Miscellaneous Notes on Linear Algebra

Simon Xiang

January 11, 2021

Who ever suffered from learning too much linear algebra? These notes will seek to fill in my linear algebra gaps. Source files: https://git.simonxiang.xyz/math_notes/files.html

Contents

Lecture 1

Inner-Product Spaces

What is an inner product?? Let's find out.

1.1 Inner Products

The length of a vector x is the **norm** of x, denoted $\|x\|$. If $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, we have $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Note that the norm is not linear. For $x, y \in \mathbb{R}^n$, the **dot product** of x and y, denoted $x \cdot y$, is defined by $x \cdot y = x_1 y_1 + \cdots + x_n y_n$. Note that this is a number, not a vector. Clearly $x \cdot x = \|x\|^2$ for all $x \in \mathbb{R}^n$, which implies $x \cdot x \geq 0$ for all $x \in \mathbb{R}^n$ ($x \cdot x = 0$ only if x is the zero vector). The map that sends $x \in \mathbb{R}^n$ to $x \cdot y$ in $\mathbb{R}$ for fixed y is linear since $\mathbb{R}$ is a field. The dot product is also commutative, since $\mathbb{R}$ is.

Inner products generalize dot products. Recall that $|\lambda|^2 = \lambda\overline{\lambda}$ for $\lambda \in \mathbb{C}$. For $z \in \mathbb{C}^n$, we define the norm of z by $\|z\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$. We take the modulus of z_i since we want the result to be nonnegative. Note that $\|z\|^2 = z_1\overline{z_1} + \cdots z_n\overline{z_n}$. We want to think of $\|z\|^2$ as the inner product of z with itself, like in $\mathbb{R}^n$. This suggests we define the inner product of $w = (w_1, \cdots, w_n) \in \mathbb{C}^n$ with z as $w_1\overline{z_1} + \cdots + w_n\overline{z_n}$. We expect the inner product of w with z equal the complex conjugate of the inner product of z with w. With this motivation in mind, let us define inner products.

Definition 1.1 (Inner product)**.** An **inner product** on an F-vector space V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in F$ such that

(i) $\langle v, v\rangle \geq 0$ for all $v \in V$; (**positivity**)

(ii) $\langle v, v\rangle = 0$ iff $v = 0$; (**definiteness**)

(iii) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$; (**additivity in first slot**)

(iv) $\langle av, w\rangle = a\langle v, w\rangle$ for all $a \in F$ and all $v, w, \in V$; (**homogeneity in first slot**)

(v) $\langle v, w\rangle = \overline{\langle w, v\rangle}$ for all $v, w \in V$. (**conjugate symmetry**).

For real numbers, condition (v) simply becomes $\langle v, w\rangle = \langle w, v\rangle$ for all $v, w \in V$. An **inner product space** is a vector space V along with an inner product on V.

Example 1.1. The most important example is the **Euclidian inner product** on $\mathbf{F}^n$ (Axler uses $\mathbf{F}$ to denote either $\mathbb{C}$ or $\mathbb{R}$). We define an inner product on $\mathbf{F}^n$ by

$$\langle (w_1, \cdots, w_n), (z_1, \cdots, z_n)\rangle = w_1\overline{z_1} + \cdots w_n\overline{z_n}.$$

An example of another inner product on $\mathbf{F}^n$ is defined by $\langle (w_1, \cdots, w_n), (z_1, \cdots, z_n)\rangle = c_1 w_1\overline{z_1} + \cdots + c_n w_n\overline{z_n}$ for c_i positive constants. The case where $c_i = 1$ for all i is simply the standard Euclidian inner product.

Example 1.2. Consider the vector space $\mathscr{P}_m(\mathbf{F})$, the polynomial ring over $\mathbf{F}$ of polynomials with degree at most m. We can define an inner product on $\mathscr{P}_m(\mathbf{F})$ by

$$\langle p, q\rangle = \int_0^1 p(x)\overline{q(x)}\,dx.$$

For fixed $w \in V$, the function that takes v to $\langle v, w\rangle$ is a linear map $V \to \mathbf{F}$. So $\langle 0, w\rangle = 0$, and by condition (v) $\langle w, 0\rangle = 0$ as well. Furthermore, $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ and $\langle u, av\rangle = \overline{a}\langle u, v\rangle$ hold as well: This second condition is known as conjugate homogeneity in the second slot.

1.2 Norms

For $v \in V$, we define the **norm** of v, denoted $\|v\|$, by $\|v\| = \sqrt{\langle v, v\rangle}$. For example, if $p \in \mathscr{P}_m(\mathbf{F})$, then $\|p\| = \sqrt{\int_0^1 |p(x)|^2\,dx}$. Some properties: $\|v\| = 0$ iff $v = 0$, and $\|av\| = |a|\|v\|$. To see this, note that $\|av\|^2 = \langle av, av\rangle = a\langle v, av\rangle = a\overline{a}\langle v, v\rangle = |a|^2\|v\|^2$, taking square roots gives us our result. This illustrates a general idea: working with norms squared is easier than working directly with norms.

Two vectors $u, v \in V$ are **orthogonal** if $\langle u, v\rangle = 0$. The zero vector is orthogonal to every vector, and the only vector orthogonal to itself. Assume $V = \mathbb{R}^2$, now let us state a 2500 year old theorem.

Pythagorean Theorem. *If u, v are orthogonal vectors in V, then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. Exercise. ⊠

Suppose $u, v \in V$. We want to write u as a scalar multiple of v plus a vector w orthogonal to v. Let $a \in \mathbf{F}$ be a scalar, then $u = av + (u - av)$. We need to choose a such that v is orthogonal to $u - av$, in other words, we want $0 = \langle u - av, v\rangle = \langle u, v\rangle - a\|v\|^2$. So we should choose $a = \langle u, v\rangle / \|v\|^2$ (where $v \neq 0$). Then

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v\rangle}{\|v\|^2} v\right).$$

Cauchy-Schwarz Inequality. *If $u, v \in V$, then*

$$|\langle u, v\rangle| \leq \|u\|\|v\|.$$

This inequality is an equality iff one of u, v is a scalar multiple of the other.

Proof. Let $u, v \in V$, and assume $v \neq 0$. Consider $u = \frac{\langle u, v\rangle}{\|v\|^2} v + w$, where w is orthogonal to v. By the Pythagorean theorem, we have

$$\|u\|^2 = \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 = \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}.$$

Multiply both sides, take a square root, and we are done. This is an equality iff $w = 0$, but this is true iff u is a multiple of v. ⊠

Triangle Inequality. *If $u, v \in V$, then*

$$\|u + v\| \leq \|u\| + \|v\|.$$

This is an equality iff one of u, v is a nonnegative multiple of the other.

Proof. Let $u, v \in V$. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle \leq \|u\|^2 + \|v^2\| + 2\|u\|\|v\| = (\|u\| + \|v\|)^2.$$

The inequality step frollows from Cauchy-Schwartz, where $2\operatorname{Re}\langle u, v\rangle \leq 2|\langle u, v\rangle|$. Taking square roots gives the triangle inequality. This is an equality iff the two inequalities above are equalities, which is true iff $\langle u, v\rangle = \|u\|\|v\|$. ⊠

Parallelogram Equality. *If $u, v \in V$, then*

$$\|u + v\|^2 + \|u - v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right).$$

Proof. Exercise. ⊠

1.3 Orthonormal Bases

A list $(e_1, \cdots, e_m)$ of vectors in V is orthonormal if $\langle e_j, e_k \rangle = 0$ when $j \neq k$ and equals 1 when $j = k$, for $j, k \in \{1, \cdots, m\}$. Orthonormal lists are nice.

Proposition 1.1. *If $(e_1, \cdots, e_m)$ is an orthonormal list of vectors in V, then*

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \cdots, a_m \in \mathbf{F}$.

Proof. Since each e_j has norm 1, this follows from repeated applications of the Pythagorean theorem. ⊠

Corollary 1.1. *Every orthonormal list of vectors is linearly independent.*

An **orthonormal basis** of V is an orthonormal list of vectors in V that forms a basis for V. The standard basis is a good example. If we find an orthonormal list of length $\dim V$, then this is automatically an orthonormal basis of V (since they must be LI). In general, given a basis $(e_1, \cdots, e_n)$ of V and a vector $v \in V$, we know there is some choice of scalars $a_1, \cdots, a_m$ such that $v = a_1 e_1 + \cdots a_n e_n$, but finding the a_j's can be difficult. This is not the case for an orthonormal basis.

Theorem 1.1. *Suppose $(e_1, \cdots, e_n)$ is an orthonormal basis of V. Then*

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$$

for every $v \in V$.