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, and 0 the zero vector. The length of $x \in X$, denoted $\|x\|$, is the distance from x to 0 (the origin).

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,

$$(x-y, x-y) = (x,x) - 2(x,y) + (y,y)$$

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

Note that this is independent of choice of coordinate system.

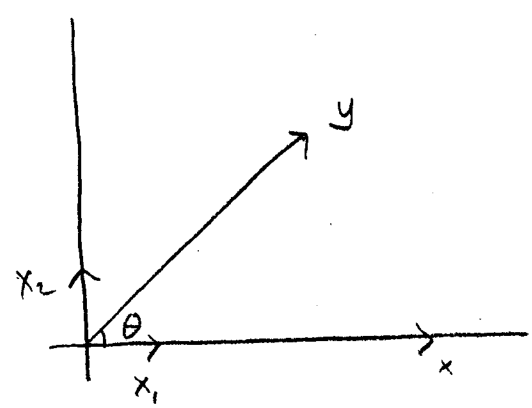
Pick x_1 so that $x = c x_1$ for some $c > 0$.

Pick x_2 so that $y \in \mathrm{Span}(x_1, x_2)$

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

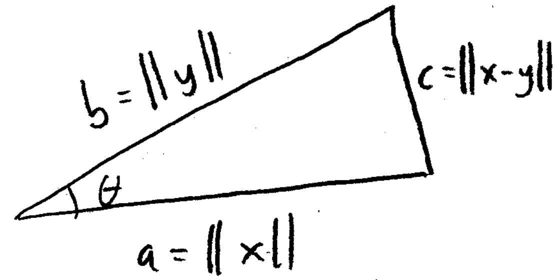
$$y = (\|y\| \cos\theta, \|y\| \sin\theta, 0, \dots, 0)$$

$$\Rightarrow (x,y) = \|x\| \, \|y\| \cos\theta$$



2

Now consider the following triangle:



We just derived the

<u>law of cosines</u>: $c^2 = a^2 + b^2 - 2ab\cos\theta$.

The dot product is what gives us this notion of geometry (i.e, lengths and angles). We can abstract this.

<u>Def</u>: A Euclidean structure 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 holding iff $x = 0$ ("positivity.")

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

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

Define the <u>norm</u> of $x \in X$ by $\|x\| = (x, x)^{1/2}$, and the <u>distance</u> between $x, y \in X$ by $\|x - y\|$.

<u>Theorem 7.1</u>: (Cauchy-Schwarz). For all $x, y \in X$,

$$|(x, y)| \leq \|x\| \cdot \|y\|.$$

Proof: Define a function $q: \mathbb{R} \to \mathbb{R}$, $q(t) = \|x + ty\|^2 \geq 0$.

We can write $q(t) = \|x\|^2 + 2t(x,y) + t^2\|y\|^2$.

Assume $y \neq 0$, put $t = \frac{-(x,y)}{\|y\|^2}$:

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

Note that the result is trivial if $y = 0$. $\square$

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 $x \perp y$.

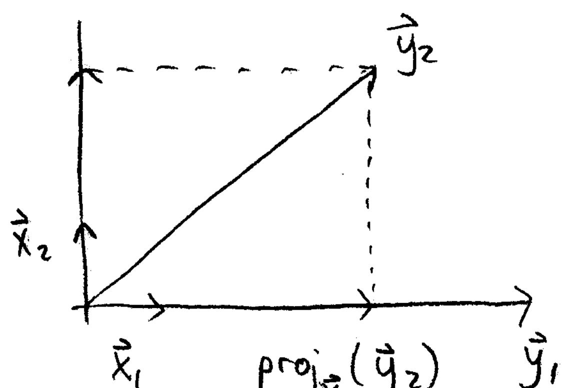
Remark: If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagorean theorem)

<u>Def</u>: A basis $\vec{x}_1, \ldots, \vec{x}_n$ for X is <u>orthonormal</u> (wrt a Euclidean structure) if $(x^{(i)}, x^{(j)}) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

<u>Theorem 7.4</u>: (Gram-Schmidt) Given an arbitrary basis $\vec{y}_1, \ldots, \vec{y}_n$ there is a related orthonormal basis $\vec{x}_1, \ldots, \vec{x}_n$ for which $\vec{x}_k \in \mathrm{Span}\{\vec{y}_1, \ldots, y_k\}$.

<u>Proof</u>: Just construct it.

If $\|e\| = 1$, define $\mathrm{proj}_e(y) = (y, e)e$



Put $\vec{x}_1 = \dfrac{\vec{y}_1}{\|\vec{y}_1\|}$

$$\vec{x}_2 = c_2\left(\vec{y}_2 - \mathrm{proj}_{\vec{x}_1}(\vec{y}_2)\right)$$

$$\vdots$$

$$\vec{x}_k = c_k\left(\vec{y}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\vec{x}_j}(\vec{y}_k)\right)$$

(we can pick c_k so that $\|\vec{x}_k\| = 1$). $\square$

Note that the condition that $\vec{x}_k \in \mathrm{Span}\{\vec{y}_1, \ldots, \vec{y}_k\}$ means that the change of basis matrix is <u>upper-triangular</u>.

If $\vec{x}_1, \dots, \vec{x}_n$ is an orthonormal basis, then for any $x \in X$,

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

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

$$(x, y) = \sum_{k=1}^{n} \sum_{j=1}^{n} a_j b_k (\vec{x}_j, \vec{x}_k) = \sum_{j=1}^{n} a_j b_j$$

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

Thus, the mapping $X \longrightarrow \mathbb{R}^n$, $x \longmapsto (a_1, \dots, a_n)$, $a_j = (x, \vec{x}_j)$ is an isomorphism that carries the inner product of X to the standard dot product of $\mathbb{R}^n$.

<u>Theorem 7.5</u>: Every linear function $\ell \in X'$ can be written as $\ell(x) = (x, y)$ for some $y \in X$.

<u>Proof</u>: Let $\vec{x}_1, \dots, \vec{x}_n$ be an orthonormal basis, and let $b_k = \ell(\vec{x}_k)$. Put $y = \sum_{k=1}^{n} b_k \vec{x}_k$.

<u>Claim</u>: This works. (Easy to check)

$\square$

<u>Corollary</u>: The mapping $R_y : X \longrightarrow X'$, $y \longmapsto f(-, y)$ is an isomorphism.

<u>Remark</u>: There is an analogous map $L_x : X \longrightarrow X'$, $x \longmapsto f(x, -)$.

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

<u>Remark</u>: 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 $Y^\perp$.

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

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

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

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

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

<u>Existence</u>: Let $\vec{x}_1, \ldots, \vec{x}_k$ be an orthonormal basis of Y.

Extend to an orthonormal basis $\vec{x}_{k+1}, \ldots, \vec{x}_n$ of X.

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

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

<u>Theorem 7.7</u>: P_Y is linear, 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>: Write $x - z = (y - z) + y^\perp$, $y = P_Y(x)$.

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

when $z = y$. $\square$

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

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

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 $X \longrightarrow \mathbb{R}$.

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 $A^* : U \longrightarrow X$, $y = A^* u$.

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

scalar product in X scalar product in U.

Theorem 7.9

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

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

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

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

Proof: 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) and 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.

Application: ("Least squares" method from numerical analysis)

In many situations, we need to solve an overdetermined system of equations, i.e, $Ax = b$, where A is an $m \times n$ matrix with $m > n$, $x, b \in \mathbb{R}^n$.

In general, there is no solution. The least squares method seeks to find the "best approximation" by minimizing $\|Ax - b\|^2$.

Remark: Such a solution is unique only if $Ay = 0 \Rightarrow y = 0$, because if x is a solution and $Ay = 0$, then $Ax = A(x + ky)$, for any $k \in \mathbb{R}$. Call such a y a nullvector of A.

Theorem 7.10: Let A be an $m \times n$ matrix, $m > n$, whose only nullvector is 0. The (unique) vector x that minimizes $\|Ax - b\|^2$ is the solution to $A^*Az = A^*b$.

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

Recall (Section 3) that this holds iff the homogeneous equation $A^*Az = 0$ has a unique sol'n (i.e, only the trivial sol'n).

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

This implies that $Ay = 0 \Rightarrow y = 0$. ✓

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 $x \in \mathbb{R}$ and let $y = x - z$. (Goal: show that $\|Ax - b\|$ is minimized when $y = 0$).

$Ax - b = A(y + z) - b = (Az - b) + Ay$, and $Az - b \perp Ay$.

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 such a vector 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 $A^* = n$.

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

□

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

<u>Proof</u>: Exercise.

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

<u>Example</u>: Any <u>translation</u> $M(x) = x + a$ is an isometry.

<u>Remark</u>: 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.

<u>Theorem 7.12</u>: Let $M: X \to X$ be an isometry that fixes 0. Then:
(i) M is linear
(ii) $M^*M = I$ (And conversely, this implies M is an isometry)
(iii) M is invertible, and M^{-1} is an isometry.
(iv) $\det M = \pm 1$.

<u>Proof</u>: Pick $x, y, z \in X$ and say $M(x) = x'$, $M(y) = y'$, $M(z) = z'$.
Note that $\|M(x)\| = \|x\|$ (take $y = 0$).

(i) We have $\|x'\| = \|x\|$, $\|y'\| = \|y\|$, and $\|x' - y'\| = \|x - y\|$.
Thus, $\|x'\|^2 - 2(x', y') + \|y'\|^2 = \|x' - y'\|^2 = \|x - y\|^2 = \|x\| - 2(x, y) + \|y\|^2$.

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

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

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

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

Since M preserves norms & inner products,

$$\|z'-x'-y'\|^2 = \|z-x-y\|^2 = 0 \implies z'-x'-y'=0. \quad ✓$$

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

Thus, $(x, M^*My-y)=0$ holds for all $x\in X$,

and so $M^*My-y=0$.

(Note: Reverse the steps for the converse.)

(iii) If $\|M(x)\|=0$ then $\|x\|=0 \implies M$ is invertible ✓

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

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

Recall that $\det M^*=\det M$, and so $\det M=\pm 1$. ✓ $\square$

The geometric meaning of Theorem 7.12 is that any map that preserves distances is linear (i), and preserves both angles (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 $O(n, K)$.

The subgroup of matrices with determinant 1 is called the <u>special orthogonal group</u>, denoted $SO(n, K)$.

<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 into itself.

But how do we measure the size of a map $X \longrightarrow U$?

<u>Def</u>: If $A: X \longrightarrow U$ is a linear map between Euclidean spaces, then define the <u>norm</u> of A to be $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$

(Recall that sup is the <u>supremum</u>, or least upper bound of a set.)

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

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

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

Proof: (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, $\|Az\| = \|Ake\| = \|k\,Ae\| = |k|\cdot\|Ae\| \le |k|\cdot\|A\|\cdot\|e\| = \|A\|\cdot\|z\|$. ✓

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

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

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

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

□

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

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

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

$\square$

<u>Theorem 7.15</u>: 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.

<u>Proof</u>: Let $C = A - B$.

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

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

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

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

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

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

<u>Review some basic real analysis:</u>

<u>Def.</u> A sequence of numbers $\{a_k\}$ <u>converges</u> to a if $|a_k - a| \to 0$.

We say $\lim_{k\to\infty} a_k = a$.

A <u>Cauchy sequence</u> 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 sequentially compact, that is, every bounded sequence contains a convergent subsequence.

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

Def: A sequence $\{x_k\}$ of vectors in a Euclidean space converges to a limit x, i.e., $\lim_{k\to\infty} x_k = x$, if $\|x_k - x\| \to 0$ as $k \to \infty$. A sequence $\{x_k\}$ is a Cauchy sequence if $\|x_k - x_j\| \to 0$ as $j, k \to \infty$. It is bounded if for some $R \geq 0$, $\|x_k\| \leq R$ for all k.

Theorem 7.16: Let X be a finite-dimensional Euclidean space.

(i) Every Cauchy sequence converges (i.e., X is complete).

(ii) Every bounded sequence contains a convergent subsequence (i.e., it is sequentially compact.)

Proof: (i) Note that 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, $x_k = (a_{k,1}, \dots, a_{k,n})$.

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

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

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

Then $|a_{k,j}| \le \|x_k\| \le R$, for all k.

Since $\mathbb{R}$ is sequentially compact, there is a subsequence of $\{x_k\}$ for which $\{a_{k,1}\} \to a_1$.

This subsequence contains a subsequence for which $\{a_{k,2}\} \to a_2$, and so on.

Thus, we can continue to get a subsequence for which each $\{a_{k,j}\} \to a_j$. Let $x = (a_1, \dots, a_n)$.

$\|x_k - x\|^2 = \sum_{i=1}^{n} |a_{k,j} - a_j|^2 = 0 \Rightarrow x_k \to x.$ ✓ $\square$

<u>Remark</u>: We defined $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$ but in this case, it's just $\max\{\|Ax\| : \|x\| = 1\}$: Take a sequence $\{x_k\}$ $(\|x_k\| = 1)$ 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, $\|Ae\| \le \|Ax\|$ 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 sequentially compact, then $\dim X < \infty$.

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

For each k, we can construct a sequence $x_1, \dots, x_n$ of orthonormal vectors. Thus we obtain an infinite sequence $x_1, x_2, \dots$ 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. $\square$

<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 $\dim X = \infty$.

Complex Euclidean structure:

Let X be a finite-dimensional space over $\mathbb{C}$.

Define the inner product as $(x, y) = \sum_{i=1}^{n} x_i \bar{y}_i$, where

$x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, and $\bar{y}_i$ denotes complex conjugation.

We want to define the adjoint of $A : X \to U$ so that $(Ax, u) = (x, A^* u)$. But it is no longer just the transpose of A.

Suppose $A = (a_{ij})$, so that $(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j$.

Now, $(Ax, u) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij} x_j\right)\bar{u}_i$

$$\begin{aligned} &= (a_{11}x_1 + \cdots + a_{1n}x_n)\bar{u}_1 \\ &+ (a_{21}x_1 + \cdots + a_{2n}x_n)\bar{u}_2 \\ &\;\vdots \qquad\qquad\qquad\qquad \vdots \\ &+ (a_{n1}x_1 + \cdots a_{nn}x_n)\bar{u}_n = \sum_{j=1}^{n} x_j\left(\overline{\sum_{i=1}^{n} \bar{a}_{ij} u_i}\right) = (x, A^* u) \end{aligned}$$

Thus, $A^* u = \sum_{i=1}^{n} \bar{a}_{ij} u_i \Rightarrow (A^* u)_i = \sum_{j=1}^{n} \bar{a}_{ji} x_j$.

So A^* in matrix form is the <u>complex conjugate transpose</u> of A.

<u>Def</u>: A complex Euclidean structure in a complex vector space is endowed by a <u>complex scalar product</u>, denoted (x, y) such that: (i) (x, y) is a linear function of x (y fixed)

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

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

<u>Remark</u>: The complex scalar product is <u>not</u> bilinear. Rather, it is a <u>skew-linear</u> function of y:

$$(x, ky) = \bar{k}(x, y) \qquad \text{for all } x, y \in X,\ k \in \mathbb{C}.$$