Inverse Problems

Sommersemester 2023

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

First lecture, April 17, 2023
(Updated: April 24, 2023)

Practical matters

- Lectures on Mondays at 10:15-12:00 in A6/025/026 by Vesa Kaarnioja.
- Exercises on Tuesdays at 10:15-12:00 in A6/007/008 by Vesa Kaarnioja starting next week.
- Weekly exercises published after each 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 conditions for completing this course are *successfully completing and submitting at least 60% of the course's exercises* and *successfully passing the course exam*.

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.

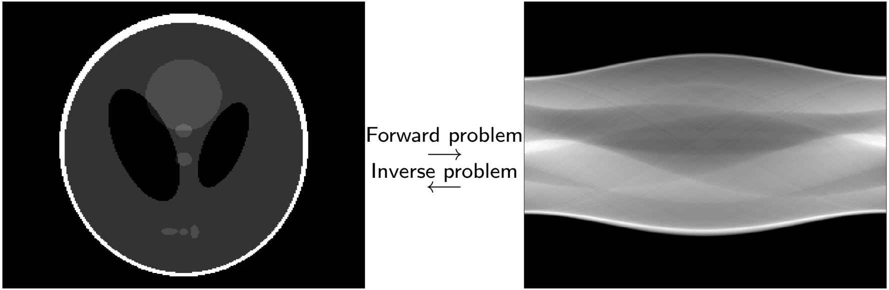


Figure: Computerized tomography (CT)

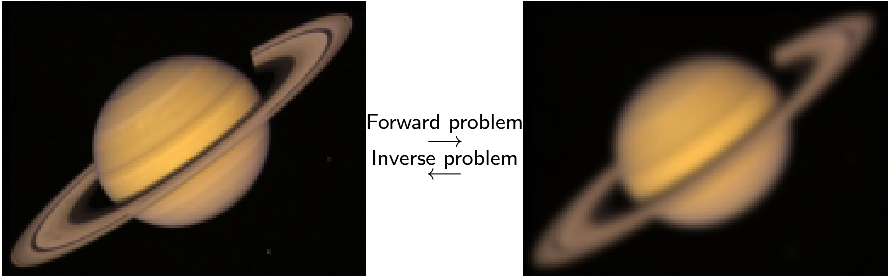


Figure: Image deblurring (deconvolution)

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

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 (in the sense of 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:

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

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

Here, x_{exact} denotes 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

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

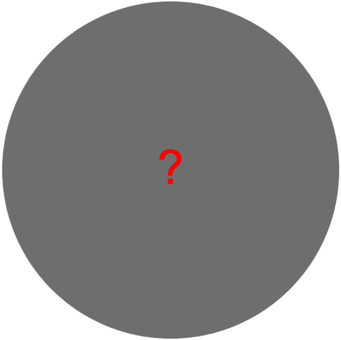
for some $\alpha > 0$.

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

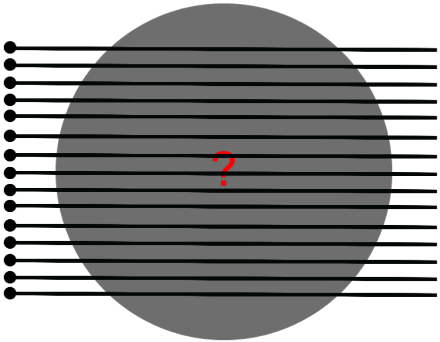
$$\int_{-\infty}^{\infty} |a(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} |\widehat{a}(\xi)|^2 \mathrm{d}\xi$$

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.

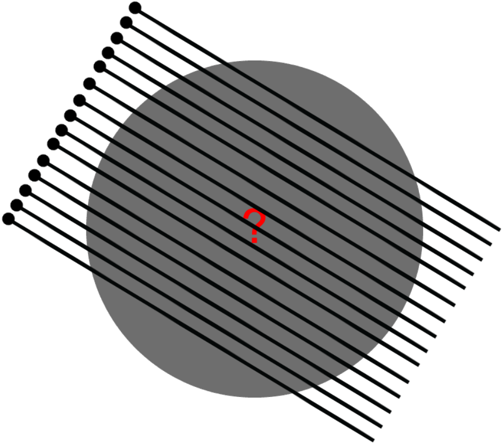
Case study: parallel-beam X-ray tomography



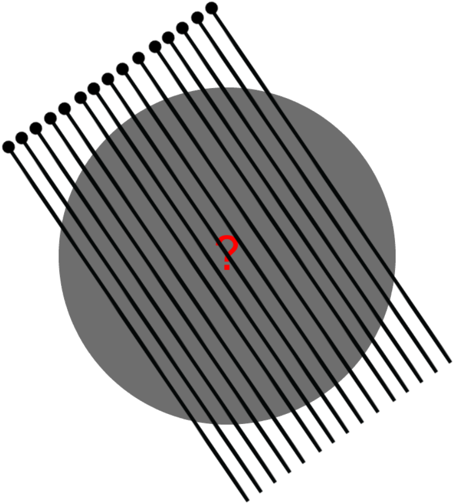
Case study: parallel-beam X-ray tomography



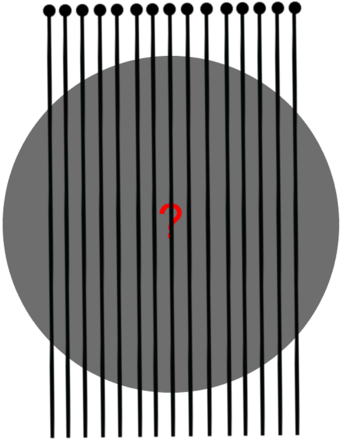
Case study: parallel-beam X-ray tomography



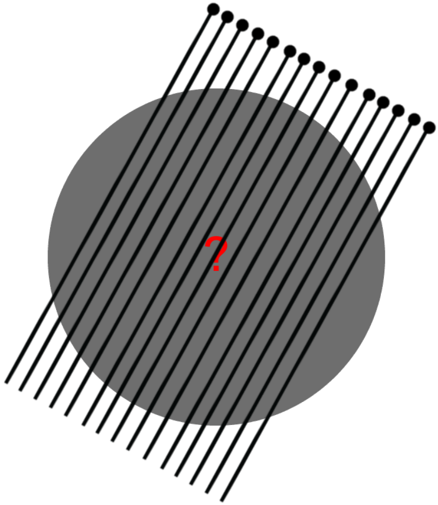
Case study: parallel-beam X-ray tomography



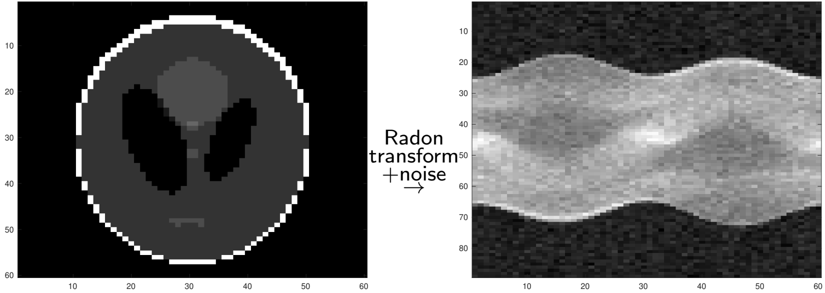
Case study: parallel-beam X-ray tomography



Case study: parallel-beam X-ray tomography



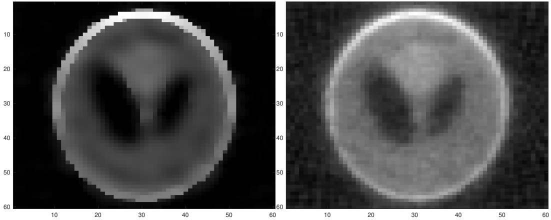
Let us consider the following phantom (botton left), which we use to simulate measurements taken from 60 angles contaminated with 5 % Gaussian noise (sinogram on the bottom right). Inverse problem: use the sinogram data (X-ray images taken from the different directions) to reconstruct the internal structure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

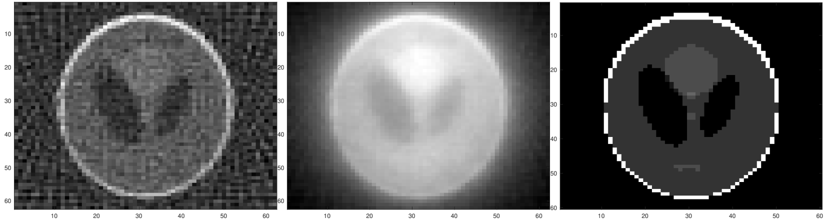
Formation of a CT sinogram (Samuli Siltanen):
https://www.youtube.com/watch?v=q7Rt_OY_7tU

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).

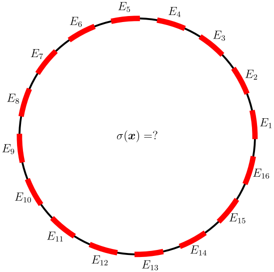


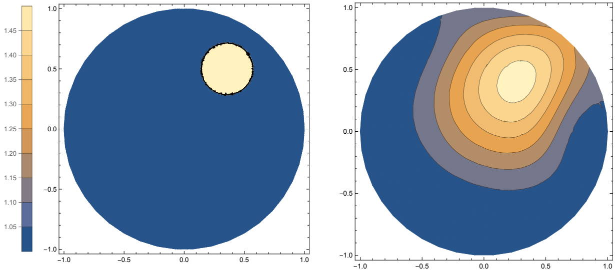
Left: filtered back projection. Middle: unfiltered back projection. Right: ground truth.

Electrical impedance tomography

Use measurements of current and voltage collected at electrodes covering part of the boundary to infer the interior conductivity of an object/body.

$$
\begin{cases}
\nabla \cdot (\sigma \nabla u) = 0 & \text{in } D, \\
\sigma \frac{\partial u}{\partial \boldsymbol{n}} = 0 & \text{on } \partial D \setminus \bigcup_{k=1}^{L} \overline{E_k}, \\
u + z_k \sigma \frac{\partial u}{\partial \boldsymbol{n}} = U_k & \text{on } E_k,\ k \in \{1, \dots, L\}, \\
\int_{E_k} \sigma \frac{\partial u}{\partial \boldsymbol{n}} \, \mathrm{d}S = I_k, & k \in \{1, \dots, L\},
\end{cases}
$$





- 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.

Preliminary functional analysis

Inner product space

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

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

A mapping $\langle\cdot,\cdot\rangle$ satisfying these conditions is called an *inner product*.

Example

i) $\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_k \in \mathbb{R}\}$. Then the inner product is the Euclidean dot product

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k y_k, \quad x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n).$$

ii) Let $X = C([a,b]) = \{f \mid f\colon [a,b] \to \mathbb{R} \text{ is continuous}\}$ and define

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

Then this is an inner product on $C([a,b])$.

iii) Let $X = \ell^2(\mathbb{R}) = \left\{(z_k)_{k=1}^{\infty} \mid \sum_{k=1}^{\infty} |z_k|^2 < \infty\right\}$. Then $\ell^2(\mathbb{R})$ is an inner product space when

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k y_k, \quad x = (x_1, x_2, \ldots),\ y = (y_1, y_2, \ldots).$$

Definition

A real vector space X is a *normed space* if there exists a mapping $\|\cdot\|: X \to \mathbb{R}$ satisfying

- $\|ax\| = |a|\|x\|$ for all $a \in \mathbb{R}$ and $x \in X$;
- $\|x\| \geq 0$ for all $x \in X$, where equality holds iff $x = 0$.
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (triangle inequality).

If X is an inner product space, then it is a normed space in a canonical way with the induced norm $\|\cdot\|: X \to \mathbb{R}$ defined by

$$\|x\| = \sqrt{\langle x, x\rangle}, \quad x \in X.$$

The first two postulates follow immediately from the properties of inner product spaces, the triangle inequality follows from the Cauchy–Schwarz inequality.

Proposition (Cauchy–Schwarz inequality)

If $(X, \langle\cdot,\cdot\rangle)$ *is an inner product space, then*

$$|\langle x, y\rangle| \leq \|x\|\|y\| \quad \textit{for all } x, y \in X.$$

Proof. Let $x, y \in X$ and $t \in \mathbb{R}$. If $x = 0$ or $y = 0$, then the claim is trivial. Suppose that $x \neq 0 \neq y$. Then

$$0 \leq \langle x + ty, x + ty \rangle = \|x\|^2 + 2t\langle x, y \rangle + t^2\|y\|^2.$$

This is a second degree polynomial w.r.t. t with at most 1 real root. Hence,

$$\begin{aligned} \text{discriminant} \leq 0 \quad &\Leftrightarrow \quad 4|\langle x, y\rangle|^2 - 4\|x\|^2\|y\|^2 \leq 0 \\ &\Leftrightarrow \quad |\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2. \end{aligned}$$

Note that if $y = ax$, $a \in \mathbb{R}$, then discriminant $= 0$ and Cauchy–Schwarz holds with equality. □

The triangle inequality is an immediate consequence of Cauchy–Schwarz:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\ &= (\|x\| + \|y\|)^2 \quad \text{for all } x, y \in X. \end{aligned}$$

For our purposes, having an inner product is not enough. We need to know that these spaces are also *complete* normed spaces.

Definition (Cauchy sequence)

A sequence $(x_k)_{k=1}^{\infty}$ of elements of $(X, \|\cdot\|)$ is called a *Cauchy sequence* if for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$m, n > N \quad \Rightarrow \quad \|x_m - x_n\| < \varepsilon.$$

Definition (Complete space)

A normed space $(X, \|\cdot\|)$ is *complete* if all Cauchy sequences in X converge to an element of X.

Definition (Banach space)

A normed space $(X, \|\cdot\|)$ which is complete with respect to $\|\cdot\|$ is a *Banach space*.

Definition (Hilbert space)

An inner product space $(H, \langle\cdot,\cdot\rangle)$ which is complete with respect to $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$ defined by the inner product is a *Hilbert space*.

Example

i) $\mathbb{R}^n$ and $\ell^2(\mathbb{R})$ are complete.

ii) $C([a,b])$ is *not* complete w.r.t. the norm

$$\|f\|^2 = \int_a^b |f(x)|^2 \,\mathrm{d}x.$$

Let $a = -1$, $b = 1$, and define

$$f_n(x) := \begin{cases} 0, & -1 \le x < 0, \\ nx, & 0 \le x \le \frac{1}{n}, \\ 1, & \frac{1}{n} < x \le 1. \end{cases}$$

Then f_n is continuous, and if $H(x) = \chi_{[0,1]}(x) = \begin{cases} 0, & -1 \le x \le 0, \\ 1, & 0 < x \le 1, \end{cases}$ we have

$$\int_{-1}^1 |f_n(x) - H(x)|^2 \,\mathrm{d}x = \int_0^{1/n} |nx - 1|^2 \,\mathrm{d}x = \int_0^{1/n} (n^2x^2 - 2nx + 1) \,\mathrm{d}x$$

$$= \left[\frac{n^2x^3}{3} - nx^2 + x \right]_{x=0}^{x=1/n} = \frac{1}{3n} - \frac{1}{n} + \frac{1}{n} = \frac{1}{3n} \xrightarrow{n\to\infty} 0.$$

We have $\|f_n - H\| \to 0$, but $H \notin C([-1,1])$.

However, note that $C([a,b])$ *is* complete w.r.t. the sup-norm $\|f\|_\infty = \sup_{a\le x\le b} |f(x)|$, but $\|\cdot\|_\infty \neq \|\cdot\|$ and *there is no inner product inducing* $\|\cdot\|_\infty$*-norm.*

Bounded linear operators in Hilbert spaces

Definition

Let X and Y be normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$. A linear operator $A\colon X \to Y$ is said to be *bounded* if there exists $C > 0$ such that

$$\|Ax\|_Y \le C\|x\|_X \quad \text{for all } x \in X.$$

Lemma

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Then a linear operator $A\colon X \to Y$ is bounded iff

$$\|A\| := \|A\|_{X\to Y} := \sup_{\|x\|_X \le 1} \|Ax\|_Y < \infty. \qquad \text{(operator norm)}$$

Proof. "$\Rightarrow$" If there is $C > 0$ s.t. $\|Ax\|_Y \le C\|x\|_X$ for all $x \in X$, then clearly $\|A\| = \sup_{\|x\|_X \le 1} \|Ax\|_Y \le C$.
"$\Leftarrow$" Let $\|A\| < \infty$. Since $\|\frac{x}{\|x\|_X}\|_X = 1$ for all $x \neq 0$, from the linearity of A we infer

$$\frac{\|Ax\|_Y}{\|x\|_X} = \|A(\tfrac{x}{\|x\|_X})\|_Y \le \|A\| \quad \text{for all } x \in X.$$

This implies the important estimate

$$\|Ax\|_Y \le \|A\|\|x\|_X \quad \text{for all } x \in X.$$

□

A linear operator is continuous precisely when it is bounded.

Proposition

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $A\colon X \to Y$ a linear operator. Then the following are equivalent:

(i) *A is a bounded operator;*

(ii) *A is continuous (in X);*

(iii) *A is continuous at one point $x_0 \in X$.*

Proof. (i) $\Rightarrow$ (ii): if $x, y \in X$ and $\varepsilon > 0$, then

$$\|x - y\|_X \le \frac{\varepsilon}{\|A\|} =: \delta \quad \Rightarrow \quad \|Ax - Ay\|_Y \stackrel{A \text{ linear}}{=} \|A(x - y)\|_Y \le \|A\|\|x - y\|_X \le \varepsilon.$$

(ii) $\Rightarrow$ (iii): trivial.
(iii) $\Rightarrow$ (i): let A be continuous at $x_0 \in X$. By definition, there exists $\delta > 0$ such that

$$\|y - x_0\|_X \le \delta \quad \Rightarrow \quad \|Ay - Ax_0\|_Y \le 1.$$

If $x \in X$ is such that $\|x\|_X \le \delta$, then by taking $y = x + x_0$:

$$\|Ax\|_Y = \|A(x + x_0) - Ax_0\|_Y \le 1.$$

On the other hand, for any $\|x\|_X \le 1$, there holds $\|\delta x\|_X = \delta\|x\|_X \le \delta$ and thus

$$\delta\|Ax\|_Y = \|A(\delta x)\|_Y \le 1, \quad \text{i.e.,} \quad \|Ax\|_Y \le \frac{1}{\delta} \quad \text{for all } \|x\|_X \le 1.$$

Therefore $\|A\| \le \frac{1}{\delta}$, meaning that A is bounded. □

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

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

We state the following easy consequences.

Lemma

For any subset $M \subset H$, $M^{\perp}$ is a closed subspace of H and $M \subset (M^{\perp})^{\perp}$.

Lemma

If M is a subspace of H, then $(M^{\perp})^{\perp} = \overline{M}$.
If M is a closed *subspace of H, then $(M^{\perp})^{\perp} = M$.*

Proposition (Hilbert projection theorem)

Let M be a nonempty, closed, and convex[†] *subset of a real Hilbert space H. Then there exists a unique element $x_0 \in M$ satisfying*

$$\|x_0\| \leq \|x\| \quad \textit{for all } x \in M.$$

Proof. Let $\delta = \inf\{\|x\| \mid x \in M\}$. We use the parallelogram identity $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$ applied to vectors $u = \frac{1}{2}x$ and $v = \frac{1}{2}y$, $x, y \in M$, to obtain

$$\frac{1}{4}\|x-y\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 - \left\|\frac{x+y}{2}\right\|^2.$$

Due to convexity $\frac{1}{2}(x+y) \in M$, so

$$\|x-y\|^2 \leq 2\|x\|^2 + 2\|y\|^2 - 4\delta^2 \quad \text{for all } x, y \in M. \tag{2}$$

Existence: let $(x_k)_{k=1}^{\infty} \subset M$ s.t. $\|x_k\| \xrightarrow{k\to\infty} \delta$. Substituting $x \leftarrow x_n$ and $y \leftarrow x_m$ in (2) yields $\|x_n - x_m\|^2 \leq 2\|x_n\|^2 + 2\|x_m\|^2 - 4\delta^2$, since $\frac{1}{2}(x_n + x_m) \in M$ for all n, m. Thus $\|x_n - x_m\| \to 0$ as $n, m \to \infty$. $(x_k)_{k=1}^{\infty}$ is Cauchy in the Hilbert space H, so there exists $x_0 := \lim_{k\to\infty} x_k \in H$. Since $\|\cdot\|$ is continuous, $\|x_0\| = \lim_{k\to\infty}\|x_k\| = \delta$. Since M is closed and $(x_k)_{k=1}^{\infty} \subset M$, the limit $x_0 \in M$.

Uniqueness: If $\|x\| = \|y\| = \delta \Rightarrow \|x-y\|^2 \leq 0$ by (2) and so $x = y$. □

[†] $tx + (1-t)y \in M$ for all $x, y \in M$, $t \in (0,1)$.

Corollary

Let H be a real Hilbert space, M a nonempty, closed, and convex subset of H, and $x \in H$. Then there exists a unique element $y_0 \in M$ such that

$$\|x - y_0\| \leq \|x - y\| \quad \textit{for all } y \in M.$$

Proof. The set $x - M := \{x - y \mid y \in M\}$ is closed and convex, and $\min\{\|x - y\| \mid x - y \in x - M\} = \min\{\|x - y\| \mid y \in M\}$. The claim follows from the previous result. □

Proposition (Orthogonal decomposition)

If M is a closed subspace of a real Hilbert space H, then

$$H = M \oplus M^{\perp},$$

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

$$y = x + x^{\perp}, \quad x \in M, \; x^{\perp} \in M^{\perp}.$$

Proof. It suffices to prove that $M \cap M^{\perp} = \{0\}$ and $M + M^{\perp} = H$.
• If $x \in M \cap M^{\perp}$, then $0 = \langle x, x \rangle = \|x\|^2$ (i.e., $x \perp x$) so $x = 0$.
$\therefore M \cap M^{\perp} = \{0\}$.
• Let $x \in H$. The Hilbert projection theorem guarantees that there exists a unique $y_0 \in M$ such that

$$\|x - y_0\| \leq \|x - y\| \quad \text{for all } y \in M. \tag{3}$$

Let $x_0 = x - y_0$ so that $x = y_0 + x_0 \in M + x_0$. It remains to show that $x_0 \in M^{\perp}$.
The inequality (3) can be written as

$$\|x_0\| \leq \|z\| \quad \text{for all } z \in x - M.$$

Since $y_0 \in M$ and M is a vector space, $y_0 + M = M$ and $M = -M$ which implies $x - M = x + M = y_0 + x_0 + M = x_0 + M$. The previous inequality can be recast as

$$\|x_0\| \leq \|z\| \quad \text{for all } z \in x_0 + M \quad \Leftrightarrow \quad \|x_0\| \leq \|x_0 + z\| \quad \text{for all } z \in M.$$

This statement is true if and only if $\langle x_0, z \rangle = 0$ for all $z \in M$. Therefore $x_0 \in M^{\perp}$.

Let M be a closed subspace. The orthogonal decomposition implies that every element $y \in H$ can be uniquely represented as

$$y = x + x^{\perp}, \; x \in M, \; x^{\perp} \in M^{\perp}.$$

Lemma

Let $M \subset H$ be a closed subspace. The mapping $P_M \colon H \to M$, $y \mapsto x$, is an orthogonal projection, i.e., $P_M^2 = P_M$ and $\mathrm{Ran}(P_M) \perp \mathrm{Ran}(I - P_M)$. It satisfies the following properties:

- P_M *is linear;*
- $\|P_M\| = 1$ *if* $M \neq \{0\}$;
- $I - P_M = P_{M^{\perp}}$;
- $\|y - P_M y\| \leq \|y - z\|$ *for all* $z \in M$;
- $y \in M \;\Rightarrow\; P_M y = y, \; (I - P_M)y = 0;$
  $y \in M^{\perp} \Rightarrow\; P_M y = 0, \; (I - P_M)y = y;$
- $\|y\|^2 = \|P_M y\|^2 + \|(I - P_M)y\|^2$ *(Pythagoras).*

Proof. Omitted; see for example [Rudin, Real and Complex Analysis, pp. 34–35]. □

Example

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

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

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

Then we have the following:

- $\mathrm{Ker}(A)$ is a *closed* subspace of H_1, and $\mathrm{Ran}(A)$ is a subspace of H_2.
- $H_1 = \mathrm{Ker}(A) \oplus (\mathrm{Ker}(A))^{\perp}$.
- $H_2 = \overline{\mathrm{Ran}(A)} \oplus (\mathrm{Ran}(A))^{\perp}$.

Proposition (Riesz representation theorem)

Let H be a real Hilbert space. If $A\colon H \to \mathbb{R}$ is a bounded linear functional, i.e., A is linear and there exists $C > 0$ such that

$$|A(x)| \leq C\|x\| \quad \textit{for all } x \in H,$$

then there exists a unique *$y \in H$ such that*

$$A(x) = \langle x, y\rangle \quad \textit{for all } x \in H.$$

Proof. If $A \equiv 0$, then $y = 0$ and this is unique. Suppose $A \neq 0$ and let

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

Since A is continuous, M is a *closed* subspace of H. Furthermore, by the orthogonal decomposition $H = M \oplus M^{\perp}$, our assumption $A \neq 0$ implies that $M \neq H \Rightarrow M^{\perp} \neq \{0\}$.

Let $x \in H$ and $z \in M^{\perp}$ with $\|z\| = 1$. Define

$$u := A(x)z - A(z)x.$$

Then

$$A(u) = A(x)A(z) - A(z)A(x) = 0.$$

meaning that $u \in M$. In particular $\langle u, z\rangle = \langle A(x)z - A(z)x, z\rangle = 0$ and

$$\begin{aligned} A(x) = A(x) \underbrace{\langle z, z\rangle}_{=\|z\|^2=1} &= \langle A(x)z, z\rangle \\ &= \langle A(z)x, z\rangle = A(z)\langle x, z\rangle = \langle x, zA(z)\rangle. \end{aligned}$$

$\therefore$ The element $y = zA(z)$ satisfies $A(x) = \langle x, y\rangle$.
To prove uniqueness, suppose that there exist $y_1, y_2 \in H$ such that

$$A(x) = \langle x, y_1\rangle = \langle x, y_2\rangle.$$

Then $\langle x, y_1 - y_2\rangle = 0$ for all $x \in H$. Choose $x = y_1 - y_2$. Then

$$0 = \langle y_1 - y_2, y_1 - y_2\rangle = \|y_1 - y_2\|^2 \quad \Leftrightarrow \quad y_1 = y_2.$$

□

Adjoint operator

Proposition

Let H_1 and H_2 be real Hilbert spaces and suppose that $A\colon H_1 \to H_2$ is a bounded linear operator. Then there exists a unique bounded linear operator $A^*\colon H_2 \to H_1$, called the adjoint *of A, satisfying $\langle Ax, y\rangle_{H_2} = \langle x, A^*y\rangle_{H_1}$. Moreover, $\|A\|_{H_1\to H_2} = \|A^*\|_{H_2\to H_1}$.*

Proof. Let $y \in H_2$ and consider $T_y\colon H_1 \to \mathbb{R}$, $x \mapsto \langle Ax, y\rangle_{H_2}$. Clearly, T_y is linear and bounded so by the Riesz representation theorem there exists a *unique* $z \in H_1$ s.t.

$$\langle Ax, y\rangle_{H_2} = T_y(x) = \langle x, z\rangle_{H_1} \quad \text{for all } x \in H_1.$$

Define $A^*y := z$.

- Let $a, b \in \mathbb{R}$ and $y_1, y_2 \in H_2$. Linearity follows from $\langle x, A^*(ay_1 + by_2)\rangle = \langle Ax, ay_1 + by_2\rangle = a\langle Ax, y_1\rangle + b\langle Ax, y_2\rangle = a\langle x, A^*y_1\rangle + b\langle x, A^*y_2\rangle = \langle x, aA^*y_1 + bA^*y_2\rangle$. Since $x \in H_1$ was arbitrary, $A^*(ay_1 + by_2) = aA^*y_1 + bA^*y_2$.
- $\|A^*\|_{H_2\to H_1} = \sup_{\|y\|_{H_2}\le 1} \|A^*y\|_{H_1} \overset{(*)}{=} \sup_{\|y\|_{H_2}\le 1} \sup_{\|x\|_{H_1}\le 1} |\langle A^*y, x\rangle|$

  $= \sup_{\|y\|_{H_2}\le 1} \sup_{\|x\|_{H_1}\le 1} |\langle y, Ax\rangle| \overset{(*)}{=} \sup_{\|x\|_{H_1}\le 1} \|Ax\|_{H_2} = \|A\|_{H_1\to H_2} < \infty.$ □

$^{(*)}$Let $\Lambda \in \mathcal{L}(H, K)$, H, K Hilbert spaces. Cauchy–Schwarz: $\sup_{\|y\|_K\le 1} |\langle \Lambda x, y\rangle_K| \le \|\Lambda x\|_K$.
Other direction: $\sup_{\|y\|_K\le 1} |\langle \Lambda x, y\rangle_K| \ge |\langle \Lambda x, \frac{1}{\|\Lambda x\|_K}\Lambda x\rangle|_K = \|\Lambda x\|_K$.
$\therefore \|\Lambda x\|_K = \sup_{\|y\|_K\le 1} |\langle \Lambda x, y\rangle_K|$.

Some properties of the adjoint operator

Proposition

Let H_1 and H_2 be real Hilbert spaces and suppose that $A, B\colon H_1 \to H_2$ are bounded linear operators. Then

(i) $\|A^*A\|_{H_1\to H_1} = \|A\|^2_{H_1\to H_2}$,

(ii) $A^{**} = A$, *where* $A^{**} = (A^*)^*$;

(iii) $(c_1A + c_2B)^* = c_1A^* + c_2B^*$, $c_1, c_2 \in \mathbb{R}$.

Proof. (i) Let $x \in H_1$, $\|x\|_{H_1} = 1$. By the Cauchy–Schwarz inequality,

$$\|Ax\|^2_{H_2} = \langle Ax, Ax\rangle_{H_2} = \langle x, A^*Ax\rangle_{H_1} \le \|A^*Ax\|_{H_1} \;\Rightarrow\; \|A\|^2_{H_1\to H_2} \le \|A^*A\|_{H_1\to H_1}.$$

Other direction: $\|A^*A\| \le \|A^*\| \cdot \|A\| = \|A\|^2$ (previous slide and exercise of week 1).
(ii) If $x \in H_1$ and $y \in H_2$, then

$$\langle A^{**}x, y\rangle_{H_2} = \langle x, A^*y\rangle_{H_1} = \langle A^*y, x\rangle_{H_1} = \langle y, Ax\rangle_{H_2} = \langle Ax, y\rangle_{H_2}.$$

Hence $\langle A^{**}x - Ax, y\rangle_{H_2} = 0$ for all $y \in H_2 \Rightarrow A^{**}x = Ax$ for all $x \in H_1 \Rightarrow A^{**} = A$.
(iii) Let $x \in H_1$ and $y \in H_2$. Then

$$\begin{aligned}\langle (c_1A + c_2B)^*y, x\rangle_{H_1} &= \langle y, (c_1A + c_2B)x\rangle_{H_2} = c_1\langle y, Ax\rangle_{H_2} + c_2\langle y, Bx\rangle_{H_2} \\ &= c_1\langle A^*y, x\rangle_{H_1} + c_2\langle B^*y, x\rangle_{H_1} = \langle (c_1A^* + c_2B^*)y, x\rangle_{H_1}.\end{aligned}$$

Similarly to the previous part, we conclude that $(c_1A + c_2B)^* = c_1A^* + c_2B^*$. □

Self-adjoint operators

Definition

Let H be a Hilbert space. A bounded, linear operator $A\colon H \to H$ is called *self-adjoint* if $A^* = A$, i.e.,

$$\langle Ax, y\rangle = \langle x, Ay\rangle \quad \text{for all } x, y \in H.$$

Example

Let H be a Hilbert space and let $A, B\colon H \to H$ be bounded, linear, self-adjoint operators. Then

(i) $A + B$ is self-adjoint.

(ii) if $c \in \mathbb{R}$, then cA is self-adjoint.

(iii) if $AB = BA$, then AB is self-adjoint.

Parts (i) and (ii) follow immediately from part (iii) on the previous slide. If $x, y \in H$, then

$$\langle ABx, y\rangle = \langle BAx, y\rangle = \langle Ax, By\rangle = \langle x, ABy\rangle \quad \Rightarrow \quad (AB)^* = AB.$$

Example

Let H be a real Hilbert space and $M \subset H$ a closed subspace. Then the orthogonal projections $P_M\colon H \to M$ and $I - P_M =: P_{M^\perp} : H \to M^\perp$ are self-adjoint.

Compact operators

Definition

Let H_1 and H_2 be real Hilbert spaces. A bounded linear operator $K\colon H_1 \to H_2$ is compact if the sets $\overline{K(U)} \subset H_2$ are compact for every bounded set $U \subset H_1$.

The following characterization will be useful.

Characterization

Let H_1 and H_2 be real Hilbert spaces. A bounded linear operator $K\colon H_1 \to H_2$ is compact if and only if $(Kx_j)_{j=1}^{\infty} \subset H_2$ contains a convergent subsequence for every bounded sequence $(x_j)_{j=1}^{\infty} \subset H_1$.

Let H, H_1, and H_2 be Hilbert spaces. We have the following properties:

- All linear maps to finite-dimensional spaces are compact.
- If $A, B\colon H_1 \to H_2$ are compact, then $A + B$ is compact.
- If $K\colon H_1 \to H_2$ is compact, then
  - AK is compact for all bounded and linear $A\colon H_2 \to H$.
  - KB is compact for all bounded and linear $B\colon H \to H_1$.
- If $K_n\colon H_1 \to H_2$ are compact operators and $K\colon H_1 \to H_2$ is a bounded, linear operator such that $\|K_n - K\| \xrightarrow{n\to\infty 0} 0$, then K is compact.
- If $K\colon H_1 \to H_2$ is compact, then so is $K^*\colon H_2 \to H_1$.

Proposition

Let H_1 and H_2 be real 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

$$(\mathrm{Ran}(A))^{\perp} = \mathrm{Ker}(A^*) \tag{4}$$

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 (4) with A replaced by A^*. □

[†]Here we use the fact that $\overline{X}^{\perp} = X^{\perp}$ for any subspace X of H; see exercise 1.

Appendix: some auxiliary results

Let X and Y be normed spaces. We denote

$$\mathcal{L}(X, Y) := \{A \mid A\colon X \to Y \text{ is bounded and linear}\}.$$

Proposition

If Y is complete, then $\mathcal{L}(X, Y)$ is complete w.r.t. operator norm (i.e., it is a Banach space).

Proof. Let $x \in X$ and assume that $A_k \in \mathcal{L}(X, Y)$, $k \in \mathbb{N}$, is a Cauchy sequence. Then for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$m, n > N \quad \Rightarrow \quad \|A_m - A_n\| < \frac{\varepsilon}{\|x\|_X}.$$

Especially,

$$\|A_m x - A_n x\|_Y \leq \|A_m - A_n\| \|x\|_X < \varepsilon \quad \text{when } m, n > N,$$

so $(A_k x)$ is a Cauchy sequence in Y and therefore the limit

$$A(x) := \lim_{k \to \infty} A_k x$$

exists.

It is easy to see that $A(x) := \lim_{k\to\infty} A_k x$ is linear. It is also bounded: there exists $N \in \mathbb{N}$ such that

$$m, n > N \quad \Rightarrow \quad \|A_m - A_n\| < 1.$$

Fix $m > N$. Then for all $n > m$,

$$\|A_n\| < 1 + \|A_m\|$$

and thus

$$\|A_n x\|_Y \leq (1 + \|A_m\|)\|x\|_X.$$

But $\|Ax\|_Y = \lim_{n\to\infty} \|A_n x\|_Y \leq (1 + \|A_m\|)\|x\|_X$. Therefore A is bounded.

Finally, we need to show that $\|A_n - A\| \to 0$ as $n \to \infty$. Since we assumed $(A_k)_{k=1}^{\infty}$ to be Cauchy, let $\varepsilon > 0$ be s.t. for $m, n > N$, there holds $\|A_m - A_n\| < \varepsilon$. Then

$$\|(A - A_n)x\|_Y = \lim_{m\to\infty} \|A_m x - A_n x\|_Y \leq \varepsilon\|x\|_X \quad \text{for all } x \in X$$
$$\Rightarrow \quad \|A - A_n\| < \varepsilon.$$

Hence $\|A - A_n\| \to 0$ as $n \to \infty$. □

If $X = H_1$ and $Y = H_2$ are Hilbert spaces, then $\mathcal{L}(H_1, H_2)$ is a complete normed space.

In general, $\mathcal{L}(H_1, H_2)$ is *not* a Hilbert space even when both H_1 and H_2 are. However, in the special case $\mathcal{L}(H, \mathbb{R})$ it turns out that indeed one can associate an inner product that induces the operator norm $\|\cdot\|$ – meaning that $\mathcal{L}(H, \mathbb{R})$ is a Hilbert space! This is a consequence of the Riesz representation theorem (details omitted).

Basic properties of vector-valued series

Definition

Let E be a normed space and $(x_k) \subset E$. Define the n^{th} partial sum $S_n := \sum_{k=1}^{n} x_k$. If there exists an element $S \in E$ such that $\lim_{n\to\infty} \|S - S_n\| = 0$, then we say that the series $\sum_{k=1}^{\infty} x_k$ is *convergent* (in E) and denote

$$S = \sum_{k=1}^{\infty} x_k.$$

Moreover, we say that the series $\sum_{k=1}^{\infty} x_k$ is *absolutely convergent* if $\sum_{k=1}^{\infty} \|x_k\| < \infty$.

Proposition

The normed space E is a Banach space iff every absolutely convergent series $\sum_{k=1}^{\infty} x_k$ is convergent in E.

Theorem (Generalized Pythagorean theorem)

Let (e_k) be an orthonormal sequence in Hilbert space H and let $(\lambda_k) \subset \mathbb{R}$. Then

$$\sum_{k=1}^{\infty} \lambda_k e_k \text{ is convergent} \quad \text{iff} \quad \sum_{k=1}^{\infty} |\lambda_k|^2 < \infty.$$

In this case, we have

$$\left\| \sum_{k=1}^{\infty} \lambda_k e_k \right\|^2 = \sum_{k=1}^{\infty} |\lambda_k|^2.$$

Neumann series: "Sufficiently small perturbations of the identity are still invertible"

The following result is a well-known generalization of the geometric series formula, named after 19$^{\mathrm{th}}$ century mathematician Carl Neumann.

Theorem (Neumann series)

Let H be a real Hilbert space and let $A \in \mathcal{L}(H) := \mathcal{L}(H,H)$ be such that $\|A\| < 1$. Then $I - A$ is invertible in $\mathcal{L}(H)$ with

$$(I - A)^{-1} = I + A + \cdots + A^n + \cdots = \sum_{k=0}^{\infty} A^k,$$

and this series converges in operator norm.

Proof. Let $B_{m,n} := \sum_{k=m}^{n} A^k$, $m < n$. Since $\|A\| < 1$, we have

$$\|B_{m,n}\| \leq \sum_{k=m}^{n} \|A\|^k = \|A\|^m \sum_{k=0}^{m-n} \|A\|^k = \|A\|^m \frac{1 - \|A\|^{n-m+1}}{1 - \|A\|} \xrightarrow{m,n\to\infty} 0.$$

$\therefore$ The partial sums $\sum_{k=0}^{n} A^k$ form a Cauchy sequence in $\mathcal{L}(H)$.

Since H is a Hilbert space, $\mathcal{L}(H)$ is a Banach space and the limit

$$B := \lim_{n\to\infty} \sum_{k=0}^{n} A^k \in \mathcal{L}(H)$$

exists. We need to prove that $(I - A)B = I = B(I - A)$. Let

$$B_n := I + A + \cdots + A^n.$$

Then

$$(I - A)B_n = I - A^{n+1},$$
$$B_n(I - A) = I - A^{n+1},$$

and since $\|A\| < 1$, $\|A^{n+1}\| \leq \|A\|^{n+1} \xrightarrow{n\to\infty} 0$, we thus obtain

$$I - A^{n+1} \xrightarrow{n\to\infty} I \quad \text{in } \mathcal{L}(H)$$

and

$$(I - A)B = \lim_{n\to\infty}(I - A)B_n = I = \lim_{n\to\infty} B_n(I - A) = B(I - A). \quad \square$$

Theorem (Bessel's inequality)

Let H be a real Hilbert space and let (e_n) be an orthonormal sequence in H. Then

$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \le \|x\|^2 \quad \text{for all } x \in H.$$

Especially $\lim_{n\to\infty} \langle x, e_n \rangle = 0$.

Proof. Let $k \in \mathbb{N}$. Noting that

$$\left\langle x - \sum_{n=1}^{k} \langle x, e_n \rangle e_n, e_j \right\rangle = \langle x, e_j \rangle - \sum_{n=1}^{k} \langle x, e_n \rangle \langle e_n, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$$

for all $j \in \{1, \ldots, k\}$, we deduce that $x - \sum_{n=1}^{k} \langle x, e_n \rangle e_n \perp \sum_{n=1}^{k} \langle x, e_n \rangle e_n$ (recall that the orthogonal complement is a subspace). By the Pythagorean theorem,

$$\|x\|^2 = \left\| x - \sum_{n=1}^{k} \langle x, e_n \rangle e_n \right\|^2 + \left\| \sum_{n=1}^{k} \langle x, e_n \rangle e_n \right\|^2 \ge \left\| \sum_{n=1}^{k} \langle x, e_n \rangle e_n \right\|^2 = \sum_{n=1}^{k} |\langle x, e_n \rangle|^2.$$

Letting $k \to \infty$ yields the assertion. □

Lax–Milgram lemma

Proposition (Lax–Milgram lemma)

Let H be a real Hilbert space and let $B\colon H \times H \to \mathbb{R}$ be a bilinear mapping[†] *with $C, c > 0$ such that*

$$|B(u,v)| \leq C\|u\| \cdot \|v\| \quad \textit{for all } u, v \in H, \qquad \text{(boundedness)}$$

$$B(u,u) \geq c\|u\|^2 \quad \textit{for all } u \in H. \qquad \text{(coercivity)}$$

Let $F\colon H \to \mathbb{R}$ be a bounded linear mapping. Then there exists a unique element $u \in H$ satisfying

$$B(u,v) = F(v) \quad \textit{for all } v \in H.$$

and

$$\|u\| \leq \frac{1}{c}\|F\|.$$

[†] $B(u+v,w) = B(u,w) + B(v,w)$, $B(au,v) = aB(u,v)$,
$B(u,v+w) = B(u,v) + B(u,w)$, $B(u,av) = aB(u,v)$
for all $u, v, w \in H$ and $a \in \mathbb{R}$.

Proof. 1) Let $v \in H$ be fixed. Then the mapping

$$T \colon w \mapsto B(v, w), \;\; H \to \mathbb{R},$$

is bounded and linear. It follows from the Riesz representation theorem that there exists a unique element $a \in H$ with

$$Tw = \langle a, w \rangle \quad \text{for all } w \in H.$$

Let us define the mapping $A \colon H \to H$ by setting

$$Av = a.$$

Then

$$B(v, w) = \langle Av, w \rangle \quad \text{for all } v, w \in H.$$

2) We show that the mapping $A\colon H \to H$ is linear and bounded. Clearly,

$$\begin{aligned}\langle A(c_1 v_1 + c_2 v_2), w\rangle &= B(c_1 v_1 + c_2 v_2, w)\\ &= c_1 B(v_1, w) + c_2 B(v_2, w)\\ &= \langle c_1 A v_1 + c_2 A v_2, w\rangle\end{aligned}$$

for all $w \in H$, so $A(c_1 v_1 + c_2 v_2) = c_1 A v_1 + c_2 A v_2$. Moreover,

$$\begin{aligned}\|Av\|^2 &= \langle Av, Av\rangle\\ &= B(v, Av)\\ &\le C\|v\|\|Av\|\end{aligned}$$

which implies that

$$\|Av\| \le C\|v\|.$$

3) We show that

$$\begin{cases} A \text{ is one-to-one,} \\ \mathrm{Ran}(A) = AH \text{ is closed in } H. \end{cases}$$

We begin by noting that

$$c\|v\|^2 \leq B(v,v) = \langle Av, v\rangle \leq \|Av\|\|v\|$$

and thus

$$\|Av\| \geq c\|v\| \quad \text{for all } v \in H. \tag{5}$$

Especially
$$Av = Aw \Rightarrow A(v-w) = 0 \Rightarrow 0 = \|A(v-w)\| \geq c\|v-w\| \geq 0 \Rightarrow v = w$$
so A is one-to-one.
To see that $\mathrm{Ran}(A)$ is closed, let $y_j = Ax_j \in \mathrm{Ran}(A)$. The goal is to show that $y := \lim_{j\to\infty} y_j \in \mathrm{Ran}(A)$. We observe that

$$\lim_{j,k\to\infty} \|x_j - x_k\| \overset{(5)}{\leq} \lim_{j,k\to\infty} \frac{1}{c}\|y_j - y_k\| = 0,$$

i.e., $(x_j)_{j=1}^{\infty}$ is Cauchy and $x := \lim_{j\to\infty} x_j \in H$ exists by completeness. Moreover,

$$\lim_{j\to\infty} \|Ax_j - Ax\| \leq \lim_{j\to\infty} \|A\|\|x_j - x\| \leq C \lim_{j\to\infty} \|x_j - x\| = 0$$

and therefore

$$y = \lim_{j\to\infty} Ax_j = Ax \in \mathrm{Ran}(A).$$

4) We show that $\mathrm{Ran}(A) = H$. We prove this by contradiction: suppose that $\mathrm{Ran}(A) = \overline{\mathrm{Ran}(A)} \neq H$. Then there exists $w \in \mathrm{Ran}(A)^\perp$, $w \neq 0$.[†] This implies that

$$\|w\|^2 \leq \frac{1}{c} B(w, w) = \frac{1}{c} \langle Aw, w \rangle = 0,$$

i.e., $w = 0$. This contradiction shows that $\mathrm{Ran}(A) = H$. Therefore $A \colon H \to H$ is a continuous bijection.

5) Existence of a solution. We use the Riesz representation theorem: since $F \colon H \to \mathbb{R}$ is linear and continuous, there exists $b \in H$ such that

$$F(v) = \langle b, v \rangle \quad \text{for all } v \in H.$$

Define $u := A^{-1} b$. Hence

$$\begin{aligned} Au = b \quad &\Leftrightarrow \quad \langle Au, v \rangle = \langle b, v \rangle \quad \text{for all } v \in H \\ &\Leftrightarrow \quad B(u, v) = F(v) \quad \text{for all } v \in H. \end{aligned}$$

[†] Since $(\mathrm{Ran}(A)^\perp)^\perp = \overline{\mathrm{Ran}(A)} \neq H \Rightarrow (\mathrm{Ran}(A))^\perp \neq \{0\}$.

6) Uniqueness. Suppose that

$$B(u_1, w) = F(w) \quad \text{for all } w \in H,$$
$$B(u_2, w) = F(w) \quad \text{for all } w \in H.$$

Let $u := u_1 - u_2$. By linearity,

$$B(u, w) = 0 \quad \text{for all } w \in H.$$

The coercivity of B implies that

$$\|u\|^2 \leq \frac{1}{c} B(u, u) = 0$$

so that $u = 0$, i.e., $u_1 = u_2$.

7) *A priori bound*. If $B(u, w) = F(w)$ for all $w \in H$, then by setting $w = u$ we obtain

$$\|u\|^2 \leq \frac{1}{c} B(u, u) = \frac{1}{c} F(u) \leq \frac{1}{c} \|F\| \|u\|$$

which immediately yields

$$\|u\| \leq \frac{1}{c} \|F\|.$$

□

Density argument

Lemma

Let X, Y be Banach spaces and let $Z \subset X$ be a dense subspace. If $T: Z \to Y$ is a linear mapping such that

$$\|Tx\|_Y \leq C\|x\|_X, \quad x \in Z, \tag{6}$$

then there exists a unique extension $\widetilde{T}: X \to Y$ with $\widetilde{T}|_Z = T$ and

$$\|\widetilde{T}x\|_Y \leq C\|x\|_X, \quad x \in X. \tag{7}$$

Moreover, if (6) holds with equality, then so does (7).

Proof. Let $x \in X$. Because $Z \subset X$ is dense, there exists a sequence $(z_k)_{k=1}^{\infty} \subset Z$ s.t. $\|z_k - x\|_X \xrightarrow{k\to\infty} 0$. Let $\varepsilon > 0$. Since $(z_k)_{k=1}^{\infty}$ is a Cauchy sequence, there exists $N \in \mathbb{N}$ s.t.

$$m, n \geq N \quad \Rightarrow \quad \|z_m - z_n\|_X < \frac{\varepsilon}{C}.$$

Then there holds

$$\|Tz_m - Tz_n\|_Y = \|T(z_m - z_n)\|_Y \leq C\|z_m - z_n\|_X < \varepsilon,$$

which means that $(Tz_k)_{k=1}^{\infty}$ is a Cauchy sequence in Y. Since Y is complete, there exists $y := \lim_{k\to\infty} Tz_k$. Hence we may define $\widetilde{T}: X \to Y$ by setting $\widetilde{T}(x) = y$.

We begin by showing that $\widetilde{T}$ is well-defined. Let $(z_k)_{k=1}^{\infty}$, $(\widetilde{z}_k)_{k=1}^{\infty}$ be two sequences in Z s.t. $z_k, \widetilde{z}_k \xrightarrow{k\to\infty} x$ in X. Then

$$\|Tz_k - T\widetilde{z}_k\|_Y = \|T(z_k - \widetilde{z}_k)\|_Y \leq C\|z_k - \widetilde{z}_k\| \leq C\|z_k - x\| + C\|\widetilde{z}_k - x\| \overset{k\to\infty}{\to} 0.$$

Recalling that $\widetilde{T}(x) := \lim_{k\to\infty} Tz_k$, we obtain

$$\|T\widetilde{z}_k - \widetilde{T}(x)\| \leq \|T\widetilde{z}_k - Tz_k\| + \|Tz_k - \widetilde{T}(x)\| \overset{k\to\infty}{\to} 0,$$

showing that $\widetilde{T}$ is well-defined.
Next we show that $\widetilde{T}$ is linear. Let $x, \widetilde{x} \in X$ and $a, b \in \mathbb{R}$. Let $Z \ni z_k \xrightarrow{k\to\infty} x$ and $Z \ni \widetilde{z}_k \xrightarrow{k\to\infty} \widetilde{x}$. Now $ax + b\widetilde{x} \in X$ and $Z \ni az_k + b\widetilde{z}_k \to ax + b\widetilde{x}$. Thus

$$\widetilde{T}(ax + b\widetilde{x}) = \lim_{k\to\infty} T(az_k + b\widetilde{z}_k) = a \lim_{k\to\infty} Tz_k + b \lim_{k\to\infty} T\widetilde{z}_k = a\widetilde{T}x + b\widetilde{T}x,$$

since the limit is linear.†
Since the norm is continuous,

$$\|\widetilde{T}x\| = \|\lim_{k\to\infty} Tx_k\| = \lim_{k\to\infty} \|Tx_k\| \leq C \lim_{k\to\infty} \|x_k\| = C\|x\|.$$

Finally, $\widetilde{T}|_Z = T$ holds by construction and the uniqueness of the limit $Tz_k \to y$ ensures that there cannot exist another mapping $L\colon X \to Y$ s.t. $L|_Z = T$ and $\|Lx\| \leq C\|x\|$. □

†Let $y := \lim_{k\to\infty} Tz_k$ and $\widetilde{y} := \lim_{k\to\infty} T\widetilde{z}_k$.
Then $\|T(az_k + b\widetilde{z}_k) - ay - b\widetilde{y}\| \leq a\|Tz_k - y\| + b\|T\widetilde{z}_k - \widetilde{y}\| \to 0$.
Hence $\lim_{k\to\infty} T(az_k + b\widetilde{z}_k) = a\lim_{k\to\infty} Tz_k + b\lim_{k\to\infty} T\widetilde{z}_k$.