Inverse Problems

Sommersemester 2022

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

Second lecture, May 2, 2022

Recap from last week

Let H_1 and H_2 be separable real Hilbert spaces and let $A\colon H_1 \to H_2$ be a compact linear operator.

- Find unknown $x \in H_1$ such that
$$y = Ax, \tag{1}$$
where data $y \in H_2$ is given.
- There exist orthonormal $\{v_n\} \subset H_1$ and $\{u_n\} \subset H_2$ and $\lambda_n \searrow 0$ s.t.
$$Ax = \sum_n \lambda_n \langle x, v_n \rangle u_n \quad \text{for all } x \in H_1.$$
- There exists a solution to (1) iff
$$y = Py \quad \text{and} \quad \sum_n \frac{1}{\lambda_n^2} |\langle y, u_n \rangle|^2 < \infty,$$
where $P\colon H_2 \to \overline{\mathrm{Ran}(A)}$ is an orthogonal projection. The solutions are of the form
$$x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle v_n \quad \text{for arbitrary } x_0 \in \mathrm{Ker}(A).$$

Recap from last week – truncated SVD

For $k \in \mathbb{N}$, $k \leq \mathrm{rank}(A)$, there exists a unique $x_k \in H_1$ such that

$$Ax_k = P_k y \quad \text{and} \quad x_k \perp \mathrm{Ker}(A),$$

where $P_k \colon H_2 \to \mathrm{span}\{u_1, \ldots, u_k\}$ is an orthogonal projection. This solution can be given as

$$x_k = \sum_{n=1}^{k} \frac{1}{\lambda_n} \langle y, u_n \rangle v_n.$$

Matrix SVD: Let the SVD of matrix $A \in \mathbb{R}^{m\times n}$ be given by

$$A = U\Lambda V^{\mathrm{T}},$$

where $\Lambda \in \mathbb{R}^{m\times n}$ has the non-negative singular values $\{\lambda_j\}_{j=1}^{\min\{m,n\}}$ on its diagonal and $V \in \mathbb{R}^{n\times n}$ and $U \in \mathbb{R}^{m\times m}$ are orthogonal matrices.[†]

The TSVD solution for $1 \le k \le p := \mathrm{rank}(A)$ is given by

$$x_k = V\Lambda_k^{\dagger}U^{\mathrm{T}}y,$$

where

$$\Lambda_k^{\dagger} = \begin{pmatrix} 1/\lambda_1 & 0 & \cdots & 0 & & \cdots & 0 \\ 0 & 1/\lambda_2 & & & & & \vdots \\ \vdots & & \ddots & & & & \\ & & & 1/\lambda_k & & & \\ & & & & 0 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & & & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{n\times m}.$$

The matrix $A^{\dagger} = V\Lambda_p^{\dagger}U^{\mathrm{T}}$ is called the *Moore–Penrose pseudoinverse* of A.

[†]This means that the columns $\{v_j\}_{j=1}^{n}$ of V form an orthonormal basis for $\mathbb{R}^n$, and similarly the columns $\{u_j\}_{j=1}^{m}$ of U are an orhonormal basis of $\mathbb{R}^m$.

Morozov discrepancy principle

Let H_1 and H_2 be separable real Hilbert spaces and $A\colon H_1 \to H_2$ a compact linear operator.

How to choose the spectral cut-off index $k \geq 1$ in the TSVD problem

$$Ax = P_k y \quad \text{and} \quad x \perp \mathrm{Ker}(A)?$$

There is a rule of thumb called the *Morozov discrepancy principle:*

Suppose that the data $y \in H_2$ is a noisy approximation of noiseless "exact" data $y_0 \in H_2$. While y_0 is unknown to us, we may have an estimate on the noise level, e.g.,

$$\|y - y_0\| \approx \varepsilon > 0.$$

We choose the smallest $k \geq 1$ such that the residual satisfies

$$\|y - Ax_k\| \leq \varepsilon.$$

Intuitively, this means that we cannot expect the approximate solution to yield a smaller residual than the measurement error without fitting the solution to noise.

Q: *When does an index $k \geq 1$ satisfying $\|y - Ax_k\| \leq \varepsilon$ exist?*
A: When $\varepsilon > \|Py - y\|$ and $\mathrm{rank}(A) = \infty$, it follows from $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}(P) \perp \mathrm{Ran}(I - P)$ that

$$\|Ax_k - y\|^2 = \|Ax_k - Py + Py - y\|^2 = \|Ax_k - Py\|^2 + \|(P - I)y\|^2$$
$$= \sum_{n=k+1}^{\infty} |\langle y, u_n\rangle|^2 + \|(P - I)y\|^2 \xrightarrow{k\to\infty} \|Py - y\|^2.$$

Due to the properties of the orthogonal projection, $\|Py - y\| = \inf_{z\in\mathrm{Ran}(A)} \|z - y\|$, so this is the best we can do. (Note however that there is no guarantee that prevents $\|x_k\|$ from exploding as $k \to \infty$.)

On the other hand, if $p = \mathrm{rank}(A) < \infty$,

$$\|Ax_p - y\| = \|P_p y - y\| = \|Py - y\|.$$

One should usually avoid choosing the spectral cut-off to be this large in practice.

Numerical example: backward heat equation

Let us consider the backward heat equation:

$$\begin{cases} \partial_t u(x,t) = \partial_x^2 u(x,t) & \text{for } (x,t) \in (0,\pi) \times \mathbb{R}_+, \\ u(0,\cdot) = u(\pi,\cdot) = 0 & \text{on } \mathbb{R}_+, \\ u(\cdot,0) = f & \text{on } (0,\pi), \end{cases}$$

where $f : (0,\pi) \to \mathbb{R}$ is the initial heat distribution.

Forward problem: Given initial data $f : (0,\pi) \to \mathbb{R}$, determine the heat distribution $u(\cdot, T)$ at time $T > 0$.

Inverse problem: Reconstruct the initial state f based on noisy measurements of $u(\cdot, T)$ at time $T > 0$.

Let us consider a simple discretization of the PDE

$$\begin{cases} \partial_t u(x,t) = \partial_x^2 u(x,t) & \text{for } (x,t) \in (0,\pi) \times \mathbb{R}_+, \\ u(0,\cdot) = u(\pi,\cdot) = 0 & \text{on } \mathbb{R}_+, \\ u(\cdot,0) = f & \text{on } (0,\pi). \end{cases}$$

Let $x_j = jh$ for $j = 0, \ldots, 100$, where $h = \pi/100$ is the step size.

Zero Dirichlet boundary conditions imply that $u(x_0,t) = u(x_{100},t) = 0$.

The spatial second derivative can be discretized using the stencils

$$\partial_x^2 u(x_1,t) = \frac{-2u(x_1,t) + u(x_2,t)}{h^2} + \mathcal{O}(h^2),$$
$$\partial_x^2 u(x_j,t) = \frac{u(x_{j-1},t) - 2u(x_j,t) + u(x_{j+1},t)}{h^2} + \mathcal{O}(h^2) \quad \text{for } j = 2, \ldots, 98,$$
$$\partial_x^2 u(x_{99},t) = \frac{u(x_{98},t) - 2u(x_{99},t)}{h^2} + \mathcal{O}(h^2).$$

Denote $U(t) = (U_j(t))_{j=1}^{99} = (u(x_j,t))_{j=1}^{99}$ and $F = (f(x_j))_{j=1}^{99}$.

$$\frac{\partial}{\partial t}\begin{bmatrix} U_1(t) \\ U_2(t) \\ U_3(t) \\ \vdots \\ U_{98}(t) \\ U_{99}(t) \end{bmatrix} = \underbrace{\frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & & \ddots & & & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}}_{=:B} \begin{bmatrix} U_1(t) \\ U_2(t) \\ U_3(t) \\ \vdots \\ U_{98}(t) \\ U_{99}(t) \end{bmatrix}.$$

After spatial discretization, our PDE has been transformed into the initial value problem

$$\dot{U}(t) = BU(t), \quad U(0) = F.$$

At time $t = T > 0$, the discretized heat distribution $U := U(T)$ is given by

$$U = AF,$$

where $A = \mathrm{e}^{TB} \in \mathbb{R}^{99\times 99}$ and

$$\mathrm{e}^{M} := \sum_{k=0}^{\infty} \frac{1}{k!} M^k$$

is the *matrix exponential* (cf. function `expm` in MATLAB).

A note on simulating measurement data and inverse crimes

When simulating measurement data, one should take care not to use the same computational model for inversion as the one which was used to generate the measurements in the first place. This would lead to unreasonably good reconstructions, since this is akin to multiplying a matrix with its own inverse. This is known as an *inverse crime*. (Similar concerns also apply to non-linear problems.)

With real-life measurement data, we do not have worry about this phenomenon – measurements that come from nature are automatically independent of any computational model we end up using for practical inverse problems simulations.

A popular technique to avoid committing an inverse crime is using a higher resolution computational model to generate the measurements and interpolating the simulated data onto a coarser grid, where we plan to carry out the actual computational inversion. Another good option is to use an analytic solution, if one is readily available. We will use this technique with the heat equation.

The forward problem of the heat equation

$$\begin{cases} \partial_t u(x,t) = \partial_x^2 u(x,t) & \text{for } (x,t) \in (0,\pi) \times \mathbb{R}_+, \\ u(0,\cdot) = u(\pi,\cdot) = 0 & \text{on } \mathbb{R}_+, \\ u(\cdot,0) = f & \text{on } (0,\pi), \end{cases}$$

has the classical series solution

$$u(x,t) = \sum_{n=1}^{\infty} \hat{f}_n \, \mathrm{e}^{-n^2 t} \sin(nx),$$

where the coefficients $\hat{f}_n$ are the Fourier sine series coefficients of the initial heat distribution f satisfying

$$f(x) = \sum_{n=1}^{\infty} \hat{f}_n \sin(nx), \quad \hat{f}_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) \, \mathrm{d}x.$$

Let us fix the ground truth

$$f(x) = \begin{cases} 1 & \text{if } x \in [1,2], \\ 0 & \text{if } x \in (0,1) \cup (2,\pi). \end{cases}$$

It is easy to see that the Fourier sine coefficients are given by

$$\hat{f}_n = \frac{2}{n\pi}(\cos n - \cos 2n).$$

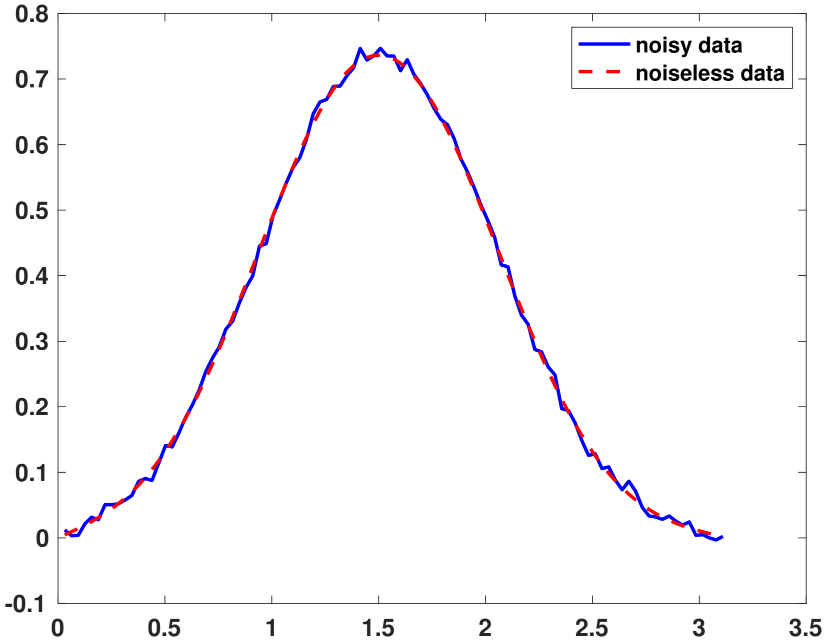
Let us plug these into the forward solution at time $t = T > 0$

$$u(x_j, T) = \sum_{n=1}^{\infty} \hat{f}_n \, \mathrm{e}^{-n^2 T} \sin(n x_j), \quad j = 1, \ldots, 99,$$

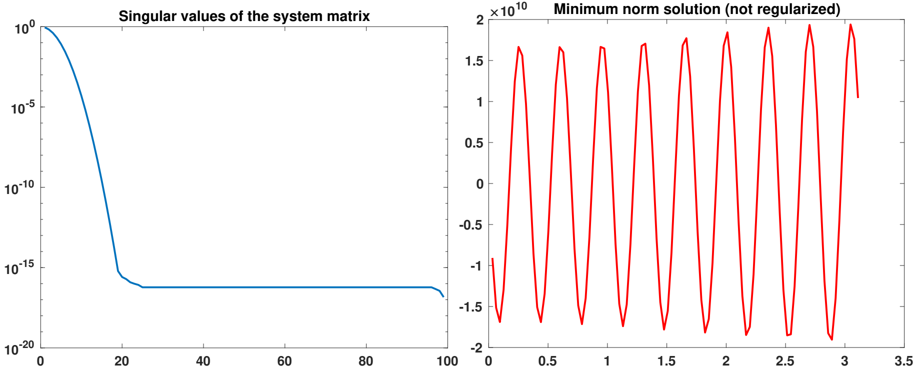
and add some simulated measurement noise!

We assume that the data $U(T) \in \mathbb{R}^{99}$ at time $T = 0.1$ is contaminated with mean-zero Gaussian noise with standard deviation 0.01, and that the discrepancy between the measured data and the underlying "exact" data equals the square root of the expected value of the squared norm of the noise vector, i.e.,

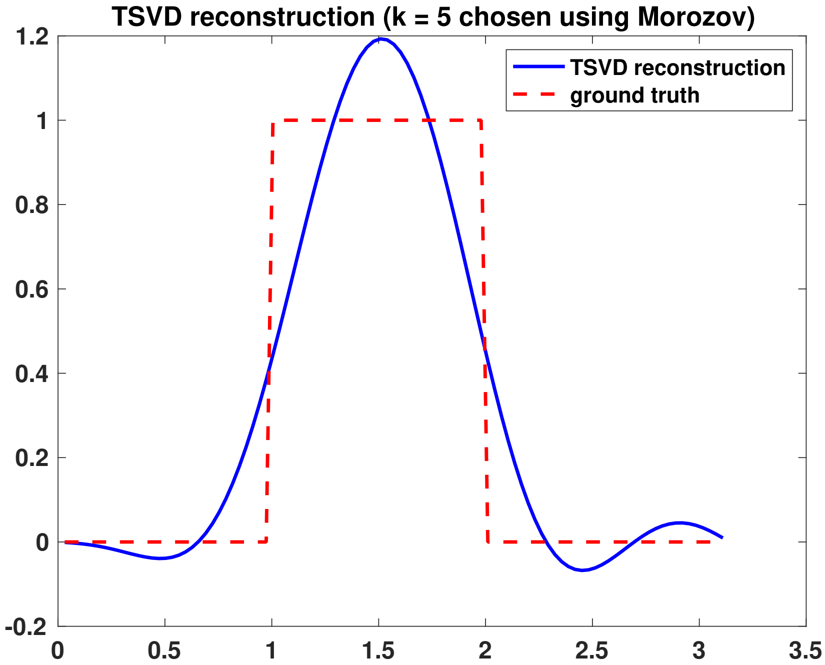
$$\varepsilon = \sqrt{99 \cdot 0.01^2} \approx 0.0995.$$



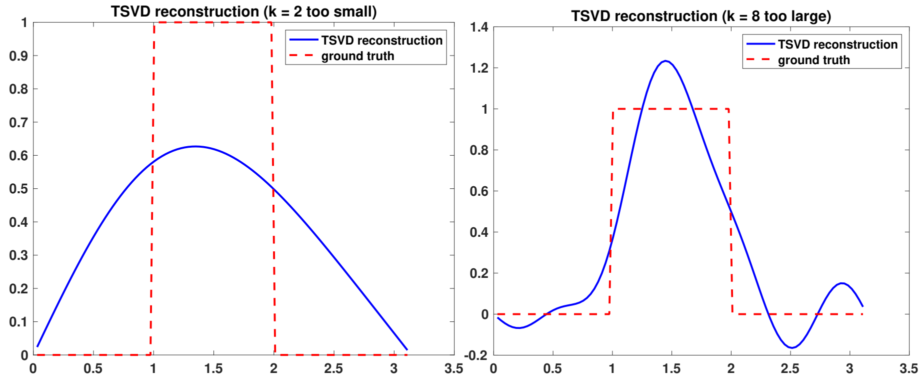
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0
0.5
1
1.5
2
2.5
3
3.5
noisy data
noiseless data



Singular values of the system matrix
10^{0}
10^{-5}
10^{-10}
10^{-15}
10^{-20}
0
20
40
60
80
100
Minimum norm solution (not regularized)
$\times 10^{10}$
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
0.5
1
1.5
2
2.5
3
3.5



TSVD reconstruction (k = 5 chosen using Morozov)
TSVD reconstruction
ground truth
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
0.5
1
1.5
2
2.5
3
3.5


TSVD reconstruction (k = 2 too small)
TSVD reconstruction
ground truth
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5
TSVD reconstruction (k = 8 too large)
TSVD reconstruction
ground truth
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
0.5
1
1.5
2
2.5
3
3.5

Tikhonov regularization

The sequence of TSVD solutions $\{x_k\}$ minimizes the norm of the residual

$$\|Ax - y\|$$

as k tends to $\mathrm{rank}(A)$. Unfortunately, when inverse/ill-posed problems are considered, it may also happen that

$$\|x_k\| \to \infty \quad \text{as } k \to \mathrm{rank}(A).$$

In consequence, it appears reasonable to try minimizing the residual and the norm of the solution *simultaneously*.

Definition

A *Tikhonov regularized solution* $x_\delta \in H_1$ is a minimizer of the Tikhonov functional

$$F_\delta(x) := \|Ax - y\|^2 + \delta\|x\|^2,$$

where $\delta > 0$ is called the *regularization parameter*.

Theorem

Let $A \colon H_1 \to H_2$ be a compact linear operator with the singular system (λ_n, v_n, u_n). Then the Tikhonov regularized solution exists, is unique, and is given by the formula

$$x_\delta = (A^*A + \delta I)^{-1}A^*y = \sum_{n=1}^{p} \frac{\lambda_n}{\lambda_n^2 + \delta}\langle y, u_n\rangle v_n,$$

where $p = \mathrm{rank}(A)$.

Remark. The Tikhonov regularized solution can be obtained without knowing the SVD of A by solving x_δ from $(A^*A + \delta I)x_\delta = A^*y$.

Proof. We make use of the *Lax–Milgram lemma*:

Lemma (Lax–Milgram)

Let H be a Hilbert space, and let $B\colon H \times H \to \mathbb{R}$ be a bilinear quadratic form such that

$$|B(x,y)| \le C\|x\|\|y\| \quad \text{for all } x, y \in H,$$
$$B(x,x) \ge c\|x\|^2 \quad \text{for all } x \in H$$

for some constants $0 < c \le C < \infty$. Then there exists a unique linear boundedly invertible operator $T\colon H \to H$ such that

$$B(x,y) = \langle x, Ty\rangle \quad \text{for all } y \in H,$$
$$\|T\| \le C \quad \text{and} \quad \|T^{-1}\| \le \frac{1}{c}.$$

In our case, we define the bilinear operator $B(x,y) := \langle x, (A^*A + \delta I)y\rangle$ and observe that $|B(x,y)| \le (\|A\|^2 + \delta)\|x\|\|y\|$ (boundedness) and $B(x,x) = \langle x, (A^*A + \delta I)x\rangle = \|Ax\|^2 + \delta\|x\|^2 \ge \delta\|x\|^2$ (coercivity).
$\therefore$ $(A^*A + \delta I)^{-1}$ exists such that $\|(A^*A + \delta I)^{-1}\| \le \frac{1}{\delta}$. In particular, $x_\delta = (A^*A + \delta I)^{-1}A^*y$ is well-defined.

Recall that $Ax = \sum_n \lambda_n \langle x, v_n \rangle u_n$ and $A^* y = \sum_n \lambda_n \langle y, u_n \rangle v_n$. Especially,

$$A^* A x = \sum_n \lambda_n^2 \langle x, v_n \rangle v_n.$$

Since $H_1 = \mathrm{Ker}(A) \oplus \mathrm{Ker}(A)^\perp$, we can write

$$x_\delta = P x_\delta + Q x_\delta = \sum_n \langle x_\delta, v_n \rangle v_n + Q x_\delta,$$

where $P \colon H_1 \to \mathrm{Ker}(A)^\perp = \overline{\mathrm{span}\{v_n\}}$ and $Q \colon H_1 \to \mathrm{Ker}(A)$ are orthogonal projections. Thus

$$(A^* A + \delta I) x_\delta = A^* y \quad \Leftrightarrow \quad \sum_n (\lambda_n^2 + \delta) \langle x_\delta, v_n \rangle v_n + Q x_\delta = \sum_n \lambda_n \langle y, u_n \rangle v_n.$$

Equating terms yields that $Q x_\delta = 0$ and

$$(\lambda_n^2 + \delta) \langle x_\delta, v_n \rangle = \lambda_n \langle y, u_n \rangle \quad \Leftrightarrow \quad \langle x_\delta, v_n \rangle = \frac{\lambda_n}{\lambda_n^2 + \delta} \langle y, u_n \rangle,$$

as desired.

Finally, to show that x_δ minimizes the quadratic functional $F_\delta(x) = \|Ax - y\|^2 + \delta\|x\|^2$, consider

$$x = x_\delta + z,$$

where $z \in H_1$ is arbitrary. Now

$$\begin{aligned} F_\delta(x) &= F_\delta(x_\delta + z) = F_\delta(x_\delta) + \langle z, (A^*A + \delta I)x_\delta - A^*y\rangle + \langle z, (A^*A + \delta I)z\rangle \\ &= F_\delta(x_\delta) + \langle z, (A^*A + \delta I)z\rangle, \end{aligned}$$

by definition of x_δ.The last term is nonnegative and vanishes only if $z = 0$. This proves the claim. □

Morozov discrepancy principle for Tikhonov regularization

Suppose that the measurement $y \in H_2$ is a noisy version of some underlying "exact" data $y_0 \in H_2$, and that

$$\|y - y_0\| \approx \varepsilon > 0.$$

In the framework of Tikhonov regularization, the Morozov discrepancy principle tells us to choose the regularization parameter $\delta > 0$ so that the residual satisfies

$$\|y - Ax_\delta\| = \varepsilon.$$

It turns out that there is a unique regularization parameter satisfying this condition if

$$\|y - Py\| < \varepsilon < \|y\|,$$

where $P \colon H_2 \to \overline{\mathrm{Ran}(A)}$ is an orthogonal projection.

Properties of the Tikhonov regularized solution

Theorem

Let $A\colon H_1 \to H_2$ be a compact linear operator with the singular system (λ_n, v_n, u_n). Let $P\colon H_2 \to \overline{\mathrm{Ran}(A)}$ be an orthogonal projection. Then we have the following:

(i) $\delta \mapsto \|Ax_\delta - y\|$ *is a strictly increasing function of* $\delta > 0$.

(ii) $\|Py - y\| = \lim_{\delta \to 0+} \|Ax_\delta - y\| \leq \|Ax_\delta - y\| \leq \lim_{\delta \to \infty} \|Ax_\delta - y\| = \|y\|$.

(iii) *If* $Py \in \mathrm{Ran}(A)$, *then* x_δ *converges to the solution of the problem*

$$Ax = Py \quad \text{and} \quad x \perp \mathrm{Ker}(A)$$

as $\delta \to 0+$.

Corollary

The equation $\|Ax_\delta - y\| = \varepsilon$ *has a unique solution* $\delta = \delta(\varepsilon)$ *iff* $\|(I - P)y\| < \varepsilon < \|y\|$.

Interpretation: $\|(I - P)y\| < \varepsilon$ means that any component in the data y orthogonal to the range of A must be due to noise; $\varepsilon < \|y\|$ means that the error level should not exceed the signal level.

Proof. Suppose that the operator A has the SVD

$$Ax = \sum_n \lambda_n \langle x, v_n \rangle u_n.$$

Then $Av_n = \lambda_n u_n$, the orthogonal projection $P \colon H_2 \to \overline{\mathrm{Ran}(A)}$ is

$$Py = \sum_n \langle y, u_n \rangle u_n,$$

and the Tikhonov regularized solution x_δ and its image under A are

$$x_\delta = \sum_n \frac{\lambda_n}{\lambda_n^2 + \delta} \langle y, u_n \rangle v_n \quad \Rightarrow \quad Ax_\delta = \sum_n \frac{\lambda_n^2}{\lambda_n^2 + \delta} \langle y, u_n \rangle u_n.$$

(i) It follows that

$$\begin{aligned} \|Ax_\delta - y\|^2 &= \|Ax_\delta - Py\|^2 + \|(I-P)y\|^2 \\ &= \sum_n \left(\frac{\lambda_n^2}{\lambda_n^2 + \delta} - 1 \right)^2 |\langle y, u_n \rangle|^2 + \|(I-P)y\|^2 \\ &= \sum_n \left(\frac{\delta}{\lambda_n^2 + \delta} \right)^2 |\langle y, u_n \rangle|^2 + \|(I-P)y\|^2. \end{aligned}$$

We arrived at

$$\|Ax_\delta - y\|^2 = \sum_n \left(\frac{\delta}{\lambda_n^2 + \delta}\right)^2 |\langle y, u_n\rangle|^2 + \|(I-P)y\|^2.$$

For each term of the sum,

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\left(\frac{\delta}{\lambda_n^2 + \delta}\right)^2 = \frac{2\delta\lambda_n^2}{(\lambda_n^2 + \delta)^3} > 0,$$

implying that the mapping $\delta \mapsto \|Ax_\delta - y\|^2$ is strictly increasing.

(ii) It is easy to see that

$$\|Ax_\delta - y\|^2 = \sum_n \left(\frac{\delta}{\lambda_n^2 + \delta}\right)^2 |\langle y, u_n\rangle|^2 + \|(I-P)y\|^2 \xrightarrow{\delta\to 0+} \|(I-P)y\|^2,$$

$$\begin{aligned}\|Ax_\delta - y\|^2 &= \sum_n \left(\frac{\delta}{\lambda_n^2 + \delta}\right)^2 |\langle y, u_n\rangle|^2 + \|(I-P)y\|^2 \\ &\xrightarrow{\delta\to\infty} \|Py\|^2 + \|(I-P)y\|^2 = \|y\|^2.\end{aligned}$$

(iii) Let $Py \in \mathrm{Ran}(A)$. This implies that there exists $x \in \mathrm{Ker}(A)^{\perp}$ such that $Ax = Py$; this is the minimum norm solution

$$x = \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle v_n,$$

for which it can be shown that

$$x_\delta = \sum_n \frac{\lambda_n}{\lambda_n^2 + \delta} \langle y, u_n \rangle v_n \xrightarrow{\delta \to 0+} \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle v_n = x. \quad \square$$

Remark. In parts (ii) and (iii), one should take care when interchanging the order of the limit and the summation, i.e., justifying the steps

$$\lim_{\lambda \to 0+} \sum_n = \sum_n \lim_{\lambda \to 0+} \quad \text{and} \quad \lim_{\lambda \to \infty} \sum_n = \sum_n \lim_{\lambda \to \infty}.$$

Standard techniques involve the monotone convergence theorem and the dominated convergence theorem (note that these apply to infinite series as well as integrals). In part (iii), it is helpful to observe that $x_\delta \xrightarrow{\delta \to 0+} x$ iff $\langle x_\delta, \phi \rangle \xrightarrow{\delta \to 0+} \langle x, \phi \rangle$ for all $\phi \in H_1$ and $\|x_\delta\| \xrightarrow{\delta \to 0+} \|x\|$.

Tikhonov regularization with matrices

Consider the special case $H_1 = \mathbb{R}^n$ and $H_2 = \mathbb{R}^m$ corresponding to the matrix equation $y = Ax$. The Tikhonov functional takes the special form

$$F_\delta(x) = \left\| \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix} x - \begin{bmatrix} y \\ 0 \end{bmatrix} \right\|^2, \quad I \in \mathbb{R}^{n\times n},\ 0 \in \mathbb{R}^n.$$

The minimizer can be found by solving the least squares problem

$$\begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix} x = \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} y \\ 0 \end{bmatrix}$$

or, equivalently,

$$(A^{\mathrm{T}} A + \delta I)x = A^{\mathrm{T}} y.$$

In MATLAB, this can be implemented simply as follows (MATLAB's `mldivide` or "backslash" operator automatically tries to solve the corresponding least squares problem for non-square matrices):

```
K = [A;sqrt(delta)*eye(n)];
z = [y; zeros(n,1)];
xdelta = K\z;
```

Numerical example: backward heat equation

Let us revisit the backward heat equation from earlier:

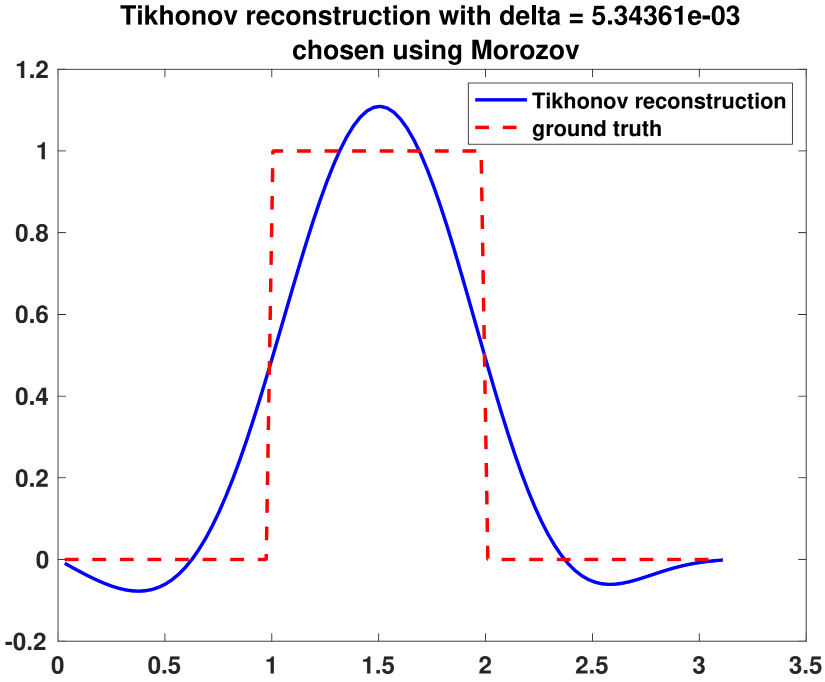
$$\begin{cases} \partial_t u(x,t) = \partial_x^2 u(x,t) & \text{for } (x,t) \in (0,\pi) \times \mathbb{R}_+, \\ u(0,\cdot) = u(\pi,\cdot) = 0 & \text{on } \mathbb{R}_+, \\ u(\cdot,0) = f & \text{on } (0,\pi), \end{cases}$$

where $f : (0,\pi) \to \mathbb{R}$ is the initial heat distribution.

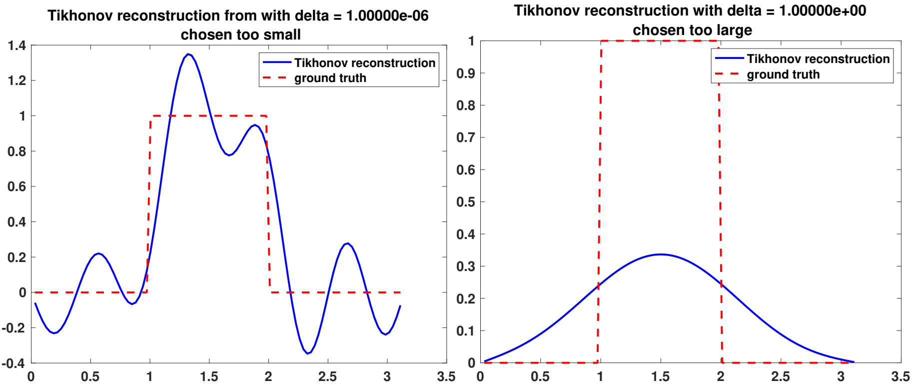
We reconstruct the initial state f based on noisy measurements of $u(\cdot, T)$ at time $T > 0$ using Tikhonov regularization.

We assume that the data $U(T) \in \mathbb{R}^{99}$ at time $T = 0.1$ is contaminated with mean-zero Gaussian noise with standard deviation 0.01, and that the discrepancy between the measured data and the underlying "exact" data equals the square root of the expected value of the squared norm of the noise vector, i.e.,

$$\varepsilon = \sqrt{99 \cdot 0.01^2} \approx 0.0995.$$



Tikhonov reconstruction with delta = 5.34361e-03
chosen using Morozov
Tikhonov reconstruction
ground truth
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
0.5
1
1.5
2
2.5
3
3.5


Tikhonov reconstruction from with delta = 1.00000e-06
chosen too small
Tikhonov reconstruction
ground truth
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3
3.5
Tikhonov reconstruction with delta = 1.00000e+00
chosen too large
Tikhonov reconstruction
ground truth
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5

Tikhonov regularization for nonlinear problems

Unlike the TSVD, Tikhonov regularization can be generalized to nonlinear problems as well. Consider a *nonlinear* operator $A\colon H_1 \to H_2$ and the problem

$$y = A(x).$$

A standard way of solving such a problem is via sequential linearizations, which leads to solving a set of linear problems involving the *Fréchet derivative* of operator A.

Definition

The function $A\colon H_1 \to H_2$ is called *Fréchet differentiable* at $x_0 \in H_1$ if there exists a continuous linear operator $A'_{x_0}\colon H_1 \to H_2$ such that

$$A(x+h) = A(x) + A'_{x_0}h + W_{x_0}(z),$$

where $\|W_{x_0}(z)\| \leq \epsilon(x_0, z)\|z\|$ and the functional $z \mapsto \epsilon(x_0, z)$ tends to zero as $z \to 0$.

The linear operator A'_{x_0} is called the *Fréchet derivative* of A at x_0.

We are interested in minimizing

$$F_\delta(x) = \|A(x) - y\|^2 + \delta\|x\|^2, \quad \delta > 0.$$

Since F_δ is no longer quadratic, it is unclear whether a unique minimizer exists and typically the minimizer cannot be given by an explicit formula even it exists.

Let A be Fréchet differentiable. The linearization of A around a given point x_0 leads to the approximation of the functional F_δ,

$$\begin{aligned} F_\delta(x) \approx \widetilde{F}_\delta(x; x_0) &= \|A(x_0) + A'_{x_0}(x - x_0) - y\|^2 + \delta\|x\|^2 \\ &= \|A'_{x_0}(x) - g(y, x_0)\|^2 + \delta\|x\|^2, \end{aligned}$$

where $g(y, x_0) := y - A(x_0) + A'_{x_0}(x_0)$.

From the previous discussion on the linear case, we know that the minimizer of $\widetilde{F}_\delta(x; x_0)$ is given by

$$x = ((A'_{x_0})^* A_{x_0} + \delta I)^{-1}(A'_{x_0})^* g(y, x_0).$$

Minimization strategy with step size control

It may happen that the solution of the linearized problem does not reflect adequately the nonlinearities of the original function. A better strategy is to implement some form of step size control. For example, we might design the following iterative method.

1. Pick an initial guess x_0 and set $k = 0$.

   Repeat:

   2. Calculate the Fréchet derivative (A'_{x_0}).
   3. Determine

      $$x = ((A'_{x_k})^* A'_{x_k} + \delta I)^{-1} (A'_{x_k})^* g(y, x_k), \quad g(y, x_k) = y - A(x_k) + A'_{x_k} x_k,$$

      and define $\Delta x = x - x_k$.
   4. Find step size $s > 0$ by minimizing the function

      $$f(s) = \|A(x_k + s\Delta x) - y\|^2 + \|x_k + s\Delta x\|^2.$$

   5. Set $x_{k+1} = x_k + s\Delta x$ and increase $k \leftarrow k + 1$.

   until convergence.

Remarks on nonlinear Tikhonov regularization

- In practice, evaluating A'_{x_k} is often the most difficult part.
- For finite-dimensional operators, the Fréchet derivative is simply the Jacobi matrix.
- Depending on the nature of the nonlinearity, one might also consider more "specialized" optimization methods (e.g., Gauss–Newton algorithm, Levenberg–Marquardt algorithm...).

More general penalty terms

A more general way of defining the Tikhonov functional is

$$F_\delta(x) = \|Ax - y\|^2 + \delta G(x),$$

where $G\colon H_1 \to \mathbb{R}_{\geq 0}$ takes non-negative values. The existence of a unique minimizer for this kind of functional depends on the properties of G, as does the workload needed for finding it.

One typical way of defining G is

$$G(x) = \|L(x - x_0)\|^2,$$

where $x_0 \in H_1$ is a given reference vector and L is some linear operator. The choice of x_0 and L reflects our prior knowledge about "feasible" solutions: Lx is some property that is known to be relatively close to the reference value Lx_0 for all reasonable solutions. (In the standard case $x_0 = 0$ and $L = I$, the solutions are "known" to lie relatively close to the origin.)

The numerical implementation of Tikhonov regularization with $G(x) = \|L(x - x_0)\|^2$ is approximately as easy as for the standard penalty term.

In the case where $H_1 = \mathbb{R}^n$ and $H_2 = \mathbb{R}^m$, the operator L is some matrix in $\mathbb{R}^{\ell+n}$ and the Tikhonov functional can be given as

$$F_\delta(x) = \left\| \begin{bmatrix} A \\ \sqrt{\delta} L \end{bmatrix} - \begin{bmatrix} y \\ \sqrt{\delta} L x_0 \end{bmatrix} \right\|^2.$$

Assuming that the singular values of K are bounded suitably far away from zero, the Tikhonov solution can be computed in MATLAB as

```
K = [A; sqrt(delta)*L];
z = [y; sqrt(delta)*L*x0];
xdelta = K\z;
```