\sigma\to\infty} w_\alpha(\sigma) = 1, \qquad \lim_{\alpha\to 0} w_\alpha(\sigma) = 1, \qquad \lim_{\alpha\to\infty} w_\alpha(\sigma) = 0.$$

- Note $p = \mathsf{R}_0 d$ (reference situation, $\alpha = 0$) recovered in the limit $\alpha \to 0$, and

$$\boxed{\lim_{\alpha\to 0} p_{\alpha,0} = p, \qquad \lim_{\alpha\to 0} p_{\alpha,\delta} = p_\delta}$$

- Splitting of reconstruction error:

$$e_{\alpha,\delta} = \mathsf{R}_\alpha d^\delta - \mathsf{R}_0 d = e^{\mathsf{noise}}_{\alpha,\delta} + e^{\mathsf{reg}}_\alpha, \qquad e^{\mathsf{noise}}_{\alpha,\delta} := \mathsf{R}_\alpha (d^\delta - d), \quad e^{\mathsf{reg}}_\alpha := \mathsf{R}_\alpha d - \mathsf{R}_0 d$$

We find
$$\boxed{e^{\mathsf{noise}}_{\alpha,\delta} = \sum_{i=1}^{n} w_\alpha(\sigma_i)\, \frac{1}{\sigma_i} (u_i^{\mathsf{H}} b) v_i, \qquad e^{\mathsf{reg}}_\alpha = \sum_{i=1}^{n} \big(w_\alpha(\sigma_i) - 1 \big)\, \frac{1}{\sigma_i} (u_i^{\mathsf{H}} d) v_i}$$

In particular, $e^{\mathsf{reg}}_\alpha \to 0$ if $\alpha \to 0$.

- Limiting values:
$$\boxed{\lim_{\alpha\to 0} \|e^{\mathsf{noise}}_{\alpha,\delta}\| = \|\mathsf{R}_0 b\|, \qquad \lim_{\alpha\to\infty} \|e^{\mathsf{reg}}_{\alpha,\delta}\| = \|\mathsf{R}_0 d\|}$$

Regularization of least squares by filtering: minimal-norm penalization

- Regularized least squares with minimal-norm penalization:

$$\mathsf{R}_\alpha d = \arg\min_{p\in\mathbb{K}^n} S_\alpha(p;d)\,, \qquad S_\alpha(p;d) = \frac{1}{2}\|Gp-d\|^2 + \frac{\alpha}{2}\|p\|^2$$

$$\boxed{\mathsf{R}_\alpha d = \big[G^{\mathsf{T}}G+\alpha I\big]^{-1}G^{\mathsf{T}}\boldsymbol{d}}$$; corresponding filter: $$\boxed{w_\alpha(\sigma) = \frac{\sigma^2}{\sigma^2+\alpha}}$$

- Modulus of continuity a decreasing function of α: we find

$$\|\mathsf{R}_\alpha\| = \max_{i\le r}\frac{\sigma_i}{\sigma_i^2+\alpha} = \begin{cases} \sigma_r/(\sigma_r^2+\alpha) & \text{if} \quad \alpha\le\sigma_r\sigma_{r-1}, \\ \sigma_i/(\sigma_i^2+\alpha) & \text{if} \quad \sigma_{i+1}\sigma_i\le\alpha\le\sigma_i\sigma_{i-1} \qquad 2\le i\le r-1, \\ \sigma_1/(\sigma_1^2+\alpha) & \text{if} \quad \sigma_2\sigma_1\le\alpha, \end{cases}$$

- Data noise and regularization contributions to reconstruction error: we find

$$e^{\mathsf{noise}}_{\alpha,\delta} = \sum_{i=1}^n \frac{\sigma_i}{\sigma_i^2+\alpha}(u_i^{\mathsf{H}}b)v_i, \qquad e^{\mathsf{reg}}_{\alpha} = \sum_{i=1}^n \frac{\alpha}{\sigma_i(\sigma_i^2+\alpha)}(u_i^{\mathsf{H}}d)v_i$$

and in particular:

$$\frac{\mathsf{d}}{\mathsf{d}\alpha}\|e^{\mathsf{noise}}_{\alpha,\delta}\|^2\Big|_{\alpha=0} < 0, \qquad \frac{\mathsf{d}}{\mathsf{d}\alpha}\|e^{\mathsf{reg}}_{\alpha,\delta}\|^2\Big|_{\alpha=0} > 0$$

Regularization (choice of α): a compromise between stability w.r.t. data noise and accuracy)

Regularization of least squares by filtering: minimal-norm penalization

- Since $\forall s > 0,\ s/(s^2+\alpha) \le \alpha^{-1/2}/2$, we have

$$\|e_\alpha^{\text{noise}}\| \le \frac{1}{2}\alpha^{-1/2}\delta, \qquad \lim_{\alpha\to 0} \|e_\alpha^{\text{reg}}\| = 0$$

  Ajusting α to δ: if $\alpha = \delta^p$ $(0 < p < 2)$

$$\|\boldsymbol{e}_\alpha\| \le \frac{1}{2}\delta^{1-p/2} + o(1) = o(1) \quad (\delta \to 0)$$

Analysis can be extended to e.g.

$$S_\alpha(p,d) = \frac{1}{2}\|Gp - d\|^2 + \frac{\alpha}{2}\|p - p_{\text{prior}}\|^2$$

Regularization of least squares by filtering: minimal-norm penalization

- Decomposition of cost function residual $S_\alpha(p^\alpha;d)$ at optimum into output residual and solution norm:

$$S_\alpha(p^\alpha;d) = D(\alpha) + \alpha R(\alpha) \qquad D(\alpha) := \|Gp^\alpha - d\|^2, \quad R(\alpha) := \|p^\alpha\|^2$$

- Evaluation of residuals (using SVD of G):

$$D(\alpha) = \sum_{i=1}^{r} \frac{\alpha^2}{(\sigma_i^2+\alpha)^2} |u_i^{\mathsf{H}} d|^2 + \sum_{i=r+1}^{m} |u_i^{\mathsf{H}} d|^2$$

$$R(\alpha) = \sum_{i=1}^{r} \frac{\sigma_i^2}{(\sigma_i^2+\alpha)^2} |u_i^{\mathsf{H}} d|^2$$

- Output residual **increases**, and regularized solution norm **decreases**, with α:

$$D'(\alpha) = 2\alpha \sum_{i=1}^{r} \frac{\sigma_i^2}{(\sigma_i^2+\alpha)^3} |u_i^{\mathsf{H}} d|^2 > 0, \qquad D'(\alpha) + \alpha R'(\alpha) = 0 \;(\implies R'(\alpha) < 0)$$

Regularization of least squares by filtering (minimal-norm penalization): L-curve

Define the L-curve (associated with cost function S_α) as the parametric curve $\alpha \geq 0 \mapsto (D(\alpha), R(\alpha))$ in (D, R)-plane.

- The L-curve is monotonic (R a decreasing function of D)

  Proof: follows at once from $D' > 0$ *and* $R' < 0$ □

- The L-curve is convex

  Proof: convexity here amounts to curvature $\kappa(\alpha)$ *being positive, with* $\kappa(\alpha) := (D'R'' - D''R') / (D'^2 + R'^2)^{3/2}$, *and hence reduces to verifying* $D'R'' - D''R' \geq 0$.
  Using property $D'(\alpha) = -\alpha R'(\alpha)$, *we deduce* $D'R'' - D''R' = R'^2 \geq 0$. □

- Extremal points $A := (D(0), R(0))$ et $B := (D(\infty), R(\infty))$ of L-curve ($p := \mathsf{R}_0 d$:

$$D(0) = \|Gp - d\|^2, \qquad R(0) = \|p\|^2$$

$$D(\infty) = \|d\|^2 > D(0) \qquad R(\infty) = 0$$

  Moreover, since (again) $D'(\alpha) = -\alpha R'(\alpha)$, the extremal slopes are

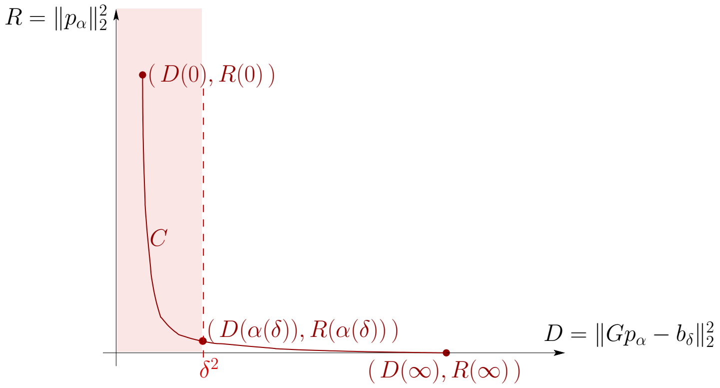
$$\frac{\mathrm{d}R}{\mathrm{d}D}\Big|_{\alpha=0} = -\infty, \qquad \frac{\mathrm{d}R}{\mathrm{d}D}\Big|_{\alpha=\infty} = 0^-$$

Regularized least squares: choice of α using L-curve

- Assume data noise level δ is known (realistic in some cases, e.g. mechanical testing using digital image correlation).
- Use that L-curve is convex, reformulate regularized least-squares:

$$\min_{p \in \mathbb{K}^n} \|p\|_2^2, \qquad \text{subject to} \quad \|Gp - d_\delta\|_2^2 \leq \delta^2$$

- Select α such that $D(\alpha) = \delta^2$ (i.e. set LS residual equal to data noise)
- Unique solution provided $\delta < \|d_\delta\|_2$



Regularization of least squares by filtering: truncated SVD

Matrices with fast decay of σ_i: truncated SVD as alternative to Tikhonov regularization.

- Ignore all singular values less than $\sqrt{\alpha} > 0$, i.e. set $\boxed{w_\alpha(\sigma) = H(\sigma^2 - \alpha)}$.
  Regularized inversion then defined by

$$p_\alpha = \mathsf{R}_\alpha d := \sum_{i=1}^{q(\alpha)} \frac{q_i^{\mathsf{H}} d}{\sigma_i} v_i, \qquad q(\alpha) := \sup\{ j,\ \sigma_j^2 - \alpha \geq 0 \}$$

$$\text{hence:} \qquad e^{\text{noise}}_{\alpha,\delta} = \sum_{i=1}^{q(\alpha)} \frac{1}{\sigma_i} (u_i^{\mathsf{H}} b) v_i, \qquad e^{\text{reg}}_{\alpha} = \sum_{i=1}^{q(\alpha)} \frac{1}{\sigma_i} (u_i^{\mathsf{H}} d) v_i$$

- Since $1/\sigma_i < \alpha^{-1/2}$ in $e^{\text{noise}}_{\alpha,\delta}$, we may adjust α to δ as before: if $\alpha = \delta^p$ $(0 < p < 2)$

$$\|e^{\text{noise}}_{\alpha}\| \leq \alpha^{-1/2}\delta, \quad \lim_{\alpha\to 0} \|e^{\text{reg}}_{\alpha}\| = 0 \quad \text{and} \quad \|\boldsymbol{e}_\alpha\| \leq \delta^{1-p/2} + o(1) = o(1) \quad (\delta \to 0)$$

- Again by analogy with regularized least squares, define

$$\boxed{D_q := \|Gp_q - d\|_2^2 = \sum_{i>q} |u_i^{\mathsf{H}} d|^2} \ \text{(decreasing)}, \qquad \boxed{R_q := \|p_q\|_2^2 = \sum_{i\leq q} \frac{|u_i^{\mathsf{H}} d|^2}{\sigma_i^2}} \ \text{(increasing)}$$

- L-curve C_n: interpolates points (D_q, R_q) $(1 \leq r \leq n)$. C_n is convex:

$$S_q := \frac{R_q - R_{q+1}}{D_q - D_{q+1}} = -\frac{|z_{q+1}|^2}{\sigma_{q+1}^2} \frac{1}{|z_{q+1}|^2} = -\frac{1}{\sigma_{q+1}^2}, \qquad q \mapsto S_q \text{ increasing}$$

- Discrete parameter $1/q$ plays role of regularization parameter α.

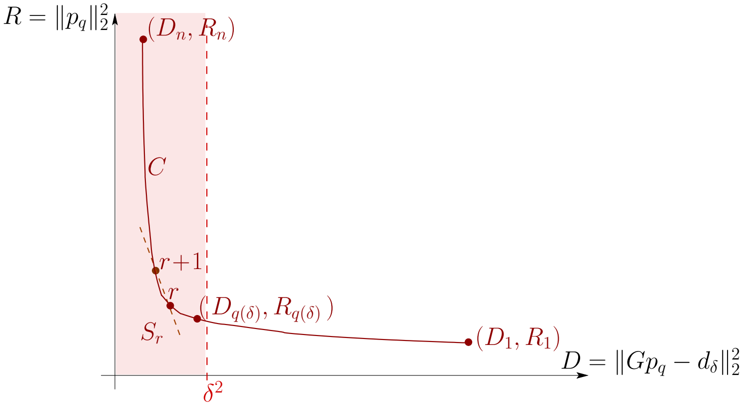
Regularized solution using truncated SVD

Eckart-Young-Mirsky theorem

Let $A \in \mathbb{K}^{m\times n}$, $q \leq n$. $\widehat{A}_q$ is best rank-q approximation of A (spectral and Frobenius norms):

$$\widehat{A}_q = \mathop{\arg\min}_{\substack{B\in\mathbb{K}^{m\times n}\\ \text{rank}(B)=q}} \Big\{ \|A-B\|_2 \text{ or } \|A-B\|_{\mathsf{F}} \Big\}; \qquad \|A-\widehat{A}_q\|_2 \leq \sigma_{q+1}, \quad \|A-\widehat{A}_q\|_{\mathsf{F}}^2 \leq \sum_{i=q+1}^{n} \sigma_i^2.$$

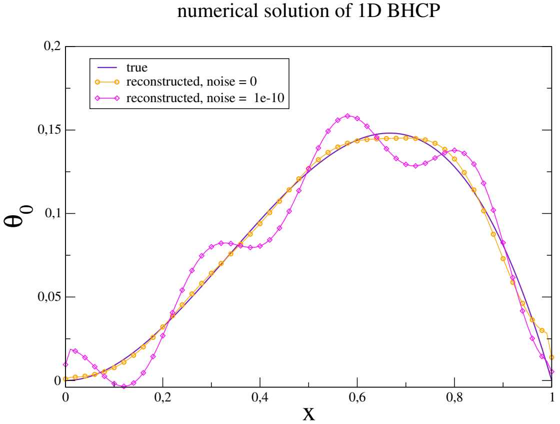
Discrete L-curve



Example: backward heat equation

Physical problem: find the temperature distribution in a system before thermal measurements are made (example: space vehicle re-entry).

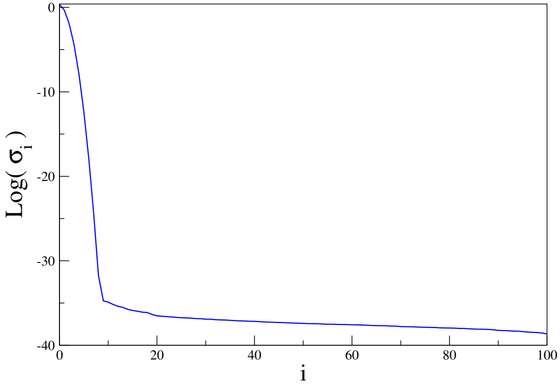
$$\underbrace{\Theta(\cdot, T)}_{\text{measurement}} = \underbrace{\mathcal{A}([0, T])}_{\text{heat eq.}}\underbrace{\Theta(\cdot, 0)}_{\text{unknown}}$$



Example: backward heat equation

Physical problem: find the temperature distribution in a system before thermal measurements are made (example: space vehicle re-entry).

$$\underbrace{\Theta(\cdot,T)}_{\text{measurement}} = \underbrace{\mathcal{A}(T)}_{\text{heat eq.}}\underbrace{\Theta(\cdot,0)}_{\text{unknown}} \implies \boxed{A(T)\Theta_0 = \Theta_T} \text{ after space discretization of } \Theta$$

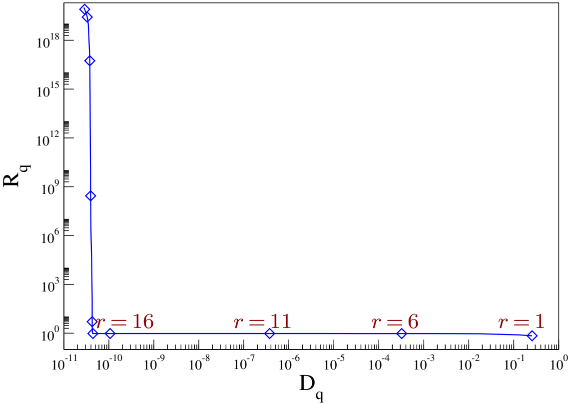


Singular values of $A(T)$ ($x \in [0,1]$, $\Delta x = 1/100$)

- Matrix A: exact rank 100, numerical rank < 10.

Example: backward heat equation

Discrete L-curve, simulated data with $\delta = 10^{-5}$,



- Optimal choice of q (L-curve for noise level $\delta = 10^{-5}$);
- Lowest actual temperature reconstruction error: $\approx 10^{-2}$ (in relative L^2 norm) for $r = 19$.

1. Inverse and ill-posed problems: examples

2. Short overview of solution approaches

3. Finite-dimensional ill-conditioned linear systems

4. **Regularization by promotion of sparsity**

5. (a glimpse of) Bayesian approach to inverse problems

Regularization by promotion of sparsity

Sparsity of (approximate) solutions of linear systems important for some applications:

- Image deblurring:

$$Gp = d + w, \qquad G = BW\ , \quad \begin{cases} W \in \mathbb{R}^{m\times n}: & \text{wavelet basis} \\ B \in \mathbb{R}^{m\times m}: & \text{models blurring} \\ p \in \mathbb{R}^{n}: & \text{restored image} \\ d \in \mathbb{R}^{m}: & \text{blurred image (data)} \\ w \in \mathbb{R}^{m}: & \text{unknown noise} \end{cases}$$

  Goal: find a sparse representation $Wp = \sum_{i=1}^{n} w_i p_i$ of restored image ($p_i = 0$ for many i)

- Reflexive idea: regularized least squares (see previous)

$$\min_{p\in\mathbb{R}^n} \|BWp - d\|_2^2 + \alpha\|p\|_2^2$$

  However, 2-norm regularizer $\|p\|_2^2$ allows many entries with small magnitude.

- Ideally, should use $p=0$ ("counting norm"). However, $\|\cdot\|_p^p$ non-smooth, non-convex if $p<1$.
- Compromise choice: use 1-norm regularizer; $\|\cdot\|_1$ convex, Lipschitz (but not smooth)

$$\min_{p\in\mathbb{R}^n} J_\alpha(p), \qquad J_\alpha(p) := \|BWp - d\|_2^2 + \alpha\|p\|_1 \quad (\text{“}L^2\text{-}L^1 \text{ functional”})$$

  Example: $p = \varepsilon(1,\ldots,1)$; $\|p\|_1 = n\varepsilon$ penalizes non-sparsity better than $\|p\|_2^2 = n\varepsilon^2$

- More-difficult minimization problem: J_α not quadratic and not differentiable

Minimization of functionals with a nonsmooth part

- Recall steepest descent update step for smooth (e.g. quadratic) functionals:

  $$\boxed{p^{(k+1)} = p^{(k)} - t^{(k)} \nabla J(p^{(k)})}\,, \qquad \text{step length } t^{(k)} \text{ found by line search}$$

  Method of explicit type (∇J evaluated at initial point).

- L^2-L^1 functional J_α not everywhere differentiable:
  $\rightarrow$ $\nabla J(p^{(k)})$ potentially not defined
  $\rightarrow$ $t \mapsto J_\alpha(p(t))$ potentially not differentiable at some $p(t) := p^{(k)} - t\nabla J(p^{(k)})$
- Update step however generalizable to

  $$J(p) = f(p) + g(p) \qquad f, g \text{ convex and } f \text{ differentiable}$$

  Idea: modified update step (explicit for f but implicit for g):

  $$\text{(a)} \;\; \hat{p}^{(k)} = p^{(k)} - t\nabla f(p^{(k)}), \qquad \text{(b)} \;\; p^{(k+1)} = \hat{p}^{(k)} - t\nabla g(p^{(k+1)}).$$

  Fix step length t, solve (b) for $p^{(k+1)}$.
  $\rightarrow$ Trivial (closed-form) if g quadratic
  $\rightarrow$ Newton's method if g twice-differentiable

Minimization of functionals with a nonsmooth part

(a) $\hat{p}^{(k)} = p^{(k)} - t\nabla f(p^{(k)})$, (b) $p^{(k+1)} = \hat{p}^{(k)} - t\nabla g(p^{(k+1)})$.

- If g convex and differentiable, (b) can be reformulated as

  $$\big(p^{(k+1)} - \hat{p}^{(k)}\big) + t\nabla g(p^{(k+1)}) = 0,$$

  equivalent (as necessary and sufficient optimality condition) to

  $$p^{(k+1)} = \arg\min_{p\in\mathbb{R}^n} \Big(\tfrac{1}{2}\|p - \hat{p}^{(k)}\|_2^2 + tg(p) \Big)$$

  Still (uniquely) solvable if g convex and **not** differentiable.

- For h any lower semicontinuous (lsc) convex function, define proximal operator $\text{prox}_h : \mathbb{R}^n \to \mathbb{R}$:

  $$\text{prox}_h(y) = \arg\min_{x\in\mathbb{R}^n} \big(\tfrac{1}{2}\|x - y\|_2^2 + h(x) \big)$$

Update rule for $J(p) = f(p) + g(p)$, g convex but possibly non-smooth:

$$p^{(k+1)} = \text{prox}_{tg}\big(p^{(k)} - t\nabla f(p^{(k)})\big)$$

h lsc: Epigraph $\{ (x,t) \in \mathbb{R}^n \times \mathbb{R},\ h(x) \le t \}$ closed in $\mathbb{R}^n \times \mathbb{R}$ (among several equivalent definitions).

Minimization of L^2-L^1 functionals

Specialize update step to L^2-L^1 functional J_α:

- Major simplification of non-smooth part:
$$g(p) = \alpha\|p\|_1 = \alpha\big(|p_1|+\ldots+|p_n|\big)$$
- Proximal-operator defining minimization
$$\mathsf{prox}_{tg}(y) := \underset{x\in\mathbb{R}^n}{\arg\min}\,\big(\tfrac{1}{2}\|x-y\|_2^2 + g(x)\big)$$
uncouples into n univariate problems
$$\min_{x_1\in\mathbb{R}^n} \tfrac{1}{2}|x_1-y_1|^2 + \alpha t|x_1|, \quad \ldots \quad \min_{x_n\in\mathbb{R}^n} \tfrac{1}{2}|x_n-y_n|^2 + \alpha t|x_n|$$
- Univariate proximal operators found in closed form:
$$\mathsf{prox}_{u\mapsto\alpha t|u|}(y) = \underset{x\in\mathbb{R}}{\arg\min}\,\big(\tfrac{1}{2}(x-y)^2 + \alpha t|x|\big) = \begin{cases} 0 & |y|\le \alpha t \\ y(1-\alpha t/|y|) & |y|\ge \alpha t \end{cases}$$

L^2-L^1 update step (given step length t):

(a) $\hat{p}^{(k)} = p^{(k)} - tA^\mathsf{T}(Ap^{(k)} - b)$, (b) $p_i^{(k+1)} = \begin{cases} 0 & |\hat{p}_i^{(k)}|\le \alpha t \\ \hat{p}_i^{(k)} - \alpha t\,\mathsf{sign}(\hat{p}_i^{(k)}) & |\hat{p}_i^{(k)}|\ge \alpha t \end{cases}$.

- $|y| < \alpha t \implies \mathsf{prox}_{u\mapsto\alpha t|u|}(y) = 0$: sparsity-promoting mechanism of L^2-L^1 minimization.
- Reduce $\alpha \implies$ weaker sparsity promotion.

Minimization of L^2-L^1 functionals

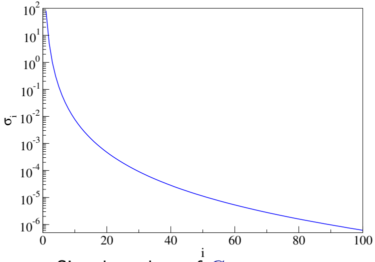
Algorithm 1 FISTA iterations for the L^2-L^1 minimization problem (Beck, Teboulle 2009)

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\alpha > 0$, $x^{(0)} \in \mathbb{R}^n$ (data and initial guess)
$L = \|A^\mathsf{T} A\|_2$ (spectral radius of $A^\mathsf{T} A$, i.e. largest eigenvalue of $A^\mathsf{T} A$)
$t = \alpha / L$ (step length, maximum permissible value)
$y^{(1)} = x^{(0)}$, $s^{(1)} = 1$ (first iteration)
for $k = 1, 2, \ldots$ **do**
    $\hat{x}^{(k)} = x^{(k-1)} - tA^\mathsf{T}(Ax^{(k-1)} - b)$ (explicit step)
    $x^{(k)} = \text{prox}_{u \mapsto t\|u\|}(\hat{x}^{(k)})$ (apply proximal operator)
    $s^{(k+1)} = \frac{1}{2}\big(1 + \sqrt{1 + 4s^{(k)2}}\big)$ (update algorcolor parameter $s^{(k)}$)
    $y^{(k+1)} = x^{(k)} + \frac{s^{(k)}-1}{s^{(k+1)}}\big(x^{(k)} - x^{(k-1)}\big)$
    If convergence test satisfied: return $x = x^{(k)}$
end for

"random" example

$$J_\alpha(p) = \frac{1}{2}\|Gp - d\|_2^2 + \alpha\|p\|_1, \quad G \in \mathbb{R}^{m\times n}\ (m = 5000,\ n = 100), \quad \text{cond}(G) \approx 1.3\,10^9.$$

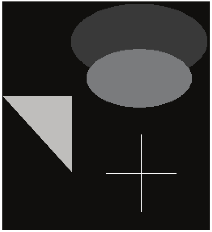
$$\boxed{\min_{x\in\mathbb{R}^n} \|Gp - d\|_2 \approx 77.51}$$



Singular values of G

α	$\# x_i^\alpha \neq 0$	$\|Gp^\alpha - d\|_2$	$\|d^\alpha\|_1$	# iters.
0	100	77.51	$3.54\,10^7$	N/A
0.002	54	78.16	1076.	436 197
0.005	8	78.19	224.0	226 218
0.01	5	78.21	53.80	81 683
0.1	2	78.22	1.538	5 564
0.2	3	78.22	1.17	5 031
0.5	2	78.22	0.9298	2 455
2	1	78.23	0.1076	873

Image restoration example (Beck, Teboulle 2009)

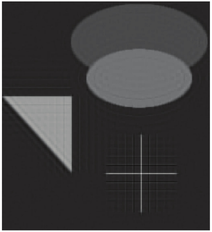


Original

Blurred

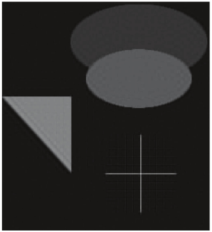
MTWIST: $F_{100} = 3.83\text{e-}1$

MTWIST: $F_{200} = 3.41\text{e-}1$

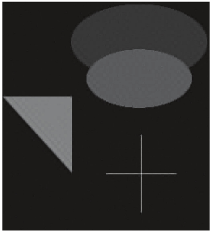


with alternative algorithm

FISTA: $F_{100} = 3.21\text{e-}1$



FISTA: $F_{200} = 3.09\text{e-}1$



with FISTA

Conditional probability

- Definition of a conditional probability $P(A|B)$:

$$\boxed{P(A|B) = \frac{P(A\cap B)}{P(B)}}$$

  $P(A\cap B) = P(B\cap A)$ (symétrie) donne la formule de Bayes:

$$\boxed{P(A|B)P(B) = P(B|A)P(A)}$$

- **Idea:** use Bayes to "invert" the parameter-data relationship between $\boldsymbol{p}$ and $\boldsymbol{d}$:

$$f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}})f_{\mathcal{D}}(\boldsymbol{d}_{\text{obs}}) = f_{\mathcal{D}|\mathcal{P}}(\boldsymbol{d}_{\text{obs}}|\boldsymbol{p})f_{\mathcal{P}}(\boldsymbol{p})$$

  $f_{\mathcal{P}}(\boldsymbol{p})$: probability density describing prior information on $\boldsymbol{p}$
  $f_{\mathcal{D}|\mathcal{P}}(\boldsymbol{d}|\boldsymbol{p})$: probability density describing the effect of the forward problem
  (probability density describing modeling and measurement uncertainties)

  $f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}})f_{\mathcal{D}}(\boldsymbol{d}_{\text{obs}})$: probability density on $\boldsymbol{p}$ given $\boldsymbol{d}$
  (defines a solution of the inverse problem)

- Estimators on $\boldsymbol{p}$ extracted by post-processing $f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}})$. In particular:

$$\boldsymbol{p}_{\text{MAP}} = \arg\max_{p} f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}}) \qquad \textit{Maximum a posteriori estimate}$$

Finite-dimensional inversion using Gaussian densities

- Probability density function of random Gaussian vector $y \in \mathbb{R}^n$:

$$f(y) = \frac{1}{\sqrt{(2\pi)^N \mathsf{det}(C)}} \exp\Big(-\tfrac{1}{2}(y-\overline{y})^{\mathsf{T}} C^{-1}(y-\overline{y}) \Big)$$

$\overline{y}$: mean value; $\boldsymbol{C}$: (SPD) covariance matrix

- Prior information on parameter, data and model taken of the form

$$\begin{aligned} f_{\mathcal{P}}(p) &= \mathcal{N}(p_0, C_{\mathsf{p}}) \\ f_{\mathcal{D}}(d) = f_{\mathcal{D}}(d|d_{\mathsf{obs}})\mu_{\mathcal{D}}(d) &= \mathcal{N}(d_{\mathsf{obs}}, C_{\mathsf{d}}) \\ f_{\mathcal{D}|\mathcal{P}}(d|p) &= \mathcal{N}(G(p), C_{\mathsf{T}}) \end{aligned}$$

- Posterior informtion found to be defined by

$$f_{\mathcal{P}}(p) = (\mathsf{cste})\, f_{\mathcal{P}}(p) \times \exp\big(-\tfrac{1}{2}(G(p)-d_{\mathsf{obs}})^{\mathsf{T}} C_{\mathcal{D}}^{-1}(G(p)-d_{\mathsf{obs}})\big)$$

$$\text{with} \quad C_{\mathcal{D}} = C_{\mathsf{d}} + C_{\mathsf{T}}$$

- MAP estimate:

$$\begin{aligned} p^\star &= \arg\max_p f_{\mathcal{P}}(p) \\ &= \arg\min_p \Big\{ (G(p)-d_{\mathsf{obs}})^{\mathsf{T}} C_{\mathcal{D}}^{-1}(G(p)-d_{\mathsf{obs}}) + (p-p_0)^{\mathsf{T}} C_{\mathsf{p}}^{-1}(p-p_0) \Big\} \end{aligned}$$

Minimization problem that resembles a (deterministic) Tiknonov regularization