MTHSC 853: Linear Algebra

7. Euclidean structure:

Goal: Abstract the concept of Euclidean distance.

Let's review basic Euclidean structure.

Let X be a real n-dimensional Euclidean space (e.g., from vector calculus), with 0 the zero vector.

The **length** of $x \in X$, denoted $\|x\|$ = distance from x to 0.

By the Pythagorean theorem, $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$.

The **dot product** of $x, y \in X$ is $(x, y) = \sum_{j=1}^{n} x_j y_j$.

Clearly, $\|x\|^2 = (x, x)$.

Since the dot product is symmetric & bilinear,

$$(x+y, x+y) = (x, x) + 2(x, y) + (y, y)$$
$$= \|x\|^2 + 2(x, y) + \|y\|^2 = \|x+y\|^2$$
$$(x-y, x-y) = \|x\|^2 - 2(x, y) + \|y\|^2 = \|x-y\|^2 \qquad (*)$$

Note that this is independent of choice of coordinate system.

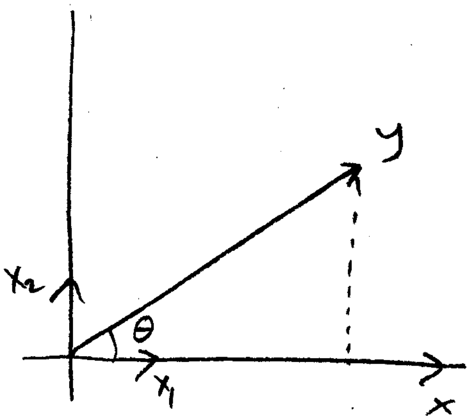
Geometrically, we understand $\|x\|$, $\|y\|$, and $\|x-y\|$.

To understand (x, y), we'll pick a "special" x & y.

Given any coordinate system (orthogonal basis);

Pick x to be a scalar of x_1

Pick $y \in \text{Span}(x_1, x_2)$.

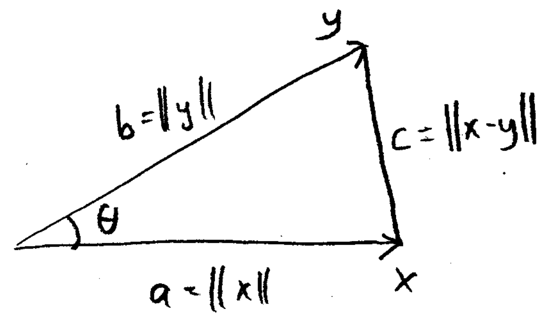


$x = (\|x\|, 0, \dots, 0)$

$y = (\|y\| \cos\theta, \|y\| \sin\theta, 0, \dots, 0)$

$\Rightarrow (x, y) = \|x\| \, \|y\| \cos\theta \quad \Rightarrow \boxed{\cos\theta = \frac{(x, y)}{\|x\| \cdot \|y\|}}$

We can derive the law of cosines from $(*)$: $c^2 = a^2 + b^2 - 2ab\cos\theta$.



The dot product gives us a notion of geometry (lengths and angles). We can abstract this

Def: A <u>Euclidean structure</u> in a real vector space is endowed by an <u>inner product</u>, which is a symmetric bilinear form with the additional property that $(x, x) \geq 0$ with equality iff $x = 0$ ("positivity").

We'll show that all of Euclidean geometry follows from having an innerproduct.

* Throughout, assume that X is an n-dimensional real inner product space.

Def: The norm of $x \in X$ is $\|x\| = (x,x)^{1/2}$, and the distance between $x, y \in X$ is $\|x-y\|$.

Theorem 7.1 (Cauchy-Schwarz). For all $x, y \in X$, $|(x,y)| \leq \|x\| \cdot \|y\|$.

Proof: Define a function $q : \mathbb{R} \to \mathbb{R}$ as $q(t) = \|x+ty\| \geq 0$.

Write $q(t) = \|x\|^2 + 2t(x,y) + t^2\|y\|$.

Assume $y \neq 0$ (resut is trivial in this case).

Put $t = \dfrac{-(x,y)}{\|y\|^2}$:

$$q(t) = \|x\|^2 - \frac{2(x,y)^2}{\|y\|^2} + \frac{(x,y)^2}{\|y\|^2} = \|x\|^2 - \frac{(x,y)^2}{\|y\|^2} \geq 0 \qquad \square$$

Cor: For any $x, y \in X$, $-1 \leq \dfrac{(x,y)}{\|x\| \cdot \|y\|} \leq 1$ $\square$

Def: The *angle* θ between $x, y \in X$ is $\theta = \cos^{-1} \dfrac{(x,y)}{\|x\| \cdot \|y\|}$.

Theorem 7.2: $\|x\| = \max \{ (x,y) : \|y\| = 1 \}$.

Proof: Exercise (HW).

Theorem 7.3 (Triangle inequality): For all $x, y \in X$, $\|x+y\| \leq \|x\| + \|y\|$.

Proof: Note that
$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2(x,y) + \|y\|^2 \\ &\leq \|x\| + 2\|x\| \cdot \|y\| + \|y\|^2 \quad \text{(Cauchy-Schwarz)} \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$
$\square$

Def: Two vectors $x, y \in X$ are *orthogonal* if $(x,y) = 0$. We write this as $x \perp y$.

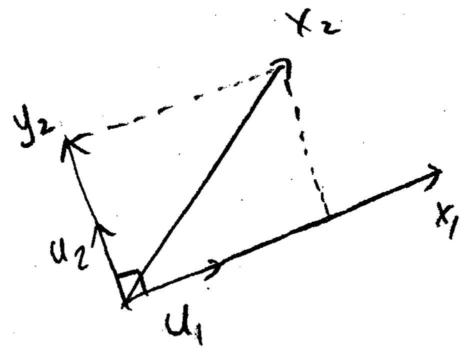
Remark: If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. (Pythagorean theorem)

Def: A basis $x_1, \dots, x_n$ for X is *orthonormal* (w.r.t. a Euclidean structure) if $(x_i, x_j) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$.

Theorem 7.4 (Gram-Schmidt): Given an arbitrary basis $x_1, \dots, x_n$, there is a related orthonormal basis $u_1, \dots, u_n$ for which $u_k \in \operatorname{Span}(x_1, \dots, x_k)$.

Proof: Just construct it.

If $\|u\| = 1$, define $\text{proj}_u(x) = (x, u)\, u$.



Put

$$y_1 := x_1 \qquad u_1 := \frac{y_1}{\|y_1\|}$$

$$y_2 := x_2 - \text{proj}_{u_1} x_2 \qquad u_2 := \frac{y_2}{\|y_2\|}$$

$$y_3 = x_3 - \text{proj}_{u_1} x_3 - \text{proj}_{u_2} x_3 \qquad u_3 := \frac{y_3}{\|y_3\|}$$

$$\vdots \qquad\qquad \vdots$$

$$y_k = x_k - \sum_{i=1}^{k-1} \text{proj}_{u_i} x_k \qquad y_k := \frac{y_k}{\|y_k\|}$$

It is easy to check that this works. $\square$

Remark: When put in matrix form, the matrix $A := [x_1\ x_2 \ldots x_n]$ can be factored as

$$\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} (x_1, u_1) & (x_2, u_1) & (x_3, u_1) & \cdots \\ 0 & (x_2, u_2) & (x_3, u_2) & \cdots \\ \vdots & 0 & (x_3, u_3) & \cdots \\ & \vdots & 0 & \ddots \\ & & \vdots & \end{bmatrix}$$

$$A \quad = \quad Q \quad \cdot \quad R$$

This is called the QR-factorization of A.

R is upper-triangular because $u_k \in \text{Span}(x_1, \ldots, x_k)$.

If $x_1, \dots, x_n$ is an orthonormal basis, then for any $x \in X$,

$$x = \sum_{j=1}^{n} a_j x_j \quad \text{where} \quad a_j = (x, x_j).$$

Moreover, if $y = \sum_{k=1}^{n} b_k x_k$, then

$$(x, y) = \left(\sum a_j x_j , \sum b_k x_k\right) = \sum_{j=1}^{n} (a_j x_j , b_j x_j) = \sum_{j=1}^{n} a_j b_j .$$

In particular, if $y = x$, then $\|x\|^2 = \sum_{j=1}^{n} a_j^2$.

Thus, the mapping $X \to \mathbb{R}^n$, $x \longmapsto (a_1, \dots, a_n)$, $a_j = (x, x_j)$

is an isomorphism that carries the inner product of X

to the standard dot product of $\mathbb{R}^n$.

Theorem 7.5: Every linear function $\ell \in X'$ can be written as

$\ell(x) = (x, y)$ for some $y \in X$.

Proof: Let $x_1, \dots, x_n$ be an orthonormal basis, and let $b_k = \ell(x_k)$.

Put $y = \sum_{k=1}^{n} b_k x_k$.

Claim: This works. (Easy to check). $\square$

Corollary: The mapping $R_y : X \to X'$, $y \mapsto (-, x)$ is an isomorphism.

Remark: There is an analogous map $L_X : X \longrightarrow X'$, $x \longmapsto (x, -)$.

Def: Let Y be a subspace of X. The orthogonal complement of Y is the set $\{x \in X : (x,y) = 0 \ \forall y \in Y\}$.

Remark: In Section 2, we defined $Y^\perp = \{\ell \in X' : (\ell, y) = 0 \ \forall y \in Y\}$.

Under the natural identification $x \xmapsto{L_X} (x, -)$, these two sets are the same, thus we will also denote the orthogonal complement of Y by $Y^\perp$.

Theorem 7.6: For any subspace Y of X, $X = Y \oplus Y^\perp$.

Proof: We must show that for any $x \in X$, we can uniquely write $x = y + y^\perp$, where $y \in Y$, $y^\perp \in Y^\perp$.

Uniqueness: Suppose $x = y + y^\perp = z + z^\perp$, $y, z \in Y$, $y^\perp, z^\perp \in Y^\perp$.

Then $y - z = z^\perp - y^\perp \in Y \cap Y^\perp \Rightarrow y - z \perp y - z$

$\Rightarrow 0 = (y-z, y-z) = \|y-z\|^2 \Rightarrow y = z$.

Existence: Let $x_1, \dots, x_k$ be an orthonormal basis of Y.

Extend to an orthormal basis $x_1, \dots, x_n$ of X.

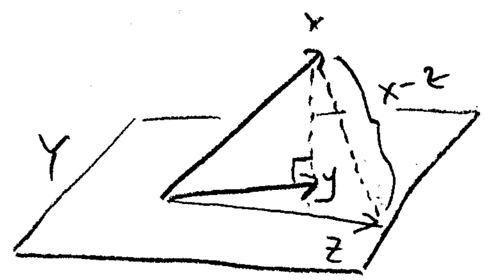
We get $x = \sum_{j=1}^{n} a_j x_j = \sum_{i=1}^{k} a_i x_i + \sum_{j=k+1}^{n} a_j x_j = y + y^\perp$. $\square$

<u>Def</u>: The map $P_Y : X \to X$, $x = y + y^\perp \mapsto y$ is called the <u>orthogonal projection</u> of x onto Y.

<u>Theorem 7.7</u>: P_Y is <u>linear</u> and <u>idempotent</u> (i.e., $P_Y^2 = P_Y$).

<u>Proof</u>: Exercise.

<u>Theorem 7.8</u>: Let $Y \subseteq X$ be a subspace. Then $P_Y(x) = z$, where $\|x - z\| = \min\{\|x - y\| : y \in Y\}$.



<u>Proof</u>: Let $y = P_Y(x)$, and write $x - z = (y - z) + y^\perp$.

By Pythagorean theorem, $\|x - z\|^2 = \|y - z\|^2 + \|y^\perp\|^2$ is minimized when $z = y$. $\square$

Now, let X, U be Euclidean spaces and $A : X \to U$ linear.

We can identify X with X' and U with U', and under this identification, the transpose of A maps $U \to X$.

We call the transpose of A the <u>adjoint</u> of A, denoted A^*.

<u>Full definition</u>: Let $A: X \to U$ be a linear map between Euclidean spaces. For any $u \in U$, $\ell(x) = (Ax, u)$ is a linear function $X \to \mathbb{R}$. By Theorem 7.5, for some $y \in X$,

$\ell(x) = (x, y) = (Ax, u)$.

The vector $y \in X$ depends (linearly) on $u \in U$, i.e., for some function $A^*: U \to X$, $y = A^* u$.

Thus, we have $(x, A^* u) = (Ax, u)$

scalar product in X ← Scalar product in U.

<u>Theorem 7.9</u>:

(i) If $A, B: X \to U$ are linear, then $(A+B)^* = A^* + B^*$.

(ii) If $A: X \to U$, $C: U \to V$ are linear, then $(CA)^* = A^* C^*$.

(iii) If $A: X \to X$ is 1-1, then $(A^{-1})^* = (A^*)^{-1}$.

(iv) $(A^*)^* = A$.

<u>Proof</u>: Let $x \in X$, $u \in U$, $v \in V$.

(i) $((A+B)x, u) = (Ax, u) + (Bx, u) = (x, A^* u) + (x, B^* u) = (x, (A^* + B^*) u)$. ✓

(ii) $(CAx, v) = (Ax, C^* v) = (x, A^* C^* v)$ ✓

(iii) By (ii) & the easy fact that $I^* = I$: $I = (A^{-1}A)^* = A^*(A^{-1})^*$. ✓

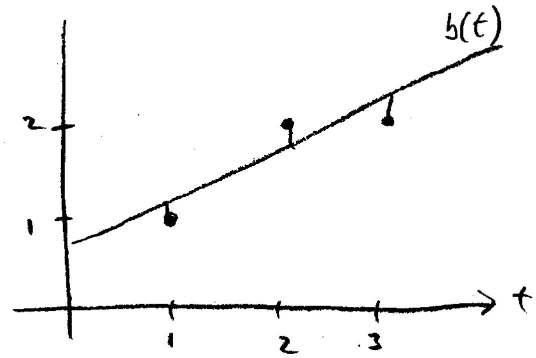
(iv) $(Ax, u) = (u, Ax) = (A^* u, x) = (u, A^{**} x) = (A^{**} x, u)$. ✓ $\square$

As before, the matrix representations of A & A^* are transposes of each other.

<u>Application</u>: "Least squares."

<u>Example</u>: Consider 3 data points in $\mathbb{R}^2$: $(1,1)$, $(2,2)$, $(3,2)$.

<u>Goal</u>: Find the "best fit" line
$b(t) = C + Dt$.

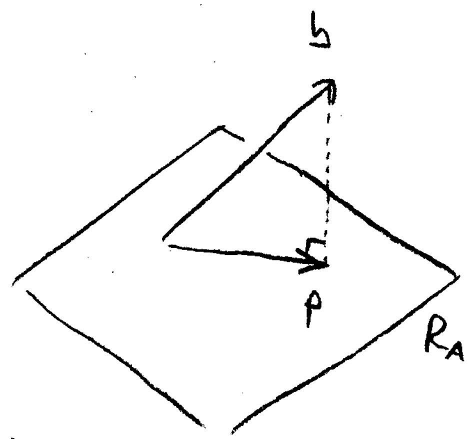


A "perfect solution" would solve the following:

$$\begin{cases} C + D = 1 \\ C + 2D = 2 \\ C + 3D = 2 \end{cases} \implies \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{b}.$$

Clearly, $Ax = b$ has no soln.

Next best thing: Solve $Az = p$,
where $p = \mathrm{proj}_{R_A}(b)$.
= "closest vector in R_A to b."



<u>How</u>: Solve $\boxed{A^*Az = A^*b}$ instead – this works!

This minimizes $\|Az - b\|^2$ too!

Least squares.

Given: Overdetermined system $Ax=b$, where A is $m\times n$, $m>n$.

Remark: A solution to $Ax=b$ (if one exists) is unique iff

$Ay=0 \Rightarrow y=0$. [Reason: All sol'ns $= x_p + N_A$.]

Theorem 7.10: Let A be $m\times n$, $m>n$, and $N_A=\{\vec{0}\}$.

The (unique) vector x that minimizes $\|Ax-b\|^2$ is

the solution to $A^*Az=A^*b$.

Proof: Step 1: Show $A^*Az=A^*b$ has a unique solution.

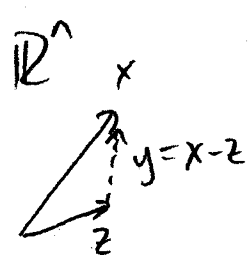
Need to show that A^*A has nullspace $\{\vec{0}\}$.

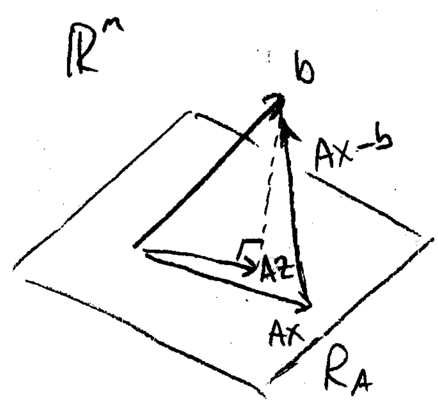
If $A^*Ay=0$ then $0=(A^*Ay,y)=(Ay,Ay)=\|Ay\|^2=0$

Thus, $Ay=0 \Rightarrow y=0$. ✓

Step 2: Claim: If $z\in\mathbb{R}^n$ has the property that $Az-b\perp R_A$,

then z minimizes $\|Ax-b\|$





Proof: Pick $x\in\mathbb{R}^n$.

Let $y=x-z$.

Goal: Show $\|Ax-b\|$ is minimized when $y=0$.

$$Ax-b = A(y+z)-b = (Az-b)+Ay \quad \text{and} \quad Az-b \perp Ay.$$

Pythagorean theorem $\Rightarrow \|Ax-b\|^2 = \|Az-b\|^2 + \|Ay\|^2$

Clearly, this is minimized if $\|Ay\|=0 \Rightarrow y=0 \Rightarrow x=z.$ ✓

Step 3: Show that this particular z satisfies $A^*Az=A^*b$.

We showed that $(Az-b, Ay)=0$ for all $y \in X$.

$\Rightarrow (A^*(Az-b), y)=0$ for all $y \in X$

By assumption, rank $A=n \Rightarrow$ rank $A^*=n$

$\Rightarrow A^*(Az-b)=0 \Rightarrow A^*Az=A^*b.$ ✓ $\square$

Theorem 7.11: If P_Y is the orthogonal projection onto Y, then $P_Y^* = P_Y$.

Proof: Exercise

Def: A function $M: X \to X$ is an isometry if for all $x, y \in X$, $\|Mx-My\| = \|x-y\|$. ("Distance-preserving.")

Example: Any translation $Mx = x+a$ is an isometry.

Remark: Given any isometry, one can compose it with a translation to get an isometry that fixes 0. Conversely, any isometry can be decomposed into one that fixes 0, followed by a translation.

Theorem 7.12: Let $M : X \to X$ be an isometry fixing 0.

Then: (i) M is linear

(ii) $M^* M = I$ (and conversely, this implies M is an isom.)

(iii) M is invertible, and M^{-1} is an isometry.

(iv) $\det M = \pm 1$.

Proof: Pick $x, y, z \in X$ and say $Mx = x'$, $My = y'$, $Mz = z'$.

Note that $\|Mx\| = \|x\|$. (Take $y = 0$).

(i) We have $\|x'\| = \|x\|$, $\|y'\| = \|y\|$, and $\|x' - y'\| = \|x - y\|$.

Thus, $\|x'\|^2 - 2(x', y') + \|y'\|^2 = \|x' - y'\|^2 = \|x - y\|^2 = \|x\|^2 - 2(x, y) + \|y\|^2$

This shows $(x, y) = (x', y')$, i.e., M preserves inner products.

Next, suppose $z = x + y$. We'll show $z' = x' + y'$.

(14)

$$\|z'-x'-y'\|^2 = \|z'\|^2 + \|y'\|^2 + \|x'\|^2 - 2(z',x') - 2(z',y') + 2(x',y')$$

$$\|z-x-y\|^2 = \|z\|^2 + \|y\|^2 + \|x\|^2 - 2(z,x) - 2(z,y) + 2(x,y).$$

Since M preserves norms & inner products:

$$\|z'-x'-y'\|^2 = \|z-x-y\|^2 = 0 \Rightarrow z'-x'-y' = 0. \checkmark$$

(ii) $(x,y) = (Mx, My) = (x, M^*My)$

Bilinearity $\Rightarrow (x,y) - (x, M^*My) = (x, M^*My - y) = 0$

Since this holds for all $x \in X$, $M^*My - y = 0$.

[Note: Reverse the steps for the converse.]

(iii) If $\|Mx\| = 0$ then $\|x\| = 0 \Rightarrow M$ is invertible. $\checkmark$

(M^{-1} is clearly an isometry.)

(iv) Since $M^*M = I$, $(\det M^*)(\det M) = 1$

Recall that $\det M^* = \det M \Rightarrow \det M = \pm 1$. $\checkmark$ $\square$

The geometric meaning of Theorem 7.12 is that any map that preserves <u>distances</u> is <u>linear</u> (i), and preserves both <u>angles</u> (see proof of (i)) and volume (iv).

<u>Def</u>: A matrix that maps $\mathbb{R}^n$ to itself isometrically is <u>orthogonal</u>. The orthogonal matrices (fixed n) form a group under multiplication, called the <u>orthogonal group</u>, denoted $O(n)$. The subgroup of matrices with determinant 1 is called the <u>special orthogonal group</u>, denoted $SO(n)$.

<u>Prop</u>: A matrix M is orthogonal iff its columns vectors form an orthonormal set.

<u>Proof</u>: Exercise.

Recall that the determinant is one way to measure the "size" of a linear map from a space X to itself. But how do we measure the size of a map $X \to U$?

<u>Def</u>: If $A: X \to U$ is a linear map between Euclidean spaces, then define the <u>norm</u> of A as $\|A\| := \sup\{\|Ax\| : \|x\| = 1\}$. (Recall that sup is the <u>supremum</u>, or least upper bound of a set.)

<u>Theorem 7.13</u>: For any linear map $A: X \to U$,

(i) $\|Az\| \le \|A\|\,\|z\|$ for all $z \in X$,

(ii) $\|A\| = \sup\{(Ax, v) : \|x\| = 1, \|v\| = 1\}$.

<u>Proof</u>: (i) By definition of $\|A\|$, $\|Az\| \le \|A\| \cdot \|z\|$ for all unit vectors $z \in X$. In general, write $z = ke$, $\|e\| = 1$.

Now, $\|Az\| = \|Ake\| = \|kAe\| = |k| \cdot \|Ae\| \le |k| \cdot \|A\| \cdot \|e\| = \|A\| \cdot \|z\|$

(ii) By Theorem 7.2 $(u = Ax)$, $\|Ax\| = \max\{(Ax, v) : \|v\| = 1\}$.

By definition, $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$

$= \sup\{\max\{(Ax, v) : \|v\| = 1, \|x\| = 1\}\}$

$= \sup\{(Ax, v) : \|v\| = \|x\| = 1\}$. $\square$

<u>Theorem 7.14</u> Suppose we have linear maps $A, B: X \to U$, $C: U \to V$.

(i) $\|kA\| = |k| \cdot \|A\|$

(ii) $\|A + B\| \le \|A\| + \|B\|$

(iii) $\|CA\| \le \|C\| \cdot \|A\|$

(iv) $\|A^*\| = \|A\|$

Proof: (i) $\|kA\| = \sup\{\|kAx\| : \|x\|=1\} = |k| \cdot \sup\{\|Ax\| : \|x\|=1\}$ ✓

(ii) $\|(A+B)x\| = \|Ax+Bx\| \le \|Ax\| + \|Bx\|$

$$\|A+B\| = \sup\{\|(A+B)x\| : \|x\|=1\} \le \sup\{\|Ax\|+\|Bx\| : \|x\|=1\}$$
$$\le \sup\{\|Ax\| : \|x\|=1\} + \sup\{\|Bx\| : \|x\|=1\}$$
$$= \|A\| + \|B\|.$$ ✓

(iii) By Theorem 7.13 (i), $\|CAx\| \le \|C\| \cdot \|Ax\| \le \|C\| \cdot \|A\| \cdot \|x\|$

Now, take the supremum of both sides, over all unit vectors. ✓

(iv) $\|A\| = \sup(Ax, v) = \sup(x, A^*v) = \sup(A^*v, x) = \|A^*\|$,

where supremums are taken over all unit vectors $x, v \in X$. ✓ $\square$

Theorem 7.15: Let $A: X \to X$ be invertible. Suppose $B: X \to X$ has the property that $\|A-B\| < \frac{1}{\|A^{-1}\|}$. Then B is invertible.

Proof: Let $C = A - B$.

$B = A - C = A(I - A^{-1}C) = A(I-S)$, where $S = A^{-1}C$.

It suffices to show that $I-S$ is invertible.

Suppose not, and pick $0 \ne x \in N_{I-S}$.

Now, $(I-S)x = 0 \implies Sx = x \implies \|S\| \geq 1$ (since $x \neq 0$).

But $\|S\| = \|A^{-1}C\| \leq \|A^{-1}\| \cdot \|C\| < 1$ (by assumption). $\lightning$ $\square$

Remark: This proof assumes $\dim X < \infty$, but it also holds for $\dim X = \infty$.

Basic real analysis (review)

Def. A sequence of numbers $\{a_k\}$ converges to a if $|a_k - a| \to 0$. We say $\lim_{k\to\infty} a_k = a$.

A Cauchy sequence is any sequence $\{a_k\}$ for which $|a_k - a_j| \to 0$ as $j, k \to \infty$.

The real numbers are complete, because every Cauchy sequence converges to a limit.

The real numbers are also locally compact, that is, every bounded sequence contains a convergent sequence.

Goal: Extend these properties from numbers to vectors in a finite-dimensional Euclidean space.

<u>Def</u>: A sequence $\{x_k\}$ of vectors in a Euclidean space <u>converges</u> to a limit x if $\|x_k - x\| \to 0$ as $k \to \infty$.

A sequence $\{x_k\}$ is a <u>Cauchy sequence</u> if $\|x_k - x_j\| \to 0$ as $j, k \to \infty$. It is <u>bounded</u> if for some $R \geq 0$, $\|x_k\| \leq R$ for all k.

<u>Theorem 7.16</u>: Let X be a finite-dimensional Euclidean space.

(i) X is <u>complete</u>

(ii) X is <u>locally compact</u>.

<u>Proof</u>: (i) For any $x = (a_1, \dots, a_n)$, $y = (b_1, \dots, b_n)$, we have $|a_j - b_j| \leq \|x - y\|$.

Let $\{x_k\}$ be a Cauchy sequence: $x_k = a_{k1}, \dots, a_{kn})$.

Then each $\{a_{kj}\}_{k=1}^{\infty}$ is a Cauchy sequence, say it converges to $a_j \in \mathbb{R}$. Put $x = (a_1, \dots, a_n)$.

By definition, $\|x_k - x\|^2 = \sum_{i=1}^{n} |a_{kj} - a_j|^2 \to 0 \Rightarrow x_k \to x$ ✓

(ii) Let $\{x_k\}$ be bounded, with $\|x_k\| \leq R$.

Then $|a_{kj}| \leq \|x_k\| \leq R$ for all k.

$\mathbb{R}$ locally compact $\Rightarrow \exists$ subsequence of $\{x_k\}$ s.t. $\{a_{k1}\} \to a_1$.

This subsequence has a subsequence for which $\{a_{k2}\} \to a_2$,

and so on.

Thus, we can continue to get a subsequence for which

each $\{a_{kj}\} \to a_j$.

Put $x = (a_1, \dots, a_n)$.

Now, $\|x_k - x\|^2 = \sum_{j=1}^{n} |a_{kj} - a_j|^2 = 0 \Rightarrow x_k \to x$. ✓ □

<u>Remark</u>: We defined $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$ but in this

case, it's just $\max\{\|Ax\| : \|x\| = 1\}$: Take a subsequence

$\{x_k\}$ (say $\|x_k\| = 1$) for which $\|Ax_k\| \to \|A\|$. By

Theorem 7.16, this sequence has a subsequence $\{x_{k_i}\}$

that converges to some $x = (a_1, \dots, a_n)$.

Now, $\|Ae\| \leq \|Ax\|$ for all unit vectors $e \in X$. ✓

The converse of Theorem 7.16 holds:

<u>Theorem 7.17</u>: If a Euclidean space X is locally compact,
then $\dim X < \infty$.

<u>Proof</u>: Suppose $\dim X = \infty$ and let $y_1, y_2, \ldots$ be an infinite set of linearly independent vectors.

For each k, we can construct a sequence $x_1, \ldots, x_n$ of orthonormal vectors. Thus, we obtain an infinite sequence $x_1, x_2, \ldots$ for which $\|x_i - x_j\|^2 = \|x_i\|^2 - 2(x_i, x_j) + \|x_j\|^2 = 2$.

Thus $\{x_k\}$ contains no convergent subsequence.

<u>Def</u>: A sequence $\{A_n\}$ of maps $X \to U$ converges to a limit A if $\lim_{n\to\infty} \|A_n - A\| = 0$.

<u>Prop</u>: If $\dim X < \infty$, then $A_n \to A$ iff $A_n x \to Ax$ for all $x \in X$.

<u>Proof</u>: Exercise (Hw).

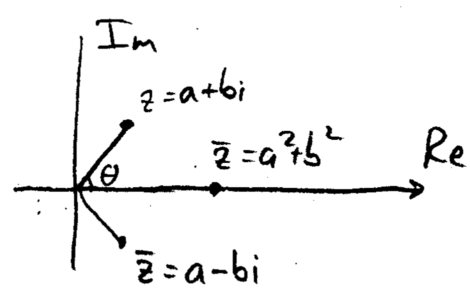
<u>Remark</u>: This does <u>not</u> hold if $\dim X = \infty$.

<u>Complex Euclidean structure</u>

<u>Review of complex numbers</u>:

Let $z = a + bi$.

$|z|^2 = z\bar{z} = a^2 + b^2$



Now, consider $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ in $\mathbb{C}^n$.

$$|z|^2 = \bar{z}^T z = [\bar{z}_1 \ \bar{z}_2 \cdots \bar{z}_n] \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2$$

If X is a finite-dimensional space over $\mathbb{C}$, then define the inner product as $(z, w) = \bar{w}^T z$

also written $\boxed{w^H z}$ "Hermitian product"

Compare to real case: $(x, y) = y^T z$.

Properties of the complex inner product:

(i) linear in x: $(kx, y) = k(x, y)$

skew-linear in y: $(x, ky) = \bar{k}(x, y)$

(ii) $\overline{(x, y)} = (y, x)$ (skew-symmetry)

(iii) $(x, x) > 0$ for all $x \neq 0$ (positivity)

In a complex Euclidean space,

$$\|x+y\|^2 = \|x\|^2 + (x, y) + (y, x) + \|y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(x, y) + \|y\|^2$$

However, most results for real spaces carry over to complex spaces – just replace $(x, y) := y^T x$ with $(x, y) := y^H x$.

Define the <u>adjoint</u> of a linear map $A: X \to U$ to be the map $A^*: U \to X$ such that $(x, A^* u) = (Ax, u)$ $\forall x \in X$, $u \in U$.

In a complex space: $(A^* u)^H x = u^H (A^*)^H x = u^H A x$

$$\Rightarrow (A^*)^H = A \quad \Rightarrow \quad \boxed{A^* = A^H = \bar{A}^T} \quad \text{"Hermitian"}$$

<u>Orthonormal</u> vectors: $q_i^T q_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ over $\mathbb{R}$

$$q_i^H q_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \text{ over } \mathbb{C}.$$

<u>Def</u>: An <u>isometry</u> of a complex space is a function $M: X \to X$ such that $\|Mx - My\| = \|x - y\|$ for all $x, y \in X$. An isometric map of a complex space is also called unitary.

<u>Prop</u>: If $M: X \to X$ is unitary, then

(i) $M^* M = I$
(ii) M^{-1} is unitary
(iii) The set of unitary maps of X forms a group.
(iv) $|\det M| = 1$.

The <u>norm</u> of a linear map between complex spaces is defined just as for real spaces: $\|A\| := \sup\{\|Ax\| : \|x\| = 1\}$.