Inverse Problems

Sommersemester 2022

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

First lecture, April 25, 2022

Practical matters

- Lectures on Mondays at 10.15-12.00 in T9/046 by Vesa Kaarnioja.
- Exercises on Mondays at 12.15-14.00 in T9/046 by Vesa Kaarnioja starting next week.
- Weekly exercises published after the lecture. Please return your written solutions to Vesa either by email (vesa.kaarnioja@fu-berlin.de) or at the beginning of the exercise session in the following week.
- The course grade is determined as a weighted average of the exercise points (25%) and the course exam (75%).
- 50% completion of all tasks ensures a passing grade, 90% completion of all tasks ensures the best grade.

Course contents

- The first part of the course will cover classical variational regularization methods. We will follow Chapters 1–4 in
  - J. Kaipio and E. Somersalo (2005). Statistical and Computational Inverse Problems. Springer, New York, NY.
- Second part of the course will cover Bayesian inverse problems. We will follow the texts
  - D. Sanz-Alonso, A. M. Stuart, and A. Taeb (2018). Inverse Problems and Data Assimilation. https://arxiv.org/abs/1810.06191
  - J. Kaipio and E. Somersalo (2005). Statistical and Computational Inverse Problems. Springer, New York, NY.
  - D. Calvetti and E. Somersalo (2007). Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing. Springer, New York, NY.

What is an inverse problem?

- **Forward problem:** Given known causes (initial conditions, material properties, other model parameters), determine the effects (data, measurements).
- **Inverse problem:** Observing the effects (noisy data), recover the cause.

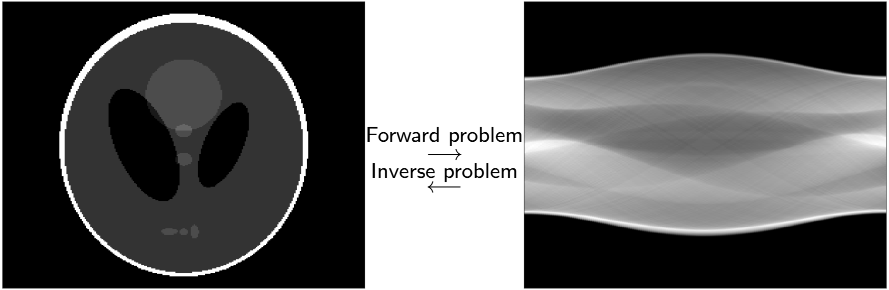


Figure: Computerized tomography (CT)

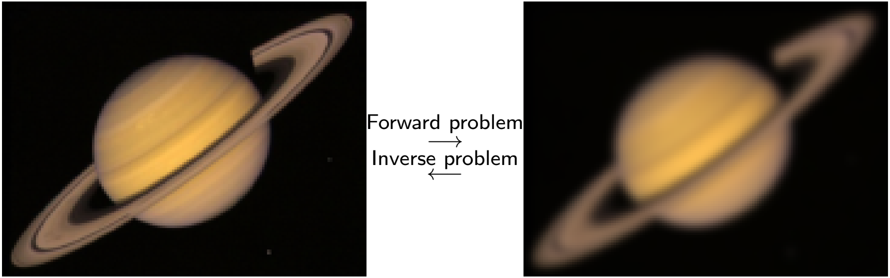


Figure: Image deblurring (deconvolution)

$$y = (K * f)(x) = \int_{\mathbb{R}^2} K(x - x')f(x')\,\mathrm{d}x'$$

Introduction: What is an inverse problem?

We consider the indirect measurement of an unknown physical quantity $x \in X$. The measurement $y \in Y$ is related to the unknown by a physical or mathematical *model*

$$y = F(x), \tag{1}$$

where $F\colon X \to Y$ is called the *forward mapping*.

- Computing y for a given x is called the *forward problem*.
- Finding x for a given measurement y (the *data*) is called the *inverse problem*.

The inverse problem is often ill-posed, making it more difficult than the corresponding direct problem.

A problem is called *well-posed (according to Hadamard)*, if

(a) a solution exists,
(b) the solution is unique, and
(c) the solution depends continuously on the data.

If one or more of these conditions are violated, the problem is called *ill-posed*.

Some examples of ill-posed inverse problems are X-ray tomography, image deblurring, the inverse heat equation, and electrical impedance tomography (EIT).

The ill-posedness of an inverse problem poses a challenge because usually, errors are present in the measurements. Incorporating these into model (1) in the form of additive *noise* η leads to a more realistic model

$$y = F(x) + \eta.$$

The violation of the above conditions leads to various difficulties.

- If condition (a) is violated, i.e., if the image $\mathrm{Ran}(F)$ of F does not cover the whole space Y, then there may not exist a solution to $F(x) = y$ for noisy data $y = F(x^\dagger) + \eta$ created by a ground truth $x^\dagger$, although a solution exists for noise free data $y = F(x^\dagger)$, since η does not need to lie in $\mathrm{Ran}(F)$.
- If condition (c) is violated, then the solution to $F(x) = y$ for noisy data $y = F(x^\dagger) + \eta$ may be far away from the solution for noise free data $y = F(x^\dagger)$, even if F is invertible and the noise η is small, due to the discontinuity of F^{-1}.

Example.

The deblurring (or deconvolution) problem of recovering an input signal x from an observed signal y (possibly contaminated by noise) occurs in many imaging as well as image and signal processing applications. The mathematical model is

$$y(t) = \underbrace{\int_{-\infty}^{\infty} a(t-s)x(s)\mathrm{d}s}_{=:(a*x)(t)},$$

where the function a is known as the blurring kernel.

If $\widehat{a}$ is "nice", we can use the Fourier transform together with the convolution theorem to solve the problem analytically:

$$y(t) = (a*x)(t) \quad \Leftrightarrow \quad \widehat{y}(\xi) = \widehat{a}(\xi)\widehat{x}(\xi) \quad \Leftrightarrow \quad \widehat{x}(\xi) = \frac{\widehat{y}(\xi)}{\widehat{a}(\xi)}$$

$$\Leftrightarrow \quad x(t) = \mathcal{F}^{-1}\left\{\frac{\widehat{y}}{\widehat{a}}\right\}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{it\xi}\frac{\widehat{y}(\xi)}{\widehat{a}(\xi)}\mathrm{d}\xi.$$

Let x_{exact} denote the solution to this problem with exact, noiseless data.

However, if we can only observe noisy measurements, we must consider

$$y(t) = (a * x)(t) + \eta(t) \quad \Leftrightarrow \quad \widehat{y}(\xi) = \widehat{a}(\xi)\widehat{x}(\xi) + \widehat{\eta}(\xi).$$

The solution formula from the previous slide gives (in the Fourier side)

$$\widehat{x}(\xi) = \frac{\widehat{y}(\xi)}{\widehat{a}(\xi)} = \widehat{x}_{\mathrm{exact}}(\xi) + \frac{\widehat{\eta}(\xi)}{\widehat{a}(\xi)};$$

then we apply the inverse Fourier transform on both sides. However, this reconstruction might not be well-defined and it is typically not stable, i.e., it does not depend continuously on the data y. The kernel a usually decreases exponentially (or has compact support). A typical example is a Gaussian kernel

$$a(t) = \frac{1}{2\pi\alpha^2} \exp\left(-\frac{t^2}{2\alpha^2}\right)$$

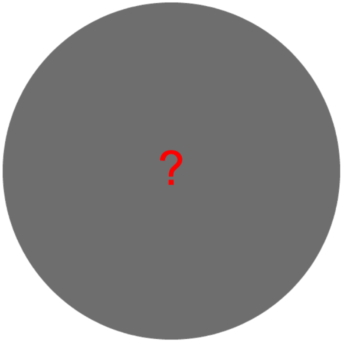
for some $\alpha > 0$.

By the Plancherel theorem, $\widehat{a} \in L^2(\mathbb{R})$ and

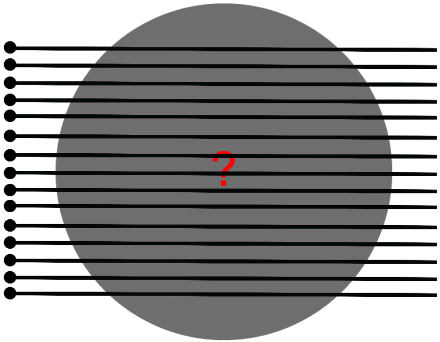
$$\int_{-\infty}^{\infty} |a(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} |\widehat{a}(\xi)|^2 \mathrm{d}\xi$$

if $a \in L^2(\mathbb{R})$. This implies in particular that $\widehat{a}(\xi) \to 0$ as $|\xi| \to \infty$. As a consequence, high frequencies $\widehat{\eta}(\xi)$ of the noise get amplified arbitrarily strong in $\widehat{x}$. Thus, even the presence of small noise can lead to large changes in the reconstruction.

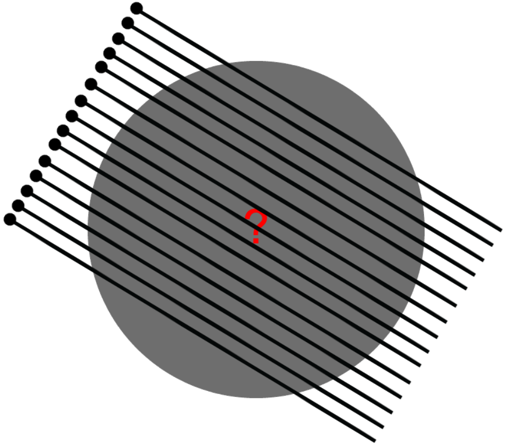
Case study: parallel-beam X-ray tomography



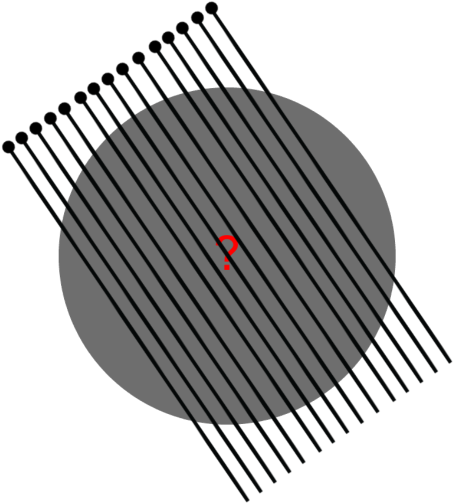
Case study: parallel-beam X-ray tomography



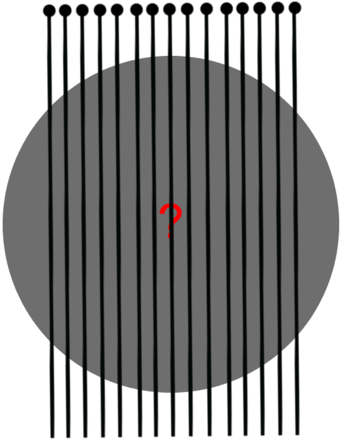
Case study: parallel-beam X-ray tomography



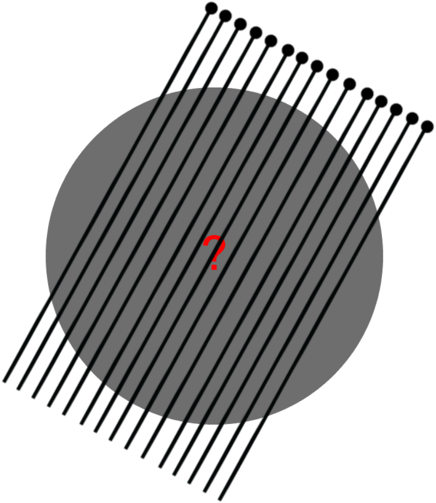
Case study: parallel-beam X-ray tomography



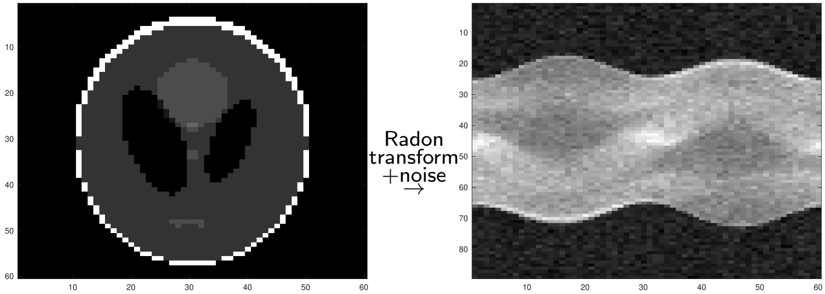
Case study: parallel-beam X-ray tomography



Case study: parallel-beam X-ray tomography



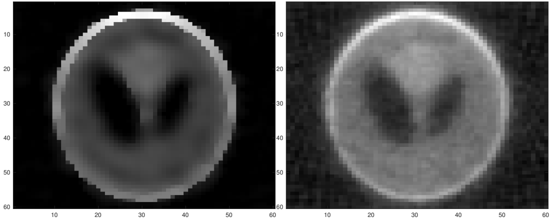
Let us consider the following phantom (botton left), which we use to simulate measurements taken from 60 angles contaminated with 5 % Gaussian noise (sinogram on the bottom right). Inverse problem: use the sinogram data (X-ray images taken from the different directions) to reconstruct the internal structure of the physical body (i.e., the phantom).



Technical (but important) note: to avoid the so-called inverse crime, the measurements for the inversion on the following page were generated using a higher resolution phantom.

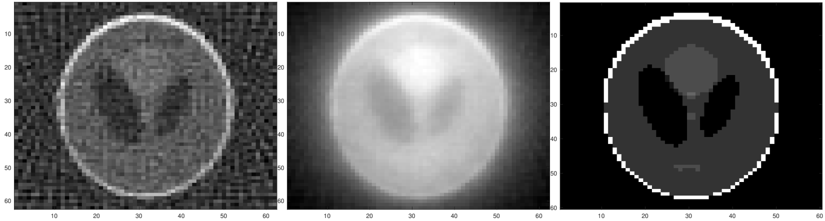
Formation of a CT sinogram (Samuli Siltanen):
https://www.youtube.com/watch?v=q7Rt_OY_7tU

Reconstructions $\arg\min_x \{\|Ax - m\|^2 + \mathcal{R}(x)\}$ from noisy measurements m with some selected penalty terms $\mathcal{R}$ are given immediately below.



Left: reconstruction with total variation regularization. Right: same with Tikhonov regularization.

Some other reconstructions for comparison (and the target phantom).



Left: filtered back projection. Middle: unfiltered back projection. Right: ground truth.

- Successful solution of inverse problems requires specially designed algorithms that can tolerate errors in measured data.
- How to incorporate all possible prior and expert knowledge about the possible solutions when solving inverse problems?
- The statistical approach to inverse problems aims to quantify how uncertainty in the data or model affects the solutions obtained in problems.

Separable Hilbert space

A vector space H is called a *real inner product space* if there exists a mapping $\langle\cdot,\cdot\rangle\colon H \times H \to \mathbb{R}$ satisfying

- $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in H$;
- $\langle ax_1 + bx_2, y\rangle = a\langle x_1, y\rangle + b\langle x_2, y\rangle$ for all $x_1, x_2, y \in H$ and $a, b \in \mathbb{R}$;
- $\langle x, x\rangle \geq 0$, where equality holds iff $x = 0$.

Moreover, H is a *Hilbert space* if

- H is *complete* with respect to the norm $\|x\| = \sqrt{\langle x, x\rangle}$, $x \in H$,

and it is said to be *separable* if, additionally,

- there exists a *countable orthonormal basis* $\{\phi_n\}$ of H with respect to the inner product $\langle\cdot,\cdot\rangle$, that is,

$$\langle \phi_j, \phi_k\rangle = \delta_{j,k} \quad \text{and} \quad x = \sum_n \langle x, \phi_n\rangle \phi_n \quad \text{for all } x \in H.$$

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

Definition

The kernel (or null space) of operator A is defined as

$$\mathrm{Ker}(A) := \{x \in H_1 \mid Ax = 0\}.$$

The range (or image) of operator A is defined as

$$\mathrm{Ran}(A) := \{y \in H_2 \mid y = Ax,\ x \in H_1\}.$$

- A linear operator A is continuous if and only if there exists a constant $C > 0$ such that $\|Ax\| \leq C\|x\|$ for all $x \in H_1$.
- There exists a unique adjoint operator $A^* \colon H_2 \to H_1$ defined by

$$\langle Ax, y \rangle = \langle x, A^* y \rangle \quad \text{for all } x \in H_1,\ y \in H_2.$$

  (This is a consequence of the Riesz representation theorem.)
- $\mathrm{Ker}(A)$ is a *closed* subspace of H_1, and $\mathrm{Ran}(A)$ is a subspace of H_2.

Let H be a real Hilbert space.

Definition

Two elements $x, y \in H$ are said to be *orthogonal* if $\langle x, y \rangle = 0$.

Let $X \subset H$ be a subset. The orthogonal complement of X in H is defined as

$$X^\perp := \{y \in H \mid \langle x, y \rangle = 0 \quad \text{for all } x \in X\}.$$

- For any set $X \subset H$, $X^\perp$ is a closed subspace of H and $X \subset (X^\perp)^\perp$.
- If X is a non-closed subspace, then $(X^\perp)^\perp = \overline{X}$.
- If X is a closed subspace, then $X = (X^\perp)^\perp$. In this case, there exists the *orthogonal decomposition*

$$H = X \oplus X^\perp,$$

  which means that every element $y \in H$ can be uniquely represented as

$$y = x + x^\perp, \quad x \in X,\ x^\perp \in X^\perp.$$

- If $X \subset H$ is a closed subspace, the mapping $P_X : H \to X$, $y \mapsto x$, is an *orthogonal projection*, i.e., $P_X^2 = P_X$ and $\mathrm{Ran}(P_X) \perp \mathrm{Ran}(I - P_X)$.

Lemma

Let $X \subset H$ be a closed subspace. The orthogonal projection $P_X : H \to X$ satisfies the following properties:

- P_X *is linear;*
- P_X *is self-adjoint:* $P_X^* = P_X$;
- $\|P_X\| = 1$ *if* $X \neq \{0\}$;
- $I - P_X = P_{X^\perp}$;
- $\|y - P_X y\| \leq \|y - z\|$ *for all* $z \in X$;
- $z = P_X y$ *iff* $z \in X$ *and* $y - z \in X^\perp$.

Proposition

Let H_1 and H_2 be Hilbert spaces and $A\colon H_1 \to H_2$ a continuous linear operator. Then

$$H_1 = \mathrm{Ker}(A) \oplus (\mathrm{Ker}(A))^\perp = \mathrm{Ker}(A) \oplus \overline{\mathrm{Ran}(A^*)},$$
$$H_2 = \overline{\mathrm{Ran}(A)} \oplus (\mathrm{Ran}(A))^\perp = \overline{\mathrm{Ran}(A)} \oplus \mathrm{Ker}(A^*).$$

Proof. $H_1 = \mathrm{Ker}(A) \oplus (\mathrm{Ker}(A))^\perp$ and $H_2 = \overline{\mathrm{Ran}(A)} \oplus (\overline{\mathrm{Ran}(A)})^\perp = \overline{\mathrm{Ran}(A)} \oplus (\mathrm{Ran}(A))^\perp$ follow immediately from the previous discussion.[†] The claim

$$(\mathrm{Ran}(A))^\perp = \mathrm{Ker}(A^*) \tag{2}$$

follows immediately by observing that $x \in \mathrm{Ker}(A^*)$ iff

$$0 = \langle A^* x, y \rangle = \langle x, Ay \rangle \quad \text{for all } y \in H_1.$$

The claim $(\mathrm{Ker}(A))^\perp = \overline{\mathrm{Ran}(A^*)}$ follows by applying (2) with A replaced by A^*. □

[†] We use the fact that $\overline{X}^\perp = X^\perp$ for any subspace X of H. The inclusion $\overline{X}^\perp \subset X^\perp$ is trivial. To see the other direction, let $x \in X^\perp \Rightarrow \langle x, y \rangle = 0$ for all $y \in X$. Let $z \in \overline{X}$ be arbitrary. We can write $z = \lim_{n\to\infty} z_n$ for a sequence $\{z_n\} \subset X$. Hence, $\langle x, z \rangle = \langle x, \lim_{n\to\infty} z_n \rangle = \lim_{n\to\infty} \langle x, z_n \rangle = 0$. Since $z \in \overline{X}$ was arbitrary, $\overline{X}^\perp = X^\perp$.

Fredholm equation

Let us formalize the problem that we will concentrate on during the first part of the course.

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. We are interested in finding $x \in H_1$ such that

$$y = Ax,$$

where $y \in H_2$ is given.

Examples:

- $H_1 = H_2 = L^2(a, b)$.
- $H_1 = \mathbb{R}^n$, $H_2 = \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$.

[†]A continuous (bounded) linear operator $A\colon H_1 \to H_2$ between Hilbert spaces H_1 and H_2 is said to be compact if the sets $\overline{A(U)} \subset H_2$ are compact for every bounded set $U \subset H_1$.

Singular value decomposition of a compact operator

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

Then there exist (possibly countably infinite) orthonormal sets of vectors $\{v_n\} \subset H_1$ and $\{u_n\} \subset H_2$, and a sequence of positive numbers $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq 0$ with $\lim_{n\to\infty} \lambda_n = 0$ in the countably infinite case such that

$$Ax = \sum_n \lambda_n \langle x, v_n \rangle u_n \quad \text{for all } x \in H_1. \tag{3}$$

In particular,

$$\overline{\mathrm{Ran}(A)} = \overline{\mathrm{span}\{u_n\}} \quad \text{and} \quad (\mathrm{Ker}(A))^{\perp} = \overline{\mathrm{span}\{v_n\}}.$$

The system (λ_n, v_n, u_n) is called a *singular system* of A, and (3) is a *singular value decomposition* (SVD) of A.

Singular value decomposition of matrices: $H_1 = \mathbb{R}^n$ and $H_2 = \mathbb{R}^m$

Let $H_1 = \mathbb{R}^n$ and $H_2 = \mathbb{R}^m$, meaning that

$$y = Ax$$

is a matrix equation with $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, and $y \in \mathbb{R}^m$.

Since this operator has finite rank ($\mathrm{rank}(A) := \dim \mathrm{Ran}(A) < \infty$), we have

$$Ax = \sum_{j=1}^{p} \lambda_j (x^{\mathrm{T}} v_j) u_j, \quad p := \mathrm{rank}(A) \leq \min\{n, m\},$$

where $\{v_j\}_{j=1}^p \subset \mathbb{R}^n$ and $\{u_j\}_{j=1}^p \subset \mathbb{R}^m$ are sets of orthonormal vectors and $\{\lambda_j\}_{j=1}^p$ are positive numbers such that $\lambda_j \geq \lambda_{j+1}$.

It is possible to complete the sequences of (orthonormal) singular vectors $\{v_j\}_{j=1}^{p} \subset \mathbb{R}^n$ and $\{u_j\}_{j=1}^{p} \subset \mathbb{R}^m$ with complementary orthonormal vectors $\{v_j\}_{j=p+1}^{n}$ and $\{u_j\}_{j=p+1}^{m}$ such that $\{v_j\}_{j=1}^{n}$ forms an orthonormal basis for $\mathbb{R}^n$ and $\{u_j\}_{j=1}^{m}$ forms an orthonormal basis for $\mathbb{R}^m$. This can be done, e.g., using the Gram–Schmidt process.

Define the matrices

$$V = [v_1, \ldots, v_n] \in \mathbb{R}^{n \times n},$$
$$U = [u_1, \ldots, u_m] \in \mathbb{R}^{m \times m}.$$

Due to the orthonormality of $\{v_j\}$ and $\{u_j\}$, the matrices V and U are orthogonal:

$$V^{\mathrm{T}} V = V V^{\mathrm{T}} = I \quad \text{and} \quad U^{\mathrm{T}} U = U U^{\mathrm{T}} = I.$$

Next, we define the matrix $\Lambda \in \mathbb{R}^{m\times n}$ as follows:

$$\Lambda = \left(\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & O_{m\times(n-m)} \\ & & \lambda_m & \end{array}\right) \quad \text{if } m < n,$$

$$\Lambda = \left(\begin{array}{ccc} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \\ \hline & O_{(m-n)\times n} & \end{array}\right) \quad \text{if } m > n,$$

and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ if $m = n$.

It is simple to check that

$$Ax = \sum_{j=1}^{p} \lambda_j u_j v_j^{\mathrm{T}} x = U\Lambda V^{\mathrm{T}} x \quad \text{for all } x \in \mathbb{R}^n,$$

which yields the matrix singular value decomposition (SVD)

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

This is the decomposition MATLAB returns with the command `svd`.

Note that in the matrix SVD, the singular values $\{\lambda_j\}_{j=1}^{\min\{m,n\}}$ are *non-negative* and

$$\begin{aligned}
\mathrm{Ran}(A) &= \mathrm{span}\{u_j \mid 1 \le j \le p\}, \\
\mathrm{Ker}(A) &= \mathrm{span}\{v_j \mid p+1 \le j \le n\}, \\
(\mathrm{Ran}(A))^{\perp} &= \mathrm{span}\{u_j \mid p+1 \le j \le m\}, \\
(\mathrm{Ker}(A))^{\perp} &= \mathrm{span}\{v_j \mid 1 \le j \le p\},
\end{aligned}$$

where $p = \mathrm{rank}(A) = \max_{1 \le k \le \min\{m,n\}}\{k \mid \lambda_k > 0\}$.

Solvability of $y = Ax$

Let us assume that H_1 and H_2 are separable real Hilbert spaces and let $A \colon H_1 \to H_2$ be a compact linear operator. Let $P \colon H_2 \to \overline{\mathrm{Ran}(A)}$ be an orthogonal projection. This can be represented using the singular system of A as

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

Theorem

The equation $y = Ax$ has a solution iff

$$y = Py \quad \text{and} \quad \underbrace{\sum_n \frac{1}{\lambda_n^2} |\langle y, u_n \rangle|^2 < \infty}_{\text{“Picard criterion”}}.$$

In this case, the solution is of the form

$$x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle v_n \quad \text{for arbitrary } x_0 \in \mathrm{Ker}(A).$$

Proof. "⇒" Suppose that $y = Ax$ has a solution $x \in H_1$. This implies that $y \in \mathrm{Ran}(A)$ (thus $y = Py$) and, moreover,

$$\langle y, u_j \rangle = \langle Ax, u_j \rangle = \langle x, A^* u_j \rangle = \lambda_j \langle x, v_j \rangle$$
$$\Rightarrow \sum_n \frac{1}{\lambda_n^2} |\langle y, u_n \rangle|^2 = \sum_n |\langle x, v_n \rangle|^2 \overset{\text{Bessel inequ.}}{\leq} \|x\|^2 < \infty.$$

"⇐" Next, suppose that $y = Py$ and the Picard criterion hold and define $x := x_0 + \sum_n \lambda_n^{-1} \langle y, u_n \rangle v_n$, where $x_0 \in \mathrm{Ker}(A)$ is arbitrary. We obtain

$$Ax = Ax_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle A v_n = \sum_n \langle y, u_n \rangle u_n = Py = y. \quad \square$$

Remark. In the above proof, it is helpful to note that if A has the SVD

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

then its adjoint A^* has the SVD

$$A^* y = \sum_n \lambda_n \langle y, u_n \rangle v_n.$$

Note that for any $x \in H_1$, we have

$$\|Ax - y\|^2 = \|Ax - Py\|^2 + \|(I - P)y\|^2 \geq \|(I - P)y\|^2.$$

Hence, if y has a nonzero component in the subspace orthogonal to the range of A, the equation $Ax = y$ cannot be satisfied exactly. Thus, the best we can do is to solve the projected equation

$$Ax = PAx = Py.$$

However, there is in general no guarantee that the Picard criterion

$$\sum_n \frac{1}{\lambda_n^2} |\langle Py, u_n \rangle|^2 < \infty$$

is satisfied for a general $y \in H_2$ if $\mathrm{rank}(A) = \dim \mathrm{Ran}(A) = \infty$.

Truncated singular value decomposition (TSVD)

Let us define a family of finite-dimensional orthogonal projections by

$$P_k \colon H_2 \to \operatorname{span}\{u_1, \ldots, u_k\}, \quad y \mapsto \sum_{n=1}^{k} \langle y, u_n \rangle u_n.$$

By the orthogonality of $\{u_n\}$,

$$P(P_k y) = \sum_n \langle P_k y, u_n \rangle u_n = \sum_{n=1}^{k} \langle y, u_n \rangle u_n = P_k y$$

and

$$\sum_n \frac{1}{\lambda_n^2} |\langle P_k y, u_n \rangle|^2 = \sum_{n=1}^{k} \frac{1}{\lambda_n^2} |\langle y, u_n \rangle|^2 < \infty.$$

Note that $k \leq \operatorname{rank}(A)$ if $\operatorname{rank}(A) < \infty$.

It follows that the problem

$$Ax = P_k y \tag{4}$$

is always solvable. Taking on both sides the inner product with u_n, we find that

$$\lambda_n \langle x, v_n \rangle = \begin{cases} \langle y, u_n \rangle, & 1 \le n \le k \\ 0, & n > k. \end{cases}$$

Hence the solutions to (4) are given by

$$x_k = x_0 + \sum_n \frac{1}{\lambda_n} \langle P_k y, u_n \rangle v_n = x_0 + \sum_{n=1}^{k} \frac{1}{\lambda_n} \langle y, u_n \rangle v_n \in H_1$$

for any $x_0 \in \mathrm{Ker}(A)$. Observe that since for increasing k,

$$\|Ax_k - Py\|^2 = \|(P - P_k)y\|^2 \xrightarrow{k \to \infty} 0,$$

the residual of the projected equation can be made arbitrarily small.

Finally, to remove the ambiguity of the sought solution due to the possible noninjectivity of A, we select $x_0 = 0$. This choice minimizes the norm of x_k since, by orthogonality,

$$\|x_k\|^2 = \|x_0\|^2 + \sum_{n=1}^{k} \frac{1}{\lambda_n^2} |\langle y, u_n \rangle|^2.$$

Definition

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 with a singular system (λ_n, v_n, u_n). By the truncated SVD approximation (TSVD) of the problem $Ax = y$, we mean the problem of finding $x \in H_1$ such that

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

for some $k \geq 1$.

Theorem

The solution to the TSVD problem has a unique solution x_k, called the truncated SVD (TSVD) solution, given by

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

The TSVD solution satisfies

$$\|Ax_k - y\|^2 = \|(I - P)y\|^2 + \|(P - P_k)y\|^2 \xrightarrow{k\to\infty} \|(I - P)y\|^2.$$

Truncated SVD for a matrix $A \in \mathbb{R}^{m\times n}$

The truncated SVD solution, i.e., solution of

$$Ax = P_k y \quad \text{and} \quad x \perp \mathrm{Ker}(A), \quad 1 \le k \le p := \mathrm{rank}(A),$$

where $P_k\colon \mathbb{R}^m \to \mathrm{span}\{u_1, \dots, u_k\}$ is an orthogonal projection, is given by

$$x_k = \sum_{j=1}^{k} \frac{1}{\lambda_j}\langle y, u_j\rangle v_j = \sum_{j=1}^{k} \frac{1}{\lambda_j} v_j(u_j^{\mathrm{T}} y) = V\Lambda_k^{\dagger} U^{\mathrm{T}} y,$$

where A has the SVD $A = U\Lambda V^{\mathrm{T}}$ and we define

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

where $\lambda_1 \ge \cdots \ge \lambda_p > 0$ are the singular values of A (i.e., diagonal of Λ).

Moore–Penrose pseudoinverse

For the largest possible cut-off $k = p = \mathrm{rank}(A)$, the matrix

$$A^\dagger := A_p^\dagger = V\Lambda_p^\dagger U^\mathrm{T} =: V\Lambda^\dagger U^\mathrm{T}$$

is called the *Moore–Penrose pseudoinverse*. It follows from the above that $x^\dagger = A^\dagger y$ is the solution of the projected (matrix) equation

$$Ax = Py,$$

where $P \colon \mathbb{R}^m \to \mathrm{Ran}(A)$ is the orthogonal projection.

The solution $x^\dagger = A^\dagger y$ is called the *minimum norm solution* of the problem $y = Ax$ since

$$\|A^\dagger y\| = \min\{\|x\| : \|Ax - y\| = \|(I - P)y\|\},$$

where P is the projection onto the range of A. The minimum norm solution is the solution that minimizes the residual error and has the minimum norm.

Since the smallest singular value λ_p is extremely small in inverse problems, the use of the pseudoinverse is usually very sensitive to inaccuracies in the data y.

Spectral regularization using TSVD, i.e., discarding singular values below a certain threshold from the forward model, is a simple and popular technique used to render linear problems less ill-posed while improving the noise robustness of the numerical inversion procedure. More on this next week...