Projections and least squares problems

Nicholas Hu · Last updated on 2025-02-22

Projections

Let H be a Hilbert space and $Y \subseteq H$. The **(orthogonal) projection operator** onto Y is defined for $x \in H$ by

$$\operatorname{proj}_Y(x) := \operatorname*{argmin}_{y \in Y} \frac{1}{2}\|y - x\|^2.$$

> **Hilbert projection theorem (first projection theorem)**
>
> If Y is nonempty, closed, and convex, then $\operatorname{proj}_Y(x)$ is a singleton (so $\operatorname{proj}_Y : H \to Y$ is well-defined).

Proof. Let $(y_n)_{n=1}^{\infty} \subseteq Y$ be such that $d_n := \frac{1}{2}\|y_n - x\|^2 \to d := \inf_{y \in Y} \frac{1}{2}\|y - x\|^2$. By the parallelogram identity,

$$\left\|\frac{y_m + y_n}{2} - x\right\|^2 + \left\|\frac{y_m - y_n}{2}\right\|^2 = 2\left\|\frac{y_m - x}{2}\right\|^2 + 2\left\|\frac{y_n - x}{2}\right\|^2 = d_m + d_n,$$

where $\|\frac{y_m + y_n}{2} - x\|^2 \geq 2d$ by convexity. Taking $m, n \to \infty$ shows that (y_n) is Cauchy and therefore convergent to some $y \in Y$ with $\frac{1}{2}\|y - x\|^2 = d$. Moreover, if $y' \in Y$ is another minimizer, replacing y_m, y_n by y, y' above shows that $y = y'$. ∎

Recall that the **polar cone** of Y is $Y^\circ := \{x \in H : \forall y \in Y\ (\Re(\langle x, y\rangle) \leq 0)\}$ and that the **orthogonal complement** of Y is $Y^\perp := \{x \in H : \forall y \in Y\ (\langle x, y\rangle = 0)\}$; clearly, if Y is a *subspace* of H, then $Y^\circ = Y^\perp$.

> **Characterization of projections (second projection theorem)**
>
> If Y is nonempty, closed, and convex, then $y = \operatorname{proj}_Y(x)$ if and only if $y \in Y$ and $x - y \in (Y - y)^\circ$.

Proof. If $y = \operatorname{proj}_Y(x)$ and $y' \in Y$, then for all $\lambda \in [0, 1]$, we have

$$\|y - x\|^2 \leq \|(1 - \lambda)y + \lambda y' - x\|^2 = \|y - x\|^2 + 2\lambda\Re(\langle y - x, y' - y\rangle) + \lambda^2\|y' - y\|^2,$$

so $\Re(\langle y - x, y' - y\rangle) \geq 0$. Conversely, if $y, y' \in Y$ and $x - y \in (Y - y)^\circ$, then setting $\lambda = 1$ in the inequality above shows that $y = \operatorname{proj}_Y(x)$. ∎

> **Firm nonexpansiveness of the projection operator**
>
> If Y is nonempty, closed, and convex, then
>
> $$\|\operatorname{proj}_Y(x) - \operatorname{proj}_Y(x')\|^2 + \|(I - \operatorname{proj}_Y)(x) + (I - \operatorname{proj}_Y)(x')\|^2 \leq \|x - x'\|^2.$$

Proof. Let $y = \operatorname{proj}_Y(x)$ and $y' = \operatorname{proj}_Y(x')$, and add the inequalities $\Re(\langle x' - y', y - y'\rangle) \leq 0$ and $\Re(\langle x - y, y' - y\rangle) \leq 0$. ∎

In particular, this implies that the projection operator is nonexpansive: $\|\mathrm{proj}_Y(x) - \mathrm{proj}_Y(x')\| \le \|x - x'\|$.

If Y is a *closed subspace* of H, it follows from the above that $y = \mathrm{proj}_Y(x)$ if and only if $y \in Y$ and $x - y \in Y^\perp$, and that $\mathrm{proj}_Y : H \to Y$ is a *linear* operator with $\|\mathrm{proj}_Y\| \le 1$, $\mathrm{im}(\mathrm{proj}_Y) = Y$, and $\ker(\mathrm{proj}_Y) = Y^\perp$. In addition, $\mathrm{proj}_{Y^\perp} = I - \mathrm{proj}_Y$.

Least squares problems

Let H_1 and H_2 be Hilbert spaces and suppose that $A : H_1 \to H_2$ is a continuous linear operator with closed image.[1] The **(linear) least squares problem** is that of finding an $x \in H_1$ that minimizes $\frac{1}{2}\|b - Ax\|^2$ for a given $b \in H_2$, or equivalently, that satisfies $Ax = \mathrm{proj}_{\mathrm{im}(A)}\, b$. Using the fact that $\mathrm{im}(A)^\perp = \ker(A^*)$, we can also write this as the **normal equation** $A^*Ax = A^*b$.

The pseudoinverse

To solve the least squares problem, we observe that $A \restriction_{\ker(A)^\perp} : \ker(A)^\perp \to \mathrm{im}(A)$ is bijective since $Ax = Ax'$ implies that $x - x' \in \ker(A)$ and $y = Ax$ implies that $y = A(x - \mathrm{proj}_{\ker(A)}\, x)$. Thus, the **pseudoinverse** $A^+ : H_2 \to H_1$ of A, defined as

$$A^+ := A \restriction_{\ker(A)^\perp}^{-1} \circ \mathrm{proj}_{\mathrm{im}(A)},$$

is a well-defined continuous linear operator, and by construction $x^* := A^+b$ is a solution to the least squares problem.

This solution need not be unique; however, it is the unique solution of *minimal norm* because $x - x^* \in \ker(A)$ for any solution x, so $\|x\|^2 = \|x - x^*\|^2 + \|x^*\|^2 \ge \|x^*\|^2$ with equality if and only if $x = x^*$.

It is straightforward to verify that:

- $A^+ = A^{-1}$ if A is bijective
- $\mathrm{im}(A^+) = \ker(A)^\perp$, $\ker(A^+) = \mathrm{im}(A)^\perp$
- $AA^+ = \mathrm{proj}_{\mathrm{im}(A)}$, $A^+A = \mathrm{proj}_{\mathrm{im}(A^+)}$ (and in fact, these characterize the pseudoinverse)
- $(A^+)^+ = A$
- $(A^*)^+ = (A^+)^*$
- $A^+ = (A^*A)^+A^* = A^*(AA^*)^+$

In the finite-dimensional case, if $A \in \mathbb{C}^{m\times n}$ has full column rank, then $A^+ = (A^*A)^{-1}A^*$ by the identities above; similarly, if it has full row rank, then $A^+ = A^*(AA^*)^{-1}$. More generally, if $\hat{U}\hat{\Sigma}\hat{V}^*$ is a compact SVD of A (that is, $\hat{\Sigma}$ is $r \times r$, where $r = \mathrm{rank}(A)$), then $A^+ = \hat{V}\hat{\Sigma}^{-1}\hat{U}^*$.

---

1. Note that this implies that A^* also has closed image, so $\mathrm{im}(A)^\perp = \ker(A^*)$ and $\ker(A)^\perp = \overline{\mathrm{im}(A^*)} = \mathrm{im}(A^*)$. ↩