1. Some Basics from Linear Algebra

With these notes, I will try and clarify certain topics that I only quickly mention in class.

First and foremost, I will assume that you are familiar with many basic facts about real and complex numbers. In particular, both $\mathbb{R}$ and $\mathbb{C}$ are fields; they satisfy the field axioms. For $z = x + iy \in \mathbb{C}$, the modulus, i.e. $|z| = \sqrt{x^2 + y^2} \geq 0$, represents the distance from z to the origin in the complex plane. (As such, it coincides with the absolute value for real z.) For $z = x + iy \in \mathbb{C}$, complex conjugation, i.e. $\overline{z} = x - iy$, represents reflection about the x-axis in the complex plane. It will also be important that both $\mathbb{R}$ and $\mathbb{C}$ are complete, as metric spaces, when equipped with the metric $d(z, w) = |z - w|$.

1.1. Vector Spaces.

One of the most important notions in this course is that of a vector space. Although vector spaces can be defined over any field, we will (by and large) restrict our attention to fields $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

The following definition is fundamental.

Definition 1.1. *Let $\mathbb{F}$ be a field. A* **vector space** V **over** $\mathbb{F}$ *is a non-empty set V (the elements of V are called* **vectors***) over a field $\mathbb{F}$ (the elements of $\mathbb{F}$ are called* **scalars***) equipped with two operations:*
i) To each pair $u, v \in V$, there exists a unique element $u + v \in V$. This operation is called **vector addition***.*
ii) To each $u \in V$ and $\alpha \in \mathbb{F}$, there exists a unique element $\alpha u \in V$. This operation is called **scalar multiplication***.*
These operations satisfy the following relations:
For all $\alpha, \beta \in \mathbb{F}$ and all $u, v, w \in V$,

(1) $u + (v + w) = (u + v) + w$ *and* $u + v = v + u$
(2) *There is a vector $0 \in V$ (called the* **additive identity***) such that $u + 0 = u$ for all $u \in V$*
(3) *For each vector $u \in V$, there is a vector $-u \in V$ (called the* **additive inverse** *of u) such that $u + (-u) = 0$*
(4) $\alpha(u + v) = \alpha u + \alpha v$
(5) $(\alpha + \beta)u = \alpha u + \beta u$
(6) $(\alpha\beta)u = \alpha(\beta u)$
(7) $1u = u$ *for all* $u \in V$

The phrase "Let V be a complex (or real) vector space." means that V is a vector space over $\mathbb{F} = \mathbb{C}$ (or $\mathbb{F} = \mathbb{R}$). It is clear that every complex vector space is a real vector space.

Example 1 (Vectors)**.** *Let $\mathbb{F}$ be a field and $n \geq 1$ be an integer. Take*

$$V = \{(v_1, v_2, \cdots, v_n) : v_j \in \mathbb{F} \text{ for all } 1 \leq j \leq n\}$$

The set V is often called the collection of n-tuples with entries in $\mathbb{F}$, and some write $V = \mathbb{F}^n$. With the usual notions of addition and scalar multiplication, i.e. for $v, w \in V$ and $\lambda \in \mathbb{F}$, set

$$v + w = (v_1 + w_1, v_2 + w_2, \cdots, v_n + w_n) \quad \text{and} \quad \lambda v = (\lambda v_1, \lambda v_2, \cdots, \lambda v_n)$$

V is a vector space over $\mathbb{F}$.

Example 2 (Matrices)**.** *Let $\mathbb{F}$ be a field and take integers $m, n \geq 1$. Take*

$$V = \{A = \{a_{ij}\} : a_{ij} \in \mathbb{F} \text{ for all } 1 \leq i \leq m \text{ and } 1 \leq j \leq n\}$$

The set V is often called the collection of $m \times n$ matrices with entries in $\mathbb{F}$, and some write $V = \mathbb{F}^{m\times n}$. Here we often visualize A as a matrix with m rows and n columns, i.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

With the usual notions of addition and scalar multiplication, i.e. for $A, B \in V$ and $\lambda \in \mathbb{F}$, set

$$A + B = \{c_{ij}\} \quad \text{with } c_{ij} = a_{ij} + b_{ij} \quad \text{for all } 1 \le i \le m \text{ and } 1 \le j \le n \quad \text{and} \quad \lambda A = \{\lambda a_{ij}\}$$

V is a vector space over $\mathbb{F}$.

One can argue that Example 2 is a special case of Example 1, however, it is often useful to think of these as two distinct examples . . .

Example 3 (Functions)**.** *Consider the set*

$$V = \{f : (0,1) \to \mathbb{C} : \ f \text{ is continuous at each } x \in (0,1)\}$$

The set V is often called the collection of complex-valued, continuous functions on $(0,1)$, and some write $V = C((0,1), \mathbb{C})$. With the usual notions of addition and scalar multiplication, i.e. for $f, g \in V$ and $\lambda \in \mathbb{C}$, set

$$(f+g)(x) = f(x) + g(x) \quad \text{and} \quad (\lambda f)(x) = \lambda f(x) \quad \text{for all } x \in (0,1),$$

V is a vector space over $\mathbb{C}$.

Definition 1.2. *Let V be a vector space over $\mathbb{F}$. A non-empty set $U \subset V$ is said to be a subspace of V if U is a vector space over $\mathbb{F}$ when it is equipped with the same addition and scalar multiplication rules that make V a vector space over $\mathbb{F}$.*

To check that a (non-empty) subset $U \subset V$ is a subspace, one need only check closure under addition and scalar multiplication, i.e. $u, v \in U$ imply $u + v \in U$ and $u \in U$ imply $\lambda u \in U$ for all $\lambda \in \mathbb{F}$.

Let V be a vector space over $\mathbb{F}$ and $n \ge 1$ be an integer. Let $v_1, v_2, \cdots, v_n \in V$ and $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{F}$. The vector $v \in V$ given by

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i$$

is called a linear combination of the vectors $v_1, v_2, \cdots, v_n$.

Definition 1.3. *Let V be a vector space over $\mathbb{F}$. Let $n \ge 1$ and $v_1, v_2, \cdots, v_n \in V$. The collection of all linear combinations of the vectors $v_1, v_2, \cdots, v_n$, regarded as a subset of V, is called the span of these vectors. Our notation for this it*

$$\text{span}(v_1, v_2, \cdots, v_n) = \{v = \sum_{i=1}^{n} \lambda_i v_i : \lambda_i \in \mathbb{F} \text{ for all } 1 \le i \le n\}$$

One readily checks that for any $n \ge 1$ and any collection of vectors $v_1, v_2, \cdots, v_n \in V$, $\text{span}(v_1, v_2, \cdots, v_n) \subset V$ is a subspace of V.

Definition 1.4. *Let V be a vector space over $\mathbb{F}$. If there is some $n \ge 1$ and vectors $v_1, v_2, \cdots, v_n \in V$ for which*

$$\text{span}(v_1, v_2, \cdots, v_n) = V \, ,$$

then V is said to be finite-dimensional. Any collection of vectors for which the above is true is called a spanning set for V. If V is not finite dimensional, then V is said to be infinite-dimensional.

Let us return to our examples.

Consider Example 1. The collection of n-tuples $\{e_j\}$ with $1 \leq j \leq n$ defined by $e_j = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$ with the multiplicative identity $1 \in \mathbb{F}$ in the j-th component and the additive identitiy $0 \in \mathbb{F}$ in all other components is a spanning set for $V = \mathbb{F}^n$. In this case, V is finite-dimensional.

Consider Example 2. The collection of matrices $\{E_{ij}\}$ defined by fixing $1 \leq i \leq m$ and $1 \leq j \leq n$ and declaring that E_{ij} has a 1 in the i, j entry and 0 in all other entries has $mn < \infty$ elements. One checks that this is a spanning set for $V = \mathbb{F}^{m \times n}$, and hence V is finite-dimensional.

Consider Example 3. This vector space is not finite-dimensional. In fact, for any $n \geq 1$, one can construct n disjoint compact intervals in $(0, 1/2)$. For each of these intervals, choose a non-zero, continuous function supported in that interval. The span of these functions will clearly not include any continuous function compactly supported in $(1/2, 1)$.

Definition 1.5. *Let V be a vector space over $\mathbb{F}$. A collection of vectors $v_1, v_2, \cdots, v_n \in V$ is said to be linearly independent if the only solution of the equation*

$$\sum_{i=1}^{n} \lambda_i v_i = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

with $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{F}$ is $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Definition 1.6. *Let V be a finite-dimensional vector space over $\mathbb{F}$. A collection of vectors $v_1, v_2, \cdots, v_n \in V$ is said to be a basis of V if the collection is a linearly independent, spanning set. In other words, the collection $v_1, v_2, \cdots, v_n$ is linearly independent and* $\text{span}(v_1, v_2, \cdots, v_n) = V$.

One can prove that every finite dimensional vector space has a basis. One can also prove that for a fixed finite-dimensional vector space V, any two bases have the same number of elements.

Definition 1.7. *Let $V \neq \{0\}$ be a finite-dimensional vector space over $\mathbb{F}$. Denote by* $\dim(V)$ *the number of elements in any basis of V. This positive integer is called the dimension of V. By convention, we take* $\dim(\{0\}) = 0$.

Consider Example 1. The collection of n-tuples $\{e_j\}$, defined previously, is a basis of $V = \mathbb{F}^n$. As such, $\dim(\mathbb{F}^n) = n$.

Consider Example 2. The collection of matrices $\{E_{ij}\}$, defined previously, is a basis of $V = \mathbb{F}^{m \times n}$. As such, $\dim(\mathbb{F}^{m \times n}) = mn$.

1.2. **Metric Spaces, Normed Spaces, and Inner-Product spaces.**

1.2.1. *Definitions.* In this class, we will encounter spaces with various *structures* on them. We will here discuss three important examples of these structures.

Metric Spaces: We start with the notion of a metric space.

Definition 1.8. *A metric on a (non-empty) set X is a function $\rho : X \times X \to [0, \infty)$, with $(x, y) \mapsto \rho(x, y)$, which satisfies the following:*

(1) $\rho(x, y) = 0$ *if and only if* $x = y$*;*
(2) $\rho(x, y) = \rho(y, x)$ *for all* $x, y \in X$*;*
(3) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ *for all* $x, y, z \in X$*.*

A set X equipped with a metric is called a metric space; this is often written as (X, ρ).

Typically, $\rho(x, y)$ is interpreted as the distance between x and y in X.

Note that, in general, a metric space need not have an additive structure; i.e. a metric space is not always a vector space. What will be important for us is that, in the context of metric spaces, one can define *completeness*.

Definition 1.9. *Let (X, ρ) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty}$ in X is said to converge to $x \in X$ if $\lim_{n\to\infty} \rho(x_n, x) = 0$. This may be written as $x_n \to x$. A sequence $\{x_n\}_{n=1}^{\infty}$ in X is said to be Cauchy if $\rho(x_n, x_m) \to 0$ as $m, n \to \infty$. The metric space (X, ρ) is said to be complete if every Cauchy sequence converges to an element of X.*

Normed Spaces: Next, we consider normed spaces. A normed space is a vector space with a notion of the *length* of a vector. This statement is made precise with the following definition.

Definition 1.10. *Let V be a complex vector space. A map: $\| \cdot \| : V \to \mathbb{R}$, $x \mapsto \|x\|$, is said to be a norm on V if it satisfies the following properties:*

(1) **Positive Definiteness:** *For each $x \in V$, $\|x\| \geq 0$ and moreover, $\|x\| = 0$ if and only if $x = 0$.*
(2) **Positive Homogeneity:** *For each $x \in V$ and $\lambda \in \mathbb{C}$,*
$$\|\lambda x\| = |\lambda| \|x\|.$$
(3) **Triangle Inequality:** *For each $x, y \in V$,*
$$\|x + y\| \leq \|x\| + \|y\|.$$

V is said to be a normed space if it is a complex vector space equipped with a norm; this is often written $(V, \| \cdot \|)$.

One can easily show that every normed space $(V, \| \cdot \|)$ is a metric space when V is equipped with the metric $\rho(x, y) = \|x - y\|$.

Definition 1.11. *Let $(V, \| \cdot \|)$ be a normed space. If the corresponding metric space (V, ρ), with $\rho(x, y) = \|x - y\|$, is complete, then V is said to be a Banach space.*

Inner-Product Spaces: Finally, we consider inner-product spaces. A inner-product space is a vector space with a notion of *angles* between vectors. This statement is made precise with the following definition.

Definition 1.12. *Let V be a complex vector space. A map: $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$, $(x, y) \mapsto \langle x, y \rangle$, is said to be an inner-product on V if it satisfies the following properties:*

(1) **Second Component Linear:** *For each $x, y, z \in V$ and any $\lambda \in \mathbb{C}$, one has that*
$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad \textit{and} \quad \langle x, \lambda y \rangle = \lambda \langle x, y \rangle$$

(2) **Positive Definiteness:** *For each* $x \in V$, $\langle x, x\rangle \geq 0$ *and moreover,* $\langle x, x\rangle = 0$ *if and only if* $x = 0$.

(3) **Conjugate Symmetry:** *For each* $x, y \in V$, $\langle y, x\rangle = \overline{\langle x, y\rangle}$.

V is said to be an inner-product space if it is a complex vector space equipped with an inner-product; this is often written $(V, \langle\cdot,\cdot\rangle)$.

Although we will not consider such examples in this class, one can also define the notion of a normed space as well as an inner-product space for *real* vector spaces.

It is not hard to show that every inner-product space is a normed space. In fact, let $(V, \langle\cdot,\cdot\rangle)$ be an inner-product space. The map $\|\cdot\| : V \to [0, \infty)$ with

$$\|x\| = \sqrt{\langle x, x\rangle} \quad \text{for each } x \in V\,,$$

is easily checked to be a norm.

In fact, the first step in this proof is the following.

Theorem 1.13 (Cauchy-Schwarz Inequality). *Let V be a complex vector space. For each $x, y \in V$, one has that*

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

where $\|x\| = \sqrt{\langle x, x\rangle}$ *as discussed above.*

Given the bound above, it is straight-forward to verify that every complex inner-product space is a normed space; this is a homework problem. As such it is also a metric space and so the notion of completeness is relevant. For clarity, the metric here is:

$$\rho(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle} \quad \text{for all } x, y \in V$$

Definition 1.14. *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner-product space. If (V, ρ), regarded as a metric space with metric coming from the norm as described above, is complete, then V is said to be a Hilbert space.*

Remark: Not all normed spaces are inner-product spaces. In fact, one can show that the norm on a normed space $(V, \|\cdot\|)$ arises from an inner-product if and only if the norm satisfies

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \quad \text{for all } x, y \in V\,.$$

This well-known relation is called the Parallelogram Law. In particular, not all Banach spaces are Hilbert spaces. Examples of this include $L^1(\mathbb{R})$ and $C((0,1))$ equipped with $\|\cdot\|_\infty$. In this homework, we will check that every finite dimensional inner-product space is a Hilbert space.

1.2.2. *Some Examples.* Consider Example 1. One readily checks that

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x_i} \cdot y_i$$

defines an inner-product on $V = \mathbb{C}^n$. In this case, $V = \mathbb{C}^n$ is a normed space when equipped with

$$\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

Consider Example 2. One readily checks that

$$\langle A, B\rangle_{\text{HS}} = \text{Tr}[A^* B]$$

defines an inner-product on $V = \mathbb{C}^{n\times n}$. This is called the Hilbert-Schmidt inner product. In this case, $V = \mathbb{C}^{n\times n}$ is a normed space when equipped with

$$\|A\|_{\text{HS}} = \sqrt{\text{Tr}[A^* A]}$$

which is called the Hilbert-Schmidt norm.

Consider Example 2 again. One readily checks that $V = \mathbb{C}^{n \times n}$ is a normed space when equipped with

$$\|A\| = \sup_{\substack{\psi \in \mathbb{C}^n: \\ \psi \neq 0}} \frac{\|A\psi\|}{\|\psi\|}$$

This norm is called the operator norm.

Consider Example 3. One readily checks that

$$\langle f, g \rangle = \int_0^1 \overline{f(x)} \cdot g(x)\, dx$$

defines an inner-product on $V = C((0,1), \mathbb{C})$. In this case, $V = C((0,1), \mathbb{C})$ is a normed space when equipped with

$$\|f\| = \sqrt{\int_0^1 |f(x)|^2\, dx}$$

Remark: One should note that $V = C((0,1), \mathbb{C})$ is not complete with respect to the norm given above.

1.2.3. *On Orthogonality.* The notion of perpendicular vectors is quite useful. This can be done quite generally in inner-product spaces.

Definition 1.15. *Let V be a complex inner-product space. Two vectors $x, y \in V$ are said to be orthogonal if $\langle x, y \rangle = 0$. This is often written $x \perp y$.*

Note: the zero vector is the only vector that is orthogonal to itself. In fact, the zero vector is orthogonal to every vector in a vector space.

The next result follows immediately.

Theorem 1.16 (Pythagorean Theorem)**.** *Let V be a complex inner-product space. Let $n \geq 1$ and suppose $v_1, v_2, \cdots, v_n \in V$ are a collection of pairwise orthogonal vectors in V, i.e. $\langle v_j, v_k \rangle = 0$ whenever $1 \leq j, k \leq n$ and $j \neq k$. In this case, one have that*

$$\|\sum_{i=1}^n v_i\|^2 = \sum_{i=1}^n \|v_i\|^2$$

From this, one is lead to the definition:

Definition 1.17. *Let V be a finite-dimensional complex inner-product space. The vectors $v_1, v_2, \cdots, v_n \in V$ are said to be an orthonormal basis if the vectors $v_1, v_2, \cdots, v_n \in V$ are a basis and*

$$\langle v_j, v_k \rangle = \delta_{jk}$$

for all $1 \leq j, k \leq n$. Here δ_{jk} is the Kronecker delta. (It is 1 if $j = k$ and 0 otherwise.)

It is often quite useful to express quantities in terms of orthonormal bases.

In fact, since an orthonormal basis is a basis, one can write any $v \in V$ as

$$v = \sum_{i=1}^n c_i v_i$$

whenever $v_1, v_2, \cdots, v_n \in V$ is an orthonormal basis. In this case, one readily checks that

$$c_i = \langle v_i, v \rangle \quad \text{and moreover} \quad \|v\|^2 = \sum_{i=1}^n |c_i|^2 = \sum_{i=1}^n |\langle v_i, v \rangle|^2$$

In other words, both the coefficient of v and the norm of v can be calculated in a relatively straightforward manner; once you specify an orthonormal basis in V.

1.3. **On Linear Maps.** Much of linear algebra is devoted to the study of linear maps.

Definition 1.18. *Let V and W be complex vector spaces. A function $T : V \to W$ is said to be linear if*

$$\begin{aligned} T(u+v) &= T(u) + T(v) \quad \text{for all } u, v \in V, \\ T(av) &= aT(v) \qquad\quad \text{for all } a \in \mathbb{C} \text{ and } v \in V. \end{aligned}$$

The set of all linear functions $T : V \to W$ will be denoted by $\mathcal{L}(V,W)$. We will often say $T \in \mathcal{L}(V,W)$ is a linear map or a linear transformation; we may also suppress the word linear and just refer to $T \in \mathcal{L}(V,W)$ as a map (or transformation) from V to W. In the special case that $V = W$, we write $\mathcal{L}(V)$ for $\mathcal{L}(V,V)$ and refer to $T \in \mathcal{L}(V)$ as a linear operator (or just an operator) on V.

It is an easy exercise to verify that $\mathcal{L}(V,W)$ is itself a vector space. In fact, for any $S, T \in \mathcal{L}(V,W)$ and $a \in \mathbb{C}$, vector addition and scalar multiplication are defined as follows:

$$(S+T)(v) = S(v) + T(v) \quad \text{for all } v \in V, \tag{1}$$

and

$$(aT)(v) = aT(v) \quad \text{for all } a \in \mathbb{C} \text{ and any } v \in V \tag{2}$$

Definition 1.19. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. A linear map $T : X \to Y$ is said to be bounded if there is a number $M \geq 0$ for which*

$$\|Tx\|_Y \leq M\|x\|_X \quad \text{for all } x \in X. \tag{3}$$

If $T : X \to Y$ is a linear map and there is no number $M \geq 0$ for which (3) holds, then T is said to be unbounded. We will denote by $\mathcal{B}(X,Y)$ the collection of all bounded linear maps $T : X \to Y$. In the special case that $X = Y$, we set $\mathcal{B}(X) = \mathcal{B}(X,X)$ and call $T \in \mathcal{B}(X)$ a bounded linear operator on X.

If $T : X \to Y$ is a bounded linear map, then we define the operator norm of T, which we denote by $\|T\|$, to be

$$\|T\| = \inf\{M \geq 0 : \|Tx\|_Y \leq M\|x\|_X \quad \text{for all } x \in X\}. \tag{4}$$

It is convenient to observe several different expressions for $\|T\|$:

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y = \sup_{\|x\|_X = 1} \|Tx\|_Y\,. \tag{5}$$

We will check these equalities in a homework problem.

Theorem 1.20. *A linear map is bounded if and only if it is continuous.*

Proof. Let $T \in \mathcal{B}(X,Y)$. For each $x, y \in X$, the bound

$$\|T(x) - T(y)\|_Y = \|T(x-y)\|_Y \leq \|T\|\|x-y\|_X \tag{6}$$

is clear. In this case, one sees continuity at $y \in X$ by taking $\epsilon > 0$ and letting $\delta = \frac{\epsilon}{\|T\|}$. Since this is true for any $y \in X$, T is continuous.

Now suppose T is continuous at $x = 0$. Let $\epsilon = 1$. Then, by the definition of continuity, there is a corresponding $\delta > 0$ for which

$$\|T(x)\|_Y = \|T(x) - T(0)\|_Y \leq 1 \quad \text{whenever} \quad \|x\|_X = \|x - 0\|_X \leq \delta\,. \tag{7}$$

Consider now any $x \in X \setminus \{0\}$. Define $\tilde{x} \in X$ by setting

$$\tilde{x} = \delta \frac{x}{\|x\|_X} \tag{8}$$

Under this scaling, $\|\tilde{x}\|_X \le \delta$ and so $\|T(\tilde{x})\|_Y \le 1$. Linearity of T then shows that

$$\|T(x)\|_Y = \frac{\|x\|_X}{\delta}\|T(\tilde{x})\|_Y \le \frac{1}{\delta}\|x\|_X \tag{9}$$

and hence T is bounded. □

The above proof actually shows that continuity at 0 implies continuity at all x for linear maps.

Theorem 1.21 (BLT Theorem)**.** *Let $(X, \|\cdot\|_X)$ be a normed space and $(Y, \|\cdot\|_Y)$ be a Banach space. If $M \subset X$ is a dense subspace and $T : M \to Y$ is a bounded linear map, then there is a unique bounded linear map $\overline{T} : X \to Y$ for which $\overline{T}(x) = x$ for all $x \in M$. Moreover, $\|\overline{T}\| = \|T\|$.*

Here the phrase BLT refers to the bounded linear transformation whose existence is the main point of this result. In words, Theorem 1.21 shows that any bounded linear map defined on a subspace of a normed space can be uniquely extended to a bounded linear map on the closure of that subspace. Moreover, it is also a crucial observation that this unique extension does not increase the operator norm of the original bounded map.

Proof. Let $x \in X$. Since $M \subset X$ is dense, there is a sequence $\{x_n\}_{n\ge1} \subset M$ with $x_n \to x$. Consider the sequence $\{T(x_n)\}_{n\ge1} \subset Y$. Since T is a bounded linear map on M,

$$\|T(x_m) - T(x_n)\|_Y = \|T(x_m - x_n)\|_Y \le \|T\|\|x_m - x_n\|_X \to 0 \quad \text{as } m, n \to \infty\,. \tag{10}$$

Here we use that convergent sequences in X are necessarily Cauchy. Since Y is a Banach space, the Cauchy sequence $\{T(x_n)\}_{n\ge1}$ converges, and so we may define $\overline{T} : X \to Y$ by setting

$$\overline{T}(x) = \lim_{n\to\infty} T(x_n)\,. \tag{11}$$

Note that $\overline{T}$ is well-defined. In fact, one readily checks that this limiting value does not depend on the particular sequence converging to x. To see this, let both $\{x_n\}_{n\ge1} \subset M$ and $\{\tilde{x}_n\}_{n\ge1} \subset M$ converge to some $x \in X$. In this case,

$$\|x_n - \tilde{x}_n\|_X \le \|x_n - x\|_X + \|x - \tilde{x}_n\|_X \to 0 \quad \text{as } n \to \infty \tag{12}$$

and therefore,

$$\|T(x_n) - T(\tilde{x}_n)\|_Y \le \|T\|\|x_n - \tilde{x}_n\|_X \to 0 \quad \text{as } n \to \infty\,. \tag{13}$$

It is clear that $\overline{T}$ is an extension of T because for any $x \in M$, one may evaluate $\overline{T}(x)$ by using the constant sequence $\{x_n\}_{n\ge1} \subset M$ with $x_n = x$ for all $n \ge 1$. Checking linearity of $\overline{T}$ is a simple exercise.

To see that $\overline{T}$ is bounded, note that for any $x \in X$:

$$\|\overline{T}(x)\|_Y = \lim_{n\to\infty} \|T(x_n)\|_Y \le \lim_{n\to\infty} \|T\|\|x_n\|_X = \|T\|\|x\|_X \tag{14}$$

and thus $\|\overline{T}\| \le \|T\|$. Using the sup expression for the operator norm, for any $\epsilon > 0$, there is some $x_\epsilon \in M \setminus \{0\}$ for which

$$\|T\| - \epsilon \le \frac{\|T(x_\epsilon)\|_Y}{\|x_\epsilon\|_X} \le \|\overline{T}\| \tag{15}$$

using again that $\overline{T}$ is a bounded extension of T.

To see that this bounded linear extension is unique, let $\tilde{T}$ be any other bounded linear extension of T. Let $x \in X$ and $\{x_n\}_{n\ge1} \subset M$ be a sequence converging to x. Since $\tilde{T}$ is bounded, it is continuous and thus:

$$\tilde{T}(x) = \lim_{n\to\infty} \tilde{T}(x_n) = \lim_{n\to\infty} T(x_n) = \overline{T}(x) \tag{16}$$

and so the two maps coincide. □

more will be added here . . .
define linear maps
define vector space of maps with norm
state result on completeness
stress example of dual
State Riesz rep theorem for Hilbert spaces
define the matrix of a linear map
define inverse
define isomorphic
define eigenvalue
define eigenvector
self-adjoint matrices
they have real eigenvalues
lable important matrices: unitary hermitian orthogonal . ..

1.4. **Some Important Theorems.** Below we discuss some important theorems from linear algebra; this needs more editing . . .

1.4.1. *The Spectral Theorem.* One of the most important theorems for quantum mechanics is the spectral theorem. We briefly discuss it below. Recall that for any complex vector space V, the set $\mathcal{L}(V)$ is the set of all linear maps $T : V \to V$.

Definition 1.22. *Let V be a complex, inner-product space. $T \in \mathcal{L}(V)$ is said to be normal if and only if $TT^* = T^*T$.*

In words, T is normal if and only if T commutes with its adjoint.

Recall: the map $[\cdot,\cdot] : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V)$, $(S,T) \mapsto [S,T]$, defined by

$$[S,T] = ST - TS \quad \text{for all } S,T \in \mathcal{L}(V)$$

is called the commutator of S and T. S and T are said to commute if $[S,T] = 0$.

It is clear from Definition 1.22 that all self-adjoint operators $T \in \mathcal{L}(V)$ are normal.

Theorem 1.23 (Spectral Theorem)**.** *Let V be a finite-dimensional, complex inner-product space and $T \in \mathcal{L}(V)$. T is normal if and only if there is an orthonormal basis of V consisting entirely of eigenvectors of T.*

Let's restrict our attention mainly to self-adjoint matrices.

Let $A \in M_d$ be self-adjoint. By the spectral theorem, there is a basis of $\mathbb{C}^d$, which we label by $u_1, u_2, \cdots, u_d \in \mathbb{C}^d$, for which:

$$\langle u_j, u_k \rangle = \delta_{jk} \quad \text{i.e. the basis is orthonormal}$$

and moreover, there are numbers $\lambda_j \in \mathbb{R}$ for which

$$Au_j = \lambda_j u_j \quad \text{for all } 1 \leq j \leq d\,.$$

With respect to this basis then, each $u \in \mathbb{C}^d$ can be written as

$$u = \sum_{i=1}^{d} c_j u_i \quad \text{with} \quad c_i = \langle u_i, u \rangle$$

since the basis is orthonormal, and moreover, A acts trivially

$$Au = \sum_{i=1}^{d} c_i A u_i = \sum_{i=1}^{d} \lambda_i \langle u_i, u \rangle u_i$$

Written differently, this shows that

$$A = \sum_{i=1}^{d} \lambda_i P_i$$

where P_i is the orthogonal projection onto the one-dimensional subspace spanned by u_i.

More is true. If we take $U \in M_d$ to be the matrix whose columns are given by the u_i, i.e,

$$U = \big(u_1 | u_2 | \cdots | u_d\big)$$

then U is unitary, i.e. $U^*U = \mathbb{1}$ and moreover,

$$U^*AU = \mathrm{diag}(\lambda_j)$$

i.e. A is unitarily diagonalizable.

The above fact inspires the definition of a function of an operator.

Let $f : \mathbb{R} \to \mathbb{C}$ be any function. For any $A \in M_d$ that is self-adjoint, let U be the unitary matrix constructed above. Define $f(A) \in M_d$ by setting

$$f(A) = U\mathrm{diag}(f(\lambda_j))U^*$$

If f is real-valued, then one easily checks that this matrix $f(A)$ is self-adjoint, and a short calculation shows that

$$f(A) = \sum_{i=1}^{d} f(\lambda_j) P_j$$

where the P_j are the orthogonal projections described above.

Of course, for the simple function $f(x) = x$ we recover what we knew; namely

$$A = f(A) = \sum_{j=1}^{d} f(\lambda_j) P_j = \sum_{j=1} \lambda_j P_j$$

as expected.

Defining functions of self-adjoint (and more generally normal) matrices is often called *functional calculus*. It has many applications.

Let $H \in M_d$ be self-adjoint. For any $t \in \mathbb{R}$, consider the function $f_t : \mathbb{R} \to \mathbb{C}$ given by $f_t(x) = e^{-itx}$. The matrix $f_t(H) = e^{-itH} \in M_d$ will come up often in quantum mechanics. This unitary matrix, with $f_t(H)^* = e^{itH}$, will be crucial in understanding the solution of Schrödinger's equation.

1.4.2. *The Singular Value Decomposition.* We state two versions of the singular value decomposition. One for matrices $A \in M_d$ and one for general matrices $A \in \mathbb{C}^{m \times n}$.

Case 1: (Square matrices) Let $A \in M_d$. Recall that for any $z \in \mathbb{C}$, we can write z in polar form. In other words, we can write $z = e^{i\theta}|z|$ where $|z|$ is the modulus of z and $\theta \in [0, 2\pi)$. A similar fact is true for matrices.

Theorem 1.24 (Polar Decomposition)**.** *Let $A \in M_d$. There is a unitary $U \in M_d$ for which $A = U|A|$. Writing A in this manner is called the polar decomposition of A.*

We will not prove this but note the following.

For any $A \in M_d$, the matrix $A^*A \in M_d$ is clearly non-negative. By the spectral theorem, this self-adjoint operator may be written as

$$A^*A = \sum_{j=1}^{d} \sigma(A)_j^2 P_{\psi_j}$$

with $\sigma(A)_j^2$ being the non-negative eigenvalues of A^*A. The self-adjoint operator

$$|A| = \sum_{j=1}^{d} \sigma(A)_j P_{\psi_j}$$

is the one appearing in the statement above. The non-negative numbers $\sigma(A)_j$ are called the singular values of A.

Theorem 1.25 ((square) Singular Value Decomposition)**.** *For any $A \in M_d$, there are two orthonormal bases $\{e_1, e_2, \cdots, e_d\}$ and $\{f_1, f_2, \cdots, f_d\}$ of $\mathbb{C}^d$ for which*

$$Ax = \sigma(A)_1 \langle e_1, x\rangle f_1 + \sigma(A)_2 \langle e_2, x\rangle f_2 + \cdots + \sigma(A)_d \langle e_d, x\rangle f_d = \sum_{j=1}^{d} \sigma(A)_j \langle e_j, x\rangle f_j$$

for all $x \in \mathbb{C}^d$.

Using the Polar decomposition statement above, the proof is simple. Let $A \in M_d$. Let $\{e_1, e_2, \cdots, e_d\}$ be the orthonormal basis of eigenvectors of $|A|$ whose existence is guaranteed by the spectral theorem. Take $\{f_1, f_2, \cdots, f_d\}$ to be $f_j = Ue_j$ where U is the unitary whose existence is guaranteed by the Polar decomposition theorem above. One easily checks that $\{f_1, f_2, \cdots, f_d\}$ is also an orthonormal basis of $\mathbb{C}^d$, and moreover, for any $x \in \mathbb{C}^d$,

$$|A|x = \sum_{j=1}^{d} \sigma(A)_j \langle e_j, x \rangle e_j \quad \Rightarrow \quad Ax = U|A|x = \sum_{j=1}^{d} \sigma(A)_j \langle e_j, x \rangle f_j$$

as claimed.

Case 2: (General matrices) We will consider integers $m \geq 1$ and $n \geq 1$ and matrices $A \in \mathbb{C}^{m \times n}$.

Let $A \in \mathbb{C}^{m \times n}$ and label the entries $A = \{a_{ij}\}$ with $a_{ij} \in \mathbb{C}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. To each such A there is a matrix $A^* \in \mathbb{C}^{n \times m}$ with entries $A^* = \{a^*_{ji}\}$ given by $a^*_{ji} = \overline{a_{ij}}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. We will refer to this matrix A^* as the adjoint of A. It is easy to check that the realtion

$$\langle Ax, y \rangle_{\mathbb{C}^m} = \langle x, A^* y \rangle_{\mathbb{C}^n} \quad \text{holds for all } x \in \mathbb{C}^n \text{ and } y \in \mathbb{C}^m \, .$$

Based on this observation, the following is clear.

Let $A \in \mathbb{C}^{m \times n}$. The matrix $B = A^*A \in \mathbb{C}^n$ is non-negative, in fact,

$$\langle x, Bx \rangle_{\mathbb{C}^n} = \langle Ax, Ax \rangle_{\mathbb{C}^m} = \|Ax\|^2_{\mathbb{C}^m} \geq 0$$

As such, B is self-adjoint and has non-negative eigenvalues. Let us write this as

$$B = \sum_{j=1}^{n} \sigma(A)_j^2 P_{\Psi_j}$$

using the spectral theorem. We will continue to refer to the non-negative numbers $\{\sigma(A)_j\}$ as the singular values of A.

Note: It is probably also useful to observe that: If $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times m}$, then $AB \in \mathbb{C}^{m \times m}$ and $BA \in \mathbb{C}^{n \times n}$ and moreove,

$$\mathrm{Tr}[AB] = \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{m} b_{ji} a_{ij} = \sum_{j=1}^{n} (BA)_{jj} = \mathrm{Tr}[BA]$$

This means that one can define the Hilbert Schmidt norm more generally . . .

Theorem 1.26 ((general) Singular Value Decomposition)**.** *For any $A \in \mathbb{C}^{m \times n}$, there are two unitary matrices $U_1 \in M_m$ and $U_2 \in M_n$ for which*

$$A = U_1 D U_2^*$$

where $D \in \mathbb{C}^{m \times n}$ is diagonal with non-negative entries. The diagonal entries of D are the singular values of A.

There is a uniqueness question . . . What about orthogonality in the real case?

The proof (from Bill Jacob) goes as follows. Let $A \in \mathbb{C}^{m \times n}$. Assume WLOG that $n \leq m$. (Otherwise, apply the argument below to A^*; must check that one gets the same singular values . . .)

As indicated above, the matrix $B = A^*A \in M_n$ is non-negative, and hence, self-adjoint. By the spectral theorem, there exists an orthonormal basis of eigenvectors of B; we label them by $\{u_1, u_2, \cdots, u_n\}$ and note that $A^*Au_j = Bu_j = \sigma(A)_j^2 u_j$. The unitary matrix $U_2 \in M_n$ in the statement of the result is the matrix with these vectors as columns, i.e. $U_2 = (u_1|u_2|\cdots|u_n)$. Note further that

$$\langle Au_j, Au_k \rangle_{\mathbb{C}^m} = \langle A^*Au_j, u_k \rangle_{\mathbb{C}^n} = \sigma(A)_j^2 \delta_{jk}$$

as the vectors $\{u_1, u_2, \cdots, u_n\}$ are orthonormal. Thus the vectors $\{Au_1, Au_2, \cdots, Au_n\} \subset \mathbb{C}^m$ are pairwise orthogonal and normalized so that $\|Au_j\|^2 = \sigma(A)_j^2$.

For any j such that $\sigma(A)_j > 0$, define normalized vectors $w_j \in \mathbb{C}^m$ by setting $\sigma(A)_j w_j = Au_j$. If $A \neq 0$, there is at least one such non-trivial vector. Denote by $\{w_1, w_2, \cdots, w_m\}$ an extension of this list of non-trivial, orthonormal vectors to a complete orthonormal basis of $\mathbb{C}^m$. Let $U_1 \in M_m$ be the unitary matrix whose columns are given by these vectors, i.e. $U_1 = (w_1|w_2|\cdots|w_m)$.

Take $D \in \mathbb{C}^{m\times n}$ to be the diagonal matrix whose entries are $d_{ii} = \sigma(A)_i$. If we denote by $\{e_1, e_2, \cdots, e_n\}$ the standard basis in $\mathbb{C}^n$ and $\{e'_1, e'_2, \cdots, e'_m\}$ the standard basis in $\mathbb{C}^m$ it is easy to check that $U_2^* u_i = e_i$ and $U_1 e'_j = w_j$. In this case, we find that

$$(U_1 D U_2^*)(u_i) = (U_1 D)(e_i) = \sigma(A)_i (U_1)(e'_i) = \sigma(A)_i w_i = Au_i$$

As we have checked this on a basis, we are done.