16. Linear Operators

- T: x → y is called a linear operator if T(x) ∈ Y and
  \[ T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) \]
  \[ \forall \alpha, \beta \in \mathbb{R}, \ x_1, x_2 \in X. \]

- If X and Y are normed spaces, then T is continuous at \( x \) if given \( \varepsilon > 0 \), there exists \( \delta = \delta_\varepsilon > 0 \) such that
  \[ \| T(x) - T(y) \| \leq \varepsilon \]
  \[ \forall y: \| y - x \| \leq \delta \]

- T is said to be bounded if
  \[ \| T(x) \| \leq M \| x \| \text{ for all } x \]
  for some \( M < \infty \). In this case,
  \[ \| T \| = \sup_{x \neq 0} \frac{\| T(x) \|}{\| x \|} \]
  \[ = \sup_{\| x \| = 1} \| T(x) \| \]

  \[ \| T \| = \sup_{\| x \| = 1} \| T(x) \| \]
The following theorem is along the lines of a similar result for linear functionals.

**Theorem:** Let \( T : X \to Y \) be a linear operator.

(i) \( T \) is continuous if \( x \in X \).

(ii) \( T \) is continuous if it is bounded.

\[ \square \]

\( \mathcal{B}(X, Y) \): Space of all bounded linear operators from \( X \) to \( Y \). It is a normed space with the operator norm defined earlier.

**Theorem:** \( \mathcal{B}(X, Y) \) is a Banach space if \( X \) is a normed space and \( Y \) is a Banach space.

**Proof:** Use the 3-step procedure to prove that a space is Banach.

\[ \square \]

**Examples:**

(i) Let \( A \) be an \( m \times n \) matrix and consider it to be a linear operator from \( \mathbb{R}^m \) to \( \mathbb{R}^n \).
with both spaces equipped with the $l_2$-norm.

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \sqrt{x^T A^T A x} = \sqrt{\lambda_{\text{max}}(A^T A)}$$

by a result from linear algebra.

$\sqrt{\lambda(A^T A)}$ are called the singular values of $A$. Thus, $\|A\|$ is equal to the largest singular value.

(2) Let $T : M \to C([0,1])$, where $M \subset C([0,1])$ is the subspace of continuously differentiable functions, be defined as

$$T x = \frac{d}{dt} x(t)$$

$T$ is linear, but is not bounded, because $\|x\|$ can be bounded, but its derivative can be unbounded.

However, if the domain (the space over which $T$ is defined) is defined to be $D([0,1])$, the
space of continuously differentiable functions with

$$\|x\| = \max_{t \in [0,1]} |x(t)| + |x'(t)|,$$

then $T$ is bounded.

$$\|T\| = \sup_{\|x\|=1} \max_{t \in [0,1]} |x'(t)|$$

$$= 1$$

since the numerator $\leq 1$ and

the norm can be made arbitrarily close to the denominator may make

$|x(t)|$ close to zero.

Some definitions and properties of linear operators

(1) Domain: $T: X \rightarrow Y$ may have a domain $D \subset X$, i.e.,

the set of vectors over which it is defined.
(2) \( R(T) \): range of \( T \)
\[ = \{ y : y = Tx \text{ for some } x \in X \} \]

(3) \( N(T) \): null space of \( T \)
\[ = \{ x : Tx = 0 \} \]

(4) Let \( S \subseteq D \).
\[ \{ y : y = Tx \text{ for some } x \in S \} \]

is called the image of \( S \).
Denoted by \( T(S) \).

(5) Let \( S \subseteq R(T) \).
\[ \{ x : y = Tx \text{ for some } y \in S \} \]

is called the pre-image or inverse image of \( S \).
Denoted by \( T^{-1}(S) \).

**Theorem:** \( T : X \to Y \; S : Y \to Z \)

Then \( ST \) is defined as
\[ STx = S(T(x)) \] \( ST \) is linear and
\[ \| ST \| \leq \| S \| \| T \| . \]
Proof: Easy.

Solving \( y = T x \) for a given \( y \)

Three cases:

- Unique solution exists
- No solution
- Too many solutions.

Unique solution case: When does a unique solution exist?

Suppose \( T \) is one-to-one (i.e., if \( y_1 = T x_1 \) and \( y_2 = T x_2 \), then \( y_1 = y_2 \) if \( x_1 = x_2 \)) and onto \((\mathbb{R}(T) = Y)\), then given \( y \), there is a unique \( y \)

Solving \( y = T x \): in this case, we denote the solution by

\[ x = T^{-1}(y). \]

\( T^{-1} \) is called the inverse of \( T \).
Theorem:

(1) $T^{-1}$ is linear

(2) If $X$ and $Y$ are Banach spaces, then $T^{-1}$ is also bounded if $T$ is bounded.

Proof:

(1) $T^{-1}(\alpha y_1 + \beta y_2) = x$

$\Rightarrow \alpha y_1 + \beta y_2 = Tx$

$\Rightarrow \alpha T x_1 + \beta T x_2 = Tx$

$\Rightarrow T(\alpha x_1 + \beta x_2) = T x$

$\Rightarrow x = \alpha x_1 + \beta x_2$

(2). This is a deep result in functional analysis called the Banach inverse theorem. We will not prove it here.

To consider the other case, i.e., when $y = Tx$ has many solutions or no solution, we have to define the concept of an adjoint of an operator, which is the equivalent of transpose of a matrix.
Adjoint: The adjoint $T^*: Y^* \to X^*$ is an operator that takes $y^* \in Y^*$ as input and produces $T^*y^* \in X^*$ as output. The definition of $T^*$ requires some thought.

(1) Note that $Tx \in Y$. Thus, given $y^*$, we can evaluate $y^*$ at $Tx$, which in our bracketed notation, is $\langle Tx, y^* \rangle$.

(2) For a fixed $y^*$, $\langle Tx, y^* \rangle$ is a functional over $X$. Also

$$|\langle Tx, y^* \rangle| \leq \|T\| \|x\| \|y^*\| \leq \|T\| \|x\| \|y^*\|$$

Thus, $f(x) = \langle Tx, y^* \rangle$ is a bounded linear functional since, from above,

$$\left\| f(x) \right\| \leq \|T\| \|y^*\|,$$

i.e., $f \in X^*$.

(3) The above facts (1) and (2) can be summarized as follows:
Consider a linear operator $T$. Then, for each $y^*$, $T$ and $y^*$ together define a functional $f(x)$, i.e., $T$ defines an operator from
\[ y^* \rightarrow x^* \] This operator is called the adjoint \( T^* \). More compactly, \( T^* \) is defined by
\[ \langle Tx, y^* \rangle = \langle x, T^* y^* \rangle \]

Example: Let \( A_{m \times n} \) be a matrix operator from \( \mathbb{R}^m \) to \( \mathbb{R}^n \), each endowed with the \( l_2 \) norm.

\( (\mathbb{R}^m)^\prime = \mathbb{R}^m \) and \( (\mathbb{R}^n)^\prime = \mathbb{R}^n \)

Then, \( A^* : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and is defined by
\[ (x, A^* y) = (Ax, y) \]
\[ x^T A^* y = x^T A^T y \]
\[ \Rightarrow \quad A^* = A^\top \] as expected.

In function notation,
\[ T^* y^* (x) = y^* (Tx) \]
\[ T^* y^* x = y^* T x \]

In \( T^* \) is also sometimes defined by
\[ T^* y^* = y^* T \]

Thus, \( T^* \) and \( y^* \) of \( T \) are
both functionals on $X$. Thus $T^\circ$ is an operator whose composition with every $y^*$ gives a function identical to the composition of $y^*$ with $T$.

**Theorem:** \[ \| T^\circ \| = \| T \| \quad \text{and} \quad T^\circ \text{ is linear.} \]

**Proof:** $T^\circ$ is linear is easy to prove.

\[ \| T^\circ \| = \sup_{\| y^* \| = 1} \| T^\circ y^* \| \]

\[ = \sup_{\| y^* \| = 1} \left\{ \sup_{\| x \| = 1} \langle x, T^\circ y^* \rangle \right\} \]

\[ = \sup_{\| y^* \| = 1} \sup_{\| x \| = 1} \langle x, T^\circ y^* \rangle \]

\[ \leq \sup_{\| y^* \| = 1} \| x \| \sup_{\| x \| = 1} \| y^* \| \]

\[ \leq \| T \| \sup_{\| y^* \| = 1} \| x \| \|

Now, we prove the reverse inequality:

\[ \| T \| \leq \| T^\circ \| \]

\[ \| T \| = \sup_{\| x \| = 1} \| T x \| \]

Let $x$ be a vector with $\| x \| = 1$.
\(Tx \in Y. \text{ By } \mathbf{H} - \mathbf{B}, \exists y^* \in Y^* \text{ s.t.}
\)
\[
\|y^*\| = 1 \quad \text{and} \quad \langle Tx, y^* \rangle = \|Tx\|
\]
\[
\iff \langle x, T^* y^* \rangle = \|Tx\|
\]
\[
\implies \|Tx\| = \langle x, T^* y^* \rangle \leq \|x\| \|T^* y^*\| = \|x\| \sup_{\|y^*\| = 1} \|T^* y^*\|
\]
\[
\implies \|T\| = \sup_{\|x\| = 1} \|Tx\| \leq \sup_{\|x\| = 1} \|T^* y^*\| = \|T^*\| = \|T^* y^*\| = \sup_{\|y^*\| = 1} \|T^* y^*\|
\]

**Properties of adjoints:**

(i) If \( I \) is the identity operator from \( X \) to \( X \), then
\[
T^* = I
\]

(ii) \((A_1 + A_2)^* = A_1^* + A_2^*\)

(iii) \((\alpha A)^* = \alpha A^* \quad \forall \alpha \in \mathbb{R}\)

(iv) \((A_1 A_2)^* = A_2^* A_1^*\)

(v) \((A^{-1})^* = (A^*)^{-1}\)

\[
\]
\( Y^* = Y \) and \( X^* = X \). In this case, \( T^* \) can be defined as

\[
(Tx \mid y) = (x \mid T^* y)
\]

Note that, in this case, \( T^{**} = T \).

\( T \) is called **self-adjoint** if \( T = T^* \)

Note: this definition makes sense only if \( Y = X \).

\( T \) is called **positive definite** if

\[
(Tx \mid x) \geq 0 \quad \forall x \in X
\]

**Example:** Let \( X = Y = L_2[0,1] \).

Let \( T x = \int_0^t K(t,s) x(s) \, ds \)

Find \( T^* \).

**Solution:** First, \( T \) must be a bounded linear operator,

\[
\|Tx\|^2 = \int_0^t \left( \int_0^t K(t,s) x(s) \, ds \right)^2 \, dt
\]

\[
\leq \int_0^t \left( \int_0^t |K(t,s)| \, ds \right)^2 \, dt \int_0^t |x(s)|^2 \, ds \, dt
\]

CS inequality

\[
\leq \|x\|^2 \int_0^t \int_0^t K^2(t,s) \, ds \, dt
\]
\[
< (\text{Constant}) \|x\|^2
\]

if
\[
\int_0^1 \int_0^1 K(t,s) \, ds \, dt < \infty.
\]

We will assume this is true.

\[
(Tx|y) = \int_0^1 \left[ \int_0^t K(t,s) x(s) \, ds \right] y(t) \, dt
\]

\[
= \int_0^1 \left( \int_s^t K(t,s) y(t) \, dt \right) x(s) \, ds
\]

\[
= (x, T^* y)
\]

\[
\Rightarrow \quad T^* y = \int_0^1 K(t,s) y(t) \, dt
\]

for some fixed
\[
0 \leq t_1 < t_2 < \ldots < t_n \leq 1.
\]

Example: \(X = C[0,1], \ Y = \mathbb{R}^n\) (i.e., \(\mathbb{R}^n\) with the \(l_2\)-norm). Define
\[
T(x) = (x(t_1), x(t_2), \ldots, x(t_n))
\]
Find $T^+(x)$

**Solution:** $T x \in E^n$ since $(E^n)^* = E^n$

because $E^n$ is a Hilbert space.

$((C[0,1])^*) = NBV[0,1].$

$$
\langle Tx, y^* \rangle = \sum_{i=1}^{n} x(k_i) y_i
$$

$$
= \int x(t) \, dv(t) = \langle x, T^+ y^* \rangle
$$

where

![Diagram](image)

so $T^+ y^*$ is $v(t)$ given above.

**Relationship between range & nullspace**

We will only consider the situation $T : X \to Y$, where $X$ and $Y$ are Hilbert spaces. The more general cases are treated in the book.

We will assume $T \in B(X, Y)$, i.e., $T$ is a bounded linear operator.
First recall relationships for a matrix $A$: suppose $y = Ax$, i.e., $y \in \text{R}(A)$

$$y^T \bar{y} = 0$$

$$\iff \quad x^T A^T \bar{y} = 0$$

Thus, if $A^T \bar{y} = 0$, then $y^T \bar{y} = 0$

i.e., $(\text{R}(A))^\perp = (\text{N}(A^T))$

A similar relationship holds for operators between Hilbert spaces.

**Theorem:**

(i) $(\text{R}(T^*))^\perp = \text{N}(T)$

(ii) $(\text{R}(T^*))^\perp = \text{N}(T)$

**Proof:**

(i) $T: X \to Y$ and $T^*: Y \to X$

$$\text{R}(T) \subseteq Y \quad \text{N}(T^*) \subseteq Y$$

$y \in \text{R}(T) \iff y = Tx \quad \forall x \in X$

$\bar{y} \in \text{N}(T^*) \iff 0 = T^* \bar{y}$

$\langle y | \bar{y} \rangle = \langle Tx | \bar{y} \rangle = \langle x | T^* \bar{y} \rangle = 0$
Thus, if \( g \in \mathcal{N}(T^*) \) and \( y \in (R(T))^\perp \),

then \( g \perp y \).

\[ \Rightarrow \quad \mathcal{N}(T^*) \subseteq (R(T))^\perp \]

Suppose \( g \in (R(T))^\perp \), then

\[ (Tx \mid g) = 0 \]

By the definition of adjoint,

\[ (Tx \mid g) = (x \mid T^* g) = 0 \]

Since \( g \in (R(T))^\perp \), this relationship is true for all \( x \). \( \Rightarrow \) \( T^* g = 0 \)

or \( g \in \mathcal{N}(T^*) \).

i.e., \( (R(T))^\perp \subseteq \mathcal{N}(T^*) \).

(ii) Similar to (i).

Now, let's take the complement of (i) above:

\( (R(T))^\perp = (\mathcal{N}(T^*))^\perp \)

It is easy to see that \( (R(T))^\perp \)

is a subspace. If it is closed,
by the decomposition theorem, $Y$ can be decomposed as

$$Y = R(T) \oplus R(T)$$

This decomposition is unique, i.e., given $y \in Y$, there is a unique $y_1 \in R(T)$ and $y_2 \in R(T)$ such that $y = y_1 + y_2$.

$$\Rightarrow R(T) \perp = R(T)$$

However, if $R(T)$ is not closed, then

$$R(T) = \overline{R(T)} (\text{the closure of } R(T)).$$

The reason is that both

$$(R(T))' \quad \text{and} \quad (\overline{R(T)})'$$

are the same, by the continuity of the inner product. And thus,

$$R(T) = (\overline{R(T)})' = \overline{R(T)}$$

by the decomposition theorem. Thus, we have the following corollary to the previous theorem.

**Corollary:**

(i) $\overline{R(T)} = N(T^*)'$

(ii) $\overline{R(T^*)} = N(T)'$
Example illustrating $R(T)$ need not be closed.

Before we present the example, we note that all infinite sequences with a finite number of non-zero elements $\in l_2$. The closure of this set is $l_2$.

Now comes the example.

$$T(x) = \left\{ \frac{x_1}{2^{1/2}}, \frac{x_2}{2^{1/2}}, \frac{x_3}{2^{1/2}}, \ldots \right\}$$

$T(x) \in l_2$ if $x \in l_2$. Further

$$||T|| = ||(1, \frac{1}{2^{1/2}}, \frac{1}{3^{1/2}}, \ldots)||_2 < \infty$$

$\lim_{n \to \infty} \frac{1}{n} = 0$.

Clearly, all sequences with a finite # non-zero elements $\notin R(T)$. Thus,

$$\overline{R(T)} \subset l_2.$$

However, $\left\{ 1, \frac{1}{2}, \frac{1}{3}, \ldots \right\} \in l_2$, but cannot be in $R(T)$ because $x$ has to be $\left\{ 1, \frac{1}{2^{1/2}}, \frac{1}{3^{1/2}}, \ldots \right\}$ for this to happen. But such an $\alpha \notin l_2$. 

$\Box$
\[ y = T \mathbf{x} \text{ has no solution} \]

\[
\begin{align*}
\min & \quad \| y - T \mathbf{x} \| \\
\text{This is equivalent of} & \quad \min \| y - \hat{y} \| \\
& \quad \hat{y} \in R(T)
\end{align*}
\]

If an optimum solution exists (it may not exist \( R(T) \) need not be closed),

\[ y - \hat{y} \in R(T) \]

\[ \Rightarrow \quad y - \hat{y} \in \mathcal{N}(T^\perp) \]

\[ \Rightarrow \quad T^\perp y - T^\perp \hat{y} = 0 \]

\[ \Rightarrow \quad T^\perp y = T^\perp \hat{y} \]

But \( \hat{y} = T \mathbf{x} \)

\[ \Rightarrow \quad T^\perp y = T^\perp T \mathbf{x} \]

If \( T^\perp T \) is invertible, then there
exists a unique solution
\[
x = (T^T T)^{-1} T^T y
\]

\[y = Tx \text{ has many solutions}
\]

\[
\min ||x||
\]

s.t.
\[y = Tx
\]

Let \(x_0\) be one solution to
\[y = Tx. \text{ Define}
\]
\[-z = x - x_0 \quad \forall x : Tx = y
\]
\[-Tz = Tx - T x_0 = y - y = 0
\]

Thus, the problem is
\[
\min ||z - x_0||
\]

s.t.
\[z \in M = \{z : Tz = 0\}
\]

or
\[
\min ||z - x_0||
\]

s.t.
\[z \in N^*(T)
\]

\(N^*(T)\) is closed since \(N^*(T) = R(T)\). Then,
(see Prop. 1, sec. 3.4 of book)
there exists a unique optimal \( \hat{x} \) and \( \hat{z} \). \( \hat{x} \) is optimal i.f.f.

\[ \hat{z} = x_0 \in N(T)^\perp = R(T^*) \]

i.e., \( \hat{z} \) is optimal i.f.f. we can find \( \hat{y} \) s.t.

\[ \hat{z} - x_0 = T^* \hat{y} \]

\[ \hat{z} - x_0 = x \], \( x \) is optimal i.f.f.

\( \forall \hat{y} \in \mathcal{Y} \)

\[ x = T^* \hat{y} \]

since \( x \) has to satisfy

\[ y = XT^* \hat{y} \]

Thus, if we can find a \( \hat{y} \) solving the above equation, then the optimal \( \hat{x} \) is given by

\[ \hat{x} = T^* \hat{y} \]

If \( TT^* \) is invertible, we have

\[ \hat{y} = (TT^*)^{-1} y \]

and

\[ \hat{x} = T^* (TT^*)^{-1} y \]