Functional Analysis

Daniele Zago

November 10, 2021

Contents

Lecture 1: Introduction **1**

1.1 Review of prerequisites 1

1.2 Measure theory 3

Lecture 2: Banach spaces **5**

2.1 L^p spaces 5

2.2 Banach spaces 7

Lecture 3: L^p spaces and bounded linear operators **8**

3.1 Operators between Banach spaces 10

Lecture 1: Introduction

2021-10-19

The course contents will be mostly on measure theory with some basic functional analysis notions about Banach and Hilbert spaces.

1.1 Review of prerequisites

Def. (Eigenvalue)
Let $A \in M^{n\times n}(\mathbb{R})$ a $n \times n$ symmetric matrix of real numbers, such that $A^\top = A$. An *eigenvalue* $\lambda \in \mathbb{C}$ of A is a number such that there exists $v \in \mathbb{R}^n$, $v \neq \mathbf{0}$ we have

$$Av = \lambda v.$$

Remark

- v is called an *eigenvector* associated to λ.
- Eigenvalues are the solution in λ to the following equation, called *characteristic polynomial*

$$p(\lambda) = \det\big(A - \lambda I_n\big) = \mathbf{0}.$$

  The equation has at most n solution when counting multiplicities.

Prop. 1
0 is an eigenvalue of A $\iff \det A = 0$.

Since by the Binet formula we have $\det(AB) = \det A \det B$, then we find that A is invertible $\iff$ $\det A \neq 0$:

$$\det I = 1 = \det A \det A^{-1}.$$

In general, we could define the eigenvalues as those numbers λ such that $A - \lambda I$ is <u>not</u> invertible. This definition is useful when considering linear operators defined on Hilbert spaces instead of classical spaces such as $\mathbb{R}^n$.

In fact, a matrix $A \in M^{n\times n}(\mathbb{R})$ is associated to linear functions $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $v \mapsto Av$. This function is linear since

$$\lambda v + \mu w \implies A(\lambda v + \mu w) = \lambda Av + \mu Aw.$$

Def. (Invertible matrix)
A is *invertible* with inverse A^{-1} if the associated linear function T admits an inverse, i.e.

$$\begin{aligned} T : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ v &\longmapsto Av \end{aligned} \iff \begin{aligned} T^{-1} : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ v &\longmapsto A^{-1}v \end{aligned}$$

Def. (Positive definite matrix)
A matrix A is *positive definite* if A is symmetric and all its eigenvalues $\lambda_1, \ldots, \lambda_k$ are strictly positive.

Remark

A linear operator T will accordingly be called positive definite when the eigenvalues of the associated matrix A are positive definite.

In particular, a positive definite matrix is such that there exists a $c > 0$ for which

$$v^\top A v \geq c|v|^2 \quad \forall v \in \mathbb{R}^n,$$

and in particular $c = \min\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$.

Def. (Random variable)
A random variable is a measurable function on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined by

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto X(\omega) \end{aligned}$$

Most of the time, instead of looking at the random variable itself we consider the law of X, which is a function μ_X defined on all Borel sets of Ω,

$$\mu_X(A) = \mathbb{P}\big(\omega \in \Omega : X(\omega) \in A\big).$$

We characterize the random variables by the measure μ_X that it is induced on the real numbers. The gaussian distribution is the distribution of a r.v. with mean 0 and variance 1 such that

$$\mu_X(A) = \frac{1}{C} \int_A e^{-|x|/2} dx$$

For convergence since sums of integers are very hard to compute we can instead look at the corresponding function integral:

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \longleftrightarrow \int_1^\infty \frac{1}{x^\alpha} dx.$$

The sum of the areas of all the rectangles corresponding to the integers $\big(n, f(n)\big)$ is such that

$$\text{Area} = f(1) + f(2) + \ldots + f(n) + \ldots,$$

and for sure the sum of all rectangles is controlled by the area of the curve $f(x)$,

$$\sum_{i=1}^{\infty} \frac{1}{n^\alpha} - f(1) \leq \int_1^{+\infty} \frac{1}{x^\alpha} dx.$$

Therefore, if the integral is finite we automatically know that the series is convergent, and we also have an upper bound on the value of the series. However, if the integral is infinite we can take the upper summation of rectangles and control the integral by the upper summation S_n

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} - 1 \leq \int_1^{\infty} \frac{1}{x^\alpha} dx \leq \sum_{n=1}^{+\infty} \frac{1}{n^\alpha}.$$

If instead we want to control the summation we can majorize and minorize it with respect to the integral by considering

$$\int_1^{+\infty} \frac{1}{x^\alpha} dx \leq \sum_{n=1}^{\infty} \frac{1}{n^\alpha} - 1 \leq \int_1^{+\infty} \frac{1}{x^\alpha} dx + 1.$$

1.2 Measure theory

Let X be a set and Σ a collection of subsets of X. Σ is called a σ-algebra if it is closed wrt to complement and countable unions:

i $A \in \Sigma \implies X \setminus A \in \Sigma$.

ii $A_1, A_2, \ldots \in \Sigma \implies \bigcup_{n \in \mathbb{N}} A_i \in \Sigma$.

The smallest σ-algebra is $\{\emptyset, X\}$, whereas the biggest is the family of all subsets of X, $\mathscr{P}(X)$. In principle we need to consider σ-algebras which are smaller than $\mathscr{P}(X)$, since coherent measures are difficult to define on it.

If $C \subseteq \mathscr{P}(X)$ is a family of subsets of X, then the *σ-algebra generated by C* is the smallest σ-algebra which contains C,

$$\sigma(C) =$$

If we consider $X = \mathbb{R}$, then the *Borelian σ-algebra $\mathscr{B}$* is the σ-algebra generated by all open intervals (a, b), $a, b \in \mathbb{R}$:

$$C = \{(a, b) : a < b, a, b \in \mathbb{R}\}.$$

$$\mathscr{B} = \sigma(C).$$

In any set X where we can define a topology we can also define its associated Borel σ-algebra.

Under the above σ-algebra, all intervals of the form

$$[a, b], [a, b), (a, b], (a, \infty), [a, \infty), (-\infty, b], (-\infty, b)$$

are all contained in B. This is easy to prove by considering countable unions and complements

$$(a, +\infty) = \bigcup_{i=1}^{\infty} (a, a+i) \in \mathscr{B}$$

$$\mathbb{R} \setminus (a, +\infty) = (-\infty, a] \in \mathscr{B}.$$

Moreover, $\mathscr{B}$ is equivalently

$$\sigma\big(\{(a,b) : \dots\}\big)$$

$$\sigma\big(\{(a,b] : \dots\}\big)$$

$$\sigma\big(\{[a,b] : \dots\}\big)$$

$$\sigma\big(\{[a,b) : \dots\}\big)$$

We can construct some subsets A of $\mathbb{R}$ such that $A \notin \mathscr{B}$, for example the Vitali set. The more general proof that $\mathscr{B} \subsetneq \mathscr{P}(\mathbb{R})$ takes into account the cardinality of the two sets, which can be shown to be different.

(X, Σ) then $\mu : \Sigma \to [0, +\infty]$ is a measure if

- $\mu(\emptyset) = 0$
- $\forall A_i \in \Sigma$, $A_i \cap A_j = \emptyset$, then $\mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \sum_i \mu(A_i)$.

Consequence

We can prove some properties

a) μ is monotone, i.e. $A \subseteq B \implies \mu(A) \leq \mu(B)$

b) μ is "continuous", i.e. if $A_i \subseteq A_{i+1} \subseteq A_{i+2}$ then

$$\mu\big(\bigcup_i A_i\big) = \lim_{i \to \infty} \mu(A_i).$$

c) If $A_{i+1} \subseteq A_i$ at each i we have

$$\mu\big(\bigcap_i A_i\big) = \lim_{i \to \infty} \mu(A_i).$$

Def. (Finiteness)
μ is said to be *finite* if $\mu(X) < \infty$, and σ-*finite* if there is a sequence of $A_i \in \Sigma$ such that $X = \bigcup_{i=1}^{\infty} A_i$ with $\mu(A_i) < \infty$.

Lecture 2: Banach spaces

2021-11-09

We start by defining some useful spaces in functional analysis, namely L^p and M^p spaces.

2.1 L^p spaces

Def. (L^p spaces)
For $p \in [1, \infty)$ we define the following vectorial space

$$L^p(A) = \Big\{ f : \mathbb{R}^n \to \mathbb{R} : f \text{ is measurable and } \int_A |f(x)|^p \, \mathrm{d}x < \infty \Big\}.$$

For $p = \infty$, we define

$$L^\infty(A) = \Big\{ f : R^n \to \mathbb{R} : f \text{ is measurable and } |f(x)| \leq c \text{ for almost every } x \in A \Big\}.$$

Def. (M^p spaces)
When we consider $A = \Omega$, we have the space of random variables with finite p^{th} moment,

$$M^p = \Big\{ X : \Omega \to \mathbb{R} : X \text{ is measurable and } \mathbb{E}[|X|^p] < \infty \Big\},$$

with an analogous definition for $p = \infty$,

$$M^\infty = \Big\{ X : \Omega \to \mathbb{R} : X \text{ is measurable and } |X(\omega)| \leq c \text{ for almost every } \omega \in \Omega \Big\}.$$

Remark

› With these definitions, it's immediate to define ***convergence in p-space*** as

$$f_n \xrightarrow{L^p} f \iff \int_{\mathbb{R}} |f_n(x) - f(x)|^p \, \mathrm{d}x \xrightarrow{n \to \infty} 0,$$

and analogously, ***convergence in p-mean*** for a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ as

$$X_n \xrightarrow{L^p} X \iff \mathbb{E}[|X_n - X|^p] \xrightarrow{n \to \infty} 0.$$

These spaces are particularly important for functional analysis since they are examples of Banach spaces, whose structure we are going to study more generally.

We now state some inequalities which are useful for studying L^p and M^p spaces.

Def. (Conjugate exponent)
Let $p > 1$, then the ***conjugate exponent*** of p is $q > 1$ such that

$$\frac{1}{p} + \frac{1}{q} = 1 \iff q = \frac{p}{p-1}.$$

Moreover, if $p = 1$ we say that its conjugate exponent is $q = +\infty$ and vice versa.

Thm. 1 (Young's inequality)
Let p, q be conjugate exponents, then for all $x, y > 0$ we have that

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Remark

- This property is a generalization of the classic identity $(a-b)^2 > 0 \implies ab < \frac{a^2}{2} + \frac{b^2}{2}$.

Thm. 2 (Hölder's inequality)
Let $p, q \in [1, \infty]$ be conjugate exponents, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, then for all measurable functions f on $O \subseteq \mathbb{R}^n$ we have

$$\|fg\|_1 \leq \|f\|_p \|g\|_q,$$

which if expanded becomes

$$\int_O |f(x)g(x)| \, \mathrm{d}x \leq \Big(\int_O |f(x)|^p \Big)^{\frac{1}{p}} \Big(\int_O |g(x)|^q \Big)^{\frac{1}{q}}.$$

If moreover $f, g \in L^p(O)$ then $fg \in L^1(O)$ and this becomes an equality $\iff$ $|f|^p$ and $|g|^q$ are linearly dependent in $L^1(O)$.

Corollary 1 (Minkowski's inequality)
Let $f, g \in L^p(O)$, then we have that $\| \cdot \|_p$ satisfies the triangle inequality, i.e.

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p,$$

which if expanded becomes

$$\Big(\int_O |f(x) + g(x)|^p \Big)^{\frac{1}{p}} \leq \Big(\int_O |f(x)|^p \Big)^{\frac{1}{p}} + \Big(\int_O |g(x)|^p \Big)^{\frac{1}{p}}.$$

2.2 Banach spaces

Let $(X, \|\cdot\|)$ be a normed vectorial space, then it is possible to define the ***distance induced by the norm*** as the function

$$d(x, y) = \|x - y\|, \quad x, y \in X.$$

This distance induces a topology (i.e. a notion of open and closed sets) by firstly defining the ***balls of radius*** $\boldsymbol{r}$ ***centered in*** $\boldsymbol{x_0}$ as

$$B(x_0, r) = \{x \in X : d(x_0, x) < r\},$$

and then defining a set $A \subseteq X$ as ***open*** if for all $x \in A$ there exists $r > 0$ such that $B(x, r) \subset A$.

Def. (Banach space)
Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence, i.e. $\lim_{n,m\to\infty} \|x_n - x_m\| = 0$. If all such sequences have limit in X, then X is called a ***Banach space***.

Remark

› This property is called ***completeness***, and the definition states that "a Banach space is a complete normed vector space".

Thm. 3 ($\boldsymbol{L^p}$ are Banach spaces)
The spaces $L^p(\mathbb{R})$, $p \in [1, \infty]$ are Banach spaces w.r. to the distance induced by the norm

$$\|f\|_p = \left(\int_{\mathbb{R}} |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}, \quad p \in [1, \infty),$$

and

$$\|f\|_\infty = \sup_x |f(x)|.$$

Remark

› The theorem can be stated in terms of M^p being Banach spaces, when endowed with the norms

$$\|X\|_p = \mathbb{E}[|X|^p]^{\frac{1}{p}}, \quad p \in [1, \infty),$$

$$\|X\|_\infty = \sup_\omega |X(\omega)|, \quad p = \infty.$$

Proof.

□

Lecture 3: L^p spaces and bounded linear operators

2021-11-10

Recall that

$$M^p = \{X : \Omega \to \mathbb{R}, \text{ such that } \mathbb{E}[|X^p|] < \infty\}$$

$$L^p = \{f : \mathbb{R} \to \mathbb{R}, \text{ such that } \int_{\mathbb{R}} |f(x)|^p \, \mathrm{d}x < \infty\}$$

These are actually the same spaces if you consider $M^p = L^p(\Omega)$, indeed you can see that

$$\mathbb{E}[|X|^p] = \int_{\Omega} |x|^p \, \mathrm{d}\mathbb{P}(\omega)$$

Prop. 2 (L^p spaces inclusion)
$M^1 \supseteq M^2 \ldots$, and in general $M^n \subseteq M^k$ if $1 \le k \le n$. In general, if X is a random variable such that $\mathbb{E}[|X|^n] < \infty$, then $\mathbb{E}[|X|^k] < \infty$ for all $k \le n$.

Proof.
By Jensen's inequality, if $f : \mathbb{R} \longrightarrow \mathbb{R}$ *convex* ($\implies$ meas. and cont.) and X is a random variable, then the random variable $f(X)$ is such that

$$\mathbb{E}[f(X)] \ge f\big(\mathbb{E}[X]\big).$$

We fix $n \ge k \ge 1$ and $X \in M^n \implies \mathbb{E}[|X|^n] < \infty$. We want to prove that also $X \in M^k$: we therefore fix $f(x) = |x|^{\frac{n}{k}}$, and we see that

$$\frac{n}{k} \ge 1 \implies f \text{ is a convex function.}$$

Applying Jensen's inequality,

$$\mathbb{E}[f(Y)] \ge f\big(\mathbb{E}[Y]\big)$$

we use the functions f and r.v. $Y = |X|^k$ to see that

$$\begin{cases} f(Y) = \big(|X|^k\big)^{\frac{n}{k}} = |X|^n \\ \mathbb{E}[f(Y)] = \mathbb{E}[|X|^n] < \infty \quad \text{by assumption} \end{cases}$$

and so we obtain $\infty > \mathbb{E}[|X|^n] \ge \mathbb{E}[|X|^k]^{\frac{n}{k}}$, by taking square roots we have

$$\infty > \mathbb{E}[|X|^n]^{\frac{1}{n}} \ge \mathbb{E}[|X|^k]^{\frac{1}{k}},$$

from which we conclude that $X \in M^n \implies X \in M^k$ for all $1 \le k \le n$ and moreover $\|X\|_k \le \|X\|_n$.

Finally, we observe that this implies a relationship w.r. to convergence for all $1 \le k \le n$:

$$X_n \xrightarrow{M_n} X \implies X_n \xrightarrow{M_k} X \quad \text{for all } 1 \le k \le n.$$

$\square$

Remark

This proof is analogous when considering more general L^p spaces, since M^p is a particular space if $A = \Omega$. Another proof can be obtained by applying Hölder's inequality:

Proof.
Consider $f \in L^n(A)$, we want to prove that $f \in L^k(A)$ for $1 \leq k \leq n$. Since $f \in L^n(A)$,

$$\int_{\mathbb{R}} |f(x)|^n \mathbb{1}_A(x)\, \mathrm{d}x < \infty,$$

which means that $f(x)\mathbb{1}_A(x) \in L^n(\mathbb{R})$. Moreover, its' trivial to see that $\mathbb{1}_A(x) \in L^p(\mathbb{R})$ for all p if $\mathcal{L}(A) < \infty$, since

$$\int_{\mathbb{R}} |\mathbb{1}_A(x)|^p\, \mathrm{d}x = \mathcal{L}(A).$$

Now, let $q = \frac{n}{n-1}$ be the conjugate exponent of n, we have that

$$\begin{cases} f\mathbb{1}_A \in L^n(\mathbb{R}) \\ \mathbb{1}_A \in L^{\frac{n}{n-1}}(\mathbb{R}) \end{cases}$$

then by Hölder's inequality, $(f\mathbb{1}_A) \cdot \mathbb{1}_A \in L^1(\mathbb{R})$ and we see that

$$\begin{aligned} \|f\|_{L^1(A)} = \int_{\mathbb{R}} |f(x)|\mathbb{1}_A(x)\, \mathrm{d}x &\leq \|f\|_{L^n(A)} \Big(\int_{\mathbb{R}} \mathbb{1}_A(x)^{\frac{n}{n-1}}\, \mathrm{d}x \Big)^{\frac{n-1}{n}} \\ &= \|f\|_{L^n(A)} \cdot \big(\mathcal{L}(A)\big)^{\frac{n-1}{n}} \end{aligned}$$

Therefore, we conclude that $L^n(A) \subseteq L^1(A)$ for any $n \geq 1$. With a similar argument, by accurately choosing the exponent k we also have that $f \in L^k(A)$ for all $1 \leq k \leq n$,

$$f \in L^k(A) \iff |f|^k \in L^1(A).$$

We have that

$$|f|^k \in L^{\frac{n}{k}}(A) \quad \frac{n}{k} > 1$$

since $\big(|f|^k\big)^{\frac{n}{k}} = |f|^n$ and by choosing the conjugate exponent of $\frac{n}{k}$,

$$q = \frac{\frac{n}{k}}{\frac{n}{k} - 1}$$

$\square$

Remark

- Recall that if $L(A) = \mathbb{R}$ then this fact is not true, since we proved that $L^2(\mathbb{R}) \not\subseteq L^1(\mathbb{R})$.
- This holds for random variables since they induce a bounded measure, $\mathbb{P}(\Omega) = 1$.
- $L^\infty(A) \subseteq \bigcap_{k\geq 1} L^k(A)$ but it is not equal to it.

We also have a way of computing the L^k norm in terms of the L^n norm:

$$\int_A |f|^k \,\mathrm{d}x \leq \|f\|_n^k \mathcal{L}(A)^{\frac{n-k}{n}} \implies \|f\|_k \leq \|f\|_n \cdot \mathcal{L}(A)^{\frac{n-k}{nk}}.$$

3.1 Operators between Banach spaces

We are working with spaces whose elements are functions (or random variables). Let X, Y be Banach spaces, e.g. $X = Y = L^p$ or M^p, $X = L^p$ or M^p and $Y = \mathbb{R}$.

Let $T : X \longrightarrow Y$ be a linear operator, i.e. an operator that maintains the vectorial structure,

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) \quad \text{for all } x_1, x_2 \in X \text{ and } \alpha_1, \alpha_2 \in \mathbb{R}.$$

Def. (Continuity)
We say that $T : X \to Y$ is a ***continuous*** operator if for every sequence $x_n \in X$ such that $x_n \to x \in X$, then $Tx_n \to Tx$, i.e. converging sequences are mapped in converging sequences.

Def. (Boundedness)
We say that $T : X \to Y$ is ***bounded*** if there exists a constant $C > 0$ such that

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

Remark

We don't require the image to be bounded, since the condition turns out to be too strong for linear operators between Banach spaces.

Prop. 3 (Continuity implies boundedness)
If X, Y are Banach spaces and $T : X \longrightarrow Y$ is a linear operator, then T is continuous if and only if T is bounded.

Remark

Because of this theorem, we will always talk about bounded operators instead of continuous operators, since they are the same.

Proof.
$\boxed{\text{Bounded} \implies \text{continuous}}$: Let $x_n \in X$ such that $x_n \xrightarrow{n\to\infty} x \in X$, which by definition means that $\|x_n - x\|_X \xrightarrow{n\to\infty} 0$. From this, we want to prove that $Tx_n \xrightarrow{n\to\infty} Tx$.

Since the operator is bounded, there exists a C such that for all $x \in X$

$$\|Tx\|_Y \leq C\|x\|_X,$$

and applying this to $y = x_n - x \in X$ we have

$$0 \leq \|T(x_n - x)\|_Y \leq C\|x_n - x\|_X \xrightarrow{n\to\infty} 0,$$

and by linearity we have that $\|Tx_n - Tx\|_Y \to 0$.

$\boxed{\text{Continuous} \implies \text{bounded}}$: We actually prove another fact which implies this property, i.e. T not bounded $\implies$ T not continuous, by constructing a sequence which is not converging under the map T. Assume that T is not bounded, then for any $C > 0$ we can always find at least a point $x \in X$ such that

$$\|Tx\|_Y > C\|x\|_X,$$

In particular, let $C = n$, then we always have that there exists $x_n \in X$ such that

$$\|Tx_n\|_Y > n\|x_n\|_X \quad \text{for } \|x_n\|_X \neq 0$$

Consider now the sequence of points $y_n = \frac{x_n}{n\|x_n\|_X}$, for which we have that $y_n \in X$ since X is a vectorial space. Then,

$$\|y_n\|_X = \left\| \frac{x_n}{n\|x_n\|_X} \right\| \overset{\text{norm.}}{=} \frac{1}{n}.$$

In particular, this means that $\|y_n - 0\|_X = \frac{1}{n} \xrightarrow{n\to\infty} 0$ and if T is continuous, then we should see that $Ty_n \to T0 = 0$ in Y since it is a linear operator. Now,

$$Ty_n = T\left(\frac{x_n}{n\|x_n\|_X}\right) \overset{\text{lin.}}{=} \frac{1}{n\|x_n\|_X} T(x_n),$$

and

$$\|Ty_n\|_Y = \left\| \frac{1}{n\|x_n\|_X} T(x_n) \right\|_Y \overset{\text{lin.}}{=} \frac{1}{n\|x_n\|_X} \|Tx_n\|_Y \overset{\text{Hp.}}{>} \frac{1}{n\|x_n\|_X} n\|x_n\|_X = 1,$$

therefore Ty_n does not converge to 0 in Y and therefore we prove the desired property.

$\square$

Thm. 4 (Space of linear operators)
Let X, Y be Banach spaces, then we have that

$$\mathcal{B}(X,Y) = \{T : X \to Y \text{ linear bounded operators}\}$$

is a Banach space with norm given by

$$\|T\| = \sup_{\|X\|_X \leq 1} \|Tx\| \overset{lin.}{=} \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Proof.
$\mathcal{B}(X,Y)$ is naturally a vectorial space, since $\alpha T+\beta S$ is a linear operator defined by

$$(\alpha T+\beta S)(x) := \alpha T(x)+\beta S(x).$$

Moreover, if T,S are bounded then $\alpha T+\beta S$ is also bounded by

$$\begin{aligned}\|\alpha Tx+\beta Sx\|_Y \leq \alpha\|Tx\|_Y+\beta\|Sx\|_Y &\leq \underbrace{|\alpha|\cdot C\|x\|_X}_{T\text{ bounded}}+\underbrace{|\beta|\cdot D\|x\|_X}_{S\text{ bounded}}\\ &= \big(|\alpha|C+|\beta|D\big)\|x\|_X.\end{aligned}$$

Moreover, $\|T\|$ defined above is a norm since it satisfies the three properties

1. $\|T\|\geq 0$ for all T.
2. $\|\alpha T\| = \sup\limits_{\|X\|_X\leq 1}\|\alpha Tx\| = |\alpha|\cdot\sup\limits_{\|x\|\leq 1}\|Tx\| = |\alpha|\cdot\|T\|$.
3. $\|T+S\| = \sup\limits_{\|x\|\leq 1}\|Tx+Sx\|_Y \leq \sup\limits_{\|x\|\leq 1}\|Tx\|_Y+\sup\limits_{\|x\|\leq 1}\|Sx\|_Y$.

We are not going to prove that $\mathcal{B}(X,Y)$ is a complete space w.r. to convergence induced by the norm, since it is a bit complex.

$\square$

Another important result for linear bounded operators is the following, which is a theorem that can be used to prove that $\mathcal{B}(X,Y)$ is a Banach space.

Thm. 5 (Banach-Steinhaus theorem)
If T_n is a sequence of bounded linear operators from X to Y and for every $x\in X$ there exists $\lim\limits_{n\to\infty}T_nx$ in Y then the operator defined by

$$Tx := \lim_{n\to\infty}T_nx$$

is a bounded linear operator.

Example (Linear bounded operators)
Consider for a fixed $X\in M^p$ and $q=\frac{p}{p-1}$ the operator defined by

$$\begin{aligned}T: M^q &\longrightarrow \mathbb{R}\\ Y &\longmapsto \mathbb{E}[X\cdot Y],\end{aligned}$$

then this is a linear bounded operator. Linearity is immediate since the $\mathbb{E}[\cdot]$ operator is linear, whereas for boundedness we have to show that $\exists C>0$ such that $\|Tx\|_{\mathcal{Y}}\leq C\|x\|_{\mathcal{X}}$ where $\mathcal{X}=M^q$ and $\mathcal{Y}=\mathbb{R}$. For all $Y\in M^q$ we therefore want to check that

$$|\mathbb{E}[X\cdot Y]| \overset{?}{\leq} C\cdot\mathbb{E}[\|Y\|^q]^{\frac{1}{q}}.$$

We can do so by applying Hölder's inequality, which allows us to write

$$|\mathbb{E}[X \cdot Y]| \overset{\text{Jens.}}{\leq} \mathbb{E}[|Z \cdot Y|] \overset{\text{Höld}}{\leq} \underbrace{\mathbb{E}[|X|^p]^{\frac{1}{p}}}_{C} \cdot \mathbb{E}[|Y|^q]^{\frac{1}{q}},$$

and the operator is bounded the constant $C = \mathbb{E}[|X|^p]^{\frac{1}{p}}$ (recall that X is fixed).

Moreover, we can actually prove (exercise) that the norm of T is equal to

$$\|T\| = \mathbb{E}[|X|^p]^{\frac{1}{p}}$$

Example

If we consider $T : \mathbb{R}^n \to \mathbb{R}^m$, then the set of bounded linear operators is the set of operators defined by the $M_{m\times n}(\mathbb{R})$ matrices,

$$x \mapsto Tx = Ax, \quad A \in M_{m\times n}(\mathbb{R}).$$

Moreover, the norm $\|T\|$ is connected to the norm of the matrix A.

Remark

From the example above, the space $\mathcal{B}(X, Y)$ is an infinite-dimensional generalization of the space of matrices.

$$\text{Space of matrices} \quad \xrightarrow{\text{Infinite dim.}} \quad \mathcal{B}(X, Y),$$

and the results we are going to prove for infinite-dimensional linear operators are analogues to those of standard finite-dimensional vector spaces.