

Chapter 3

Duality in Banach Space

Modern optimization theory largely centers around the interplay of a normed vector space and its corresponding dual. The notion of duality is important for the following reasons:

1. Dual space plays a role in the Banach space analogous to the inner product in a Hilbert space. With suitable interpretation, we can generalize the notion of “projection” to arbitrary normed vector space.
2. Dual space is essential for the concept of “gradient” which, in turn, leads to the variational analysis of Lagrange multipliers.

This chapter will focus on the both the algebraic and the geometric properties of dual spaces.

- Linear Functional
- Hahn-Banach Theorem
- Applications
- Adjoints

Linear Functional

- Suppose X and Y are normed spaces. Let $\Lambda \in \mathcal{B}(X, Y)$ denote the collection of all bounded linear operators from X to Y .
 - ◊ If a linear operator is continuous at a single point, then it is continuous throughout the entire X .
 - ◊ A linear operator is bounded if and only if it is continuous.

- Associated to each $\Lambda \in \mathcal{B}(X, Y)$ the number

$$\|\Lambda\| := \sup\{\|\Lambda\mathbf{x}\| \mid \mathbf{x} \in X, \|\mathbf{x}\| = 1\}.$$

- ◊ $\|\Lambda\|$ is a norm on the space $\mathcal{B}(X, Y)$.
- ◊ If Y is a Banach space, so is $\mathcal{B}(X, Y)$.
- Some trivial examples of linear functionals:
 - ◊ On $C[0, 1]$, $f(\mathbf{x}) = \mathbf{x}(\frac{1}{2})$ is a bounded linear functional. [What is \$\|f\|\$?](#)
 - ◊ On $\mathcal{L}_2[0, 1]$, $f(\mathbf{x}) = \int_0^1 \mathbf{y}(t)\mathbf{x}(t)dt$, where $\mathbf{y} \in \mathcal{L}_2[0, 1]$ is fixed, is a bounded linear functional.
 - ◊ The space of finitely nonzero sequences with

$$\|\{\xi_i\}_{i=1}^\infty\| = \max_{1 \leq i \leq n} |\xi_i|$$

is a Banach space. The function

$$f(\mathbf{x}) = \sum_{\xi_k \neq 0} k\xi_k$$

is an unbounded functional.

- The Banach space $\mathcal{B}(X, \mathbb{F})$ is called the *dual space* of X and is denoted by X^* .
 - ◇ Show that every Cauchy sequence in X^* converges and, hence, X^* is a Banach space.
 - ◇ For the symmetry (or duality) that exists between X and X^* , elements in X^* are often denoted by \mathbf{x}^* . The action of \mathbf{x}^* on any element $\mathbf{z} \in X$ is denoted by

$$\mathbf{x}^*(\mathbf{z}) = \langle \mathbf{z}, \mathbf{x}^* \rangle.$$

- ◇ The symmetric notation $\langle \mathbf{z}, \mathbf{x}^* \rangle$ is meant to suggest that the bounded linear functional in a Banach space plays a similar role as the inner product in a Hilbert space.

Basic Dual Spaces

- The dual of \mathbb{R}^n .
 - ◇ Assume that the 2-norm is used.
 - ◇ Any linear functional Λ has a unique vector representation $\boldsymbol{\eta}$ via
 - ▷ $\Lambda(\mathbf{e}_i) = \eta_i$.
 - ▷ If $\mathbf{x} = \sum_{i=1}^n \xi_i \mathbf{e}_i$, then $\Lambda(\mathbf{x}) = \sum_{i=1}^n \xi_i \eta_i = \langle \mathbf{x}, \boldsymbol{\eta} \rangle$.
 - ◇ By Cauchy-Schwarz inequality,

$$|\Lambda(\mathbf{x})| = \langle \mathbf{x}, \boldsymbol{\eta} \rangle \leq \|\mathbf{x}\| \|\boldsymbol{\eta}\|.$$

- ▷ $\|\Lambda\| = \|\boldsymbol{\eta}\|$
 - ◇ $(\mathbb{R}^n)^* = \mathbb{R}^n$ with the same 2-norm.
 - ◇ What will it be if other norms are used?

- The dual of ℓ_p , $1 \leq p < \infty$.

◇ $\ell_p := \{\mathbf{x} = \{\xi_i\}_{i=1}^\infty \mid \sum_{i=1}^\infty |\xi_i|^p < \infty\}$ is a normed vector space with

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^\infty |\xi_i|^p \right)^{\frac{1}{p}}.$$

◇ (The Hölder Inequality) If $1 \leq p, q \leq \infty$ are such that $1/p + 1/q = 1$, then given any $\mathbf{x} = \{\xi_i\}_1^n \in \ell_p$ and $\mathbf{y} = \{\eta_i\}_{i=1}^\infty \in \ell_q$, it is true that

$$\sum_{i=1}^n |\xi_i \eta_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

◇ Any bounded linear functional Λ on ℓ_p can be uniquely represented via

▷ $\Lambda(\mathbf{e}_i) = \eta_i$.

▷ If $\mathbf{x} = \{\xi_i\}_{i=1}^\infty$, then $\Lambda \mathbf{x} = \sum_{i=1}^\infty \xi_i \eta_i$ (by continuity of Λ).

◇ Need to show that $\{\eta_i\}_{i=1}^\infty \in \ell_q$ (by the fact that Λ bounded).

▷ Show that $\|\Lambda\| = \|\{\eta_i\}_{i=1}^\infty\|_q$.

◇ $(\ell_p)^* = \ell_q$ with the q -norm.

◇ Note that $(\ell_1)^* = \ell_\infty$, but $(\ell_\infty)^* \neq \ell_1$ (Why?).

- The dual of $\mathcal{L}_p[a, b]$.
 - ◇ There is a one-to-one correspondence between bounded linear functionals $\Lambda \in \mathcal{L}_p[a, b]$ and elements $\mathbf{y} \in \mathcal{L}_q[a, b]$ such that

$$\Lambda \mathbf{x} = \int_a^b \mathbf{x}(t) \mathbf{y}(t) dt$$

and $\|\Lambda\| = \|\mathbf{y}\|_q$.

- ▷ How to prove that every bounded linear functional is representable in the above inner product form?
- ▷ Weierstrass Approximation Theorem: The space of polynomials is dense in $C[a, b]$.

- The dual of c_0 .

◇ The space $c_0 = \{\mathbf{x} = \{\xi_i\}_{i=1}^{\infty} \mid \lim_{n \rightarrow \infty} \xi_n = 0\}$ is a normed vector space with norm

$$\|\mathbf{x}\| = \max_i |\xi_i|.$$

◇ $l_1 \subset c_0 \subset l_{\infty}$.

◇ $c_0^* = l_1$ in the sense that

$$\Lambda \mathbf{x} = \sum_{i=1}^{\infty} \xi_i \eta_i$$

for some $\mathbf{y} = \{\eta_i\}_{i=1}^{\infty} \in l_1$.

◇ Let $c = \{\mathbf{x} = \{\xi_i\}_{i=1}^{\infty} \mid \lim_{n \rightarrow \infty} \xi_n < \infty\}$ with $\|\mathbf{x}\| = \max_i |\xi_i|$. Characterize the dual space c^* .

- The dual of a general Hilbert space.
 - ◇ (Riesz Representation Theorem) Every bounded linear functional Λ on a Hilbert space H has a unique representation $\mathbf{y} \in H$ such that

$$\Lambda \mathbf{x} = \langle \mathbf{x}, \mathbf{y} \rangle$$

and $\|\Lambda\| = \|\mathbf{y}\|$.

- ◇ If H is a Hilbert space, then $H^* = H$.

Hahn-Banach Theorem, I

- The Hahn-Banach Theorem is one of the most critical result for optimization in a normed space.
- The Hahn-Banach Theorem can be stated in several equivalent ways each of which has its own particular advantages and applications.
- Two fundamental versions of the theorem, the extension form and the geometric form, will be discussed in this note.
- The extension form serves as an appropriate generalization of the projection theorem from the Hilbert space to the normed vector space.
- The geometric form, together with the notion of hyperplanes, serves to separate a point and a disjoint convex subset in the space.

Extension Form

- (Hahn-Banach Extension Theorem) Let V be a normed separable vector space over \mathbb{R} .

- ◇ Assume that

- ▷ There exists a functional $p : V \rightarrow \mathbb{R}$

$$\begin{aligned} p(\mathbf{x} + \mathbf{y}) &\leq p(\mathbf{x}) + p(\mathbf{y}), \\ p(\alpha\mathbf{x}) &= \alpha p(\mathbf{x}), \end{aligned}$$

for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \geq 0$.

- ▷ Suppose that f is a real-valued linear functional on a subspace S and

$$f(\mathbf{s}) \leq p(\mathbf{s}) \text{ for all } \mathbf{s} \in S.$$

- ◇ Then there exists a linear functional $F : V \rightarrow \mathbb{R}$ such that $F(\mathbf{x}) \leq p(\mathbf{x})$ for all $\mathbf{x} \in V$ and $F(\mathbf{s}) = f(\mathbf{s})$ for all $\mathbf{s} \in S$.

- Shall assume that V is separable, that is, V has a countable dense set.

- ◇ The Hahn-Banach theorem is true in arbitrary normed space where the proof is done by using [Zorn's lemma](#).

- Assume $\mathbf{y} \notin S$.

- ◊ Consider the subspace $\text{span}\{S, \mathbf{y}\}$.
- ◊ Note that for arbitrary $\mathbf{s}_1, \mathbf{s}_2 \in S$,

$$f(\mathbf{s}_1) + f(\mathbf{s}_2) \leq p(\mathbf{s}_1 - \mathbf{y}) + p(\mathbf{s}_2 + \mathbf{y}).$$

Hence, there exist a constant c such that

$$\sup_{\mathbf{s} \in S} (f(\mathbf{s}) - p(\mathbf{s} - \mathbf{y})) \leq c \leq \inf_{\mathbf{s} \in S} (p(\mathbf{s} + \mathbf{y}) - f(\mathbf{s})).$$

- ◊ Extend f by defining

$$g(\mathbf{s} + \alpha \mathbf{y}) = f(\mathbf{s}) + \alpha g(\mathbf{y})$$

where $g(\mathbf{y}) := c$, as is determined from above.

- Claim that $g(\mathbf{s} + \alpha \mathbf{y}) \leq p(\mathbf{s} + \alpha \mathbf{y})$ for all $\alpha \in \mathbb{R}$.

- ◊ If $\alpha > 0$, then

$$g(\mathbf{s} + \alpha \mathbf{y}) = \alpha \left(c + f\left(\frac{\mathbf{s}}{\alpha}\right) \right) \leq \alpha \left(p\left(\frac{\mathbf{s}}{\alpha} + \mathbf{y}\right) - f\left(\frac{\mathbf{s}}{\alpha}\right) + f\left(\frac{\mathbf{s}}{\alpha}\right) \right) = p(\mathbf{s} + \alpha \mathbf{y}).$$

- ◊ If $\alpha < 0$, then

$$g(\mathbf{s} + \alpha \mathbf{y}) = -\alpha \left(-c + f\left(\frac{\mathbf{s}}{-\alpha}\right) \right) \leq -\alpha \left(p\left(\frac{\mathbf{s}}{-\alpha} - \mathbf{y}\right) - f\left(\frac{\mathbf{s}}{-\alpha}\right) + f\left(\frac{\mathbf{s}}{-\alpha}\right) \right) = p(\mathbf{s} + \alpha \mathbf{y}).$$

- Let $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ be a countable dense of V .
 - ◇ Select a subset of vectors $\{\mathbf{y}_1, \mathbf{y}_2, \dots\}$, one at a time, which is mutually independent and independent of S , in the way described above.
 - ◇ Extend f to subspaces $\text{span}\{\text{span}\{S, \mathbf{y}_1\}, \mathbf{y}_2\}, \dots\}$, one at a time, as described above.
 - ◇ The space $\text{span}\{S, \mathbf{y}_1, \mathbf{y}_2, \dots\}$ is dense in V on which the extension g is defined.
- The final extension F is defined by continuity.

- Corollaries:

◇ If f is a bounded linear functional on a subspace S of a real-valued normed vector space V , then there exists a bounded linear functional F on V which is the extension of f and $\|F\| = \|f\|$.

◇ Let $\mathbf{x} \in V$. Then there is a nontrivial $\mathbf{x}^* \in V^*$ such that

$$\langle \mathbf{x}, \mathbf{x}^* \rangle = \|\mathbf{x}\| \quad \text{and} \quad \|\mathbf{x}^*\| = 1.$$

▷ Assume $\mathbf{x} \neq \mathbf{0}$. Consider the bounded linear functional $f(\alpha\mathbf{x}) = \alpha\|\mathbf{x}\|$ on the subspace $\text{span}\{\mathbf{x}\}$.

▷ $\|f\| = 1$.

▷ There exists an extension F on V

▷ **Show that the converse of the above statement is not true.**

- Recall that $\ell_1^* = \ell_\infty$.

- If $\mathbf{x} = \{\xi_i\}_{i=1}^\infty$, define

$$f(\mathbf{x}) = \sum_{i=1}^{\infty} \left(1 - \frac{1}{i}\right) \xi_i.$$

- f is a bounded linear functional. Indeed, $f \longleftrightarrow \{0, \frac{1}{2}, \frac{2}{3}, \dots\}$.

- $\|f\| = \sup \frac{|f(\mathbf{x})|}{\|\mathbf{x}\|} = 1$.

- $f(\mathbf{x}) < \|\mathbf{x}\|$.

◇ Let B^* be the closed unit ball of V^* . Then for every $\mathbf{x} \in V$,

$$\|\mathbf{x}\| = \{\langle \mathbf{x}, \mathbf{x}^* \rangle \mid \mathbf{x}^* \in B^*\}.$$

Applications

- An Idea of Application: Existence of a solution to a minimization problem.
 - ◇ Given a functional f on a subspace M of a normed space, an arbitrary extension of f to the entire space in general would be “bad”, i.e., the resulting extension might be unbounded or have larger norm than $\|f\|$.
 - ◇ Which extension will give the minimum norm and what minimum norm can be attained?
 - ◇ The Hahn-Banach theorem guarantees the existence of a minimum norm extension and prescribes the norm of the “best” extension.
- The second dual space.
- Orthogonality in Banach space.
- Extension of the projection theorem.
- Duality principles.

The Dual of $C[a, b]$

- $(C[a, b])^* = NBV[a, b]$. That is, f is a bounded linear functional on $C[a, b]$ if and only there is a function of bounded variation v on $[a, b]$ such that

$$f(x) = \int_a^b x(t)dv(t).$$

In this case, $\|f\| = T_a^b(v)$.

- ◇ This has been a major application of the Hahn-Banach theorem.
- ◇ Define $B := \{x : [a, b] \rightarrow \mathbb{R} \mid x \text{ is bounded on } [a, b]\}$ and $\|x\|_B := \sup_{a \leq t \leq b} |x(t)|$.
- ◇ $C[a, b] \subset B$.
- ◇ By the Hahn-Banach theorem, an extension F of f from $C[a, b]$ to B exists and $\|F\| = \|f\|$.

◇ Try to understand how F should look like.

▷ For any $s \in (a, b]$, define

$$u_s(t) = \begin{cases} 1, & \text{if } a \leq t \leq s, \\ 0, & \text{if } s < t \leq b, \end{cases}$$

and $u_a \equiv 0$. Obviously, $u_s \in B$.

▷ Define

$$v(s) = F(u_s).$$

▷ Claim that $v \in BV[a, b]$. Indeed, $T_a^b \leq \|f\|$.

$$\begin{aligned} \sum_{i=1}^n |v(t_i) - v(t_{i-1})| &= \sum_{i=1}^n (\operatorname{sgn}\{v(t_i) - v(t_{i-1})\})(v(t_i) - v(t_{i-1})) \\ &= \sum_{i=1}^n (\operatorname{sgn}\{v(t_i) - v(t_{i-1})\})(F(u_{t_i}) - F(u_{t_{i-1}})) \\ &= F\left(\sum_{i=1}^n (\operatorname{sgn}\{v(t_i) - v(t_{i-1})\})(u_{t_i} - u_{t_{i-1}})\right) \\ &\leq \|F\| \left\| \sum_{i=1}^n (\operatorname{sgn}\{v(t_i) - v(t_{i-1})\})(u_{t_i} - u_{t_{i-1}}) \right\| = \|f\|. \end{aligned}$$

▷ Given any $x \in C[a, b]$, consider the function

$$z(t) := \sum_{i=1}^n x(t_{i-1})(u_{t_i} - u_{t_{i-1}}).$$

▷ Note that the difference

$$\|z - x\|_B = \max_i \max_{t_{i-1} \leq t \leq t_i} |x(t_{i-1}) - x(t)|$$

goes to zero as the partition is made finer.

▷ By continuity, $F(z) \rightarrow F(x) = f(x)$.

▷ Observe that

$$F(z) = \sum_{i=1}^n x(t_{i-1})(v(t_i) - v(t_{i-1})) \rightarrow \int_a^b x(t)dv(t).$$

▷ It is clear that $|F(x)| \leq \|x\|T_a^b(v)$. Hence, $\|F\| = T_a^b(v)$.

◇ The bounded variation v corresponding to f is not unique.

▷ How could the functional $f(x) = x(1/2)$ be represented in terms a bounded variation $v(t)$?

◇ Define $NBV[a, b] := \{v \in BV[a, b] | v(a) = 0, v \text{ is right-continuous}\}$.

▷ The corresponding between $C[a, b]$ and $NBV[a, b]$ is one-to-one.

Second Dual Space

- The function $f(\mathbf{x}^*) = \langle \mathbf{x}, \mathbf{x}^* \rangle$, where $\mathbf{x} \in V$ is fixed, is bounded linear functional on V^* with $\|f\| \leq \|\mathbf{x}\|$.
 - ◊ By the Hanh-Banach theorem, there is one functional \mathbf{x}^* such that $\|\mathbf{x}^*\| = 1$ and $\langle \mathbf{x}, \mathbf{x}^* \rangle = \|\mathbf{x}\|$.
 - ◊ $\|f\| = \|\mathbf{x}\|$.
 - ◊ Depending upon whether \mathbf{x} or \mathbf{x}^* is fixed, both V and V^* define bounded functionals on each other.
- The dual space V^{**} of V^* is called the second dual of V .
- The mapping $\phi : V \longrightarrow V^{**}$ via

$$\langle \mathbf{x}^*, \phi(\mathbf{x}) \rangle = \langle \mathbf{x}, \mathbf{x}^* \rangle$$

is called the *natural mapping* of V into V^{**} .

- ◊ Note that $\|\phi\| = \|\mathbf{x}\|$. We denote $\phi(\mathbf{x}) = \mathbf{x}^{**}$.
- ◊ The mapping ϕ is not necessarily onto.
- ◊ If ϕ is onto, we say that the space V is reflexive and denote $V = V^{**}$.
 - ▷ $\ell_p^{**} = \ell_q^* = \ell_p$, if $1 < p < \infty$.
 - ▷ ℓ_1 and L_1 is not reflexive.
 - ▷ Any Hilbert space is reflective.

Orthogonality in Banach Space

- The symmetric operator $\langle \mathbf{x}, \mathbf{x}^* \rangle$ in a Banach space plays the role of inner product in a Hilbert space.

◇ (Cauchy-Schwarz Inequality) $\langle \mathbf{x}, \mathbf{x}^* \rangle \leq \|\mathbf{x}\| \|\mathbf{x}^*\|$.

- A vector \mathbf{x}^* is said to be *aligned* with a vector \mathbf{x} if and only if $\langle \mathbf{x}, \mathbf{x}^* \rangle = \|\mathbf{x}\| \|\mathbf{x}^*\|$.
- A vector \mathbf{x}^* is said to be *orthogonal* to a vector \mathbf{x} if and only if $\langle \mathbf{x}, \mathbf{x}^* \rangle = 0$.
- Orthogonal complement:

- ◇ Given $S \in V$, the orthogonal complement S^\perp of S is defined to be

$$S^\perp := \{\mathbf{x}^* \in V^* \mid \langle \mathbf{x}, \mathbf{x}^* \rangle = 0, \text{ for every } \mathbf{x} \in S\}.$$

- ◇ Given $U \in V^*$, the orthogonal complement ${}^\perp U$ of U is defined to be

$${}^\perp U := \{\mathbf{x} \in V \mid \langle \mathbf{x}, \mathbf{x}^* \rangle = 0, \text{ for every } \mathbf{x}^* \in U\}.$$

- ▷ Think about U as a collection of normal vectors.
- ▷ Think about ${}^\perp U$ as the intersection of several (not necessarily finite) hyperplanes.

- If M is a closed subspace of a normed space, then

$${}^\perp(M^\perp) = M.$$

- ◊ It is clear that $M \subset {}^\perp(M^\perp)$.
- ◊ Prove the other direction by contradiction.
 - ▷ Suppose $\mathbf{x} \notin M$. Define linear functional

$$f(\alpha\mathbf{x} + \mathbf{m}) = \alpha$$

on the space spanned by $\mathbf{x} + M$.

- ▷ f is bounded since

$$\|f\| = \sup_{\mathbf{m} \in M, \alpha \in \mathbb{R}} \frac{|f(\alpha\mathbf{x} + \mathbf{m})|}{\|\alpha\mathbf{x} + \mathbf{m}\|} = \sup_{\mathbf{m} \in M} \frac{f(\mathbf{x} + \mathbf{m})}{\|\mathbf{x} + \mathbf{m}\|} = \frac{1}{\inf_{\mathbf{m} \in M} \|\mathbf{x} + \mathbf{m}\|} < \infty$$

- Why is it that the inf cannot be zero?

- ▷ By the Hahn-Banach theorem, f can be extended to $\mathbf{x}^* \in V^*$.
- ▷ Because $\langle \mathbf{m}, \mathbf{x}^* \rangle = 0$ for all $\mathbf{m} \in M$, $\mathbf{x}^* \in M^\perp$.
- ▷ By construction, $\langle \mathbf{x}, \mathbf{x}^* \rangle = 1$. So, $\mathbf{x} \notin {}^\perp(M^\perp)$.

Characterizing Alignment in $(C[a, b])^*$

- Corresponding to $\mathbf{x}^* \in (C[a, b])^*$, recall that
 - ◊ The functional must be of the form $\langle \mathbf{x}, \mathbf{x}^* \rangle = \int_a^b \mathbf{x}(t) d\mathbf{v}(t)$ for some $\mathbf{v} \in NBV$.
 - ◊ Indeed, $\mathbf{v}(s) = F(u_s)$, where F is the extension of \mathbf{x}^* and u_s is a collection of step functions.
- Recall also that, for a fixed $\mathbf{x} \in V$, a functional $\mathbf{x}^* \in V^*$ is aligned with \mathbf{x} if and only if

$$\langle \mathbf{x}, \mathbf{x}^* \rangle = \int_a^b \mathbf{x}(t) d\mathbf{v}(t) = \|\mathbf{x}\| \|\mathbf{x}^*\| = \|\mathbf{x}\|_\infty \sup_{a=t_0 < t_1 < \dots < t_n = b} \sum_i |v(t_i) - v(t_{i-1})|.$$

- Let $\Gamma = \{t \in [a, b] \mid |\mathbf{x}(t)| = \|\mathbf{x}\|_\infty\}$.
 - ◊ Γ may be finite or infinite, but is always nonempty.
- (Prove the following facts:) \mathbf{x}^* is aligned with \mathbf{x} if and only if
 - ◊ \mathbf{v} is constant outside Γ .
 - ◊ \mathbf{v} is nondecreasing if $\mathbf{x}(t) > 0$.
 - ◊ \mathbf{v} is nonincreasing if $\mathbf{x}(t) < 0$.
 - ◊ If Γ is finite, then an aligned functional must consist of a finite number of step discontinuities.

Minimum Norm Problems

- The minimum norm problems in Hilbert space can be extended into Banach space with the exceptions that
 - ◇ The solution in a Banach space is not necessarily unique.
 - ▷ Give an example of non-uniqueness.
 - ◇ The orthogonality condition should be modified in the sense of dual space.
 - ◇ The system for determining the optimizing vector in a Banach space generally is nonlinear.
- There are remarkable analogues in the theory for minimum norm problem between Hilbert space and Banach space.

Extension of Projection Theorem

- Given \mathbf{x} in a real normed space X and a subspace M , then

$$d = \inf_{\mathbf{m} \in M} \|\mathbf{x} - \mathbf{m}\| = \max_{\|\mathbf{x}^*\| \leq 1, \mathbf{x}^* \in M^\perp} \langle \mathbf{x}, \mathbf{x}^* \rangle.$$

- ◇ What is the meaning of the characterization on the right if X is a Hilbert space?
 - ◇ Note that there is a difference in the meaning of inf and min.
- By the definition of inf, for any $\epsilon > 0$, there exists $\mathbf{m}_\epsilon \in M$ such that

$$\|\mathbf{x} - \mathbf{m}_\epsilon\| \leq d + \epsilon.$$

- ◇ For any $\mathbf{x}^* \in M^\perp$, $\|\mathbf{x}^*\| \leq 1$, it is true that

$$\langle \mathbf{x}, \mathbf{x}^* \rangle = \langle \mathbf{x} - \mathbf{m}_\epsilon, \mathbf{x}^* \rangle \leq \|\mathbf{x}^*\| \|\mathbf{x} - \mathbf{m}_\epsilon\| \leq d + \epsilon.$$

- ◇ It is thus proved that $\langle \mathbf{x}, \mathbf{x}^* \rangle \leq d$. It only remains to show that $\langle \mathbf{x}, \mathbf{x}_0^* \rangle = d$ for some \mathbf{x}_0^* .
 - ▷ Note that we are going to prove that the max is attained at some functional.

◇ Define linear functional

$$f(\alpha \mathbf{x} + \mathbf{m}) = \alpha d$$

on the space spanned by $\mathbf{x} + M$.

▷ f is bounded since

$$\|f\| = \sup_{\mathbf{m} \in M} \frac{|f(\alpha \mathbf{x} + \mathbf{m})|}{\|\alpha \mathbf{x} + \mathbf{m}\|} = \frac{d}{\inf_{\mathbf{m} \in M} \|\mathbf{x} + \frac{\mathbf{m}}{\alpha}\|} = 1.$$

◇ By the Hahn-Banach theorem, there exists an extension \mathbf{x}_0^* of f with $\|\mathbf{x}_0^*\| = 1$.

◇ Because $\langle \mathbf{m}, \mathbf{x}_0^* \rangle = 0$ for all $\mathbf{m} \in M$, $\mathbf{x}_0^* \in M^\perp$.

◇ By construction (**What construction?**), $\langle \mathbf{x}, \mathbf{x}_0^* \rangle = d$.

• Note that the minimum distance d is achieved by $\mathbf{x}_0^* \in M^\perp$.

• If the infimum on the first equality holds for some $\mathbf{m}_0 \in M$, then \mathbf{x}_0^* is alligned with $\mathbf{x} - \mathbf{m}_0$.

◇ If $\|\mathbf{x} - \mathbf{m}_0\| = d$ for some $\mathbf{m}_0 \in M$, let $\mathbf{x}_0^* \in M^\perp$ be the functional satisfying $\|\mathbf{x}_0^*\| = 1$ and $\langle \mathbf{x}, \mathbf{x}_0^* \rangle = d$.

◇ Then

$$\langle \mathbf{x} - \mathbf{m}_0, \mathbf{x}_0^* \rangle = d = \|\mathbf{x}_0^*\| \|\mathbf{x} - \mathbf{m}_0\|.$$

Duality Principle

- The previous result states the equivalence of two optimization problems: the minimization in X as the *primal problem* and the maximization in X^* as the *dual problem*.
- A vector $\mathbf{m}_0 \in M$ (*existence*) satisfies

$$\|\mathbf{x} - \mathbf{m}_0\| \leq \|\mathbf{x} - \mathbf{m}\| \quad \text{for all } \mathbf{m} \in M$$

if and only if there exists a nonzero $\mathbf{x}^* \in M^\perp$ aligned with $\mathbf{x} - \mathbf{m}_0$.

- ◇ The “only if” part is already proved.
- ◇ Suppose that $\mathbf{x}^* \in M^\perp$ aligned with $\mathbf{x} - \mathbf{m}_0$. Then

$$\|\mathbf{x} - \mathbf{m}_0\| = \langle \mathbf{x} - \mathbf{m}_0, \mathbf{x}^* \rangle = \langle \mathbf{x}, \mathbf{x}^* \rangle = \langle \mathbf{x} - \mathbf{m}, \mathbf{x}^* \rangle \leq \|\mathbf{x} - \mathbf{m}\|$$

Another View

- Given $\mathbf{x}^* \in V^*$, then

$$\min_{\mathbf{m}^* \in M^\perp} \|\mathbf{x}^* - \mathbf{m}^*\| = \sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* \rangle.$$

- ◊ For any $\mathbf{m}^* \in M^\perp$,

$$\|\mathbf{x}^* - \mathbf{m}^*\| = \sup_{\|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* - \mathbf{m}^* \rangle \geq \sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* - \mathbf{m}^* \rangle = \sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* \rangle.$$

- ◊ Only need to show equality holds for some $\mathbf{m}_0^* \in M^\perp$.

- ▷ Consider the restriction of \mathbf{x}^* to M (with norm $\sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* \rangle$).
- ▷ Let \mathbf{y}^* be the Hahn-Banach extension of the restricted \mathbf{x}^* .
- ▷ Let $\mathbf{m}_0^* = \mathbf{x}^* - \mathbf{y}^*$.
- ▷ Then $\mathbf{m}_0^* \in M^\perp$ and $\|\mathbf{x}^* - \mathbf{m}_0^*\| = \|\mathbf{y}^*\| = \sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \mathbf{x}^* \rangle$.

- The minimum of the left is achieved by some $\mathbf{m}_0^* \in M^\perp$.
- If the supremum on the right is achieved by some $\mathbf{x}_0 \in M$, then $\mathbf{x}^* - \mathbf{m}_0^*$ is aligned with \mathbf{x}_0 .
 - ◊ Obviously $\|\mathbf{x}_0\| = 1$.
 - ◊ $\|\mathbf{x}^* - \mathbf{m}_0^*\| = \langle \mathbf{x}_0, \mathbf{x}^* \rangle = \langle \mathbf{x}_0, \mathbf{x}^* - \mathbf{m}_0^* \rangle$.

Applications

In all applications of the duality theory,

- Use the alignment properties of the space and its dual to characterize optimum solutions.
- Guarantee the existence of a solution by formulating minimum norm problems in a dual space.
- Examine to see if the dual problem is easier than the primal problem.

Chebyshev Approximation

- Given $f \in C[a, b]$, want to find a polynomial p of degree less than or equal to n that “best” approximates f in the sense of the sup-norm over $[a, b]$.
- Let M be the space of all polynomials of degree less than or equal to n .
 - ◊ M is a subspace of $C[a, b]$ of dimension $n + 1$.
- The infimum of $\|f - p\|_\infty$ is achievable by some $p_0 \in M$ since M is closed.
- Want to characterize p_0 (This is the essence of the Chebyshev theorem.)
 - ◊ The set $\Gamma = \{t \in [a, b] \mid |f(t) - p_0| = \|f - p_0\|_\infty\}$ contains at least $n + 2$ points. (Why?)
- By the previous theory, $f - p_0$ must be aligned with some element in $M^\perp \subset (C[a, b])^* = NBV[a, b]$.
 - ◊ Assume Γ contains $m < n + 2$ points $a \leq t_1 < \dots < t_m \leq b$.
 - ◊ If $\mathbf{v} \in NBV[a, b]$ is aligned with $f - p_0$, then \mathbf{v} is a piecewise continuous function with jump discontinuities only at these t_i 's.
 - ◊ Let t_k be a point of jump discontinuity of \mathbf{v} .
 - ▷ The polynomial $q(t) = \prod_{i \neq k} (t - t_i) \in M$, but $\langle q, \mathbf{v} \rangle \neq 0$ and hence $\mathbf{v} \notin M^\perp$.
 - ◊ Γ must contain at least $n + 2$ points.

Characterizing Constrained Equalities

- Given $\mathbf{y}_1, \dots, \mathbf{y}_n \in X$, consider the problem

$$\begin{aligned} & \text{Minimize} \quad \|\mathbf{x}^*\| \\ & \text{Subject to} \quad \langle \mathbf{y}_i, \mathbf{x}^* \rangle = c_i, \quad i = 1, \dots, n. \end{aligned}$$

- The minimum can be achieved.

◊ Let $M = \text{span}\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$. Then

$$d = \min_{\langle \mathbf{y}_i, \mathbf{x}^* \rangle = c_i, i=1, \dots, n} \|\mathbf{x}^*\| = \min_{\mathbf{m}^* \in M^\perp} \|\bar{\mathbf{x}}^* - \mathbf{m}^*\|,$$

for any $\bar{\mathbf{x}}^*$ satisfying the equality constraints. (Why?)

- By the duality principle,

$$d = \min_{\mathbf{m}^* \in M^\perp} \|\bar{\mathbf{x}}^* - \mathbf{m}^*\| = \sup_{\mathbf{x} \in M, \|\mathbf{x}\| \leq 1} \langle \mathbf{x}, \bar{\mathbf{x}}^* \rangle.$$

- Write $\mathbf{x} = Y\mathbf{a}$ where $Y = [\mathbf{y}_1, \dots, \mathbf{y}_n]$ and $\mathbf{a} \in \mathbb{R}^n$. Then

$$d = \min_{\langle \mathbf{y}_i, \mathbf{x}^* \rangle = c_i, i=1, \dots, n} \|\mathbf{x}^*\| = \max_{\|Y\mathbf{a}\| \leq 1} \langle Y\mathbf{a}, \bar{\mathbf{x}}^* \rangle = \max_{\|Y\mathbf{a}\| \leq 1} \mathbf{c}^\top \mathbf{a}.$$

- The optimal solution \mathbf{x}_0^* must be aligned with $Y\mathbf{a}$.

Control Problem

- Want to select the field current $u(t)$, $t \in [0, 1]$, so as to drive a motor

$$\ddot{\theta} + \dot{\theta} = u(t)$$

from $\theta(0) = \dot{\theta}(0) = 0$ to $\theta(1) = 1$ and $\dot{\theta}(1) = 0$ while maintaining that $\|u\|_\infty$ is minimized.

- Cast the problem in $X = \mathcal{L}_1[0, 1]$ with $X^* = \mathcal{L}_\infty[0, 1]$.

◊ Seek $u \in X^*$ with minimum norm.

- By the variation of constants formula, a general solution is

$$\theta(t) = \alpha + \beta e^{-t} + \int_0^t (1 - e^{s-t})u(s)ds.$$

- Boundary conditions imply the equality constraints

$$\begin{aligned} \int_0^1 e^{t-1}u(t)dt &= 0, \\ \int_0^1 (1 - e^{t-1})u(t)dt &= 1. \end{aligned}$$

Note that these functionals are *not* inner products.

- By the previous theory,

$$\min_{\langle y_i, u \rangle = c_i, i=1,2} \|u\|_\infty = \max_{\|a_1 y_1 + a_2 y_2\|_1 \leq q} a_2$$

- The control problem now is reduced to finding constants a_1 and a_2 satisfying

$$\int_0^1 |(a_1 - a_2)e^{t-1} + a_2| dt \leq 1.$$

- The optimal solution $u \in \mathcal{L}_\infty[0, 1]$ should be aligned with $(a_1 - a_2)e^{t-1} + a_2 \in \mathcal{L}_1[0, 1]$.
 - ◇ Note that $(a_1 - a_2)e^{t-1} + a_2$ can change sign at most once.
 - ◇ The alignment condition implies that u must have values $\pm \|u\|_\infty$ and can change sign at most once. (This is the so called *bang-bang control*.)

Rocket Problem

- Want to select a thrust program $u(t)$ that propel a rocket vertically with altitude $x(t)$ governed by

$$\ddot{x}(t) = u(t) - 1,$$

from $x(0) = \dot{x}(0) = 0$ to $x(T) = 1$ with minimum fuel expenditure.

- With given $u(t)$, the altitude is given by

$$x(T) = \int_0^T (T-t)u(t)dt - \frac{T^2}{2}.$$

- The final time T is not specified, the cost function for fuel is a function of T . Want to

$$\begin{aligned} &\text{Minimize} && \int_0^T |u(t)|dt, \\ &\text{Subject to} && \int_0^T (T-t)u(t)dt = 1 + \frac{T^2}{2}. \end{aligned}$$

- ◇ $\mathcal{L}_1[0, T]$ is not the dual of any normed vector space.
- ◇ Embed $\mathcal{L}_1[0, T]$ in $NBV[0, 1]$ and associate control u with the derivative of $v \in NBV[0, 1]$.
- ◇ The objective function becomes

$$\text{Minimize} \quad \int_0^T |dv(t)| = T_0^T(v) = \|v\|.$$

- By the previous theory,

$$\min_{\langle y_1, v \rangle = c_1} \|v\| = \max_{\|(T-t)a\|_\infty \leq 1} a \left(1 + \frac{T^2}{2}\right).$$

- The optimal value is given by $a = \frac{1}{T}$ at which the optimal fuel is given by

$$\min \|v\| = \left(1 + \frac{T^2}{2}\right) \frac{1}{T}.$$

- The thrust program u must be such that its derivative v is aligned with $(T - t)a$.
 - ◊ The alignment condition implies that v can only vary at $t = 0$ and hence is a step function. (Why?)
 - ◊ u is a delta function at $t = 0$ (an impulse). (In what sense?)
- The best final time is obtained by minimizing the fuel expenditure function with respect to T .
 - ◊ $T = \sqrt{2}$.
 - ◊ The derivative of the optimal thrust program is given by

$$v(t) = \begin{cases} 0 & \text{if } t = 0 \\ \sqrt{2} & \text{if } 0 < t \leq \sqrt{2}. \end{cases}$$

Hahn-Banach Theorem, II

- The notion of linear functionals in the dual space can be interpreted as that of hyperplanes in the primal space.
- The notion of extension can be interpreted as that of separation.

Hyperplanes

- A *hyperplane* H in a linear vector space X is the largest proper linear variety in X .
 - ◊ A hyperplane H is a linear variety such that $H \neq X$, and if V is any linear variety containing H , then either $V = X$ or $V = H$.

- If H is a hyperplane, then there is a linear functional f and a constant c such that

$$H = \{\mathbf{x} \in X \mid f(\mathbf{x}) = c\}.$$

- ◊ $H = \mathbf{x}_0 + M$ for some linear subspace M .
- ◊ If $\mathbf{x}_0 \notin M$, then $X = \text{span}\{\mathbf{x}_0 + M\}$. Define $f(\alpha\mathbf{x}_0 + \mathbf{m}) = \alpha$, if $\mathbf{x} = \alpha\mathbf{x}_0 + \mathbf{m}$ with $\mathbf{m} \in M$.
- ◊ If $H = M$, there exists $\mathbf{x}_1 \notin M$ and $X = \text{span}\{\mathbf{x}_1 + M\}$. Define $f(\alpha\mathbf{x}_1 + \mathbf{m}) = \alpha$, if $\mathbf{x} = \alpha\mathbf{x}_1 + \mathbf{m}$ with $\mathbf{m} \in M$.
- ◊ H is called the c -level hyperplane determined by f .
- Given any nonzero linear functional f , the set $\{\mathbf{x} \in X \mid f(\mathbf{x}) = c\}$ is a hyperplane in X .
 - ◊ Let $M = \{\mathbf{x} \in X \mid f(\mathbf{x}) = 0\}$ and \mathbf{x}_0 be such that $f(\mathbf{x}_0) = 1$.
 - ◊ For any $\mathbf{x} \in X$, $\mathbf{x} - f(\mathbf{x})\mathbf{x}_0 \in M$.
 - ▷ M is a maximal proper subspace.
 - ▷ $X = \text{span}\{\mathbf{x}_0 + M\}$.
 - ◊ Let \mathbf{x}_1 be such that $f(\mathbf{x}_1) = c$. Then $f(\mathbf{x}) = c$ if and only if $\mathbf{x} - \mathbf{x}_1 \in M$.

- If H is a hyperplane *not* containing the origin, then there exists a *unique* linear functional f such that

$$H = \{\mathbf{x} \in X \mid f(\mathbf{x}) = 1\}.$$

- ◊ What is the meaning of this unique linear functional in \mathbb{R}^n ?
- Let f be a nonzero linear functional on a normed space X . Then the c -level hyperplane H of f is closed if and only if f is continuous (bounded).
 - ◊ There is a correspondence between the closed hyperplanes and elements in the dual space X^* .

Separation of Convex Sets

- Given a convex set K containing an interior point and a linear variety V containing no interior points of K , there exists a closed hyperplane containing V but containing no interior point of K .

- ◇ There exists an element $\mathbf{x}^* \in X^*$ and a constant c such that $\langle \mathbf{v}, \mathbf{x}^* \rangle = c$ for all $\mathbf{v} \in V$ and $\langle \mathbf{k}, \mathbf{x}^* \rangle < c$ for all $\mathbf{k} \in K$.

- ◇ Consider the case that K is the unit sphere. Why is the above statement obvious?

- Given a convex set K with the origin as an interior point, define the *Minkowski functional* p of K on X by

$$p(\mathbf{x}) := \inf\left\{r \mid \frac{\mathbf{x}}{r} \in K, r > 0\right\}.$$

- ◇ $p(\mathbf{x})$ is the factor by which K must be expanded so as to include \mathbf{x} .

- ◇ Prove that $p(\mathbf{x})$ satisfies the sublinear functional conditions required by the Hahn-Banach theorem.

- ▷ $0 \leq p(\mathbf{x}) < \infty$.

- ▷ $p(\alpha\mathbf{x}) = \alpha p(\mathbf{x})$ for $\alpha > 0$.

- ▷ $p(\mathbf{x}_1 + \mathbf{x}_2) \leq p(\mathbf{x}_1) + p(\mathbf{x}_2)$.

- ▷ p is continuous.

- ▷ $\bar{K} = \{\mathbf{x} \mid p(\mathbf{x}) \leq 1\}$.

- Assume $\mathbf{0}$ is an interior point of K . Let $M = \text{span}(V)$.
 - ◇ V is a hyperplane in M containing no origin.
 - ◇ There is a linear functional f on M such that $V = \{\mathbf{x} \in M | f(\mathbf{x}) = 1\}$. (Why?)
 - ◇ Since V contains no interior point of K , $f(\mathbf{x}) = 1 \leq p(\mathbf{x})$ for all $\mathbf{x} \in V$
 - ◇ Note that $f(\mathbf{x}) \leq p(\mathbf{x})$ for all $\mathbf{x} \in M$ because
 - ▷ $f(\alpha\mathbf{x}) = \alpha \leq p(\alpha\mathbf{x})$ for $\mathbf{x} \in V$ and $\alpha > 0$.
 - ▷ $f(\alpha\mathbf{x}) = \alpha \leq 0 \leq p(\alpha\mathbf{x})$ for $\alpha < 0$.
 - ◇ By the Hahn-Banach theorem, there is an extension F of f from M to X with $F(\mathbf{x}) \leq p(\mathbf{x})$.

- Define

$$H := \{\mathbf{x} \in X | F(\mathbf{x}) = 1\}.$$

- ◇ $F(\mathbf{x}) \leq p(\mathbf{x})$. Hence F is continuous.
- ◇ For all interior point \mathbf{x} of K , $F(\mathbf{x}) < 1$.
- If K is a closed convex set in a normed space, then K is the intersection of all closed half-spaces containing K .
 - ◇ Associating closed hyperplanes with elements in X^* , a closed convex set in X may be regarded as a collection of elements in X^* .

Minimum Distance Problems

- The minimum distance from a point to a convex set K is equal to the maximum of the distances from the point to hyperplanes separating the point and K .
- The theory can be included in more general machinery to be introduced later.

Adjoint

- The constraints imposed in many optimization problems by differential equations, matrix equations, and so on can be described by linear operators.
- The solution of these problems almost invariably calls for the notion of an associated operator — the adjoints.
- Adjoint provide a convenient mechanism for describing the orthogonality and duality relations.

Basic Properties

- Let $B(X, Y)$ be the collection of all continuous linear operators from the normed space X to the normed space Y .

◊ Recall that if Y is complete, then $B(X, Y)$ is a Banach space with the induced norm

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

- Given $A \in B(X, Y)$, the adjoint operator $A^* : Y^* \rightarrow X^*$ is defined by the equation

$$\langle \mathbf{x}, A^* \mathbf{y}^* \rangle = \langle A\mathbf{x}, \mathbf{y}^* \rangle.$$

◊ $A^* \mathbf{y}^*$ is a linear functional on X .

◊ Since $|\langle A\mathbf{x}, \mathbf{y}^* \rangle| \leq \|\mathbf{y}^*\| \|A\mathbf{x}\| \leq \|\mathbf{y}^*\| \|A\| \|\mathbf{x}\|$, $\|A^* \mathbf{y}^*\| \leq \|A\| \|\mathbf{y}^*\|$. So $\|A^*\| \leq \|A\|$. The functional $A^* \mathbf{y}^*$ is an element in X^* .

◊ Indeed $\|A^*\| = \|A\|$. (Why?)

▷ For any nonzero $\mathbf{x}_0 \in X$, there exists $\mathbf{y}_0^* \in Y^*$, $\|\mathbf{y}_0^*\| = 1$, such that $\langle A\mathbf{x}_0, \mathbf{y}_0^* \rangle = \|A\mathbf{x}_0\|$.

▷ $\|A\mathbf{x}_0\| = \langle \mathbf{x}_0, A^* \mathbf{y}_0^* \rangle \leq \|A^* \mathbf{y}_0^*\| \|\mathbf{x}_0\| \leq \|A^*\| \|\mathbf{x}_0\|$. So $\|A\| \leq \|A^*\|$.

- If $A \in B(H, H)$ where H is a Hilbert space is such that $A = A^*$, then A is said to be *self-adjoint*.
- Some examples:

◇ Any linear transformation from \mathbb{R}^n to \mathbb{R}^n can be represented as a square matrix A . The adjoint of A is the usual transpose A^\top of A .

◇ If $X = Y = \mathcal{L}_2[0, 1]$ and

$$Ax = \int_0^1 K(t, s)x(s)ds$$

where $\int_0^1 \int_0^1 |K(s, t)|^2 ds dt < \infty$. Then

$$\langle Ax, y \rangle = \int_0^1 y(t) \int_0^1 K(s, t)x(s)ds dt = \int_0^1 x(s) \int_0^1 K(t, s)y(t)dt ds.$$

Thus

$$A^*y = \int_0^1 K(s, t)y(s)ds.$$

◇ If $X = C[0, 1]$, $Y = \mathbb{R}^n$ and

$$Ax = [x(t_1), x(t_2), \dots, x(t_n)].$$

For any given $y^* = [y_1, y_2, \dots, y_n] \in \mathbb{R}^n$,

$$\langle Ax, y^* \rangle = \sum_{i=1}^n y_i x(t_i) = \langle x, A^*y^* \rangle.$$

Note that $v = A^*y^* \in NBV[0, 1]$ and should be written as

$$\langle x, A^*y^* \rangle = \int_0^1 x(t)dv(t).$$

▷ v is constant except at the point t_i where it has a jump y_i .

Fredholm Alternative Theorem

- Let X and Y be normed space and $A \in B(X, Y)$. Then

$$[\mathcal{R}(A)]^\perp = \mathcal{N}(A^*).$$

- ◊ $\langle A\mathbf{x}, \mathbf{y}^* \rangle = \langle \mathbf{x}, A^*\mathbf{y}^* \rangle$.
- ◊ $\mathbf{y}^* \in \mathcal{R}(A)^\perp \implies A^*\mathbf{y}^* = 0^*$.
- ◊ $\mathbf{y}^* \in \mathcal{N}(A^*) \implies \mathbf{y}^* \perp A\mathbf{x}$ for all \mathbf{x} .

- Let X and Y be normed space and $A \in B(X, Y)$. Then

$$R(A^*) \subset (N(A))^\perp.$$

- If $R(A)$ is closed, then

$$R(A^*) = (N(A))^\perp.$$

- ◊ [The Hahn-Banach theorem is needed in the proof.](#)

Normal Equations

- Given two Hilbert spaces X and Y and $A \in B(X, Y)$, then for each given $\mathbf{y} \in Y$ the vector $\mathbf{x} \in X$ minimizes $\|\mathbf{y} - A\mathbf{x}\|$ if and only if

$$A^*A\mathbf{x} = A^*\mathbf{y}.$$

- ◇ The existence of a solution \mathbf{x} is not guaranteed since $R(A)$ is not necessarily closed.
 - ◇ The solution is not expected to be unique.
 - ◇ [How is the above result related to the projection theorem? Where is the idea of orthogonality?](#)
- Suppose $R(A) \subset Y$ is closed and $\mathbf{y} \in \mathcal{R}(A)$. The vector \mathbf{x} of minimum norm satisfying $A\mathbf{x} = \mathbf{y}$ is given by $\mathbf{x} = A^*\mathbf{z}$ where \mathbf{z} is any solution of $AA^*\mathbf{z} = \mathbf{y}$.

An Example

- Consider a system governed by the differential equations

$$\dot{\mathbf{x}}(t) = F\mathbf{x}(t) + \mathbf{b}u(t).$$

- ◊ It is desired to drive $\mathbf{x}(0) = \mathbf{0}$ to $\mathbf{x}(T) = \mathbf{x}_1$ via a suitable scalar control $u(t)$ while the energy

$$\int_0^T u^2(t) dt$$

is kept at minimum.

- ◊ By the variation of constant formula,

$$\mathbf{x}(T) = \int_0^T e^{F(T-t)} \mathbf{b}u(t) dt.$$

- Define the linear operator $A : \mathcal{L}_2[0, T] \rightarrow \mathbb{R}^n$ by

$$Au(t) = \int_0^T e^{F(T-t)} \mathbf{b}u(t) dt.$$

- ◊ The problem is equivalent to that of minimizing the norm of u while satisfying $Au = \mathbf{x}_1$.

- Since $R(A)$ is finite dimensional, it is closed.
 - ◊ The optimal solution is $u = A^*z$ where $AA^*z = \mathbf{x}_1$.
 - ◊ It remains to calculate the operator A^* and AA^* .
- For any $u \in \mathcal{L}_2[0, T]$ and $\mathbf{y} \in \mathbb{R}^n$,

$$\langle \mathbf{y}, Au \rangle = \mathbf{y}^\top \int_0^T e^{F(T-t)} \mathbf{b}u(t) dt = \int_0^T \mathbf{y}^\top e^{F(T-t)} \mathbf{b}u(t) dt = \langle A^*\mathbf{y}, u \rangle$$

- ◊ $A^*\mathbf{y} = \mathbf{b}^\top e^{F^\top(T-t)} \mathbf{y}$.
 - ◊ $AA^* = \int_0^T e^{F(T-t)} \mathbf{b}\mathbf{b}^\top e^{F^\top(T-t)} dt \in \mathbb{R}^{n \times n}$.
- The optimal control is given by

$$u = A^*(AA^*)^{-1} \mathbf{x}_1.$$