Applied Functional Analysis

Jinhoon Choi

Department of Chemical Engineering
Sogang University

August 1, 1999

Chapter 1

Introduction

To analyze and solve the science and engineering problems in a systematic way, they need to be converted into mathematical problems. The underlying physical, chemical or biological system is written in mathematical terms through mathematical modeling and the problem statement is translated into mathematical terms through mathematical formulation of the problem. In this lecture the most general mathematical framework for the analysis and solution of science and engineering problems will be presented with special emphasis on the application to optimization problems.

Chapter 2

Linear Space

Linear (or Vector) Space X over a Scalar Field $\mathcal{S}$ ($\mathbf{R}^n$, $\mathbf{C}^n$): set with the algebraic structure defined by the following axioms:

axioms for linear space

- Addition
$$+ : X \times X \to X : (x, y) \to x + y;$$
  1. Associative
  $$(x + y) + z = x + (y + z)$$
  2. Commutative
  $$x + y = y + x$$
  3. $\exists$ identity 0 such that
  $$x + 0 = 0 + x = x$$
  4. $\exists$ inverse $-x$ such that
  $$x + (-x) = 0$$
- Scalar Multiplication:
$$\cdot : \mathcal{S} \times X \to X : (a, x) \to ax;$$
  1. Associative
  $$(ab)x = a(bx)$$

- Distributive law:

$$(a+b)x = ax + bx$$
$$a(x+y) = ax + ay$$

<u>Linear Space Examples</u>

Canonical Example I:

$\mathcal{S}^n$: n-tuples $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ of scalars

Addition and Scalar Multiplication:

$$x + y = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad ax = \begin{bmatrix} ax_1 \\ \vdots \\ ax_n \end{bmatrix}.$$

Ex: $\mathbf{R}^n$, $\mathbf{C}^n$

Canonical Example II:

Function space: all functions $f(d)$ from a domain D to X

Addition and Scalar Multiplication

$$(f+g)(d) = f(d) + g(d)$$
$$(af)(d) = af(d).$$

Ex 1: Set of all sequences (functions from I to $\mathcal{S}^n$)

Ex 2: $C[a,b]$ Set of continuous functions on $[a,b]$.

Product Space $X \times Y$: linear space of (x,y) with addition and scalar Multiplication:

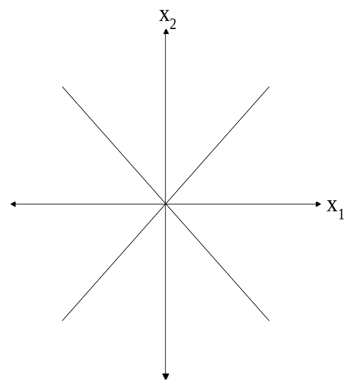
$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$a(x,y) = (ax, ay).$$

Subspace: A subset M of linear space X such that

$$\alpha x + \beta y \in M, \quad \forall x, y \in M, \quad \forall \alpha, \beta \in \mathcal{S}.$$

Notice that a subspace is a linear space by itself.

Ex 1: Let $X = \mathbf{R}^2$.



$\{(0,0\}$, any straight line and $\mathbf{R}^2$ are subspaces.

$\{(x_1,x_2):x_2=x_1\}\cup\{(x_1,x_2):x_2=-x_1\}$ is not a subspace.

Ex 2: Let $X=C[a,b]$.

$M=\{f\in X: f(a)=0\}$, subspace

$N=\{f\in X: f(b)=1\}$, not subspace

Fact: Let $\mathcal{C}$ be a collection of subspaces. Then $\cap_{M\in\mathcal{C}}M$ is a subspace.

Proof: Let $\alpha,\beta\in\mathcal{S}$ and $x,y\in\cap_{M\in\mathcal{C}}M$. Then $\alpha x+\beta y\in M$ for all $M\in\mathcal{C}$. Hence, $\alpha x+\beta y\in\cap_{M\in\mathcal{C}}M$. □

Fact: Let M and N be subspaces of X. Then $M+N=\{z:z=x+y, x\in M, y\in N\}$ is a subspace.

Linear Independence and Dimension

Linear Combination: Let $\alpha_k\in\mathcal{S}$ and $x_k\in X$. Then

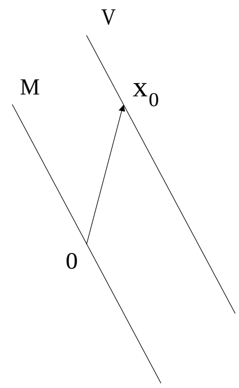
$$y=\alpha_1x_1+\cdots+\alpha_nx_n$$

is called a linear combination.

Fact: The set of all linear combinations of a set S of vectors is a subspace. Indeed it is the smallest subspace containing S and is called the subspace generated by S, denoted $span(S)$.

Linear Variety: $V=x_0+M$ where $x_0\in X$ and M is a subspace.

Ex: Let $X=C[a,b]$ and $M=\{f\in X: f(a)=0\}$.



$N = \{f \in X : f(a) = 1\} = 1 + M$, linear variety.

Def.: A set S of vectors is linearly dependent if there exists $a \in S$ such that

$$a \in span(S - \{a\}).$$

Otherwise it is called linearly independent.

Def.: A linear space is said to be finite dimensional if there exists a finite subset of vectors S such that $span(S) = X$. The number of linearly independent vectors in S is called the dimension of the space.

Notice that $\mathbf{R}^n$ and $\mathbf{C}^n$ are finite dimensional whereas the sequence space and $C[a, b]$ are infinite dimensional.

Fact: Let S and $\hat{S}$ be two linearly independent subset of vectors such that $span(S) = span(\hat{S}) = X$. Then S and $\hat{S}$ contain the same number of elements (i.e. the dimension is unique).

Proof: Let $S = \{x_1, \cdots, x_n\}$ and $\hat{S} = \{y_1, \cdots, y_m\}$. Suppose $m \geq n$ WLOG. Notice that

$$y_1 = \sum_{k=1}^{n} \alpha_k x_k.$$

Assume $\alpha_1 \neq 0$ WLOG. Then

$$x_1 = \frac{1}{\alpha_1} y_1 - \frac{1}{\alpha_1} \sum_{k=2}^{n} \alpha_k x_k.$$

Hence, $span\{y_1, x_2, \cdots, x_n\} = X$. Then

$$y_2 = \beta_1 y_1 + \sum_{k=2}^{n} \hat{\alpha}_k x_k.$$

Assume $\hat{\alpha}_2 \neq 0$ WLOG. Then

$$x_2 = \frac{1}{\hat{\alpha}_2} y_2 - \frac{1}{\hat{\alpha}_2} \beta_1 y_1 - \frac{1}{\hat{\alpha}_2} \sum_{k=3}^{n} \hat{\alpha}_k x_k$$

Hence, $span\{y_1, y_2, x_3, \cdots, x_n\} = X$.

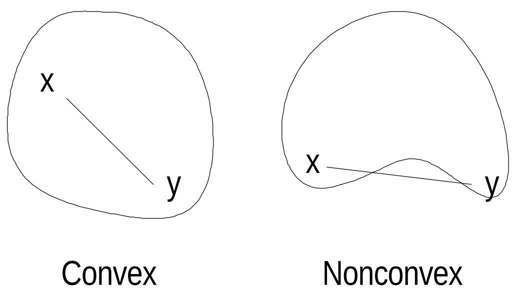
$$\vdots$$

$span\{y_1, \cdots, y_n\} = X$ and the fact follows. □

Convex Sets and Cones

A set C is said to be convex if

$$\alpha x + (1-\alpha) y \in C, \quad \forall \alpha \in [0,1],\ x, y \in C.$$



Fact: Let C be a convex set. Then

- aC is a convex set.
- If C, D are convex, so is $C + D$.

Proof: i) Let $x, y \in aC$. Then $\frac{x}{a}, \frac{y}{a} \in C$.

$$\alpha x + (1-\alpha) y = a \underbrace{\left(\alpha \frac{x}{a} + (1-\alpha) \frac{y}{a} \right)}_{\in C} \in aC.$$

ii) left as an exercise. □

Fact: Let $\{C_\alpha\}_{\alpha\in\mathcal{C}}$ be a collection of convex sets. Then $\cap_{\alpha\in\mathcal{C}}C_\alpha$ is a convex set.

Proof: Suppose $x, y \in \cap_{\alpha\in\mathcal{C}}C_\alpha$. Then $x, y \in C_\alpha$, for all $\alpha \in \mathcal{C}$. Hence, $\alpha x + (1-\alpha)y \in C_\alpha$, for all $\alpha \in \mathcal{C}$. Therefore, $\alpha x + (1-\alpha)y \in \cap_{\alpha\in\mathcal{C}}C_\alpha$. □

The convex hull (or convex cover) of a set S is the smallest convex set containing S and is denoted $co(S)$.

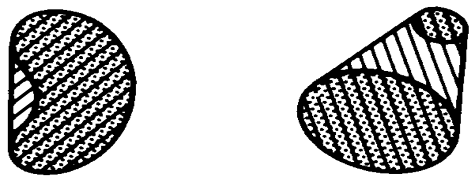


Figure 2.4 Convex hulls

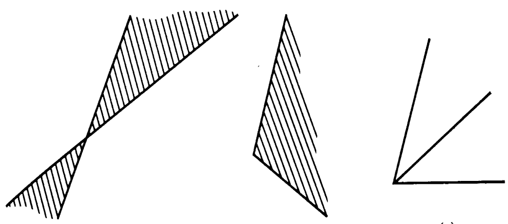
Since a linear space is convex, a subset of a linear space is always contained in a convex set.

Let $\mathcal{C}$ be the set of all convex set containing S. Then

$$co(S) = \cap_{C\in\mathcal{C}}C.$$

A set C is said to be a cone with vertex at the origin if

$$ax \in C, \quad \forall a \geq 0, \; x \in C.$$



(a) (b) (c)

Figure 2.5 Cones

Chapter 3

Topological Spaces

A topology of a linear space X is a family of subsets of X that contains $\emptyset$ and X and is closed under arbitrary unions and finite intersections.

A linear space with a topology is called a topological linear space.

Notice that a linear space has many different topology.

Ex: $\mathcal{T} = \{\emptyset, X\}$ (trivial topology)

$\mathcal{P}(X) = \{S : S \subset X\}$ (discrete topology)

Different topology defines different topological space.

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two different topology of X. Then if $\mathcal{T}_1 \subset \mathcal{T}_2$, $\mathcal{T}_1$ is said to be weaker (or coarser) than $\mathcal{T}_2$, or $\mathcal{T}_2$ is said to be stronger (or finer) than $\mathcal{T}_1$.

Notice that the discrete topology is the strongest topology a linear space can have whereas the trivial topology is the weakest.

The elements in a topology are called open sets whereas their complements are called closed sets.

The union of all open sets contained in a set S is called its interior (denoted as $\mathring{S}$) whereas the intersection of all closed sets containing the set is called its closure (denoted as $\bar{S}$). The difference between the closure and interior is called the boundary ($\partial S = \bar{S} \setminus \mathring{S}$).

Fact: A set S is closed iff $S = \bar{S}$.

Def.: A set S is called dense if $\bar{S} = X$.

Def.: A topological space is called separable if it has a countable dense subset.

Def.: A set is compact if every open cover has a finite subcover (Heine-

Borel property).

An operator (transformation or mapping) T from a subset D of a linear space X into a linear space Y is a rule that associates every elements in D to an element of Y.

Let A be an operator from a topological space X to a topological space Y. The mapping A is called continuous if the inverse image of an open set in Y is an open set in X.

Chapter 4

Metric and Complete Metric Spaces

A metric (or distance function) on X is a function from $X \times X \to \mathbf{R}$ such that for all $x, y, z \in X$,

1. $d(x,y) \geq 0$
2. $d(x,y) = 0$ iff $x = y$
3. $d(x,y) = d(y,x)$
4. $d(x,y) \leq d(x,z) + d(z,y)$ (Triangular Inequality)

A linear space equipped with a metric is called a metric space.
Notice that a linear space has many different metric.
Ex: Consider $\mathbf{R}$.
$d(x,y) = |x-y|$
$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$ (discrete metric)
Different metric defines different metric space.
Metric Space Examples
Ex 1: $\mathbf{R}^n$, $\mathbf{C}^n$ with the metric

$$d(x,y) = \sqrt{|x_1 - y_1|^2 + \cdots + |x_n - y_n|^2}.$$

Ex 2: $C[a,b]$ = the set of all continuous functions on $[a,b]$ with the metric

$$d(x,y) = \max_{a \leq t \leq b} |x(t) - y(t)|.$$

Notice that the triangular inequality follows from the fact:

$$\max_{a\le t\le b}|x(t)-y(t)| \le \max_{a\le t\le b}|x(t)-z(t)| \max_{a\le t\le b}|z(t)-y(t)|.$$

Ex 3: $C[a,b]$ with the metric

$$d(x,y)=\int_a^b |x(t)-y(t)|dt.$$

Ex 4: $l_p =$ the set of all sequences such that $\begin{cases} \sum_{j=1}^{\infty}|x_j|^p<\infty & \text{if } 1\le p<\infty \\ \sup_{1\le j<\infty}|x_j|<\infty & \text{if } p=\infty \end{cases}$

$$d(x,y)=\begin{cases} \left(\sum_{j=1}^{\infty}|x_j-y_j|^p\right)^{\frac{1}{p}} & \text{if } 1\le p<\infty \\ \sup_{1\le j<\infty}|x_j-y_j| & \text{if } p=\infty \end{cases}.$$

Ex 5: $c_0 =$ the set of all sequences that converges to zero

$$d(x,y)=\sup_{1\le j<\infty}|x_j-y_j|.$$

Notice that $c_0\subset l_\infty$.

Ex 6: $L_p[a,b] =$ the set of all measurable (integrable in the sense of Lebesque) functions on $[a,b]$ such that $\begin{cases} \int_a^b |x(t)|^p dt<\infty & \text{if } 1\le p<\infty \\ ess\sup_{a\le t\le b}|x(t)|<\infty & \text{if } p=\infty \end{cases}$

$$d(x,y)=\begin{cases} \left(\int_a^b |x(t)-y(t)|^p dt\right)^{\frac{1}{p}} & \text{if } 1\le p<\infty \\ ess\sup_{a\le t\le b}|x(t)-y(t)| & \text{if } p=\infty \end{cases}.$$

Ex 7: $H_p(D) =$ the set of all analytic functions on D such that $\begin{cases} \int_D |x(t)|^p dt<\infty & \text{if } 1\le p<\infty \\ ess\sup_{t\in D}|x(t)|<\infty & \text{if } p=\infty \end{cases}$

$$d(x,y)=\begin{cases} \left(\int_a^b |x(t)-y(t)|^p dt\right)^{\frac{1}{p}} & \text{if } 1\le p<\infty \\ ess\sup_{t\in D}|x(t)-y(t)| & \text{if } p=\infty \end{cases}.$$

Open and Closed Sets

- Open Ball: $\overset{\circ}{B}_r(x_0)=\{x: d(x,x_0)<r\}$

- Closed Ball: $B_r(x_0) = \{x : d(x, x_0) \leq r\}$

- Sphere: $S_r(x_0) = \{x : d(x, x_0) = r\}$

A set is called open if it contains an open ball (also called a neighborhood) about each of its points. A set is called closed if its complement is open.

Fact: Let X be a metric space.

1. $\emptyset$ and X are open.

2. Any union of open sets is open.

3. Any finite intersection of open sets is open.

Proof: 1) $\emptyset$ has no elements and thus is open. X is clearly open.

2) Let $x \in \cup_{a \in A} O_a$ where O_a is open. Then there exists $a' \in A$ such that $x \in O_{a'}$. This implies there exists $r > 0$ such that $\overset{\circ}{B}_r(x) \subset O_{a'} \subset \cup_{a \in A} O_a$.

3) Let $x \in \cap_{i=1}^N O_i$ where O_i is open. Then there exists $r_i > 0$ such that $\overset{\circ}{B}_{r_i}(x) \subset O_i$. Let $r = \min_{1 \leq i \leq N} r_i$. Then $\overset{\circ}{B}_r(x) \subset O_i$ for all $i = 1, \cdots, N$. Hence $\overset{\circ}{B}_r(x) \subset \cap_{i=1}^N O_i$. □

From this fact, the set of all open sets defines a topology. Hence, a metric space is a topological space but not vice versa. Hence all the properties of topological spaces carry over to the metric spaces. A topology that can be generated from a metric is called metrizable.

Ex: Consider $\mathbf{R}$ equipped with the discrete metric. Suppose S be a subset of $\mathbf{R}$. For all $x \in S$, $\overset{\circ}{B}_{\frac{1}{2}}(x) = \{x\} \in S$. Hence, all the subsets of $\mathbf{R}$ are open. To this end, the discrete metric defines the discrete topology.

$\overset{\circ}{B}_\epsilon(x_0)$ is called the ϵ neighborhood of x_0. Moreover a set containing an ϵ neighborhood of x_0 is called the neighborhood of x_0.

Fact: The interior of a set M is the set of all points for which M is its neighborhood.

A point x_0 is called a limit point of a set M if every neighborhood of x_0 contains a point in M.

Fact: The closure of a set M is the set of all points that is a limit point of M (Otherwise $x_0 \in \bar{M}^c$).

Fact: A subset M of a metric space X is dense iff every point in X is a limit point in M.

Ex: From the Taylor's theorem, the set of polynomials in a space of bounded and continuous functions on a compact set is a dense subset.

Fact: A metric space X is separable iff there exists a countable set M such that every point in X is a limit point in M.

Fact: Let X and Y be metric spaces. A mapping $T : X \to Y$ is continuous iff, at each point x_0 of X, for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$d_X(x, x_0) < \delta \quad \text{implies} \quad d_Y(Tx, Tx_0) < \epsilon.$$

Proof: ($\Leftarrow$) Let $S \subset Y$ be open and S_0 its inverse image. If $S_0 = \emptyset$, then it is open. Suppose $S_0 \neq \emptyset$. For any $x_0 \in S_0$, let $y_0 = Tx_0$. Since S is open, it contains an ϵ neighborhood N of y_0. Then x_0 has a δ neighborhood N_0 which is mapped into N. Since $N \subset S$, we have $N_0 \subset S_0$. Hence, S_0 is open.

($\Rightarrow$) For every $x_0 \in X$ and any ϵ neighborhood N of Tx_0, the inverse image N_0 of N is open and contains x_0. Hence N_0 contains a δ neighborhood of x_0 whose image is contained in N. □

Convergence

A sequence in a metric space X is said to converge or to be convergent if there exists an $x \in X$ (called a limit) such that

$$\lim_{n \to \infty} d(x, x_n) = 0.$$

Fact: A sequence is convergent iff for every $\epsilon > 0$ there exists N such that

$$d(x, x_n) < \epsilon, \quad \forall n > N.$$

A nonempty set is bounded if

$$\sup_{x,y \in M} d(x, y) < \infty.$$

Fact:

1. A convergent sequence is bounded and its limit point is unique.
2. If $x_n \to x$ and $y_n \to y$ in X, then $d(x_n, y_n) \to d(x, y)$.

Proof: 1) Suppose $x_n \to x$. Let $\epsilon = 1$. Then there exists N such that $d(x_n, x) < \epsilon$ for all $n > N$. Let $a = \max\{d(x_1, x), \cdots, d(x_N, x)\}$. Then

$$d(x_n, x) < \max\{1, a\}, \quad \forall n \geq 1$$

and thus the sequence is bounded.

Suppose $x_n \to x$ and $x_n \to z$. Then

$$0 \leq d(x,z) \leq d(x,x_n) + d(x_n,z) \to 0$$

and thus $x = z$.

2) Notice that

$$d(x_n,y_n) \leq d(x_n,x) + d(x,y) + d(y,y_n).$$

Hence

$$d(x_n,y_n) - d(x,y) \leq d(x_n,x) + d(y,y_n).$$

Similarly

$$d(x,y) - d(x_n,y_n) \leq d(x_n,x) + d(y,y_n).$$

To this end, as $n \to \infty$,

$$|d(x_n,y_n) - d(x,y)| \leq d(x_n,x) + d(y,y_n) \to 0.$$

□

Fact: Let $\bar{M}$ be the closure of a nonempty set M. Then

1. $x \in \bar{M}$ iff there exists a sequence $\{x_n\}$ in M such that $x_n \to x$.

2. M is closed iff $x_n \in M$ and $x_n \to x$ implies $x \in M$.

Proof: 1) ($\Rightarrow$) If $x \in M$, then pick the sequence $\{x, x, x, \cdots\}$. Otherwise it is an accumulation point of M. Hence for each $n = 1, 2, \cdots$, $B_{\frac{1}{n}}(x)$ contains an $x_n \in M$ and thus $x_n \to x$.

($\Leftarrow$) $x \in M$ or every neighborhood of x contains $x_n \neq x$. Hence x is an accumulation point of M and thus $x \in \bar{M}$.

2) follows from the fact that M is closed iff $M = \bar{M}$. □

Hiene-Borel Theorem: A subset M of a metric space X is compact iff every sequence in M has a convergent subsequence with its limit in M.

Proof: See any real analysis book. □

Fact: A compact subset M of a metric space X is closed and bounded but not vice versa. However, the converse is true for any finite dimensional X.

Proof: See Kreyszig's book (p.77). □

Fact: Let X and Y be metric spaces. A mapping $T : X \to Y$ is continuous at $x_0 \in X$ iff $x_n \to x_0$ implies $Tx_n \to Tx_0$.

Proof: ($\Rightarrow$) For a given $\epsilon > 0$, there exists $\delta > 0$ such that

$$d_X(x, x_0) < \delta \quad \text{implies} \quad d_Y(Tx, Tx_0) < \epsilon.$$

Let $x_n \to x_0$. Then there exists N such that for all $n > N$, we have

$$d_X(x_n, x_0) < \delta.$$

Hence for all $n > N$,

$$d_Y(Tx_n, Tx_0) < \epsilon.$$

($\Leftarrow$) Suppose the contrary. Then there exists $\epsilon > 0$ such that for every $\delta > 0$ there exists $x \neq x_0$ such that

$$d_X(x, x_0) < \delta \quad \text{but} \quad d_Y(Tx, Tx_0) \geq \epsilon.$$

Pick $\delta = \frac{1}{n}$. Then there exists $x_n \neq x_0$ such that

$$d_X(x_n, x_0) < \frac{1}{n} \quad \text{but} \quad d_Y(Tx_n, Tx_0) \geq \epsilon.$$

Clearly $x_n \to x_0$ but $Tx_n \not\to Tx$ (Contradiction!). □

Cauchy Sequence and Completeness

A sequence in a metric space is Cauchy if for every $\epsilon > 0$ there exists N such that

$$d(x_n, x_m) < \epsilon \quad \forall n, m \geq N.$$

The space is complete if all its Cauchy sequences are convergent.

We now show that the converse is always true.

Fact: Every convergent sequence is Cauchy.

Proof: Suppose $x_n \to x$. Then for every $\epsilon > 0$ there exists N such that

$$d(x_n, x) < \frac{\epsilon}{2} \quad \forall n \geq N.$$

Hence,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Fact: A subspace M of a complete metric space is complete iff M is closed in X.

Proof: ($\Rightarrow$) For every $x \in \bar{M}$, there exists $\{x_n\}$ in M such that $x_n \to x$. Since $\{x_n\}$ is Cauchy and M is complete, it converges to the unique limit $x \in M$.

($\Leftarrow$) Let $\{x_n\}$ is Cauchy in M. Then $x_n \to x \in X$ and thus $x \in \bar{M} = M$.

□

Examples of Complete Metric Spaces

Before we proceed to the examples, we present a lemma.

Lemma: Let $\{x_l\}$ be a sequence in $\mathbf{R}$. Suppose

$$\lim_{l\to\infty} \sup_{j\geq l} x_j = \lim_{l\to\infty} \inf_{j\geq l} x_j = c \in \mathbf{R}.$$

Then $\{x_l\}$ is convergent and it limit is c.

Proof: For every $\epsilon > 0$, there exists N_1 such that

$$|\sup_{j\geq l} x_j - c| < \epsilon, \qquad l \geq N_1.$$

Hence,

$$x_l - c < \epsilon, \qquad l \geq N_1.$$

Similarly for every $\epsilon > 0$, there exists N_2 such that

$$|\inf_{j\geq l} x_j - c| < \epsilon, \qquad l \geq N_2.$$

Hence,

$$x_l - c > -\epsilon, \qquad l \geq N_2.$$

Setting $N = \max\{N_1, N_2\}$,

$$|x_l - c| < \epsilon, \qquad l \geq N.$$

□

Ex 1: **R**

Proof: Consider a Cauchy sequence $\{x_l\}$ in $\mathbf{R}$. First notice that

$$\lim_{l\to\infty} \inf_{j\geq l} x_j \leq \lim_{l\to\infty} \sup_{j\geq l} x_j.$$

Hence from the lemma, it suffices to show

$$\lim_{l\to\infty} \inf_{j\geq l} x_j \geq \lim_{l\to\infty} \sup_{j\geq l} x_j.$$

Now for every $\epsilon > 0$, there exists an N such that

$$d(x_m, x_l) < \frac{\epsilon}{2}, \qquad m, l \geq N$$

and thus

$$d(x_N, x_l) < \frac{\epsilon}{2}, \qquad l \geq N.$$

Hence, $x_N + \frac{\epsilon}{2}$ and $x_N - \frac{\epsilon}{2}$ are upper and lower bounds for $\{x_N, x_{N+1}, \cdots\}$ and thus for $\{x_l, x_{l+1}, \cdots\}$, $\forall l \geq N$. Hence for all $l \geq N$,

$$x_N - \frac{\epsilon}{2} \leq \inf_{j\geq l} x_j \leq \sup_{j\geq l} x_j \leq x_N + \frac{\epsilon}{2}.$$

Therefore

$$\sup_{j\geq l} x_j \leq \inf_{j\geq l} x_j + \epsilon$$

and thus

$$\lim_{l\to\infty} \sup_{j\geq l} x_j \leq \lim_{l\to\infty} \inf_{j\geq l} x_j + \epsilon.$$

Since ϵ can be arbitrarily small,

$$\lim_{l\to\infty} \sup_{j\geq l} x_j \leq \lim_{l\to\infty} \inf_{j\geq l} x_j.$$

□

Ex 2: **C**

Ex 3: $\mathbf{R}^n$

Proof: Consider a Cauchy sequence $\{x_l = [x_1^{(l)} \ \cdots \ x_n^{(l)}]^T\}$ in $\mathbf{R}^n$. Then for every $\epsilon > 0$, there exists an N such that

$$d(x_m, x_l) = \sqrt{|x_1^{(m)} - x_1^{(l)}|^2 + \cdots + |x_n^{(m)} - x_n^{(l)}|^2} < \epsilon, \qquad m, l \geq N.$$

Then we have for $j = 1, \cdots, n$,

$$|x_j^{(m)} - x_j^{(l)}| < \epsilon, \qquad m, l \geq N.$$

To this end, $\{x_j^{(l)}\}$ is a Cauchy sequence in $\mathbf{R}$ and thus converges to a limit, say x_j. Let $x = [x_1 \ \cdots \ x_n]^T \in \mathbf{R}^n$. Then

$$d(x_m, x) \leq \epsilon, \quad m \geq N.$$

□

Ex 4: $\mathbf{C}^n$

Ex 5: l_∞

Proof: Consider a Cauchy sequence $\{x_l = (x_1^{(l)}, \ x_2^{(l)} \ \cdots)\}$ in l_∞. Then for every $\epsilon > 0$, there exists an N such that

$$d(x_m, x_l) = \sup_j |x_j^{(m)} - x_j^{(l)}| < \epsilon, \quad m, l \geq N.$$

Hence for every j

$$|x_j^{(m)} - x_j^{(l)}| < \epsilon, \quad m, l \geq N.$$

To this end, $\{x_j^{(l)}\}$ is a Cauchy sequence in $\mathbf{R}$ and thus converges to a limit, say x_j. Let $x = (x_1 \ x_2 \ \cdots)$. We now show that $x \in l_\infty$ and $x_l \to x$. As $l \to \infty$,

$$|x_j^{(m)} - x_j| \leq \epsilon, \quad m \geq N.$$

Since $x_l \in l_\infty$, there exists K_l such that $|x_j^{(l)}| \leq K_l$ for all j. Hence for every j

$$|x_j| \leq |x_j - x_j^{(l)}| + |x_j^{(l)}| \leq \epsilon + K_l, \quad l \geq N.$$

Hence $\{x_j\}$ is a bounded sequence in $\mathbf{R}$. Hence $x \in l_\infty$. Also

$$d(x_m, x) = \sup_j |x_j^{(m)} - x_j| \leq \epsilon, \quad m \geq N.$$

Hence, $x_l \to x$. □

Ex 6: c_0

Ex 7: l_p

Ex 8: $C[a, b]$

Proof: Consider a Cauchy sequence $\{x_l\}$ in $C[a, b]$. Then for every $\epsilon > 0$, there exists an N such that

$$d(x_m, x_l) = \max_{a \leq t \leq b} |x_m(t) - x_l(t)| < \epsilon, \quad m, l \geq N.$$

Hence for every $t_0 \in [a,b]$

$$|x_m(t_0) - x_l(t_0)| < \epsilon, \qquad m, l \geq N.$$

To this end, $\{x_l(t_0)\}$ is a Cauchy sequence in $\mathbf{R}$ and thus converges to a limit, say $x(t_0)$. We now show that $x \in C[a,b]$ and $x_l \to x$. As $l \to \infty$,

$$\max_{a \leq t \leq b} |x_m(t) - x(t)| \leq \epsilon, \qquad m \geq N.$$

Hence for every $t \in [a,b]$

$$|x_m(t) - x(t)| \leq \epsilon, \qquad m \geq N$$

and thus $x_l \to x$ uniformly on $[a,b]$. From calculus, the limit of uniformly converging sequence of continuous function is continuous. Hence, $x \in C[a,b]$ and $x_l \to x$. □

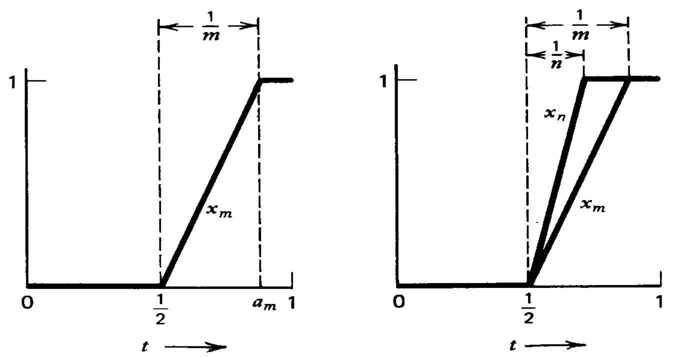
Ex 9: $L_p[a,b]$

Ex 10: $H_p(D)$

Example of Incomplete Metric Spaces

Ex: $C[0,1]$ with $d(x,y) = \int_0^1 |x(t) - y(t)| dt$

Proof: Consider a sequence $\{x_l\}_{l=2}^{\infty}$:



Notice that this sequence is Cauchy since for every $\epsilon > 0$,

$$d(x_m, x_l) = \frac{1}{2}\frac{1}{m} + \left(\frac{1}{l} - \frac{1}{m}\right) - \frac{1}{2}\frac{1}{l} = \frac{1}{2}\left(\frac{1}{l} - \frac{1}{m}\right) \leq \frac{1}{l} < \epsilon \qquad \forall m, l > \frac{1}{\epsilon}.$$

We now show that the sequence doesn't converge. Notice that

$$x_l(t) = \begin{cases} 0 & \text{if } t \in [0, \frac{1}{2}] \\ 1 & \text{if } t \in [a_l, 1] \end{cases}$$

where $a_l = \frac{1}{2} + \frac{1}{l}$. Hence for every $x \in C[a, b]$,

$$d(x_l, x) = \int_0^1 |x(t) - y(t)| dt$$

$$= \int_0^{\frac{1}{2}} |x(t)| dt + \int_{\frac{1}{2}}^{a_l} |x_l(t) - x(t)| dt = \int_{a_l}^1 |1 - x(t)| dt.$$

Hence $d(x_l, x) \to 0$ implies each integral approaches to zero and, since x is continuous, we should have

$$x(t) = \begin{cases} 0 & \text{if } t \in [0, \frac{1}{2}] \\ 1 & \text{if } t \in (\frac{1}{2}, 1] \end{cases} .$$

Hence the limit is not in $C[a, b]$ and thus the sequence doesn't converge. □

Chapter 5

Normed and Banach Spaces

A norm (or size function) on X is a function from $X \to \mathbf{R}$ such that for all $x, y \in X$,

1. $\|x\| \geq 0$
2. $\|x\| = 0$ iff $x = 0$
3. $\|ax\| = |a|\|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$

$d(x, y) = \|x - y\|$ is a metric. However, a norm may not be obtained from a metric.

Ex: Consider the discrete metric. Let

$$\|x\| = d(x, 0).$$

Then for $x \neq 0$ and $a \notin \{0, 1\}$, $a > 0$,

$$\|ax\| = d(ax, 0) = 1 \neq a = a\|x\|.$$

Hence norm cannot be obtained from the discrete metric.

A linear space equipped with a norm is called a normed space and a complete normed space is called a Banach space.

A subspace Y of a Banach space is complete iff Y is closed in X.

Notice that a normed space is a metric space but the converse is not true in general. Hence all the properties of metric and topological spaces carry over to the normed spaces.

Notice that a linear space has many different norms.
Ex: Consider $\mathbf{R}^2$.
$\|x\|_2 = \sqrt{x_1^2 + x_2^2}$
$\|x\|_\infty = \max\{|x_1|, |x_2|\}$
Different norm defines different normed space.
Fact: A norm is a continuous mapping.
Proof: From the triangular inequality, it holds that

$$|\|y\| - \|x\|| \leq \|y - x\|$$

and thus the fact follows. □

Example of incomplete normed space: $C[0,1]$ with $\|x\| = \int_0^1 |x(t)|dt$.
Examples of Banach spaces:
Ex 1: $\mathbf{R}^n$, $\mathbf{C}^n$ with the p-norm:

$$\|x\|_p = \begin{cases} (\sum_k |x_k|^p)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \max_k |x_k| & \text{if } p = \infty \end{cases}.$$

$\|\cdot\|_2$ is called the Euclidean norm.
Holder's Inequality: Let $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\sum_{k=1}^{n} |x_k y_k| \leq \|x\|_p \|y\|_q.$$

$p = q = 2$ case of Holder's inequality is called Cauchy-Schwarz inequality.
Minkowski Inequality: If $x, y \in l_p$ for $p \in [1, \infty]$,

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Ex 2: $C[a, b]$ with

$$\|x\| = \max_{a \leq t \leq b} |x(t)|.$$

Ex 3: c_0 with

$$\|x\| = \sup_{1 \leq j < \infty} |x_j|.$$

Ex 4: l_p with

$$\|x\|_p = \begin{cases} \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \sup_{1 \leq j < \infty} |x_j| & \text{if } p = \infty \end{cases}.$$

Holder's Inequality: Let $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Suppose $x \in l_p$, $y \in l_q$. Then

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_p \|y\|_q.$$

Proof: $p = 1, \infty$ cases are trivial. Let $p \in (1, \infty)$. Suppose $a > 0, b > 0$. Since log is concave, for any $0 < \lambda < 1$,

$$\lambda \log a + (1-\lambda) \log b \leq \log(\lambda a + (1-\lambda) b)$$

$$\Updownarrow$$

$$a^{\lambda} b^{(1-\lambda)} \leq \lambda a + (1-\lambda) b, \quad \forall a, b \geq 0.$$

Let

$$a = \left(\frac{|x_k|}{\|x\|_p}\right)^p, \quad b = \left(\frac{|y_k|}{\|y\|_q}\right)^q, \quad \lambda = \frac{1}{p}, \quad 1-\lambda = \frac{1}{q}.$$

Then

$$\left(\frac{|x_k|}{\|x\|_p}\right)\left(\frac{|y_k|}{\|y\|_q}\right) = a^{\lambda} b^{(1-\lambda)} \leq \lambda a + (1-\lambda) b = \frac{1}{p}\left(\frac{|x_k|}{\|x\|_p}\right)^p + \frac{1}{q}\left(\frac{|y_k|}{\|y\|_q}\right)^q$$

and thus

$$\sum_{k=1}^{\infty}\left(\frac{|x_k y_k|}{\|x\|_p \|y\|_q}\right) \leq \sum_{k=1}^{\infty}\left[\frac{1}{p}\left(\frac{|x_k|}{\|x\|_p}\right)^p + \frac{1}{q}\left(\frac{|y_k|}{\|y\|_q}\right)^q\right] = \frac{1}{p} + \frac{1}{q} = 1.$$

□

$p = q = 2$ case of Holder's inequality is called Cauchy-Schwarz inequality.

Minkowski Inequality: If $x, y \in l_p$ for $p \in [1, \infty]$,

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Proof: Consider the finite sum:

$$\sum_{k=1}^{n} |x_k + y_k|^p \leq \sum_{k=1}^{n} |x_k + y_k|^{p-1} |x_k| + \sum_{k=1}^{n} |x_k + y_k|^{p-1} |y_k|.$$

By Holder's inequality,

$$\sum_{k=1}^{n} |x_k + y_k|^p \leq \left[\sum_{k=1}^{n} \left(|x_k + y_k|^{p-1}\right)^q\right]^{\frac{1}{q}} \left[\left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}}\right].$$

Since

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Rightarrow \quad p + q = pq \quad \Rightarrow \quad pq - q = p,$$

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}}.$$

This holds for all $n < \infty$. □

Ex 5: $L_p[a,b]$ with

$$\|x\|_p = \begin{cases} \left(\int_a^b |x(t)|^p dt\right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ ess \sup_{a \leq t \leq b} |x(t)| & \text{if } p = \infty \end{cases}.$$

Holder's Inequality:

$$\int_a^b |x(t)y(t)| \leq \|x\|_p \|y\|_q.$$

Minkowski Inequality:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Ex 6: $H_p(D)$ with

$$\|x\|_p = \begin{cases} \left(\int_a^b |x(t)|^p dt\right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ ess \sup_{t \in D} |x(t)| & \text{if } p = \infty \end{cases}.$$

Holder's Inequality:

$$\int_D |x(t)y(t)| \leq \|x\|_p \|y\|_q.$$

Minkowski Inequality:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Fact: Every finite dimensional normed space is complete and thus is a Banach space.

Corollary: Every finite dimensional subspace Y of a normed space X is complete. Since Y is contained in a larger finite dimensional and thus complete subspace, it is closed in X (because Y is complete iff it is closed).

A norm $\|\cdot\|$ is said to be equivalent to another norm $\|\cdot\|_0$ if there exists a, b such that

$$a\|x\|_0 \leq \|x\| \leq b\|x\|_0.$$

Notice that equivalent norms defines the same topology.

Fact: On a finite dimensional space, any norm is equivalent to any other.

Existence of Optimal Solution

Theorem (Weierstrass): A continuous functional on a compact subset K of a normed space X achieves a maximum on K.

Proof: Let $M = \sup_{x \in K} f(x)$. There exists a sequence $\{x_k\}_{k=1}^{\infty}$ such that $f(x_k) \to M$. Since K is compact, there exists a converging subsequence $\{x_{k_i}\}$. Let the limit of subsequence be $x \in K$. Since f is continuous, $f(x_{k_i}) \to M$. $\square$

Chapter 6

Inner Product and Hilbert Spaces

An inner product on X is a function from $X \times X \to \mathbf{R}$ such that for all $x, y, z \in X$,

1. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
2. $\langle ax, y\rangle = a\langle x, y\rangle$
3. $\langle x, y\rangle = \overline{\langle y, x\rangle}$
4. $\langle x, x\rangle \geq 0$
5. $\langle x, x\rangle = 0$ iff $x = 0$

From the definition of inner product, it is clear that $\langle x, x\rangle$ is real and

$$\langle x, ay\rangle = \bar{a}\langle x, y\rangle.$$

A linear space equipped with an inner product is called an inner product (or pre-Hilbert) space and a complete inner product space is called a Hilbert space.

Def.: $\|x\| = \sqrt{\langle x, x\rangle}$.

Cauchy-Schwartz Inequality: Let X be an inner product space. For any $x, y \in X$,

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Proof: If $y = 0$, trivial. Suppose $y \neq 0$. Consider

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \langle x, \lambda y \rangle - \langle \lambda y, x \rangle + \langle \lambda y, \lambda y \rangle.$$

Define $\lambda = \frac{\langle x,y \rangle}{\langle y,y \rangle}$.

$$0 \leq \langle x, x \rangle - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle}\langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle}\langle y, x \rangle + \left| \frac{\langle x, y \rangle}{\langle y, y \rangle} \right|^2 \langle y, y \rangle = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}.$$

□

Fact: $\|x\| = \sqrt{\langle x, x \rangle}$ is a norm.
Proof:

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$$
$$\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

□

Examples of Hilbert Spaces:

1. $\mathbf{R}^n$, $\mathbf{C}^n$

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i.$$

2. l_2

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i.$$

3. $L_2[a, b]$

$$\langle x, y \rangle = \int_a^b x(t)y(t)dt.$$

4. $H_2(D)$

$$\langle x, y \rangle = \int_D x(t)\bar{y}(t)dt.$$

Parallelogram law: In an inner product space,

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

Proof: Exercise. □

Continuity of Inner Product: If $x_n \to x$ and $y_n \to y$ in an inner product space, $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

Proof: Read the textbook. $\square$

Orthogonality and Projection Theorems

Def.: Two elements of an inner product space are said to be orthogonal each other if $\langle x, y \rangle = 0$.

Def.: Let M be a subset of an inner product space X, and $x \in X$. Then x is said to be orthogonal to M if $x \perp m$ for all $m \in M$.

Pythagorean Theorem: If $x \perp y$, $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Proof:

$$\|x+y\|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + 2Re \underbrace{\langle x, y \rangle}_{=0} + \langle y, y \rangle = \|x\|^2 + \|y\|^2.$$

$\square$

Theorem: Let X be an inner product space. Let M be a linear subspace. Let $x \in X$. Suppose there exists an elements $m_0 \in M$ such that

$$\|x - m_0\| \leq \|x - m\|, \qquad \forall m \in M.$$

Then m_0 is the unique minimum of $\|x - m\|$. Moreover m_0 is the unique minimum of $\|x - m\|$ iff

$$(x - m_0) \perp M.$$

Proof: To prove the theorem, it suffices to show

1) If m_0 is a minimum, $(x - m_0) \perp M$.

2) $(x - m_0) \perp M$ implies m_0 is the unique minimum.

(1) ($\Rightarrow$) Suppose there exists $m \in M$ such that $\langle x - m_0, m \rangle = \delta \neq 0$. WLOG assume $\|m\| = 1$. Define $m_1 = m_0 + \delta m$. Then

$$\|x - m_1\|^2 = \|x - m_0 - \delta m\|^2 = \|x - m_0\|^2 - \langle x - m_0, \delta m \rangle - \langle \delta m, x - m_0 \rangle + |\delta|^2$$

$$= \|x - m_0\|^2 - |\delta|^2 < \|x - m_0\|^2.$$

Hence, m_0 is not the minimum. This is a contradiction.

(2) For any $m \in M$, Pythagorean theorem dictates that

$$\|x - m\|^2 = \|x - m_0 + \underbrace{m_0 - m}_{\in M}\|^2 = \|x - m_0\|^2 + \|m_0 - m\|^2.$$

Hence $\|x - m_0\| < \|x - m\|$ for $m \neq m_0$. $\square$

Nonexistence of m_0: Let $H = l_2$ and

$$M = \{x \in l_2 \text{ with finite number of nonzero elements}\} \quad \text{(not closed)} .$$

Consider $x = \left\{\frac{1}{2^k}\right\}$. We cannot find such m_0.

Theorem: Let X be a Hilbert space and M be a closed linear subspace. Then for any $x \in X$, there exists unique $m_0 \in M$ such that

$$\|x - m_0\| \leq \|x - m\|, \qquad \forall m \in M.$$

Moreover, m_0 is the unique minimum of $\|x - m\|$ iff

$$(x - m_0) \perp M.$$

Proof: In view of the above theorem, it suffices to show the existence of minimum. Let $\{m_k\} \subset M$ such that

$$\|x - m_k\| \to \delta = \inf_{m \in M} \|x - m\| \leq \|x\|.$$

From the parallelogram law,

$$\|(x-m_j)-(x-m_k)\|^2 + \|(x-m_j)+(x-m_k)\|^2 = 2\|x-m_j\|^2 + 2\|x-m_k\|^2.$$

Hence

$$\|m_k - m_j\|^2 = 2\|x - m_j\|^2 + 2\|x - m_k\|^2 - \underbrace{\|2x - m_j - m_k\|^2}_{4\|x - \frac{m_j + m_k}{2}\|}$$

$$\leq 2\|x - m_j\|^2 + 2\|x - m_k\|^2 - 4\delta^2 \to 0.$$

To this end, $\{m_k\}$ is a Cauchy sequence. Since X is a Hilbert space, there exists $m_0 \in X$ such that $\lim_{k\to\infty} m_k = m_0$. Since M is closed, $m_0 \in M$. Moreover,

$$\|x - m_0\| = \lim_{k \to \infty} \|x - m_k\| = \delta.$$

$\square$

<u>Orthogonal Complement</u>

Def.: Let S be a subset of an inner product space H. The collection of all vectors orthogonal to S is called the orthogonal complement of S denoted by $S^\perp$.

Fact: Let S, T be subsets of an inner product space H. Then

1. $S^{\perp}$ is a closed subspace.

2. $S \subset S^{\perp\perp}$.

3. If $S \subset T$, $T^{\perp} \in S^{\perp}$.

4. $S^{\perp\perp\perp} = S^{\perp}$.

5. If H is a Hilbert space, $S^{\perp\perp} = \overline{span(S)}$.

Proof: Exercise. □

Def.: Let H be a linear space and let M, N be two subspaces of H. Then $M \oplus N$ denotes the subspace given by $M + N$ if $M \cap N = \{0\}$.

Theorem: Let H be a Hilbert space and M be its closed subspace. Then

1. $H = M \oplus M^{\perp}$.

2. $M = M^{\perp\perp}$.

Proof: 1) It is clear that $M \oplus M^{\perp} \subset H$. Hence we only show $H \subset M \oplus M^{\perp}$. By the projection theorem, for all $x \in H$, there exists $m_0 \in M$ such that $x - m_0 \perp M$. Then let $y = x - m_0 \in M^{\perp}$. Hence $x = y + m_0$.

2) From the above fact, it suffices to show $M^{\perp\perp} \subset M$. Suppose $x \in M^{\perp\perp} subset H$. By the first part, $x = m + n$ where $m \in M \subset M^{\perp\perp}$ and $n \in M^{\perp}$. Then $n = x - m \in M^{\perp\perp}$. However $n \in M^{\perp}$ and thus $n = 0$. □

<u>Gram-Schmidt Orthonormalization</u>

Fact: An orthogonal set S of nonzero vectors is also linearly independent.

Proof: If S is not linearly independent, there exists $y \in S$ such that

$$y = \sum_{k=1}^{n} a_k y_k$$

where $y_k \in S \setminus \{y\}$ for all k, and $a_k \neq 0$ for some k. Then for all k,

$$0 = \langle y, y_k \rangle = a_k \langle y_k, y_k \rangle.$$

Since $y_k \neq 0$, this is a contradiction. □

Theorem: Let $\{x_k\}_{k=1}^{\infty}$ be a linearly independent collection of vectors in a Hilbert space H. There exists an orthonormal collection $\{e_k\}_{k=1}^{\infty}$ of vectors such that for all n,

$$span\{x_k\}_{k=1}^{n} = span\{e_k\}_{k=1}^{n}.$$

Proof: Let

$$e_1 = \frac{x_1}{\|x_1\|}$$

and for all $k > 1$,

$$z_k = x_k - \sum_{i=1}^{k-1} \langle x_k, e_i \rangle e_i,$$

$$e_k = \frac{z_k}{\|z_k\|}.$$

Then

$$x_1 = \|x_1\| e_1$$

and for all $k > 1$,

$$x_k = \sum_{i=1}^{k-1} \langle x_k, e_i \rangle e_i + \|z_k\| e_k.$$

Hence,

$$span\{x_k\}_{k=1}^n = span\{e_k\}_{k=1}^n.$$

Notice that

$$\|e_k\| = 1.$$

Moreover,

$$\langle e_2, e_1 \rangle = \frac{1}{\|z_2\|} \langle x_2 - \langle x_2, e_1 \rangle e_1, e_1 \rangle = 0.$$

Assume that we have orthogonality of $\{e_i\}_{i=1}^k$. Then for $i = 1, \cdots, k$,

$$\langle e_{k+1}, e_i \rangle = \frac{1}{\|z_{k+1}\|} \langle x_{k+1} - \sum_{j=1}^{k} \langle x_{k+1}, e_j \rangle e_j, e_i \rangle = 0.$$

Hence by induction, $\{e_k\}_{k=1}^n$ is orthonormal. □

Least Square Approximation

Consider $\{y_1, \cdots, y_n\}$ of vectors in a Hilbert space H. Let $M = span\{y_1, \cdots, y_n\}$.

Question: Given $x \in H$, find m_0 such that

$$\|x - m_0\| \leq \|x - m\|, \qquad \forall m \in M.$$

From the projection theorem, it suffices to find $m_0 = \sum_{i=1}^n a_i y_i$ such that

$$\langle x - m_0, m \rangle = 0, \qquad \forall m \in M.$$

$$\Updownarrow$$

$$\langle x - m_0, y_i \rangle = 0, \qquad i = 1, \cdots, n.$$

$$\Updownarrow$$

$$\langle x, y_i \rangle - \sum_{j=1}^{n} a_j \langle y_j, y_i \rangle = 0, \qquad i = 1, \cdots, n.$$

Define

$$[G(y_1, \cdots, y_n)]_{ij} = \langle y_i, y_j \rangle.$$

This matrix is called the normal matrix for $y_1, \cdots, y_n$. Let

$$g(y_1, \cdots, y_n) = det G(y_1, \cdots, y_n).$$

Then it is clear that $g(y_1, \cdots, y_n) \neq 0$ iff $y_1, \cdots, y_n$ are linearly independent. Assuming $y_1, \cdots, y_n$ are linearly independent

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = G^{-1}(y_1, \cdots, y_n) \begin{bmatrix} \langle x, y_1 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix}.$$

Fact: Let $x \in H$. Then

$$\|x - m_0\| = \delta^2 = \frac{g(y_1, \cdots, y_n, x)}{g(y_1, \cdots, y_n)}.$$

Proof: Read the textbook. □

<u>Fourier Series</u>

Def.: The series $\sum_{i=1}^{\infty} x_i$ is said to converge if $\{\sum_{i=1}^{n} x_i\}$ is convergent.

Theorem: Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal sequence in a Hilbert space H. Then the series $\sum_{i=1}^{\infty} a_i e_i$ converges iff $\sum_{i=1}^{\infty} |a_i|^2$ converges. Moreover if $x = \sum_{i=1}^{\infty} a_i e_i$, then $a_i = \langle x, e_i \rangle$ that is called Fourier coefficient.

Proof: Notice that for $n > m$,

$$\sum_{i=1}^{n} a_i e_i - \sum_{i=1}^{m} a_i e_i = \sum_{i=m+1}^{n} a_i e_i.$$

Then

$$\left\| \sum_{i=m+1}^{n} a_i e_i \right\|^2 = \langle \sum_{i=m+1}^{n} a_i e_i, \sum_{i=m+1}^{n} a_i e_i \rangle = \sum_{i=m+1}^{n} |a_i|^2. \qquad (*)$$

($\Leftarrow$) From ($*$), $\{\sum_{i=1}^n a_i e_i\}$ is a Cauchy sequence since $\{\sum_{i=1}^n |a_i|^2\}$ is convergent and thus is Cauchy. Since H is a Hilbert space, $\sum_{i=1}^n a_i e_i$ converges.

($\Rightarrow$) From ($*$), $\{\sum_{i=1}^n |a_i|^2\}$ is Cauchy since $\{\sum_{i=1}^n a_i e_i\}$ is convergent and thus is Cauchy. Hence $\{\sum_{i=1}^n |a_i|^2\}$ converges.

(Last Statement) By the continuity of inner product,

$$\langle x, e_i \rangle = \lim_{n\to\infty} \langle \sum_{j=1}^n a_j e_j, e_i \rangle = a_i.$$

$\square$

Remark 1: As a consequence, all separable Hilbert spaces are "analogous" to l_2. In other words, provided that $\{e_k\}$ is the set of orthonormal dense set, $x = \sum_{i=1}^\infty a_i e_i \in H$ iff $\{a_i = \langle x, e_i \rangle\} \in l_2$.

Ex: Let $H = L_2[0,1]$. Let $e_i = \sin(i\pi t)$ and $f_i = \cos(i\pi t)$. Then $\{e_i, f_i\}$ is an orthonormal dense set in H. Hence for all $f \in H$,

$$f(t) = \sum_{i=1}^\infty [a_i e_i + b_i f_i]$$

where a_i, b_i are Fourier coefficients.

Remark 2: If H is not separable, the orthogonal projection $\sum_{i=1}^\infty \langle x, e_i \rangle e_i$ of $x \in H$ on $M = span\{e_k\}$ has norm smaller than x and is the solution of the minimum norm problem:

$$\min_{m\in M} \|x - m\|.$$

Ex: $H = L_2[0,\infty)$ needs inverse Fourier transform where uncountably many sine and cosine are necessary.

Minimum Norm Problems

Theorem: Let M be a closed subspace of a Hilbert space H. Let $x \in H$. Consider the linear variety $V = x + M$. Then there exists a unique vector $v^* \in V$ with the smallest norm, i.e.,

$$\|v^*\| \leq \|v\|, \quad \forall v \in V.$$

Moreover, v^* is characterized by $v^* \perp M$.

Proof: Any vector in V has the form

$$v = x - m, \quad m \in M.$$

Hence $\|\hat{x}\| \leq \|v\|$, $\forall v \in V$, is equivalent to $\|x - m_0\| \leq \|x - m\|$, $\forall m \in M$. Now the theorem follows from the projection theorem. □

Linear Varieties:

- $N = span\{y_1, \cdots, y_n\}$ is a subspace. Hence, $U = x + N$ is a linear variety.
- $M = \{m : \langle m, y_k \rangle = 0, k = 1, \cdots, n\} = N^\perp$ is a subspace. Given $\{c_k\}$, let $v_0 \in V = \{v : \langle v, y_k \rangle = c_k, k = 1, \cdots, n\}$. Then $V = v_0 + M$ and thus is a linear variety.

Theorem: Let $\{y_k\}_{k=1}^n$ be a collection of linearly independent vectors. Then the minimum norm vector $v^* \in V$ has the form:

$$v^* = \sum_{k=1}^{n} a_k y_k.$$

Moreover

$$G(y_1, \cdots, y_n) \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Proof: N is finite dimensional and thus is closed. Hence

$$N = N^{\perp\perp} = M^\perp.$$

Since $v^* \in M^\perp$, the form of v^* follows. The rest of the proof is obvious from the definition of V. □

Ex: Find u with minimum $L_2[0,1]$ norm subject to

$$\dot{w}(t) + w(t) = u(t) \quad (1)$$

$$\dot{\theta} = w(t) \quad (2)$$

with

$$\theta(0) = w(0) = 0, \quad \theta(1) = 1, \quad w(1) = 0.$$

Let V be the subset of $L_2[0,1]$ such that every element in V satisfies these constraints.

Let $u, \hat{u} \in V$. Let $(w(t; u), \theta(t; u))$ be the solution of DE's. Then by linearity of DE's,

$$(w(t; u + \hat{u}), \theta(t; u + \hat{u})) = (w(t; u), \theta(t; u)) + (w(t; \hat{u}), \theta(t; \hat{u})).$$

Then $\theta(1; u + \hat{u}) = 2$ and thus V is not a subspace. Indeed V is a linear variety. Let $u_0 \in V$. Let M be the collection of u such that the above DE's are satisfied with the boundary conditions

$$\theta(0) = w(0) = \theta(1) = w(1) = 0.$$

Then $V = u_0 + M$.

Let $u \in V$. Then

$$w(t) = w(0)e^{-t} + \int_0^t e^{-(t-s)} u(s) ds$$

and thus

$$\int_0^1 e^{-(1-s)} u(s) ds = 0 \quad or \quad \int_0^1 e^s u(s) ds = 0.$$

On the other hand, substituting (2) into (1),

$$w(1) - w(0) + \theta(1) - \theta(0) = \int_0^1 u(s) ds = 1.$$

Hence V is the set of all elements such that

$$\int_0^1 e^s u(s) ds = 0, \quad \int_0^1 u(s) ds = 1.$$

Let $y_1(t) = e^t$, $y_2(t) = 1$. Then the problem becomes to find u such that

$$\langle u, y_1 \rangle = 0, \quad \langle u, y_2 \rangle = 1$$

with the minimum L_2 norm. By the above theorem,

$$u(t) = a_1 y_1(t) + a_2 y_2(t),$$

where a_i's are obtained from

$$\langle a_1 y_1 + a_2 y_2, y_1 \rangle = 0$$

$$\langle a_1 y_1 + a_2 y_2, y_2 \rangle = 1.$$

Minimum Distance to a Convex Set

Theorem: Let $x \in H$ and K be a closed convex subset of H. Then there exists a unique vector $k_0 \in K$ such that

$$\|x - k_0\| \leq \|x - k\|, \quad \forall k \in K.$$

Moreover k_0 is the unique minimum iff

$$\langle x - k_0, k - k_0 \rangle \leq 0, \quad \forall k \in K.$$

Proof: Let $\{k_i\}_{i=1}^{\infty}$ be a minimizing sequence, i.e.,

$$\lim_{i \to \infty} \|x - k_i\| = \delta = \inf_{k \in K} \|x - k\|.$$

By parallelogram law,

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

Since K is convex, $\frac{k_i+k_j}{2} \in K$ and thus

$$\lim_{i,j \to \infty} \|k_i - k_j\|^2 \leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0.$$

Hence, $\{k_i\}_{i=1}^{\infty}$ is a Cauchy sequence. since H is a Hilbert space, there exists $k_0 = lim_{i \to \infty} k_i \in H$. Since $\| \cdot \|$ is continuous and K is closed, $\|x - k_0\| = \delta$ and $k_0 \in K$.

To show the uniqueness, let k_1 be such that $\|x - k_1\| = \delta$. Then the sequence $\{\kappa_n\} = \{k_0, k_1, k_0, k_1, \cdots\}$ results in $\|x - \kappa_n\| \to \delta$. Hence as shown above $\{\kappa_n\}$ is Cauchy and convergent. Therefore it must hold that $k_0 = k_1$.

($\Rightarrow$) Consider the following function of $\lambda \in [0, 1]$.

$$g(\lambda) = \|x - [(1 - \lambda)k_0 + \lambda k]\|^2 = \|(1 - \lambda)(x - k_0) + \lambda(x - k)\|^2$$

$$= (1 - \lambda)^2\|x - k_0\|^2 + 2\lambda(1 - \lambda)\langle x - k_0, x - k\rangle + \lambda^2\|x - k\|^2.$$

Hence

$$\frac{d}{d\lambda} g(\lambda) = -2(1 - \lambda)\|x - k_0\|^2 + (2 - 4\lambda)\langle x - k_0, x - k\rangle + 2\lambda\|x - k\|^2.$$

If k_0 is the minimizing vector, then $\frac{d}{d\lambda} g(\lambda)\Big|_{\lambda=0} \geq 0$ and thus

$$-2\|x - k_0\|^2 + 2\langle x - k_0, x - k\rangle \geq 0.$$

$$\Downarrow$$

$$\langle x - k_0, -(x - k_0) + (x - k)\rangle \geq 0.$$

$$\Downarrow$$

$$\langle x - k_0, k - k_0\rangle \leq 0.$$

($\Leftarrow$) For any $k \in K$, $k \neq k_0$,

$$\|x - k\|^2 = \|x - k_0 + k_0 - k\|^2 = \|x - k_0\|^2 + 2\langle x - k_0, k_0 - k\rangle + \|k_0 - k\|^2$$

$$= \|x - k_0\|^2 - 2\langle x - k_0, k - k_0\rangle + \|k_0 - k\|^2 > \|x - k_0\|^2.$$

□

Ex: Let $\{y_1, \cdots, y_n\}$ be vectors in a Hilbert space. Find $a_1, \cdots, a_n \geq 0$ such that $\|x - \sum_{k=1}^n a_k y_k\|$ is minimized.

Solution: Let K be the collection of vectors given by

$$K = \left\{ \sum_{k=1}^n a_k y_k, a_k \geq 0 \right\}.$$

K is a closed convex set. From the theorem, we have $k_0 = \sum_{k=1}^n a_k^* y_k$ which has the property that

$$\langle x - k_0, k - k_0\rangle \leq 0.$$

1. If $k = k_0 + y_j$, $\langle x - k_0, y_j\rangle \leq 0$.
2. If $k = k_0 - a_i^* y_i$, $\langle x - k_0, -a_i^* y_i\rangle \leq 0$. Hence $a_i^*\langle x - k_0, y_i\rangle \geq 0$.

As a result, we have

$$\langle x - k_0, y_j\rangle \leq 0$$

and, if $a_j^* > 0$, then

$$\langle x - k_0, y_j\rangle = 0.$$

Let b be a vector given by

$$b_i = \langle x, y_i\rangle.$$

Then

$$b - G(y_1, \cdots, y_n)a^* \leq 0, \quad a^* = [a_1^*, \cdots, a_n^*]^T.$$

Let $z = b - G(y_1, \cdots, y_n)a^*$. Then $a_j^* z_j(a^*) = 0$ because either $\langle x - k_0, y_j\rangle$ or a_j^* is zero. Find a vector a^* such that $a_j^* z_j(a^*) = 0$ and $b - G(y_1, \cdots, y_n)a^* \leq 0$.

Chapter 7

Dual Spaces

Dual Space

Def.: An operator f from a linear space X into $\mathbf{R}$ is said to be a functional. If it holds

$$f(ax + by) = af(x) + bf(y), \quad a, b \in \mathbf{R},\ x, y \in X,$$

it is called a linear functional.

Fact: A linear functional on a normed space X is continuous at a point $x_0 \in X$ iff it is continuous on X.

Proof: ($\Leftarrow$) Obvious.

($\Rightarrow$) Let $x_n \to x$ as $n \to \infty$. Then

$$|f(x_n) - f(x)| = |f(x_n - x + x_0) - f(x_0)|.$$

Since $x_n - x + x_0 \to x_0$, $f(x_n - x + x_0) \to f(x_0)$ and thus $|f(x_n) - f(x)| \to 0$. $\square$

Def.: A linear functional on a normed space X is said to be bounded if there exists $M \geq 0$ such that for all $x \in X$,

$$|f(x)| \leq M\|x\|.$$

Fact: A linear functional on X is continuous iff it is bounded.

Proof: ($\Leftarrow$) If $x_n \to 0$,

$$|f(x_n)| \leq M\|x_n\| \to 0.$$

and thus continuous by the above fact.

($\Rightarrow$) There is $\delta > 0$ such that $|f(x)| < 1$, for $\|x\| \leq \delta$. Then for all $x \in X$,

$$|f(x)| = \left| f\left(\frac{\delta x}{\|x\|}\right) \right| \frac{\|x\|}{\delta} < \frac{\|x\|}{\delta}.$$

Hence $M = \frac{1}{\delta}$. □

Def.: Let X be a normed linear space. The collection of all bounded linear functional on X is called the dual space of X, denoted by X^*.

Algebraic Structure of X^*

For all $f, g \in X^*$, $a, b \in \mathbf{R}$,

$$(f+g)(x) = f(x) + g(x), \quad \forall x \in X$$

$$(af)(x) = af(x), \quad \forall x \in X$$

$$f = 0 \quad \Leftrightarrow \quad f(x) = 0,\ \forall x \in X$$

$$af + bg \in X^*.$$

Hence X^* is a linear space.

Topological Structure of X^*

For an element $f \in X^*$, we define

$$\|f\|_{X^*} = \sup_{x \in X, x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\| \leq 1} |f(x)| = \sup_{\|x\|=1} |f(x)|.$$

Then

$$\|f+g\|_{X^*} \leq \sup_{x \in X, x \neq 0} \frac{|f(x)| + |g(x)|}{\|x\|} \leq \sup_{x \in X, x \neq 0} \frac{|f(x)|}{\|x\|} + \sup_{x \in X, x \neq 0} \frac{|g(x)|}{\|x\|}$$

$$= \|f\|_{X^*} + \|g\|_{X^*}.$$

Hence $\|f\|_{X^*}$ defines a norm on X^*.

Theorem: X^* is a Banach space.

Proof: It suffices to show the completeness. Let $\{f_n\} \subset X^*$ be Cauchy. Then for any $x \in X$, $\{f_n(x)\} \subset \mathbf{R}$ is Cauchy and thus is convergent since $|f_n(x) - f_m(x)| \leq \|f_n - f_m\| \|x\|$. Define f as $f(x) = \lim_{n \to \infty} f_n(x)$. f is linear since

$$f(ax+by) = \lim_{n \to \infty} f_n(ax+by) = \lim_{n \to \infty} \{af_n(x) + bf_n(y)\}$$

$$= a \lim_{n\to\infty} f_n(x) + b \lim_{n\to\infty} f_n(y) = af(x) + bf(y).$$

For $x \neq 0$, given $\epsilon > 0$, there exists $M > 0$ such that

$$|f_n(x) - f_m(x)| < \epsilon \|x\|, \quad n, m > M, \ \ \forall x, \ x \neq 0.$$

$$\Downarrow$$

$$|f(x) - f_m(x)| \leq \epsilon \|x\|, \quad m > M, \ \ \forall x.$$

Hence

$$|f(x)| = |f(x) - f_m(x) + f_m(x)| \leq |f(x) - f_m(x)| + |f_m(x)| \leq (\epsilon + \|f_m\|)\|x\|.$$

and thus f is a bounded linear functional. Finally from $|f(x) - f_m(x)| < \epsilon\|x\|$, $m > M$, $\forall x \in X$, $\|f - f_m\| < \epsilon$ and thus $f_n \to f$. □

Ex 1: Let $X = \mathbf{R}^n$ and consider an element in X^*. Define a_k by

$$(x^*, e_k) = a_k, \quad k = 1, \cdots, n.$$

By linearity, we have

$$(x^*, x) = \sum_{k=1}^{n} a_k x_k.$$

Hence, the vector $[a_1, \cdots, a_n]^T$ is the representation of x^*. Indeed we can establish a one-to-one correspondence between X^* and $\mathbf{R}^n$.

$$\|x^*\|_{X^*} = \sup_{x \in X, x \neq 0} \frac{|(x^*, x)|}{\|x\|}.$$

Then for $1 < p < \infty$,

$$\|x^*\|_{X^*} = \sup_{x \in X, x \neq 0} \frac{|\sum_{k=1}^{n} a_k x_k|}{(\sum_k |x_k|^p)^{\frac{1}{p}}} \leq \sup_{x \in X, x \neq 0} \frac{(\sum_k |a_k|^q)^{\frac{1}{q}} (\sum_k |x_k|^p)^{\frac{1}{p}}}{(\sum_k |x_k|^p)^{\frac{1}{p}}}$$

$$= \left(\sum_k |a_k|^q \right)^{\frac{1}{q}} = \|a\|_q$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Consider x given by $x_k = |a_k|^{\frac{q}{p}} sign(a_k)$. Then

$$\|x\| = \left(\sum_k |a_k|^{\frac{q}{p}p} \right)^{\frac{1}{p}} = \left(\sum_k |a_k|^q \right)^{\frac{1}{p}} = \|a\|_q^{\frac{q}{p}}.$$

Moreover,

$$|(x^*, x)| = \sum_k |a_k|^{1+\frac{q}{p}} = \sum_k |a_k|^q = \|a\|_q^q.$$

Hence

$$\frac{|(x^*, x)|}{\|x\|} = \frac{\|a\|_q^q}{\|a\|_q^{\frac{q}{p}}} = \|a\|_q^{q\left(1-\frac{1}{p}\right)} = \|a\|_q.$$

Therefore

$$\|x^*\|_{X^*} \geq \|a\|_q$$

and thus

$$\|x^*\|_{X^*} = \|a\|_q.$$

To this end, $\mathbf{R}^n$ with q-norm is the representation space of the dual of $\mathbf{R}^n$ with p-norm.

Special case $p = 2$: Euclidean norm case

$X = \mathbf{R}^n$ is a Hilbert space

$\mathbf{R}^n$ with the Euclidean norm is the representation space of X^*

Ex 2: $\mathbf{R}^n$ with 1-norm is the representation space of the dual of $\mathbf{R}^n$ with ∞-norm and vice versa.

Ex 3: For $1 \leq p < \infty$, l_q is the representation space of the dual of l_p.

Ex 4: For $1 \leq p < \infty$, L_q is the representation space of the dual of L_p.

Ex 5: For $1 \leq p < \infty$, H_q is the representation space of the dual of H_p.

Ex 6: Dual of c_0 is l_1.

Extension Version of Hahn-Banach Theorem

Def.: Let S be a subspace of a Banach space B. Let f be a linear functional on S. A linear functional F on B is said to be an extension of f iff $F(x) = f(x),\ x \in S$.

Def.: A functional p on a Banach space is said to be sublinear if the following properties hold:

a) $p(x + y) \leq p(x) + p(y)$

b) $p(ax) = ap(x)$, $a \geq 0$.

Notice that a norm is sublinear.

Hahn-Banach Theorem: Let X be a normed linear space. Let M be a subspace of X. Let f be a linear functional on M such that

$$f(x) \leq p(x), \quad \forall x \in M$$

where p is a sublinear functional on X. Then there exists a linear extension F on X such that

$$F(x) \leq p(x), \quad \forall x \in X.$$

Proof: Read the textbook. $\square$

Corollary: Let f be a linear functional on a subspace M of a normed linear space. Then there exists a linear extension F on X that extends f and

$$\|F(x)\| = \sup_{x \in M,\, x \neq 0} \frac{|f(x)|}{\|x\|}.$$

Proof: $p(x) = \|f\|\|x\|$. $\square$

Corollary: Let $x \in X$ where X is a normed linear space. Then there exists a bounded linear functional F such that

$$F(x) = \|F\|\|x\|.$$

Proof: If $x = 0$, trivial. Suppose $x \neq 0$. Define f on $span\{x\}$ by $f(ax) = a\|x\|$. Then $\|f\| = 1$. Now extend f to X using $p(x) = \|x\|$. Then

$$F(x) = f(x) = \|x\| = \|F\|\|x\|.$$

$\square$

Dual of $C[0,1]$

Riesz Representation Theorem: For any element $x^* \in X^*$ with $X = C[0,1]$, there exists a function f of bounded variation on $[0,1]$ such that

$$x^*(x) = \int_0^1 x(t)df(t).$$

Moreover $\|x^*\| = \|f\|_{TV[0,1]}$.

Conversely, every element in BTV[0,1] defines a bounded linear functional on X.

Second Dual Space

The dual space of X^* is called the second dual space of X, and it is denoted by X^{**}.

Let (x, x^*) denote the action of x^* on an element $x \in X$. Then x can be viewed as a functional on X^* and thus $X \subset X^{**}$.

Def.: A Banach space is said to be reflexive if $X^{**} = X$.

Ex1: l_p, $1 < p < \infty$, are reflexive.

Ex2: L_p, $1 < p < \infty$, are reflexive.
Ex3: H_p, $1 < p < \infty$, are reflexive.
Ex4: l_1, l_∞, L_1, L_∞, H_1, H_∞, c_0 and $C[0,1]$ are not reflexive.

Alignment and Orthogonal Complement

Def.: Let X be a normed linear space. An element $x^* \in X^*$ is said to be aligned with an element $x \in X$ if

$$(x, x^*) = \|x^*\|\|x\|.$$

Ex: $x \in l_p$ (Then $x^* \in l_q$). Let $x_k^* = sign(x_k)|x_k|^{\frac{p}{q}}$. Then

$$(x, x^*) = \sum_{k=1}^{\infty} |x_k|^{1+\frac{p}{q}} = \sum_{k=1}^{\infty} |x_k|^p = \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{q}} = \|x^*\|\|x\|.$$

Def.: Let $x \in X$ where X is a normed linear space. $x^* \in X^*$ is said to be orthogonal to x if $(x, x^*) = 0$.

Def.: Let M be a subspace of a normed linear space X. Then the orthogonal complement of M is a subspace of X^*, denoted by $M^\perp$, defined as

$$M^\perp = \{x^* : (m, x^*) = 0,\ \forall m \in M\}.$$

Def.: Let M^* be a subspace of X^*. The orthogonal complement of M^* in X is a subset of X, denoted by ${}^\perp M^*$ (to distinguish from $(M^*)^\perp$), defined as

$${}^\perp M^* = \{x : (x, m^*) = 0,\ \forall m^* \in M^*\}.$$

Theorem: Let M be a closed subspace of a normed linear space. Then

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

Proof: It is clear that $M \subset {}^\perp(M^\perp)$. Let $x \notin M$ and $N = span\{x + M\}$. Define the linear functional on N:

$$f(ax + m) = a$$

where $m \in M$. Then

$$\|f\| = \sup_{ax+m \in N, ax+m \neq 0} \frac{|f(ax+m)|}{\|ax+m\|} = \sup_{ax+m \in N, ax+m \neq 0} \frac{|a|}{\|ax+m\|}$$

$$= \sup_{ax+m\in N, ax+m\neq 0, a\neq 0} \frac{1}{\|sign(a)x + \frac{m}{|a|}\|} = \sup_{m'\in M} \frac{1}{\|x+m'\|} = \frac{1}{\inf_{m'\in M}\|x+m'\|}.$$

Since M is closed, $\|f\| < \infty$ and thus f is bounded. Otherwise there exists $\{m'_k\}$ in M such that $\|x - m'_k\| \to 0$, and thus $x \in M$. Then by Hahn-Banach theorem, there exists an extension $x^* \in X^*$ of f. Since f, and thus x^*, vanishes on M, $x^* \in M^\perp$. However $(x, x^*) = 1$ and thus $x \notin {}^\perp(M^\perp)$. To this end, ${}^\perp(M^\perp) \subset M$ and the theorem follows. □

Minimum Norm Problems in Normed Space

Problem 1: Let M be a subspace of X. Let $x \in X$. Find $m_0 \in M$ such that

$$\|x - m_0\| \leq \|x - m\|, \quad \forall m \in M.$$

Problem 2: Let M^* be a subspace of X^*. Let $x^* \in X^*$. Find $m_0^* \in M^*$ such that

$$\|x^* - m_0^*\| \leq \|x^* - m^*\|, \quad \forall m^* \in M^*.$$

Although the problem formulation is similar to both cases, the second gives stronger results because X^* is Banach as long as X is a normed space.

Theorem: Let x be an element in a real normed linear space X and let d be the distance between $x \in X$ and a subspace M of X. Then

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

Let m_0^* be the solution of the maximization. If the infimum is achieved at $m_0 \in M$, then m_0^* is aligned with $x - m_0$.

Proof: Given $\epsilon > 0$, let $m_\epsilon \in M$ be such that $\|x - m_\epsilon\| \leq d + \epsilon$. Then for $m^* \in M^\perp$, $\|m^*\| \leq 1$,

$$(x, m^*) = (x - m_\epsilon, m^*) \leq \|x - m_\epsilon\|\|m^*\| \leq d + \epsilon.$$

Since ϵ is arbitrary,

$$(x, m^*) \leq d.$$

Let $N = span\{x + M\}$. Define the linear functional on N:

$$f(ax + m) = ad$$

where $m \in M$. Then similar to the previous theorem,

$$\|f\| = \frac{d}{\inf_{m'\in M}\|x + m'\|} = 1.$$

By Hahn-Banach theorem, there exists an extension $m_0^* \in X^*$ of f. Since f, and thus m_0^*, vanishes on M, $m_0^* \in M^\perp$. Moreover $(x, m_0^*) = d$. Hence the first part of the theorem follows.

Notice that $\|m_0^*\| = 1$. Hence

$$(x - m_0, m_0^*) = (x, m_0^*) = d = \|x - m_0\| = \|m_0^*\|\|x - m_0\|.$$

□

Corollary: Let x be an element in a real normed linear space X and let M be a subspace of X. Then $m_0 \in M$ is such that $\|x - m_0\| \leq \|x - m\|$ for all $m \in M$ iff there exists nonzero $m_0^* \in M^\perp$ aligned with $x - m_0$.

Proof: ($\Rightarrow$) Obvious.

($\Leftarrow$) WLOG assume $\|m_0^*\| = 1$. For all $m \in M$,

$$\|x - m_0\| = (x - m_0, m_0^*) = (x, m_0^*) = (x - m, m_0^*) \leq \|x - m\|.$$

□

Theorem: Let M be a subspace in a real normed linear space X and let d be the distance between $x^* \in X^*$ and a subspace $M^\perp$ of X^*. Then

$$d = \min_{m^* \in M^\perp} \|x^* - m^*\| = \sup_{\|m\| \leq 1, m \in M} (m, x^*).$$

Let m_0^* be the solution of the minimization. If the supremum is achieved at $m_0 \in M$, then m_0 is aligned with $x^* - m_0^*$.

Proof: Notice that

$$\|x^* - m^*\| = \sup_{x \in X, \|x\| \leq 1} |(x, x^* - m^*)| \geq \sup_{x \in M, \|x\| \leq 1} |(x, x^* - m^*)| = \sup_{x \in M, \|x\| \leq 1} |(x, x^*)|.$$

Let $x^*|_M$ be the restriction of x^* on M. Using Hahn-Banach Theorem, let y^* be the extension of $x^*|_M$ such that $\|y^*\| = \|x^*|_M\|$. Define $m_0^* = x^* - y^*$. Then $m_0^* \in M^\perp$ and

$$\|x^* - m_0^*\| = \|y^*\| = \|x^*|_M\| = \sup_{x \in M, \|x\| \leq 1} |(x, x^*|_M)| = \sup_{x \in M, \|x\| \leq 1} |(x, x^*)|.$$

Notice that $\|m_0\| = 1$. Hence

$$(m_0, x^* - m_0^*) = (m_0, x^*) = d = \|x^* - m_0^*\| = \|m_0\|\|x^* - m_0^*\|.$$

□

l_1 Optimal Control

Given the transfer function $H(z)$ of a linear time-invariant discrete time system,

$$H(z) = \sum_{i=0}^{\infty} h_i z^i$$

where $\{h_i\}$ is the sequence of unit pulse response of the system. In time-domain, the system is described by the convolution:

$$y = h * u.$$

This system is called BIBO stable if any $u \in l_\infty$ results in $y \in l_\infty$. It is well known in linear system theory that the linear system is BIBO stable iff $\{h_i\}$ is an l_1 sequence. Indeed

$$\|H\| = \sup_{\|u\|_\infty \leq 1} \|h * u\|_\infty = \|h\|_1.$$

Youla-Jabr-Bongiorno Parametrization of All stabilizing Controllers (Stable Plant Case): A controller C internally stabilizes the stable plant iff

$$C = \frac{Q}{1 - PQ}$$

where Q is a stable transfer function.

Then the sensitivity and complementary sensitivity become

$$S(z) = \frac{1}{1 + PC} = \frac{1 - PQ}{1 - PQ + PQ} = 1 - PQ$$

$$T(z) = \frac{PC}{1 + PC} = \frac{PQ}{1 - PQ + PQ} = PQ,$$

respectively. Notice that these closed loop maps are affine w.r.t. Q. In general, even for unstable plants, the closed loop map is given as $H - GQ$ where H and G are stable transfer functions.

Before we proceed, notice that if $x \in l_1$ and $y \in l_\infty$ are aligned, then

$$x_i = 0, \quad \text{if } y_i < \|y\|_\infty$$

$$x_i y_i \geq 0.$$

Given the closed loop map from disturbance to the output, $H - GQ$, the controller that internally stabilizes the plant and minimizes the effects of the bounded disturbances can be obtained from the following l_1 optimal control problem:

$$\inf_{\text{stable } Q} \|H - GQ\|.$$

For simplicity assume $G(z)$ have n distinct zeros inside the open unit disk, $\{a_j\}_{j=1}^n$. Define $K = GQ$. Then K can be any stable transfer function such that $K(a_j) = 0$. Hence the l_1 optimal control problem is equivalent to

$$\inf_{\text{stable } K} \|H - K\|$$

subject to

$$K(a_j) = 0, \quad j = 1, \cdots, n.$$

Notice that the constraints are equivalent to

$$\sum_{i=0}^{\infty} k_i a_j^i = 0, \quad j = 1, \cdots, n.$$

$$\Updownarrow$$

$$(k, ar_j) = 0, \ \ (k, ai_j) = 0, \quad j = 1, \cdots, n$$

where

$$ar_j = Re(1, a_j, a_j^2, \cdots), \quad ai_j = Im(0, a_j, a_j^2, \cdots).$$

To this end, the l_1 optimal control problem reduces to a minimum norm problem in l_1:

$$\inf_{k \in M} \|h - k\|_1$$

where M is the subspace defined by

$$M = \{k \in l_1 : (k, ar_j) = 0, \ (k, ai_j) = 0, \ j = 1, \cdots, n\}.$$

Existence of Optimal Solution: Notice that l_1 is the dual of c_0. Hence by the minimum norm problem theorem, the optimal solution always exists.

Finding Minimum Norm: By the minimum norm problem theorem,

$$\inf_{k \in M} \|h - k\|_1 = \max_{r \in M^{\perp}, \|r\|_{\infty} \leq 1} (h, r)$$

However, r has the representation:

$$r = \sum_{i=1}^{n} a_i ar_i + \sum_{j=1}^{n} a_{n+i} ai_j.$$

Hence

$$(h, r) = \left[\sum_{i=1}^{n} a_i ReH(a_i) + \sum_{j=1}^{n} a_{n+j} ImH(a_j) \right]$$

and

$$\|r\|_\infty \leq 1 \quad \Leftrightarrow \quad -1 \leq \sum_{i=1}^{n} a_i Re(a_i^j) + \sum_{i=1}^{n} a_{n+i} Im(a_i^j) \leq 1, \;\; j = 0, 1, 2, \cdots.$$

To this end,

$$\inf_{k \in M} \|h - k\|_1 = \max_{a_i} \left[\sum_{i=1}^{n} a_i ReH(a_i) + \sum_{j=1}^{n} a_{n+j} ImH(a_j) \right]$$

subject to

$$-1 \leq \sum_{i=1}^{n} a_i Re(a_i^j) + \sum_{i=1}^{n} a_{n+i} Im(a_i^j) \leq 1, \;\; j = 0, 1, 2, \cdots.$$

Remark: Since $\|a_j\| < 1$, only a finite set of constraints are necessary. This is due to the fact that the constraints are redundant for all j greater than some N. The jth constraint is exactly the jth coefficient of r. Hence the optimal solution r^* has the following properties:

$$\|r^*\| = 1$$

$$|r_j^*| < 1, \;\; \forall j > N.$$

Construction of Optimal Solution: Suppose K^* is the optimal solution and the optimum is μ. Then $b = h - k^* \in l_1$ is aligned with $r^* \in l_\infty$, i.e.,

$$(b, r^*) = \|b\|_1 \|r^*\|_\infty = \mu.$$

$$\Updownarrow$$

$$\sum_{i=0}^{\infty} b_i r_i^* = \|b\|_1 \|r^*\|_\infty.$$

$$\Updownarrow$$

$$b_i = 0 \quad \text{if } |r_i^*| \neq \|r^*\|_\infty$$

$$b_i r_i^* \geq 0$$

$$\|b\|_1 = \sum_{i=0}^{\infty} |b_i| = \mu.$$

From the above remark, b will have only finitely many nonzero b_i. However, there are many such b satisfying the alignment condition and not all of them leads to $k^* = h - b \in M$. For $k^* = h - b \in M$, b_i has to satisfy

$$\sum_{i=0}^{\infty} b_i a_j^i = (b, A_j) = (h, A_j) = H(a_j), \quad j = 1, \cdots, n$$

where

$$A_j = (1, a_j, a_j^2, a_j^3, \cdots).$$

To this end, $k^* \in M$ iff $b = h - k^*$ satisfies

$$b_i = 0 \quad \text{if } \|r_i^*\| \neq \|r^*\|_\infty$$

$$b_i r_i^* \geq 0$$

$$\sum_{i=0}^{\infty} |b_i| = \mu$$

$$\sum_{i=0}^{\infty} b_i a_j^i = H(a_j), \quad j = 1, \cdots, n.$$

From the solution b of these equalities and inequalities, k^*, Q^* and C^* can be found.

Ex: Suppose there are two interpolation points:

$$a_1 = jx, \quad a_2 = -jx, \qquad -1 < x < 1.$$

Step 1 (find minimum norm): Let

$$H_r = ReH(jx), \quad H_i = ImH(jx).$$

We need to solve

$$\max_{a_1,a_2}(a_1 H_r + a_2 H_i)$$

subject to

$$-1 \le a_1 \le 1$$

$$-1 \le a_2 x \le 1$$

where the rest of constraints are redundant. The optimal solution is clearly

$$a_1 = sign(H_r)$$

$$a_2 = \frac{1}{x} sign(H_i)$$

and the minimum norm is

$$\mu = |H_r| + \frac{|H_i|}{x}.$$

Moreover, the extremal functional is given by:

$$r^* = (a_1, x a_2, -x^2 a_1, \cdots).$$

Step 2 (Construction of Optimal Solution): Notice that

$$b = \left(H_r, \frac{H_i}{x}, 0, 0, \cdots \right).$$

Chapter 8

Linear Operator

An operator (transformation or mapping) T from a subset D of a linear space X into a linear space Y is a rule that associates every element in D to an element of Y.

Terminologies:

1. D is called the domain of T.

2. $\mathcal{R}(T) = \{y : y = Tx,\ x \in D\}$ is called the range of T.

Def.: T is one-to-one (injective) if $\forall y \in Y$, $\exists$ at most one $x \in X$ such that $Tx = y$.

Def.: T is onto (surjective) if $\mathcal{R}(T) = Y$.

Def.: T is bijective if it is injective and surjective.

If $Y = \mathbf{R}$ or $\mathbf{C}$, T is called a functional.

An operator from a linear space X into a linear space Y is said to be linear if

$$T(\alpha x + \beta \hat{x}) = \alpha T x + \beta T \hat{x}, \qquad \forall x, \hat{x} \in X.$$