MATH 304
Linear Algebra

Lecture 27:
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) $\alpha(x) \geq 0$, $\alpha(x) = 0$ only for $x = 0$ (positivity)
(ii) $\alpha(rx) = |r| \alpha(x)$ for all $r \in \mathbb{R}$ (homogeneity)
(iii) $\alpha(x + y) \leq \alpha(x) + \alpha(y)$ (triangle inequality)

Notation. The norm of a vector $x \in V$ is usually denoted $\|x\|$. Different norms on $V$ are distinguished by subscripts, e.g., $\|x\|_1$ and $\|x\|_2$. 
Examples. \( V = \mathbb{R}^n, \; \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n. \)

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

Positivity and homogeneity are obvious.

The triangle inequality:

\[
|x_i + y_i| \leq |x_i| + |y_i| \leq \max_j |x_j| + \max_j |y_j|
\]

\[
\Rightarrow \max_j |x_j + y_j| \leq \max_j |x_j| + \max_j |y_j|
\]

- \( \| \mathbf{x} \|_1 = |x_1| + |x_2| + \cdots + |x_n| . \)

Positivity and homogeneity are obvious.

The triangle inequality:

\[
|x_i + y_i| \leq |x_i| + |y_i|
\]

\[
\Rightarrow \sum_j |x_j + y_j| \leq \sum_j |x_j| + \sum_j |y_j|
\]
Examples. $V = \mathbb{R}^n$, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

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

Remark. $\|x\|_2 =$ Euclidean length of $x$.

Theorem $\|x\|_p$ 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:

$$\left( |x_1 + y_1|^p + |x_2 + y_2|^p + \cdots + |x_n + y_n|^p \right)^{1/p} \leq \left( |x_1|^p + \cdots + |x_n|^p \right)^{1/p} + \left( |y_1|^p + \cdots + |y_n|^p \right)^{1/p}.$$
Definition. A **normed vector space** is a vector space endowed with a norm.

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

Then we say that a sequence \( x_1, x_2, \ldots \) *converges* to a vector \( x \) if \( \text{dist}(x, x_n) \to 0 \) as \( n \to \infty \).

Also, we say that a vector \( x \) is a good *approximation* of a vector \( x_0 \) if \( \text{dist}(x, x_0) \) is small.
Unit circle: \( \| x \| = 1 \)

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

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

\[
\| x \| = |x_1| + |x_2| \quad \text{blue}
\]

\[
\| x \| = \max(|x_1|, |x_2|) \quad \text{red}
\]
Examples. \( V = C[a, b], \ f : [a, b] \rightarrow \mathbb{R}. \)

- \( \|f\|_{\infty} = \max_{a \leq x \leq b} |f(x)|. \)
- \( \|f\|_1 = \int_a^b |f(x)| \, dx. \)
- \( \|f\|_p = \left( \int_a^b |f(x)|^p \, dx \right)^{1/p}, \ p > 0. \)

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 $\beta : V \times V \to \mathbb{R}$, usually denoted $\beta(x, y) = \langle x, y \rangle$, is called an *inner product* on $V$ if it is positive, symmetric, and bilinear. That is, if

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

An *inner product space* is a vector space endowed with an inner product.
Examples. \( V = \mathbb{R}^n \).

- \( \langle x, y \rangle = x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \).
- \( \langle x, y \rangle = d_1 x_1 y_1 + d_2 x_2 y_2 + \cdots + d_n x_n y_n \), where \( d_1, d_2, \ldots, d_n > 0 \).
- \( \langle x, y \rangle = (Dx) \cdot (Dy) \), where \( D \) is an invertible \( n \times n \) matrix.

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

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

- \( \langle x, y \rangle = x_1 y_1 - x_2 y_2 \).

Let \( v = (1, 2) \), then \( \langle v, v \rangle = 1^2 - 2^2 = -3 \).
\( \langle x, y \rangle \) is symmetric and bilinear, but not positive.

- \( \langle x, y \rangle = 2x_1 y_1 + x_1 x_2 + 2x_2 y_2 + y_1 y_2 \).

\( v = (1, 1) \), \( w = (1, 0) \) \( \implies \) \( \langle v, w \rangle = 3 \), \( \langle 2v, w \rangle = 8 \).
\( \langle x, y \rangle \) is positive and symmetric, but not bilinear.

- \( \langle x, y \rangle = x_1 y_1 + x_1 y_2 - x_2 y_1 + x_2 y_2 \).

\( v = (1, 1) \), \( w = (1, 0) \) \( \implies \) \( \langle v, w \rangle = 0 \), \( \langle w, v \rangle = 2 \).
\( \langle x, y \rangle \) is positive and bilinear, but not symmetric.
**Problem.** Find an inner product on \( \mathbb{R}^2 \) such that
\[
\langle e_1, e_1 \rangle = 2, \quad \langle e_2, e_2 \rangle = 3, \quad \text{and} \quad \langle e_1, e_2 \rangle = -1,
\]
where \( e_1 = (1, 0), \ e_2 = (0, 1) \).

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

Then \( x = x_1 e_1 + x_2 e_2, \ y = y_1 e_1 + y_2 e_2 \).

Using bilinearity, we obtain
\[
\langle x, y \rangle = \langle x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2 \rangle
\]
\[
= x_1 \langle e_1, y_1 e_1 + y_2 e_2 \rangle + x_2 \langle e_2, y_1 e_1 + y_2 e_2 \rangle
\]
\[
= x_1 y_1 \langle e_1, e_1 \rangle + x_1 y_2 \langle e_1, e_2 \rangle + x_2 y_1 \langle e_2, e_1 \rangle + x_2 y_2 \langle e_2, e_2 \rangle
\]
\[
= 2x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2.
\]

It remains to check that \( \langle x, x \rangle > 0 \) for \( x \neq 0 \).
\[
\langle x, x \rangle = 2x_1^2 - 2x_1 x_2 + 3x_2^2 = (x_1 - x_2)^2 + x_1^2 + 2x_2^2.
\]
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 \textbf{weight} function.