Engr210a Lecture 8: The projection theorem

- Motivation via controllability
- Orthogonal complements
- The projection theorem
- The image and the kernel
- Projection operators
- Minimum-norm approximation
- Dual approximation
- Controllability

Controllability

- Suppose we have the state-space system, with $x(t) \in \mathbb{R}^n$,

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad x(0) = 0$$

- This defines a map $\Upsilon_t : L_2[0,t] \to \mathbb{R}^n$ from input signals u to final state $x(t)$,

$$\Upsilon_t u = \int_0^t e^{A(t-\tau)} B u(\tau)\, d\tau$$

- We know which states are reachable:

$$\text{image}(\Upsilon_t) = \text{image}\big(\begin{bmatrix} B & AB & \dots A^{n-1}B \end{bmatrix}\big)$$

- Questions:
  - How would we find an input to drive the system to a particular state $\xi \in \mathbb{R}^n$?
  - What is the input of smallest norm that will do so?
- Given ξ, we will solve

$$\begin{array}{rl} \text{minimize} & \|u\| \\ \text{subject to} & \Upsilon_t u = \xi \end{array}$$

Goals of controllability analysis

- Given a desired final state $\xi \in \mathbb{R}^n$, we will solve

$$\begin{array}{ll} \text{minimize} & \|u\| \\ \text{subject to} & \Upsilon_t u = \xi \end{array}$$

  That is, find the *smallest* input $u \in L_2[0,t]$ which will drive the state so that $x(t) = \xi$.

- The norm of the minimal norm u_{opt} gives a measure of *how much energy* is required to reach a final state.

- This will give us a quantitative and practical notion of controllability; much more useful than the rank test.

- This question will turn out to be deeply linked to the problem of *model reduction*.

Minimum-norm solution

- In general there are many solutions to the equation $\Upsilon_t u = \xi$.

- These solutions live in an *affine set* in $L_2[0,T]$.

Closed sets

Let S be a subset of a Hilbert space H. Recall that a point $x \in H$ is called a *closure point* of S if

$$B(x, \varepsilon) \cap S \neq \emptyset \text{ for all } \varepsilon > 0$$

where $B(x, \varepsilon)$ is the open-ball of radius ε.

Theorem:

$$S \text{ is closed} \quad \Longleftrightarrow \quad \text{Every convergent sequence } \{x_0, x_1, \ldots\} \subset S \text{ converges to a point in } S$$

Proof: Let $x = \lim_{i \to \infty} x_i$.

$\Rightarrow$ Note that x is a closure point of S, since $x_i \in B(x, \varepsilon)$ for i large enough. Hence x must be contained in S if S is closed.

$\Leftarrow$ Suppose S is not closed. We construct a sequence in S whose limit is not in S.

Let $\bar{S}$ be the closure of S (the set of closure points.) Pick $y \in \bar{S}$ with $y \notin S$.

Since $y \in \bar{S}$

$$B(y, \varepsilon) \cap S \neq \emptyset \text{ for all } \varepsilon > 0$$

so pick $y_n \in B(y, n^{-1}) \cap S$ for each $n > 0$.

Clearly this sequence converges to y and $y \notin S$.

The orthogonal complement

Suppose S is a subspace of a Hilbert space H.

$$S^\perp = \Big\{ x \in H \ ; \ \langle x, y \rangle = 0 \text{ for all } y \in S \Big\}$$

$S^\perp$ is called the *orthogonal complement* of S in H. Write $x \perp y$ if $\langle x, y \rangle = 0$.

Notes

- $S^\perp$ is a subspace of H.
- $S \subset S^{\perp\perp}$. Proof: if $x \in S$, then $x \perp y$ for all $y \in S^\perp$, therefore $x \in S^{\perp\perp}$.

Theorem: $S^\perp$ is closed.

Proof: Suppose $\{x_0, x_1, \dots\} \subset S^\perp$ is a convergent sequence. We show that the limit

$$x = \lim_{i \to \infty} x_i$$

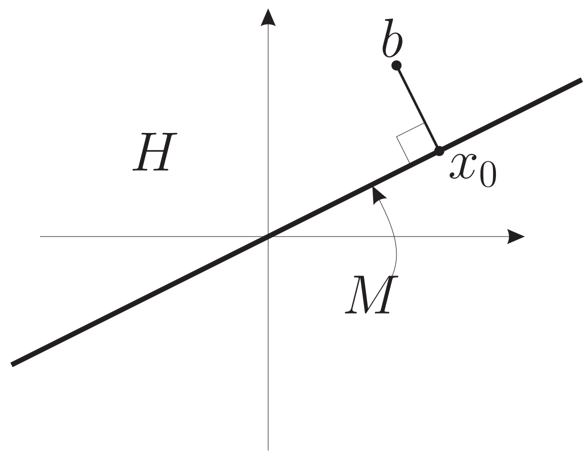
is also in $S^\perp$.

For all $y \in S$, $\langle x_i, y \rangle = 0$ for all i. For any continuous function, $\lim_{i \to \infty} f(x_i) = f(x)$. In particular the inner-product is continuous, so $\langle x, y \rangle = 0$ for all $y \in S$. Hence $x \in S^\perp$.

The projection theorem

Suppose H is a Hilbert space, $b \in H$, and M is a closed subspace of H.

$$\begin{array}{ll} \text{minimize} & \|x - b\| \\ \text{subject to} & x \in M \end{array}$$



Theorem

- *Existence:* There exists a vector x_{opt} which achieves

$$\|x_{\text{opt}} - b\| = \inf \left\{ \|x - b\| \; ; \; x \in M \right\}$$

- *Uniqueness:* The minimizing vector x_{opt} is unique.
- *Orthogonality:*

$$b - x \in M^\perp \qquad \Longleftrightarrow \qquad x \text{ is optimal}$$

Proof of the projection theorem

Orthogonality:

$$\|x_{\text{opt}} - b\| \le \|x - b\| \text{ for all } x \in M \qquad \Longleftrightarrow \qquad \langle x, b - x_{\text{opt}}\rangle = 0 \text{ for all } x \in M$$

$\Rightarrow$ Suppose to the contrary that there exists $x \in M$ such that $\langle x, b - x_{\text{opt}}\rangle = c$, and $c \neq 0$. Without loss of generality, assume $\|x\| = 1$.

Let $y = x_{\text{opt}} + cx$. Then

$$\begin{aligned}\|y - b\|^2 &= \|b - x_{\text{opt}} - cx\|^2 \\ &= \|b - x_{\text{opt}}\|^2 + \langle b - x_{\text{opt}}, -cx\rangle + \langle -cx, b - x_{\text{opt}}\rangle + \langle cx, cx\rangle \\ &= \|b - x_{\text{opt}}\|^2 - |c|^2\end{aligned}$$

Hence if $b - x_{\text{opt}}$ is not orthogonal to M, then x_{opt} is not minimizing.

$\Leftarrow$ For any $x \in M$

$$\|b - x\|^2 = \|b - x_{\text{opt}} + x_{\text{opt}} - x\|^2 = \|b - x_{\text{opt}}\|^2 + \|x_{\text{opt}} - x\|^2$$

Hence $\|b - x\| > \|b - x_{\text{opt}}\|$ if $x \neq x_{\text{opt}}$, hence x_{opt} is minimizing.

This also shows *uniqueness*.

Proof of the projection theorem

Existence:

- Suppose $b \notin M$, otherwise the optimal x is $x_{\text{opt}} = b$ and we are done.
- Let $\delta = \inf \left\{ \|x - b\| \; ; \; x \in M \right\}$. We wish to find $x \in M$ with $\|x - b\| = \delta$.
- Let $\{x_0, x_1, \dots\}$ be a sequence such that $\|x_i - b\| \to \delta$ as $i \to \infty$.
  We will show that $\lim_{i \to \infty} x_i \in M$.
- First we show this sequence is Cauchy. Recall the parallelogram law

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

which implies

$$\|(x_j - b) + (b - x_i)\|^2 + \|(x_j - b) - (b - x_i)\|^2 = 2\|x_j - b\|^2 + 2\|x_i - b\|^2$$

So

$$\begin{aligned} \|x_j - x_i\|^2 &= 2\|x_j - b\|^2 + 2\|x_i - b\|^2 - \|2b - (x_i + x_j)\|^2 \\ &= 2\|x_j - b\|^2 + 2\|x_i - b\|^2 - 4\left\|b - \frac{(x_i + x_j)}{2}\right\|^2 \end{aligned}$$

Existence, continued

- Recap: $\delta = \inf \left\{ \|x - b\| \; ; \; x \in M \right\}$. We wish to find $x \in M$ with $\|x - b\| = \delta$.
- We know $\|x_j - x_i\|^2 = 2\|x_j - b\|^2 + 2\|x_i - b\|^2 - 4\left\| b - \frac{(x_i + x_j)}{2} \right\|^2$
- M is a subspace implies that $\frac{(x_i + x_j)}{2} \in M$. Hence $\left\| b - \frac{(x_i + x_j)}{2} \right\|^2 \geq \delta$.
- Hence $\|x_j - x_i\|^2 \leq 2\|x_j - b\|^2 + 2\|x_i - b\|^2 - 4\delta$.
- $\|x_i - b\| \to \delta$ as $i \to \infty$, so we can make $\|x_j - x_i\|^2$ as small as we like by choosing i and j large enough. Hence $\{x_0, x_1, \ldots\}$ is a Cauchy sequence.
- Recall that every closed subset of a Hilbert space is complete.
- M is closed, therefore M is complete, and this Cauchy sequence therefore converges to a limit in M. That is,

  $$x_{\text{opt}} = \lim_{i \to \infty} x_i$$

  and $x_{\text{opt}} \in M$.
- Since the norm is continuous, $\|x_{\text{opt}} - b\| = \delta$.

Closed subspaces

If S is a closed subspace of a Hilbert space H, then every element $x \in H$ has a unique representation

$$x = s + t \qquad \text{where } s \in S \text{ and } t \in S^\perp$$

Proof

Given x, choose s as the unique minimizer

$$\begin{array}{rl} \text{minimize} & \|s - x\| \\ \text{subject to} & s \in S \end{array}$$

by the projection theorem, which then implies $x - s \in S^\perp$. Let $t = x - s$.

Corollary: If S is closed, then $S^{\perp\perp} = S$.

Proof: We already know $S \subset S^{\perp\perp}$. We need to show that $S^{\perp\perp} \subset S$. That is, if $x \in S^{\perp\perp}$ then $x \in S$.

Applying the above result to $x \in S^{\perp\perp}$, we have $x = s + t$, where $s \in S$ and $t \in S^\perp$. Since $S \subset S^{\perp\perp}$, this implies that $s \in S^{\perp\perp}$.

Since $t = x - s$, this implies $t \in S^{\perp\perp}$ also. But $t \in S^\perp$ also, so $t \perp t$, that is $\langle t, t \rangle = 0$, hence $t = 0$. Hence $x \in S$.

The image and the kernel of an operator

Suppose $\mathcal{U}$ and $\mathcal{V}$ are Hilbert spaces, and $A : \mathcal{U} \to \mathcal{V}$ is a bounded linear operator. Then

$$(\mathrm{image}(A))^\perp = \ker(A^*)$$

Proof

- First we prove $\ker(A^*) \subset (\mathrm{image}(A))^\perp$. Suppose $y \in \ker(A^*)$, and $z \in \mathrm{image}(A)$. Then $z = Ax$ for some x, and

$$\begin{aligned} \langle z, y \rangle &= \langle Ax, y \rangle \\ &= \langle x, A^* y \rangle = 0 \end{aligned}$$

  Hence $y \in (\mathrm{image}(A))^\perp$.

- Now we prove $(\mathrm{image}(A))^\perp \subset \ker(A^*)$. Suppose $y \in (\mathrm{image}(A))^\perp$; then ,

$$\begin{aligned} & \langle y, Ax \rangle = 0 \quad \text{for all } x \in \mathcal{U} \\ \Longrightarrow \quad & \langle A^* y, x \rangle = 0 \quad \text{for all } x \in \mathcal{U} \\ \Longrightarrow \quad & A^* y = 0 \end{aligned}$$

  which implies that $y \in \ker(A^*)$.

The image and the kernel of an operator

Suppose $\mathcal{U}$ and $\mathcal{V}$ are Hilbert spaces, and $A : \mathcal{U} \to \mathcal{V}$ is a bounded linear operator. Then

$$(\mathrm{image}(A))^{\perp} = \ker(A^*)$$

Corollary

Applying the above theorem to A^* gives

$$(\mathrm{image}(A^*))^{\perp} = \ker(A)$$

Caveat

It is not true in general that

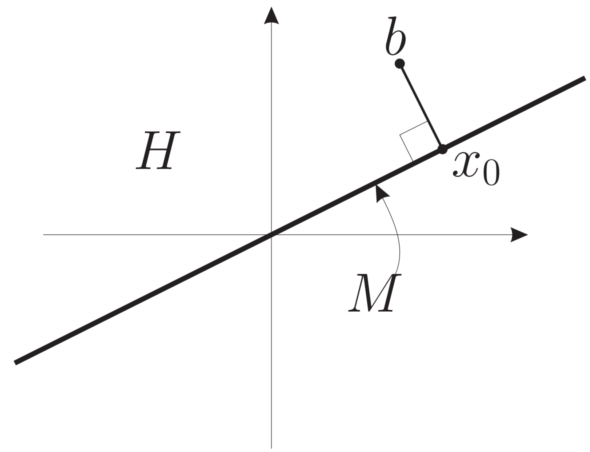
$$\mathrm{image}(A^*) = \ker(A)^{\perp}$$

although this holds for matrices.

Projection Operators

The operator P on a Hilbert space H is called a *projection operator* if it is

- *idempotent:* $P^2 = P$
- *self-adjoint:* $P^* = P$



Notes

- $M = \mathrm{image}(P)$.
- $x \in M \implies x = Px$.

Orthogonality

- For any $x \in M$ and $b \in H$,

$$\begin{aligned}\langle b - Pb, x\rangle &= \langle (I-P)b, x\rangle \\ &= \langle (I-P)b, Px\rangle \\ &= \langle (P-P^2)b, x\rangle = 0\end{aligned}$$

- $\|P\| \leq 1$, since $\|b\|^2 = \|b - Pb + Pb\|^2 = \|b - Pb\|^2 + \|Pb\|^2 \geq \|Pb\|^2$

Projection theorem revisited

Suppose $b \in H$. Then

$$\|b - x\| \geq \|b - Pb\| \qquad \text{for all } x \in S$$

That is, $z = Pb$ is a minimizing solution to

$$\min_{z \in M} \|b - z\|$$

Proof: Since $z \in S$, we have

$$\|b - x\|^2 = \|b - Pb + Pb - x\|^2 = \|b - Pb\|^2 + \|Pb - x\|^2$$

This is just the same proof we used for the sufficiency of orthogonality in the projection theorem.

Finite-dimensional subspaces

Suppose $M = \operatorname{span}\{y_1, y_2, \ldots, y_n\}$ where the y_i are orthonormal.

The linear map $P : H \to H$ defined by

$$Px = \sum_{i=1}^{n} \langle y_i, x \rangle y_i.$$

is a projection operator.

Exercise

- Verify $P^2 = P$ and $P^* = P$.

A finite dimensional subspace of a Hilbert space is closed

Suppose M is a subspace of a Hilbert space H, defined by

$$M = \operatorname{span}\{y_1, y_2, \ldots, y_n\}$$

Then M is closed.

Proof

Without loss of generality assume $\{y_1, y_2, \ldots, y_n\}$ is an orthonormal set; if the y_i are not orthonormal, we can replace them with an orthonormal set without changing M via the Gram-Schmidt procedure.

We need to show that if $\{x_0, x_1, \ldots\}$ is a convergent sequence in M, then its limit is also in M. Let $x = \lim_{j\to\infty} x_j$. Then

$$\|x - Px\| \leq \|x - x_j\| \qquad \text{for all } j$$

But we know

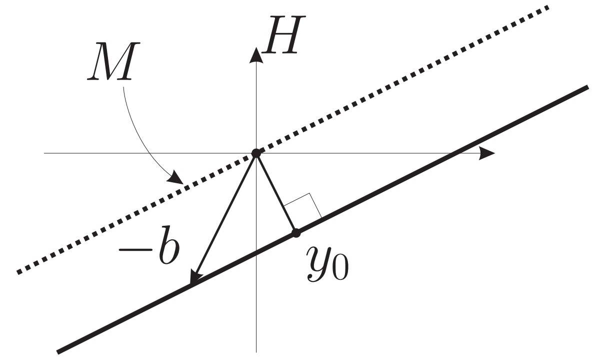
$$\lim_{j\to\infty} \|x - x_j\| = 0$$

hence $\|x - Px\| = 0$ and therefore $x = Px$, which is an element of M.

Affine projection theorem

Suppose we would like to solve

$$\begin{aligned} \text{minimize} \quad & \|y\| \\ \text{subject to} \quad & y = v - b \\ & v \in M \end{aligned}$$



That is, we are looking for the minimum norm element of the affine set

$$\{y \in H \ ; \ y = v - b \text{ for some } v \in M\}$$

Subspace form

Substituting $y = x - b$ leads to the equivalent problem

$$\begin{aligned} \text{minimize} \quad & \|x - b\| \\ \text{subject to} \quad & x \in M \end{aligned}$$

The optimality conditions, $x_{\text{opt}} \in M$ and $x_{\text{opt}} - b \in M^{\perp}$ become

$$y_{\text{opt}} \in M^{\perp}$$
$$y_{\text{opt}} + b \in M$$

Minimum-norm approximation

Suppose $A : U \to V$ is a map between Hilbert spaces.

$$\begin{array}{ll} \text{minimize} & \|Ay - b\| \\ \text{subject to} & y \in U \end{array}$$

In finite dimensions with the Euclidean norm, this is a *least-square-error* problem.

Subspace form

Substituting $x = Ay$ leads to the equivalent problem

$$\begin{array}{ll} \text{minimize} & \|x - b\| \\ \text{subject to} & x \in \mathrm{image}(A) \end{array}$$

Note

- Note that to apply the projection theorem, we need the subspace $\mathrm{image}(A)$ to be closed.

Minimum-norm approximation

y solves			x solves	
minimize	$\|Ay - b\|$	$x = Ay$ $\iff$	minimize	$\|x - b\|$
subject to	$y \in \mathcal{U}$		subject to	$x \in \mathrm{image}(A)$

- Suppose $A : \mathcal{U} \to \mathcal{V}$, and $\mathrm{image}(A)$ is closed.
- The projection theorem then implies that x_{opt} is the unique solution to the equations

$$x_{\text{opt}} \in \mathrm{image}(A) \qquad \text{feasibility}$$

$$x_{\text{opt}} - b \in \mathrm{image}(A)^{\perp} \qquad \text{optimality}$$

- Substituting $x = Ay$ implies

$$y_{\text{opt}} \text{ is optimal} \quad \iff \quad Ay_{\text{opt}} - b \in (\mathrm{image}(A))^{\perp}$$

  There may be many such y_{opt}, even though x_{opt} is unique.

- We know $\ker(A^*) = \mathrm{image}(A)^{\perp}$. Hence

$$y_{\text{opt}} \text{ is optimal} \iff A^*Ay_{\text{opt}} = A^*b$$

The dual approximation problem

Suppose $A : U \to V$ is a map between Hilbert spaces.

$$\begin{array}{rl} \text{minimize} & \|y\| \\ \text{subject to} & Ay = b \\ & y \in U \end{array}$$

Subspace form

Substituting $y = x - c$ leads to the equivalent problem

$$\begin{array}{rl} \text{minimize} & \|x - c\| \\ \text{subject to} & x \in \ker(A) \end{array}$$

where c is any vector such that $Ac = -b$.

Note

- Note that to apply the projection theorem, we need the subspace $\ker(A)$ to be closed.

The dual approximation problem

y solves		x solves
minimize $\|y\|$ subject to $Ay = b$	$Ac = -b$ $x = y + c$ $\Longleftrightarrow$	minimize $\|x - c\|$ subject to $x \in \ker(A)$

- Suppose $A : U \to V$ and $\mathrm{image}(A^*)$ is closed.
- We know $\ker(A) = \mathrm{image}(A^*)^\perp$, hence $\ker(A)$ is closed. Also

$$\begin{aligned} \ker(A)^\perp &= \mathrm{image}(A^*)^{\perp\perp} \\ &= \mathrm{image}(A^*) \end{aligned}$$

- The projection theorem then implies that x_{opt} is the unique solution to

$$x_{\mathrm{opt}} \in \ker(A)$$

$$x_{\mathrm{opt}} - c \in (\ker(A))^\perp$$

The dual approximation problem, continued

y solves		x solves
minimize $\|y\|$ subject to $Ay = b$	$Ac = -b$ $x = y + c$ $\iff$	minimize $\|x - c\|$ subject to $x \in \ker(A)$

- The projection theorem implies that x_{opt} is the unique solution to

$$x_{\text{opt}} \in \ker(A)$$
$$x_{\text{opt}} - c \in (\ker(A))^{\perp}$$

- We saw $\ker(A)^{\perp} = \text{image}(A^*)$, and substituting $x = y + c$ implies

$$y_{\text{opt}} \text{ is optimal} \quad \iff \quad \begin{aligned} y_{\text{opt}} &= A^* z \text{ for some } z \in V \\ A y_{\text{opt}} &= b \end{aligned}$$

- This is equivalent to

$$y_{\text{opt}} \text{ is optimal} \quad \iff \quad \begin{gathered} y_{\text{opt}} = A^* z \\ \text{for some } z \in V \text{ such that } AA^* z = b \end{gathered}$$

Kernels

Suppose $A : \mathcal{U} \to \mathcal{V}$ is a bounded linear operator on Hilbert spaces $\mathcal{U}, \mathcal{V}$. Then

$$\ker M^* = \ker MM^*$$

Proof

- Clearly $\ker(M^*) \subset \ker(MM^*)$.
- We need to show $\ker(MM^*) \subset \ker(M^*)$.

  Suppose $MM^*x = 0$. Then

$$\begin{aligned} & & \langle x, MM^*x \rangle = 0 \\ & \implies & \langle M^*x, M^*x \rangle = 0 \\ & \implies & M^*x = 0 \end{aligned}$$

Corollary

Suppose $M : \mathcal{U} \to \mathbb{R}^n$ is a bounded linear operator. Then

$$\text{image}(M) = \mathbb{R}^n \quad \implies \quad MM^* \text{ is invertible.}$$

Proof: $\ker(MM^*) = \ker(M^*) = \text{image}(M)^\perp = \{0\}$.

Controllability

- Suppose we have the stable state-space system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

- The *controllability operator* is the map $\Psi_c : L_2(-\infty, 0] \to \mathbb{R}^n$ from input signals u to final state $x(0)$ given by

$$\Psi_c u = \int_{-\infty}^{0} e^{A(t-\tau)} Bu(\tau)\, d\tau$$

- Given ξ, we wish to solve

$$\begin{array}{rl} \text{minimize} & \|u\| \\ \text{subject to} & \Psi_c u = \xi \end{array}$$

Theorem

Suppose (A, B) is controllable. Then

- the matrix $X_c = \Psi_c \Psi_c^*$ is nonsingular.
- The optimal u_{opt} is given by

$$u_{\text{opt}} = \Psi_c^* X_c^{-1} x_0$$

Theorem

Suppose (A, B) is controllable. Then

- the matrix $X_c = \Psi_c \Psi_c^*$ is nonsingular.
- The optimal u_{opt} is given by

$$u_{\text{opt}} = \Psi_c^* X_c^{-1} x_0$$

Proof

- All we need to show is that $\operatorname{image}(\Psi_c^*)$ is closed, then apply the projection theorem to the minimum-norm approximation problem.
- $\Psi_c^* : \mathbb{R}^n \to L_2(-\infty, 0]$. For $x \in \mathbb{R}^n$, we have

$$\begin{aligned} \Psi_c x &= \Psi_c(x_1 e_1 + \cdots + x_n e_n) \\ &= x_1 \Psi_c e_1 + \cdots + x_n \Psi_c e_n \end{aligned}$$

  hence $\operatorname{image}(\Psi_c^*) = \operatorname{span}\{\Psi_c e_1, \ldots, \Psi_c e_n\}$, which is finite dimensional, hence it it closed.

- Since (A, B) is controllable, $\operatorname{image}(\Psi_c) = \mathbb{R}^n$, hence $X_c = \Psi_c \Psi_c^*$ is nonsingular.