MATh 255 Lecture Notes

Sam Kirkiles

January 2021

Contents

1 Metric Spaces and Normed Spaces **1**

2 Holder and Minkowski Inequalities **5**

3 Inner Product Spaces **9**

1 Metric Spaces and Normed Spaces

Definition 1 *A pair (X, d) where X is a set and d is a function $d : X \times X \to R$ is called a metric space if the following holds 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) (Triangle Inequality) $d(x, y) \leq d(x, z) + d(z, y)$

The following is called a metric or a distance function on X.

A metric space (X, d) is a natural abstract "minimal" structure for discussions of the notions like convergence and continuity. To discuss differentiation, one needs more regularity and that is where normed spaces enter. In this course, we will not discuss in much detail thee differentiation theory of normed-spaces-we will only enter the framework to this study.

Example 1 *The most basic one, ande the entire Analayis 1 was devoted to it: $d(x, y) = |x - y|$. This example is the building block fo the mostexamples we will discuss.*

Example 2 $X = \mathbb{R}^n; x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$

$$d(x, y) = \left(\sum_{k=1}^{n} (x_k - y_k)^2 \right)^{1/2}$$

(X, d) is sometimes called Euclidean space and d is called Euclidean metric. Another possible choice of metric on R^n is

$$d(x,y) = \max_{1 \leq k \leq n} |x_k - y_k|$$

We will prove later that these functions d are indeed metrics on $\mathbb{R}^{\ltimes}$.

Example 3 *Let $[a,b]$ be a bounded closed interval in $\mathbb{R}$ and let $C([a,b])$ be the collection of all continuous functions*

$$f : [a,b] \to \mathbb{R}$$

$C([a,b])$ is a vector space over $\mathbb{R}$, with operations of addition and scalar multiplication defined by

$$j(f+g)(f) = f(f) + g(f)$$

$$(\alpha f)(f) = \alpha f(f)$$

The vector space $C[a,b]$ is infinite dimensional.
For $f, y \in C[a,b]$ we set

$$\begin{aligned} d(f,g) &= \sup_{a \leq f \leq b} \mid f(f) - g(f) \mid \\ &= \sup\{|f(f) - g(f)| : f \in [a,b]\} \\ &= \max\{|f(f) - g(f)| : f \in [a,b]\} \\ &= \max_{a \leq f \leq b} |f(f) - g(f)| \end{aligned}$$

Let us prove now that d is a metric on $C[a,b]$.

Property 1 $d(f,g) \geq 0$ *obvious*

Property 2

$$\begin{aligned} & d(f,g) = 0 \\ \Leftrightarrow & \sup_{a \leq f \leq b} |f(\xi) - g(f)| = 0 \\ \Leftrightarrow & |f(f) - g(f)| = 0 \quad \forall f \in [a,b] \mid \\ \Leftrightarrow & f = g \end{aligned}$$

Property 3 $d(f,g) = d(g,f)$ *follows from* $|f(f) - g(f) = |(g(f) - f(f)|$ *for all* $f \in [a,b]$

Property 4 *The triangle inequalty. Let $f, g, g \in C([a,b])$.*

Then, for any $t \in [a,b]$,

$$|f(t) - g(t)| \leq |f(t) - g(t)|_h(t) - g(t)|$$

impiles that

$$|f(t) - g(t)| \leq |f(t) - g(t)| \leq \sup_{a \leq s \leq b} |f(s) - f(s)| + \sup_{a \leq s \leq b} |f(s) - g(s)|$$

$$= d(f,h) + d(h,g)$$

So,

$$\sup_{a} |f(t) - g(t)| \leq d(f,h) + d(h,g)$$

or

$$d(f,g) \leq d(f,h) + d(h,g)$$

Example 4 *Let X be an arbitrary set and let $d : X \times X \to [0, \infty]$ be defined by*

$$d(x,y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

Then d is a metric on X (sometimes called a discrete metric on X.

To check $(1)-(3)$ of the definition of a metric is trivial. To verify the triangle inequality,

$$d(x, y \leq d(x,z) + d(z,y)$$

Note that this inequality is trivial such that $x = y$. If $x \neq y$, then we must have that either $z \neq x$ or $z \neq y$. (Otherwise, $z = x$ and $z = y$, so $x = y$).

Hence, if x and $d(x,y) = 1$, then either $d(x,z) = 1$ or $d(z,y) = 1$ (of course, they could be also both 1).

so $d(x,y) \leq d(x,z) + d(z,y)$.

We now describe normed spaces.

Definition 2 *Let X be a vector space (over $\mathbb{R}$ or over $\mathbb{C}$)*

A function

$$x \in X \to \|x\| \in \mathbb{R}$$

is called a norm function, or just a norm on X if the following holds for all $x, x \in X$ and all scalars α.

(1) $\|x\| \geq 0$

(2) $\|x\| = 0 \iff x = 0_x$ *(zero vector in X)*

(3)

$$\|\alpha x\| = |\alpha| \|x\|$$

(4) The triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$

A pair $(X, \|\dot{|}\|)$, where X is a vector space and $\|\dot{|}\|$ is a norm on X is called a normed space.

Proposition 1 *Let* $(X, \|\dot{|}\|)$ *be a normed space and let* $d : X \times X \to \mathbb{R}$ *be defined by* $d(x, y) = \|x - y\|$

Then d *is a metric on* X.

proof:

Property 1: $d(x, y) \geq 0$ *is obvious.*

Property 2: $d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0_x \Leftrightarrow x = y$

Property 3: $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |(-1)|\|y - x\| = \|y - x\| = d(y, x)$

Property 4: $d(x, y) = \|x - y\| \leq \|(x - z) + (z - y)\| = \|x - z\| + \|z - y\| = d(x, z), d(z, y)$

Proposition 2 *Let* $x = \mathbb{R}^n$ *and* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ *For* $1 \leq p < \infty$ *we set*

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}$$

and for $p = \infty$, *one sets* $\|x\|_\infty = \max_{1 \leq k \leq n} |x_k|$.

Then, for any $1 \leq p \leq \infty$, $\| \cdot \|_p$ *is a norm on* X.

proof.

For any $1 \leq p \leq \infty$, *the properties (1)-(3) in the definition of the norm*

$$(\|x\|_p \geq 0, \|x\|_p = 0 \Leftrightarrow x = 0_x, \|\alpha x\|_p = |\alpha| \cdot \|x\|_p)$$

are obvious. So one needs to verify only the triangle inquality.

In the case of $p = 1$,

$$\|x + y\|_1 = \sum_{k=1}^{n} |x_k + y_k| \leq \sum_{k=1}^{n} (|x_k| + |y_k|) = \sum_{k=1}^{n} |x_k| + \sum_{k=1}^{n} |y_k| = \|x\|_1 + \|x\|_1$$

and so indeed, $\| \cdot \|_1$ *is a norm.*

In the case $p = \infty$, *one verifies triangle inequality as follows:*

For any $1 \leq k \leq n$,

$$\begin{aligned} |x_k + y_k| &\leq |x_k| + |y_k| \\ &\leq \max_{1 \leq k \leq n} |x_k| + \max_{1 \leq k \leq n} |y_k| \\ &= \|x\|_\infty + \|x\|_\infty \end{aligned}$$

So

$$max_{1 \leq k \leq n} |x_k + y_k| \, \|x\|_\infty + \|y\|_\infty$$

or

$$\|1 \leq k \leq n\|_\infty \leq \|x\|_\infty + \|y\|_\infty$$

So $\| \cdot \|_\infty$ *is indeed a norm on* $X = \mathbb{R}^n$.

To verify the triangle inequality in the case $1 < p < \infty$ is more tricky. Before doing so, we will make a detour and discuss first the so called Holder and Minkowsky inequalities.

2 Holder and Minkowski Inequalities

Definition 3 *Let $1 < p < \infty$ which we will call exponent. The number q is called the exponent conjugate to p if*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Obviously, $q = \frac{p}{p-1}$ and $1 < q < \infty$. The exponent conjugate to q is p; Note that if $p = 2$ then $q = 2$.

Lemma 1 *Let $a, b \geq 0$ and let $1 < p, q < \infty$ be conjugate exponents. Then*

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

.

Remark: Note that in the case p=q=2, the inequality follows from $(a-b)^2 \geq$ 0.

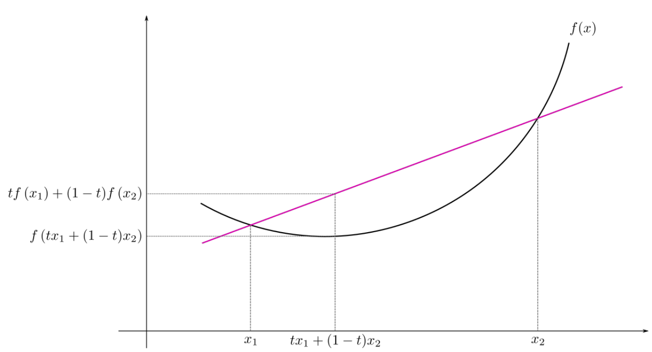
proof: The inequality is obvious if $a = 0$ or $b = 0$. So we will assume that $a > 0$ and $b > 0$.

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is called convex if for all $\alpha, \beta \in \mathbb{R}$ and all $Q \in [a, b]$,

$$f(Q\alpha + (1-Q)\beta) \leq Qf(\alpha) + (1-Q)f(\beta)$$

If f is twice differentiable, then f is convex if and only if $f''(x) \geq 0$ for all $x \in \mathbb{R}$

The geometric interpretation of convexity can be seen in this picture.



f is convex if and only if its graph on the segment $[\alpha, \beta]$ is below the straight line connecting the points $(\alpha, f\alpha))$ and $\beta, f(\beta)$.

The function $f(x) = e^x$ is convex since $f''(x) = e^x$ is convex since $f''(x) = e^x > 0$ for all x.

Let now α and β be real numbers such that

$$a = e^{\frac{\alpha}{p}} \quad and \quad b = e^{\frac{beta}{q}}$$

The numbers α and β are given by

$\alpha = pln(a), \beta = qln(b)$ (ln for $\log_e$ in the base)

Recall that

$$\frac{1}{p} + \frac{1}{q} = 1$$

and. set $Q = \frac{1}{p}$ so that $\frac{1}{q} = 1 - Q$

The convexity of $f(x) = e^x$, gives

$$f(\frac{1}{p}\alpha + \frac{1}{q}\beta) \leq \frac{1}{p}f(\alpha) + \frac{1}{q}f(\beta)$$

or

$$e^{\frac{1}{p}\alpha + \frac{1}{q}\beta} \leq \frac{1}{p}e^{\alpha} + \frac{1}{q}e^{\beta}$$

or

$$e^{\frac{1}{\rho}\alpha} \cdot e^{\frac{1}{q}\beta} \leq \frac{1}{p}\left(e^{\frac{\alpha}{p}}\right)^p + \frac{1}{q}\left(e^{\frac{B}{q}}\right)^q$$

Since

$$a = e^{\frac{\alpha}{p}} \text{ and } b = e^{\frac{b}{a}}$$

we get

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Theorem 1 *Hölder Inequality Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$ Let $1 < p, q < \infty$ be cojugate expoents. Then*

Remark: Recalling our p-functions

$$\sum_{k=1}^{n} ||x_k||y_k| \leq (\sum_{k=1}^{n} |x_k|^p)^{\frac{1}{p}} (\sum_{k=1}^{n} |y_k|^q)^{\frac{1}{q}}$$

The Hölder Inequality as

$$\sum_{k=1}^{n} |x_k||y_k| \leq \|x_k\|\|y_k\|$$

proof:

$$A = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}, B = \left(\sum_{k=1}^{n} \mid x_k|^q\right)^{\frac{1}{q}}$$

If $A = 0$, then $x_k = 0$ for all $1 \leq k \leq n$ and the Holder Inequality holds trivially. Similarly, if $B = 0$, the inequality again holds trivially. So we can assume that $A > 0, B > 0$.

For fixed k $1 \leq k \leq n$ Let

$$a = \frac{|x_k|}{A}, \quad b = \frac{|x_k|}{B}$$

Applying our lemma,

$$ab \leq \frac{[}{1}]p^p + \frac{1}{q}b^q$$

we get

$$\frac{|x_k|}{A} \cdot \frac{|y_k|}{B} \leq \frac{1}{P}\frac{|x_k|^p}{A^p} + \frac{1}{q}\frac{|y_k|^2}{B^q}$$

Summarizing the last inequality over k we get

$$\sum_{k=1}^{n} \frac{|x_k|\,|y_k|}{A \cdot B} \frac{1}{p}\frac{1}{A^p}\left(\sum_{k=1}^{n} |x_k|^p\right) + \frac{1}{q}\frac{1}{B^q}\left(\sum_{k=1}^{n} |y_k|^2\right)$$

But

$$\frac{1}{A^p}\sum_{k=1}^{n} |x_k|^p = 1, \quad \frac{1}{B^q}\sum_{k=1}^{n} |y_k|^q = 1$$

and

$$\frac{1}{p} + \frac{1}{q} = 1$$

So

$$\sum_{k=1}^{n} \frac{|x_k||y_k|}{AB} \leq 1$$

or

$$\sum_{k=1}^{n} |x_k||y_k| \leq AB$$

and that ispreciesely the statment of the Holder Inequality.

Theorem 2 *(Minkowski Inequality) Let $X = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be two vectors in R^n and $1 \leq p < \infty$. Then*

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

proof: The statement is obvious if $p = 1$ so we may assume that $1 < p < \infty$. Let q be the exponent conjugate to p. We write

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

$$= \underbrace{\sum_{k=1}^{n} (|x_k| + |y_k|)^{p-1} |x_k|}_{I} + \underbrace{\sum_{k=1}^{n} (|x_k| + |y_k|)^{p-1} |y_k|}_{II}$$

We apply Holder's inequality to I:

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

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

Since p and q are conjugate exponenents, we have that $p = q(p-1)$

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

In identical way, one derives that

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

So

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

So

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

Since $p = 1 - \frac{1}{q}$, we have proven the Minkowski inequality.

We now return to the last proposition of the last lecture and complete the proof that the p-function

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}$$

is indeed a norm on $\mathbb{R}^n$ for $1 < p < \infty$

As w have already remarked, the only statement that requires the proof is the triangle inequality 9write a quick proof for yourself.

To prove the triangle inequality, ntoice that the function $t \to t^p \qquad (p > 1)$ is increasing on $[0, \infty]$ and so

$$|x_k + y_k|^p \leq (|x_k| + ||y_k|)^p$$

$$\begin{aligned}
\|x + y\|_p &= \left(\sum_{k=1}^{n} |x_k + y_k|^p \right)^{\frac{1}{p}} \\
&\leq \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{1}{p}} \\
&\leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |x_k|^p_{:} \right)^{\frac{1}{p}} \\
&= \|x\|_p + \|y\|_p
\end{aligned}$$

So our p-function $\|x\|_p$ is indeed a norm on $\mathbb{R}^n$. From now on we will refer to it as the p-norm.

3 Inner Product Spaces

Definition 4 *Let X be a vector space over $\mathbb{R}$. A function*

$$\langle \cdot, \cdot \rangle : \quad X \times X \to \mathbb{R}$$

is called an inner product space on X if the following holds for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$

1. $\langle x, y \rangle \geq 0$
2. $\langle x, x \rangle = 0$ *if and only if* $x = 0_x$
3. $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$
4. $\langle x, y \rangle = \langle y, x \rangle$

Remark: (3) and (4) imply

$$\langle \alpha y + \beta y, z \rangle = \alpha \langle y, x \rangle + \beta \langle z, x \rangle$$

Remark: If $\mathbb{R}$is replaced with $\mathbb{C}$ then (4) is changed to

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

Same change applies to the previous remark: the inner product is anti-lineear with rspect to the first variable.

A pair $(X, \langle , \rangle)$ where X is a vector space and $\langle \cdot, \rangle$ is an inner product on X is called an inner product space.

Proposition 3 *(Cauchy-Schwartz inequality) Let $(x, \langle \cdot, \cdot \rangle$ be an inner product space. Then for all $x, y \in X$*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}$$

If either $x = 0_x$ or $y = 0_x$then both sides are equal to zero and the statemnet trivially holds. So wee can assume that $x \neq 0_x$. and $y \neq 0_x$.

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$f(t) = \langle x + ty, x + ty\rangle$

be the linearity and symmetry of the inner product

$$f(t) = \langle x, x\rangle + 2t\langle x, y\rangle + t^2\langle y, y\rangle$$

So f is a quadratic function.

By property (1) of inner product, $f(t) \geq 0$ for all t the relation

$$f(t) = 2\langle x, y\rangle + 2t\langle y, y\rangle$$

Gives that $f'(t) = 0$ if

$t = t_{min} = -\frac{\langle x,y\rangle}{\langle y,y\rangle}$

So,

$$f(t_{min}) = \langle x, x\rangle - \frac{2(\langle x, y\rangle)^2}{\langle y, y\rangle} + \frac{(\langle x, y\rangle^2\langle y, y\rangle}{(\langle y, y\rangle)}$$

$$= \langle x, x\rangle - \frac{(\langle x, x\rangle)^2}{\langle y, y\rangle} \geq 0$$

So,

$$\langle x, x\rangle\langle y, y\rangle \geq (\langle x, y\rangle)^2$$

and

$$\sqrt{\langle x, y\rangle}\sqrt{\langle y, y\rangle} \geq |\langle x, y\rangle|$$

Proposition 4 *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space. Then*

$$\|x\| = \sqrt{\langle x, x\rangle}$$

is a norm on X.

We need to verify properties (1)-(4) of the norm.

$$\|x\| \geq 0 \text{ and } \|x\| - 0 \iff x = 0_x$$

To prove (3), take $\alpha \in \mathbb{R}$ and

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = \alpha\|x\|$$

To verify the triangle inequality for $x, y \in X$ we have

$$\|x + y\|^2 = \left(\sqrt{\langle x + y, x + y\rangle}\right)^2 = \langle x + y, x + y\rangle$$

$$= \langle x, x\rangle + \langle y, y\rangle + 2\langle x, y\rangle$$
$$\leq \langle x, x\rangle + \langle y, y\rangle + 2|\langle x, y\rangle|$$
$$\leq \langle x, x\rangle + y, y\rangle + 2\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}$$
$$= \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$
$$\leq (\|x\| + \|y\|)^2$$

So

$$\|x + y\| \leq \|x\| + \|y\|$$

and so the triangle inequality holds and

$$\|\| = \sqrt{\langle \cdot, \cdot \rangle}$$

is indeed a norm on X*.*

Definition 5 *If* $(X, \|\|)$ *is a normed space, we say that norm* $\|\dot{\|}$ *is indeed an inner product space if there exist inner product* $\langle \cdot, \cdot \rangle = \sqrt{\langle \cdot, \cdot \rangle}$

We have seen that inner product implies norm. If it happens that we start with a norm, and we can find an inner product such that $\| \cdot \|$ is related to it by the above equation then we say that the norm is indeed by the inner product.

If the norm is induced by an inner product, our space has Euclidean geometry.

For example, $x \neq 0_x$ and $y \neq 0_x$, we can define the angle between the vectors x and y by the formula

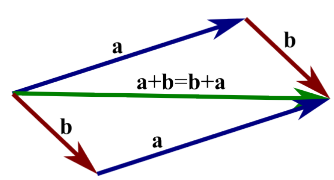
$$\cos(\alpha) = \frac{\langle x, y\rangle}{\|x\|\|y\|}, \qquad Q \in [0, \pi]$$

Now we can define the orthogonality
$X \perp y \iff \langle x, y\rangle = 0 \qquad (\iff Q = \frac{\pi}{2})$
With this definition one has pythagoras theorem:

$$\|x + y\|^2 = \|x\|^2 = \|y\|^2 \text{ if and only if } \langle x, y\rangle = 0$$

.

if and only if $\langle x, y\rangle = 0 \Leftrightarrow x \perp y$



The geometric interpretation of Pythagoras's Theorem:

The square of the length of one side of a triangle is equal to the sum of the squares of the lengths of the other two sides if and only if the triangle is right.

Proposition 5 *(Paralelogarm Law)*

Let $(X, \|\dot{\|})$ be a normed space and suppose thaat the norm $\|\dot{\|}$ is induced by an inner product $\langle\dot{|}\dot{\rangle}$ in other words

$$\|x\| = \sqrt{\langle x|x\langle}$$

Then the parallelogram law holds: For all $x, y \in X$

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

Remark: The diagonals of parallelogram determined by vectors x and y are the vectors $x-y$ and $x+y$. So the. parallelogram law states: the sum of the squares of the lengths of the diagonals is twice the sum of the squares of the length of the sides.

$$\|x+y\|^2 + \|x-y\|^2 = x+y|x+y\rangle + \langle x-y|x-y\rangle$$

$$= \langle x|x\rangle + 2\langle x|y\rangle + y|y\rangle + \langle x|x - 2\langle x|y\rangle + \langle y|y\rangle = 2(\|x\|^2 + \|y\|^2)$$

The paraelleogram has a converse that is much harder to prove. The proof will be given in the tutorial (and that is of course optional material)

Theorem 3 *Let $(x, \|\dot{\|}$. be a normed space such that for all. $x, y \in X$ the paraellogram law holds:*

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

.

Then the norm $\|\dot{\|}$ is induced by an inner product.

Thus, the validity of the parallelogram law is a sufficient and necessary condition for the norm to be induced by an inner product.

Next class we will illustrate this point on the examples of p-norms we introduced in the first part of the lecture.

At the end of the last lecture, we discussed the relation inner products and the parallelogram law. The following result emerged:

Theorem 4 *Let $(X, \|\dot{\|})$ be a normed space. Then the following statements are equivalent.*

1. *The norm $\|\dot{\|}$ is induced by inner product, namely there exists an inner product on X such that*

$$\|x\| = \sqrt{\langle x|x\rangle} \qquad \forall x \in X$$

2. *The parallelogram holds in* $(X, \|\dot{\|}$, *namely for any two vectors* $x, y \in X$ *we have*
$$\|x-y\|^2 + \|x+y\|^2 = 2(\|x\|^2 + \|y\|^2$$
.

In last lecture, we have proven the direction (1) $\implies$ (2)*, which is the easy part of the theorem. The hard part* (2) $\implies$ (1) *will be done in a tutorial, and of course this proof is optional matereial but you should know the statment.*

We have introudced the p-norms:

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}, \quad 1 \leq p < \infty$$

$$\|x\|_\infty = \max_{1 \leq k \leq n} |x_k|$$

on $\mathbb{R}^n$. If $n = 1$, then obviosuly all these norms are equal and are induced by inner product

$$\langle x|y\rangle = xy \text{ on } \mathbb{R}$$

What happens if $n \geq 2$?
The euclidean $p = 2$ is induced by inner product

$$\langle x|y\rangle = \sum_{k=1}^{n} x_k y_k$$

Check that this is indeed an inner product on $\mathbb{R}^n$. We will call it standrad inne r product and write

$$\langle x|y\rangle_{\text{standard}} = \sum_{k=1}^{n} x_k y_k$$

when there is a danger of confusion. $\|x\|_2 = \sqrt{\langle x|x\angle}$ obviously. holds. What about $p \neq 2$. Take

$$x = (1, -1, 0, \ldots, 0), y = (1, 1, 0, \ldots, 0)$$

Then

$$x + y = (2, 0, \ldots, 0$$
$$x - y = (0, -2, 0, \ldots, 0)$$

Hence,

$$\|x+y\|_{p=2}, \quad \|x-y\|_{p=2}$$

$$\|x\|_p = 2^{\frac{1}{p}} \quad \|y\|_p = 2^{\frac{1}{p}}$$

When $1 \leq p < \infty$

It follows that

$$\|x+y\|_p^2 + \|x-y\|_p^2 = p$$

and that

$$2(\|x\|_p^2 + \|y\|_p^2) = 4 * 2^{\frac{2}{p}}$$

Hence the parallelogram holds for the vectors x and y if and only if we have

$$2^{\frac{2}{p}} = 2$$

or $p = 2$.

Hence, the parallelogram fails in $(\mathbb{R}^n, \|\dot{\|}_p$ if $p \neq 2$ and $1 \leq p < \infty$.

If $p = \infty$, we have $\|x+y\|_\infty = 2, \quad \|x-y\|_\infty = 2, \|x\|_\infty = 1, \quad \|y\|_\infty = 1$ and the parallelogram fails again.

So among the normed space $(\mathbb{R}^2, \|\dot{\|}_p)$, $1 \leq p \leq \infty$, the only one whos norm induced by an inner product is $p = 2$ case. In this case, the inner product is the standard one,

$$\langle x|y\rangle = \sum_{k=1}^{n} x_k y_k$$

Classification of inner products on $\mathbb{R}^n$ (in part a review of some topics in linear algebra and multi-variable calculus).

Let A be a matrix on $\mathbb{R}^n$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

We denote its row vectors by r_1 so $r_1 = (a_{11}, \dots, a_{1n}$ $r_2 = (a_{21}, a_{22}, \dots, a_{2n})$, and so on. If $x = (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$x^\top = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Then

$$Ax^\top = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn^*n} \end{pmatrix} = \begin{pmatrix} \langle r_1 \mid x\rangle \\ \langle r_2 \mid x\rangle \\ \vdots \\ \langle r_n \mid x\rangle \end{pmatrix}$$

Matrix A is is called symmetric if $a_{ij} = a_{ji}$ for all i, j or equivalently, if $A = A^T$. A symmetric matrix has only real eigenvalues.

A symmetric matrix is positive definite if its eigenvalues are non-negative (≥ 0). A symmetric matrix is called strictly positive definite if its eigenvalues are positive.

A symmetric matrix is positive definite if for all $x \in \mathbb{R}^n$ $\qquad (Ax = Ax^T)$.

A symetric matrix is strictly positive definite if there exists a constant $\Lambda > 0$ such that for all $x \in \mathbb{R}^n$,

$$\langle x | Ax \rangle \geq \Lambda \langle x | y \rangle$$

One can show that a positive definite matrix is strictly positive definite if and only if

$$\langle x | Ax \rangle = 0 \iff x = 0_{\mathbb{R}^n} = (0, \ldots, 0)$$

If A is strictly positive definite, one can take for Λ the avlue

$$\Lambda = \min\{\lambda_1, \ldots, \lambda_n\}$$

where $\lambda_1, \ldots \lambda_n$ are eigenvalues of A.

Where you have seen this before? You should have seen in calculus several variables (40 calculus sequence, or MATH 248)

There you are given functions

$$f \colon \mathbb{R}^n \to \mathbb{R}$$

and we will assume that this function is three times differrentiable. The goal is to find and classify its extreme point (minima, maxima).

Any extremeum point x_0 is a critical point, namely it must satisfy

$$(x_0) = (\frac{\delta f}{\delta x_1}(x_0), \ldots, \frac{\delta f}{\delta x_n}(x_0) = 0_{\mathbb{R}} = (0, \ldots, 0)$$

. The usual Extremum criterion are based on the Hessian matrix

$$\text{Hess } f\,(x_n) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1}\,(x_0) & \frac{\partial^2 f}{\partial x_n \partial x_1}\,(x_0) \\ \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}\,(x_0) & \cdots \frac{\partial^2 f}{\partial x_n^2}\,(x_0) \end{bmatrix}$$

The entries of the Hessian matrix are $a_{ij} = \frac{\delta^2 f}{\delta x_j \delta x_i}(x_0)$

And the equality of mixed partial gives that

$$a_{ij} = a_{ji} \qquad \text{for all } i, j$$

which impiles that $\text{Hess} f(x_0)$ is a symmetric matrix.

The Taylor formula gives

$$f(x_0 + h) = f(x_0) + \langle h | \nabla f(x_0) \rangle + \frac{1}{2} \langle h | \text{Hess} f(x_i) h \rangle + R(h)$$

where remainder is such that

$$\lim_{\|h\|_2 \to 0} \frac{|R(h)|}{\|h\|_2} = 0$$

The above expansion is valid for any $x_0 \in \mathbb{R}^n$ But if x_0 is a critical point, namely if $(x_0) = 0$,

Then the expansion becomes

$$f(x_0 + h) = f(x_0) + \frac{1}{2}\langle h|\text{Hess}f(x_0)h\rangle + R(h)$$

This expansion gives:

1. x_0 is a point of local minimum if $\text{Hess}f(x_0)$ is positive definite.

2. x_0 is a point of strict local minimum if $\text{Hess}f(x_0)$ is strictly positive definite.

3. x_0 is a point of strict local maximum if $\text{Hess}f(x_0)$ is positive definite definite.

4. x_0 is a point of strict local maximum if $\text{Hess}f(x_0)$ is strictly positive definite.

5. If $\text{Hess}f(x_0)$ has both negative and positive eigenvalues, x_0 is called a saddle point.

The above calculus discussion should be standard and either you have seen it already or you will see it in MATH 248.

The point we wish to bring is that strictly positive definite matrices classify all possible inner products on $\mathbb{R}^n$.

Theorem 5 *(Classification of inner products on $\mathbb{R}^n$) The follwoing statements are equivalent:*

1. *$\langle\cdot|\cdot\rangle$ is an inner product on $\mathbb{R}^n$*

2. *There exists a strictly positive definite matrix A such that*

$$\langle x|y\rangle = \langle x|Ay\rangle_{standard} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

The point of implication $(2) \to (1)$ is an easy exercise. To prove that $(2) \implies (1)$ one argues as follows.

Let $a_k = (0, \dots, 0, 1, 0, \dots, 0) \qquad 1 \le k \le n$

be the staandard basis of $\mathbb{R}^n$. Any $x \in \mathbb{R}^n, x = (x_1, \dots, x_n)$ can be written as

$$x = \sum_{k=1}^{n} x_k a_k$$

.

Set $a_{ij} = \langle a_i | a_j \rangle$ *The symmetry of inner products gives that*

$$a_{ij} = a_{ji} \qquad \forall i, j$$

Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Then by linearity of the inner product, we have

$$\langle x \mid y \rangle = \left\langle \sum_{k=1}^{n} x_k q_k \mid \sum_{k=1}^{n} y_k a_k \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_i \langle a_i | a_i \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_i a_{ij} = \langle x \mid A_y \rangle_{\ standard}$$

Note that A *is symmetric matrix.*
Since $\langle x|x \rangle \geq 0$ *with equality if and only if* $x = 0_{\mathbb{R}^n} = (0, \dots, 0)$
we have that
$\langle x, Ax \rangle_{standard} \geq 0$ *with equality if and only if* $x = 0_{\mathbb{R}^n} = (0, \dots, 0)$

The (*) implies that A *is strictly positive definite. This can be proven in two ways; Either using linear algebra (diagonalization procedure for symmetric matrices) or using analysis. We will prove that (*) implies that* A *is strictly positive definite by using tools of analysis, but it is a great exercise to do it also using tools of linear algebra.*

A this point of the course, I used to give a seminar to students on linear algebra mathematical foundations of quantum mechanics, where the above exercise would be solved with linear algebra tools. This year. we can't do this, but you can access the notes that were taken by students by going to my web page /seminar.html.

The quantum statistical notes are at the bottom of thee page, and the first part of the notes deal with quantum mechanics on final dimensional inner product spaces.

In the next lecture we continue with metric space and start with a discussion of topology.