Functional analysis

Lecture 8: Hilbert spaces and the projection theorem;
Riesz representation theorem for functionals on Hilbert spaces

5 Hilbert spaces

In this section, we deal with normed spaces whose norm is given by an inner product, that is, a sesquilinear form. We start with some basic definitions.

Definition 5.1. Let X be a vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. By an *inner product* (or *scalar product*) we mean any function of two variables $(\cdot\,,\cdot)\colon X \times X \to \mathbb{K}$ satisfying the following axioms:

(i) for every $y \in X$, the map $X \ni x \mapsto (x, y)$ is linear;
(ii) $(x, y) = \overline{(y, x)}$ for all $x, y \in X$;
(iii) $(x, x) \geq 0$ for every $x \in X$;
(iv) $(x, x) = 0$ if and only if $x = 0$.

Proposition 5.2 (Cauchy–Schwarz inequality). *Let $(\cdot\,,\cdot)$ be an inner product on a linear space X. Then:*

(a) $|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}$ *for all $x, y \in X$;*
(b) *the formula $\|x\| = \sqrt{(x, x)}$ defines a norm on X.*

Proof. (a) Observe that for every $\lambda \in \mathbb{K}$ we have

$$\begin{aligned} 0 \leq (x + \lambda y, x + \lambda y) &= (x, x) + \lambda(y, x) + \overline{\lambda}(x, y) + |\lambda|^2(y, y) \\ &= (x, x) + \lambda\overline{(x, y)} + \overline{\lambda}(x, y) + |\lambda|^2(y, y). \end{aligned}$$

Hence, taking

$$\lambda = -\frac{(x, y)}{(y, y)}$$

we obtain the desired inequality.

(b) The first two axioms of norm are obvious. To see that the triangle inequality is satisfied, note that for all $x, y \in X$ we have

$$\begin{aligned} \|x + y\|^2 &= (x, x) + (x, y) + \overline{(x, y)} + (y, y) \\ &= \|x\|^2 + 2\mathrm{Re}\,(x, y) + \|y\|^2 \\ &\leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Remark. Notice that in view of the Cauchy–Schwarz inequality, the map $(\cdot\,,\cdot)$ is jointly continuous. Indeed, for any $x, x', y, y' \in X$ we have

$$\begin{aligned} |(x, y) - (x', y')| &\leq |(x, y) - (x', y)| + |(x', y) - (x', y')| \\ &= |(x - x', y)| + |x', y - y')| \leq \|x - x'\|\|y\| + \|x'\|\|y - y'\|. \end{aligned}$$

In particular, for any $y \in X$ the map $x \mapsto (x, y)$ is a continuous linear functional on X.

Definition 5.3. A normed space $(X, \|\cdot\|)$ equipped with an inner product $(\cdot\,,\cdot)$ which generates the norm $\|\cdot\|$, i.e. $\|x\| = \sqrt{(x,x)}$ for $x \in X$, is called an *inner product space*. If X is a complete inner product space, then it is called a *Hilbert space*.

We will use the following notation (x, y are elements of an inner product space X and $F \subseteq X$ is any set):

- $\mathcal{H}$, $\mathcal{K}$ will typically stand for Hilbert spaces,
- $x \perp y$ if $(x, y) = 0$, and such elements are called *orthogonal*,
- $x \perp F$ if $x \perp y$ for every $y \in F$,
- $F^{\perp}$ is the set of all elements that are orthogonal to F, i.e. $\{x \in X : x \perp F\}$; in particular, if $M \subseteq X$ is a subspace, then $M^{\perp}$ is called an *orthogonal complement* of M.

Observe that since

$$F^{\perp} = \bigcap_{y \in F} \{x \in X : x \perp y\} = \bigcap_{y \in F} \ker\,(\cdot\,, y),$$

the set $F^{\perp}$ is always closed. Plainly, we have $F \cap F^{\perp} = \{0\}$ hence, in particular, every closed subspace M of an inner product space X is complemented and $M \oplus M^{\perp} = X$ (the orthogonal complement $M^{\perp}$ is always a closed subspace).

Let X be an inner product space over $\mathbb{K}$ and $x, y \in X$. The following elementary identities are quite useful:

- $\|x\|^2 + \|y\|^2 = \|x + y\|^2$ provided that $x \perp y$ (the Pythagorean theorem),
- $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (the *parallelogram law*),
- $(x, y) = \dfrac{1}{4}\big(\|x + y\|^2 - \|x - y\|^2\big)$ if $\mathbb{K} = \mathbb{R}$,
- $(x, y) = \dfrac{1}{4}\big(\|x + y\|^2 - \|x - y\|^2 + \mathrm{i}\|x + \mathrm{i}y\|^2 - \mathrm{i}\|x - \mathrm{i}y\|^2\big)$ if $\mathbb{K} = \mathbb{C}$.

In fact, there is a Jordan–von Neumann theorem which says that the parallelogram law characterizes norms which come from an inner product; see **Problem 5.3**). Consequently, in order to verify that a given norm comes from an inner product, it is enough to consider only 2-dimensional subspaces. The last two identities are called *polarization formulas* and, as we see, they express the inner product in terms of its symmetrization; see **Problem 5.2**.

Example. Finite-dimensional spaces equipped with the Euclidean norm, that is ℓ_2^n, are canonical examples of Hilbert spaces, where the inner product is given by

$$(\xi, \eta) = \sum_{j=1}^{n} \xi_j \overline{\eta_j} \quad \text{for } \xi = (\xi_j)_{j=1}^n,\ \eta = (\eta_j)_{j=1}^n \in \mathbb{K}^n.$$

More generally, for any measure space $(X, \mathfrak{M}, \mu)$, the Banach space $L_2(\mu)$ is a Hilbert space and its norm comes from the inner product

$$(f, g) = \int_X f(x)\overline{g(x)}\,\mathrm{d}\mu(x) \quad \text{for } f, g \in L_2(\mu).$$

In particular, ℓ_2 and $\ell_2(\Gamma)$ for any index set Γ are Hilbert spaces. In fact, as we will see, it follows from the Riesz–Fischer theorem that every Hilbert space can be canonically identified with one of such spaces.

Proposition 5.4. *Let $\mathcal{H}$ be a Hilbert space. Every nonempty closed convex set $E \subseteq \mathcal{H}$ has a unique element of minimal norm.*

Proof. Define $\delta = \inf\{\|x\| : x \in E\}$. Note that the parallelogram law can be rewritten in the form

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

For any $x, y \in E$, the convexity of E implies that $\frac{1}{2}(x+y) \in E$, thus the above equation yields

$$\|x-y\|^2 \leq 2\|x\|^2 + 2\|y\|^2 - 4\delta^2 \quad \text{for all } x, y \in E. \tag{5.1}$$

If we had two points $x, y \in E$ with $\|x\| = \|y\| = \delta$, then (5.1) would imply $\|x - y\| \leq 0$, i.e. $x = y$ which proves the uniqueness part of our assertion.

For the existence, pick any sequence $(y_n)_{n=1}^\infty \subset E$ with $\|y_n\| \to \delta$. Using (5.1) we obtain

$$\|y_m - y_n\|^2 \leq 2\|y_m\|^2 + 2\|y_n\|^2 - 4\delta^2 \xrightarrow[m,n\to\infty]{} 0$$

which shows that $(y_n)_{n=1}^\infty$ is a Cauchy sequence. Hence, since E is closed, there exists $y_0 = \lim_{n\to\infty} y_n \in E$ and since the norm is continuous, we have $\|y_0\| = \delta$. □

Theorem 5.5 (Projection theorem). *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. There exists a uniquely determined pair of bounded linear operators $P\colon \mathcal{H} \to \mathcal{M}$ and $Q\colon \mathcal{H} \to \mathcal{M}^\perp$ with the following properties:*

(a) *$x = Px + Qx$ for every $x \in \mathcal{H}$;*

(b) *if $x \in \mathcal{M}$, then $Px = x$ and $Qx = 0$;*

(c) *if $x \in \mathcal{M}^\perp$, then $Px = 0$ and $Qx = x$;*

(d) *$\|x - Px\| = \operatorname{dist}(x, \mathcal{M})$.*

Proof. For any $x \in \mathcal{H}$ the coset $x + \mathcal{M}$ is a closed convex subset of $\mathcal{H}$, so we use Proposition 5.4 to define $Qx \in x + \mathcal{M}$ to be the unique element of this coset having minimal norm. Let also $Px = x - Qx$ so that property (a) is satisfied.

Claim. $Qx \in \mathcal{M}^\perp$

Fix any $y \in \mathcal{M}$ with $\|y\| = 1$ and notice that for any $\alpha \in \mathbb{K}$ we have

$$(Qx, Qx) = \|Qx\|^2 \leq \|Qx - \alpha y\|^2 = (Qx - \alpha y, Qx - \alpha y),$$

thus $0 \leq -\alpha(y, Qx) - \overline{\alpha}(Qx, y) + |\alpha|^2$. Taking $\alpha = (Qx, y)$ we see that $0 \leq -|(Qx, y)|^2$ which means that $Qx \perp y$ and proves our Claim.

From the very definition it follows that $Px \in \mathcal{M}$ for every $x \in \mathcal{H}$, so P takes values in $\mathcal{M}$ and Q takes values in $\mathcal{M}^\perp$. Hence, assertions (b) and (c) follow from (a), whereas assertion (d) follows from the definition of Q. It remains to show that P and Q are uniquely determined and linear.

Suppose we have another decomposition $x = x_0 + x_1$, where $x_0 \in \mathcal{M}$ and $x_1 \in \mathcal{M}^\perp$. Then $x_0 - Px = x_1 - Qx$ and the left-hand side belongs to $\mathcal{M}$, while the right-hand one belongs to $\mathcal{M}^\perp$. Since $\mathcal{M} \cap \mathcal{M}^\perp = \{0\}$, we thus have $x_0 = Px$ and $x_1 = Qx$. Linearity is proved in a similar fashion. For any $x, y \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{K}$ we have

$$P(\alpha x + \beta y) - \alpha Px - \beta Py = \alpha Qx + \beta Qy - Q\alpha x + \beta y),$$

and the same reasoning as above shows that both of the above differences are zero. □

The maps P and Q are called *orthogonal projections* onto $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively. Note that these linear operators are indeed projections, i.e. idempotent maps: $P^2 = P$ and $Q^2 = Q$. This can be seen from assertions (b) and (c) and it plainly follows that $\|P\| \geq 1$ and $\|Q\| \geq 1$ unless $\mathcal{M} = \{0\}$ or $\mathcal{M} = \mathcal{H}$. Note also that by the Pythagorean theorem, we have

$$\|x\|^2 = \|Px\|^2 + \|Qx\|^2 \quad \text{for every } x \in \mathcal{H}.$$

Hence, $\|P\| = \|Q\| = 1$ whenever $\mathcal{M}$ is a proper subspace. In particular, P is a norm one projection onto $\mathcal{M}$. This is the first assertion of the following corollary. The second one follows directly from Theorem 5.5.

Corollary 5.6. *Every closed subspace $\mathcal{M}$ of a Hilbert space $\mathcal{H}$ is 1-complemented. Moreover, if $\mathcal{M} \subsetneq \mathcal{H}$, then $\mathcal{M}^\perp \neq \{0\}$.*

Continuous linear functionals on Hilbert spaces have particularly simple form and they are all given by the inner product, as in the remark after Proposition 5.2.

Theorem 5.7 (Riesz representation theorem). *Let $\mathcal{H}$ be a Hilbert space. For every $\varphi \in \mathcal{H}^*$ there exists a unique vector $y \in \mathcal{H}$ such that $\varphi(x) = (x, y)$ for every $x \in \mathcal{H}$.*

Proof. Uniqueness is trivial in view of axiom (iv) from Definition 5.1. Also, if $\varphi = 0$, then take $y = 0$. From now on, assume $\varphi \neq 0$ and define $\mathcal{M} = \ker \varphi$. In view of Corollary 5.6, we have $\mathcal{M}^\perp \neq \{0\}$.

Pick any $z \in \mathcal{M}^\perp$ with $\|z\| = 1$. Since $\mathcal{M}$ is of codimension one and y must be orthogonal to the kernel of φ, we want to find $\alpha \in \mathbb{K}$ for which $y \coloneqq \alpha z$ does the job. Let $x \in \mathcal{H}$ and define

$$x_1 = x - \frac{\varphi(x)}{\varphi(y)} y \quad \text{and} \quad x_2 = \frac{\varphi(x)}{\varphi(y)} y.$$

Plainly, $x = x_1 + x_2$, $x_1 \in \mathcal{M}$ and

$$(x, y) = (x_2, y) = \frac{\varphi(x)}{\varphi(y)} \|y\|^2.$$

In order to have $\varphi(x) = (x, y)$, we must have $\varphi(y) = \|y\|^2$ which means $\alpha\varphi(z) = |\alpha|^2$, equivalently: $\alpha = \overline{\varphi(z)}$. ☐

From the Cauchy–Schwarz inequality it follows easily that $\|\varphi\| = \|y\|$. Hence, for every Hilbert space $\mathcal{H}$ we have the isometric isomorphism

$$\mathcal{H} \cong \mathcal{H}^*$$

which associates any vector $y \in \mathcal{H}$ with a functional $\varphi \in \mathcal{H}^*$ via Theorem 5.7.

Definition 5.8. Let X be an inner product space. A set $\{u_\alpha \colon \alpha \in A\} \subset X$ is called *orthogonal* if $x_\alpha \perp x_\beta$ for all $\alpha, \beta \in A$, $\alpha \neq \beta$. It is called *orthonormal* if it is orthogonal and consists of unit vectors, i.e. $\|x_\alpha\| = 1$ for each $\alpha \in A$. It is called a *maximal orthonormal set* if it is orthonormal and maximal among orthonormal subsets of X with respect to inclusion. Any maximal orthonormal subset of X is called an *orthonormal basis* of X.

Finally, for any $x \in X$ we define scalars

$$\widehat{x}(\alpha) = (x, u_\alpha) \qquad (\alpha \in A)$$

and we call them *Fourier coefficients* of x with respect to the orthonormal set $\{u_\alpha \colon \alpha \in A\}$.

Remark. Observe that if $(u_j)_{j=1}^n$ is an orthonormal set and $x \in \operatorname{lin}\{u_j : 1 \le j \le n\}$, then necessarily

$$x = \sum_{j=1}^{n} \widehat{x}(j) u_j \quad \text{and} \quad \|x\|^2 = \sum_{j=1}^{n} |\widehat{x}(j)|^2.$$

For the first formula, assume that $x = \sum_{i=1}^n \alpha_i u_i$ and observe that by orthonormality we get $\widehat{x}(j) = (\sum_{i=1}^n \alpha_i u_i, u_j) = \alpha_j$. The second formula follows trivially by induction from the Pythagorean theorem.

Proposition 5.9. *Every Hilbert space has an orthonormal basis.*

Proof. Consider the collection $\mathscr{O}$ of all orthonormal sets in a given Hilbert space $\mathcal{H}$, partially ordered by inclusion. Obviously, every chain has an upper bound in $\mathscr{O}$ which is the union of the chain. By the Kuratowski–Zorn lemma, $\mathscr{O}$ contains a maximal element which is thus an orthonormal basis of $\mathcal{H}$. □

The following procedure in finite dimension is well-known from linear algebra.

Theorem 5.10 (Gram–Schmidt orthogonalization). *Let X be an inner product space and $(v_1, v_2, \ldots)$ be a finite or countably infinite sequence of linearly independent vectors in X. There exists an orthonormal sequence $(u_1, u_2, \ldots)$ such that for every $k \in \mathbb{N}$ not exceeding the length of the sequence,* $\operatorname{lin}\{v_1, \ldots, v_k\} = \operatorname{lin}\{u_1, \ldots, u_k\}$.

Proof. For any nonzero vector $u \in X$ and any $v \in X$ we define an *orthogonal projection* of v onto the subspace spanned by u by the formula

$$\operatorname{proj}_u(v) = \frac{(v, u)}{\|u\|^2} u.$$

(It is not difficult to see that this definition coincides with Pv given by Theorem 5.5 for $\mathcal{M} = \operatorname{lin}\{u\}$.)

Define $(u_1, u_2, \ldots)$ recursively by

$$\begin{aligned}
\widetilde{u}_1 &= v_1, & u_1 &= \frac{\widetilde{u}_1}{\|\widetilde{u}_1\|} \\
\widetilde{u}_2 &= v_2 - \operatorname{proj}_{u_1}(v_2), & u_2 &= \frac{\widetilde{u}_2}{\|\widetilde{u}_2\|} \\
&\vdots & &\vdots \\
\widetilde{u}_n &= v_n - \sum_{k=1}^{n-1} \operatorname{proj}_{u_k}(v_n), & u_n &= \frac{\widetilde{u}_n}{\|\widetilde{u}_n\|} \\
&\vdots & &\vdots
\end{aligned}$$

Assuming that $\{u_1, \ldots, u_{n-1}\}$ is orthonormal, we have $\operatorname{proj}_{u_k}(\widetilde{u}_n) = 0$ for all $k < n$, which means that $u_k \perp u_n$. Hence, by induction, all the vectors $u_1, u_2, \ldots$ are pairwise orthogonal. That $\operatorname{lin}\{v_1, \ldots, v_k\} = \operatorname{lin}\{u_1, \ldots, u_k\}$ follows from the fact that $(u_1, \ldots, u_k)$ is the image of $(v_1, \ldots, v_k)$ under a linear map given by an upper triangular matrix with nonzero entries at the diagonal. □