

MATH 311

Topics in Applied Mathematics I

Lecture 28:

Norms and inner products.

Norm

The notion of *norm* generalizes the notion of length of a vector in \mathbb{R}^n .

Definition. Let V be a vector space. A function $\alpha : V \rightarrow \mathbb{R}$ is called a **norm** on V if it has the following properties:

- (i) $\alpha(\mathbf{x}) \geq 0$, $\alpha(\mathbf{x}) = 0$ only for $\mathbf{x} = \mathbf{0}$ (positivity)
- (ii) $\alpha(r\mathbf{x}) = |r| \alpha(\mathbf{x})$ for all $r \in \mathbb{R}$ (homogeneity)
- (iii) $\alpha(\mathbf{x} + \mathbf{y}) \leq \alpha(\mathbf{x}) + \alpha(\mathbf{y})$ (triangle inequality)

Notation. The norm of a vector $\mathbf{x} \in V$ is usually denoted $\|\mathbf{x}\|$. Different norms on V are distinguished by subscripts, e.g., $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|_2$.

Examples. $V = \mathbb{R}^n$, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

- $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|, \dots, |x_n|)$.

Positivity and homogeneity are obvious. Let $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$. Then $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n)$.

$$|x_i + y_i| \leq |x_i| + |y_i| \leq \max_j |x_j| + \max_j |y_j|$$

$$\implies \max_j |x_j + y_j| \leq \max_j |x_j| + \max_j |y_j|$$

$$\implies \|\mathbf{x} + \mathbf{y}\|_\infty \leq \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty.$$

- $\|\mathbf{x}\|_1 = |x_1| + |x_2| + \dots + |x_n|$.

Positivity and homogeneity are obvious.

The triangle inequality: $|x_i + y_i| \leq |x_i| + |y_i|$

$$\implies \sum_j |x_j + y_j| \leq \sum_j |x_j| + \sum_j |y_j|$$

Examples. $V = \mathbb{R}^n$, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

- $\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{1/p}$, $p > 0$.

Remark. $\|\mathbf{x}\|_2 =$ Euclidean length of \mathbf{x} .

Theorem $\|\mathbf{x}\|_p$ is a norm on \mathbb{R}^n for any $p \geq 1$.

Positivity and homogeneity are still obvious (and hold for any $p > 0$). The triangle inequality for $p \geq 1$ is known as the **Minkowski inequality**:

$$\begin{aligned} (|x_1 + y_1|^p + |x_2 + y_2|^p + \dots + |x_n + y_n|^p)^{1/p} &\leq \\ &\leq (|x_1|^p + \dots + |x_n|^p)^{1/p} + (|y_1|^p + \dots + |y_n|^p)^{1/p}. \end{aligned}$$

Normed vector space

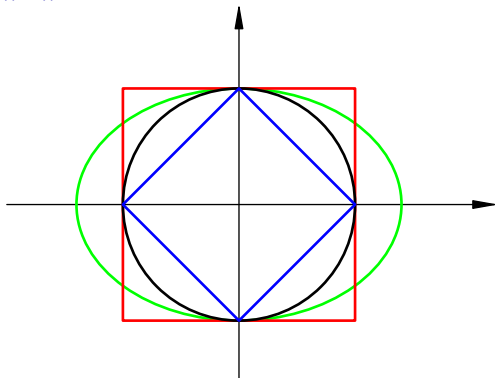
Definition. A **normed vector space** is a vector space endowed with a norm.

The norm defines a distance function on the normed vector space: $\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Then we say that a vector \mathbf{x} is a good *approximation* of a vector \mathbf{x}_0 if $\text{dist}(\mathbf{x}, \mathbf{x}_0)$ is small.

Also, we say that a sequence $\mathbf{x}_1, \mathbf{x}_2, \dots$ *converges* to a vector \mathbf{x} if $\text{dist}(\mathbf{x}, \mathbf{x}_n) \rightarrow 0$ as $n \rightarrow \infty$.

Unit circle: $\|\mathbf{x}\| = 1$



$$\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2} \quad \text{black}$$

$$\|\mathbf{x}\| = \left(\frac{1}{2}x_1^2 + x_2^2\right)^{1/2} \quad \text{green}$$

$$\|\mathbf{x}\| = |x_1| + |x_2| \quad \text{blue}$$

$$\|\mathbf{x}\| = \max(|x_1|, |x_2|) \quad \text{red}$$

Examples. $V = C[a, b]$, $f : [a, b] \rightarrow \mathbb{R}$.

- $\|f\|_{\infty} = \max_{a \leq x \leq b} |f(x)|.$

- $\|f\|_1 = \int_a^b |f(x)| dx.$

- $\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{1/p}, \quad p > 0.$

Theorem $\|f\|_p$ is a norm on $C[a, b]$ for any $p \geq 1$.

Inner product

The notion of *inner product* generalizes the notion of dot product of vectors in \mathbb{R}^n .

Definition. Let V be a vector space. A function $\beta : V \times V \rightarrow \mathbb{R}$, usually denoted $\beta(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle$, is called an **inner product** on V if it is positive, symmetric, and bilinear. That is, if

- (i) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ only for $\mathbf{x} = \mathbf{0}$ (positivity)
- (ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (symmetry)
- (iii) $\langle r\mathbf{x}, \mathbf{y} \rangle = r\langle \mathbf{x}, \mathbf{y} \rangle$ (homogeneity)
- (iv) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ (distributive law)

An **inner product space** is a vector space endowed with an inner product.

Examples. $V = \mathbb{R}^n$.

- $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$.

- $\langle \mathbf{x}, \mathbf{y} \rangle = d_1x_1y_1 + d_2x_2y_2 + \cdots + d_nx_ny_n$,
where $d_1, d_2, \dots, d_n > 0$.

- $\langle \mathbf{x}, \mathbf{y} \rangle = (D\mathbf{x}) \cdot (D\mathbf{y})$,

where D is an invertible $n \times n$ matrix.

Remarks. (a) Invertibility of D is necessary to show that $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$.

(b) The second example is a particular case of the third one when $D = \text{diag}(d_1^{1/2}, d_2^{1/2}, \dots, d_n^{1/2})$.

Problem. Find an inner product on \mathbb{R}^2 such that $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 2$, $\langle \mathbf{e}_2, \mathbf{e}_2 \rangle = 3$, and $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = -1$, where $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$.

Let $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$.

Then $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$, $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$.

Using bilinearity, we obtain

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle &= \langle x_1\mathbf{e}_1 + x_2\mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2 \rangle \\ &= x_1\langle \mathbf{e}_1, y_1\mathbf{e}_1 + y_2\mathbf{e}_2 \rangle + x_2\langle \mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2 \rangle \\ &= x_1y_1\langle \mathbf{e}_1, \mathbf{e}_1 \rangle + x_1y_2\langle \mathbf{e}_1, \mathbf{e}_2 \rangle + x_2y_1\langle \mathbf{e}_2, \mathbf{e}_1 \rangle + x_2y_2\langle \mathbf{e}_2, \mathbf{e}_2 \rangle \\ &= 2x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2.\end{aligned}$$

It remains to check that $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ for $\mathbf{x} \neq \mathbf{0}$.

Indeed, $\langle \mathbf{x}, \mathbf{x} \rangle = 2x_1^2 - 2x_1x_2 + 3x_2^2 = (x_1 - x_2)^2 + x_1^2 + 2x_2^2$.

Example. $V = \mathcal{M}_{m,n}(\mathbb{R})$, space of $m \times n$ matrices.

- $\langle A, B \rangle = \text{trace}(AB^T)$.

If $A = (a_{ij})$ and $B = (b_{ij})$, then $\langle A, B \rangle = \sum_{i=1}^m \sum_{j=1}^n a_{ij}b_{ij}$.

Examples. $V = C[a, b]$.

- $\langle f, g \rangle = \int_a^b f(x)g(x) dx$.

- $\langle f, g \rangle = \int_a^b f(x)g(x)w(x) dx$,

where w is bounded, piecewise continuous, and $w > 0$ everywhere on $[a, b]$.

w is called the **weight** function.

Theorem Suppose $\langle \mathbf{x}, \mathbf{y} \rangle$ is an inner product on a vector space V . Then

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \quad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Proof: For any $t \in \mathbb{R}$ let $\mathbf{v}_t = \mathbf{x} + t\mathbf{y}$. Then

$$\begin{aligned} \langle \mathbf{v}_t, \mathbf{v}_t \rangle &= \langle \mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} + t\mathbf{y} \rangle + t\langle \mathbf{y}, \mathbf{x} + t\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + t\langle \mathbf{x}, \mathbf{y} \rangle + t\langle \mathbf{y}, \mathbf{x} \rangle + t^2\langle \mathbf{y}, \mathbf{y} \rangle. \end{aligned}$$

Assume that $\mathbf{y} \neq \mathbf{0}$ and let $t = -\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}$. Then

$$\langle \mathbf{v}_t, \mathbf{v}_t \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + t\langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

Since $\langle \mathbf{v}_t, \mathbf{v}_t \rangle \geq 0$, the desired inequality follows. In the case $\mathbf{y} = \mathbf{0}$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{y} \rangle = 0$.

Cauchy-Schwarz Inequality:

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

Corollary 1 $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Equivalently, for all $x_i, y_i \in \mathbb{R}$,

$$(x_1y_1 + \cdots + x_ny_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

Corollary 2 For any $f, g \in C[a, b]$,

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \int_a^b |f(x)|^2 dx \cdot \int_a^b |g(x)|^2 dx.$$

Norms induced by inner products

Theorem Suppose $\langle \mathbf{x}, \mathbf{y} \rangle$ is an inner product on a vector space V . Then $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is a norm.

Proof: Positivity is obvious. Homogeneity:

$$\|r\mathbf{x}\| = \sqrt{\langle r\mathbf{x}, r\mathbf{x} \rangle} = \sqrt{r^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |r| \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Triangle inequality (follows from Cauchy-Schwarz's):

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \langle \mathbf{x}, \mathbf{x} \rangle + |\langle \mathbf{x}, \mathbf{y} \rangle| + |\langle \mathbf{y}, \mathbf{x} \rangle| + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Examples. • The length of a vector in \mathbb{R}^n ,

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

is the norm induced by the dot product

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

• The norm $\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$ on the vector space $C[a, b]$ is induced by the inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx.$$

Angle

Since $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$, we can define the *angle* between nonzero vectors in any vector space with an inner product (and induced norm):

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Then $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos \angle(\mathbf{x}, \mathbf{y})$.

In particular, vectors \mathbf{x} and \mathbf{y} are **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.