Advanced Linear Algebra

Week 14

Jamie Gabe

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

From now on $\mathcal{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
Complex conjugation is denoted by $\overline{\alpha}$ i.e. if $\alpha = \beta + i\gamma$ with $\beta, \gamma \in \mathbb{R}$, then $\overline{\alpha} = \beta - i\gamma$.
In particular, if $\alpha \in \mathbb{R}$ then $\overline{\alpha} = \alpha$.
Let V be a vector space over $\mathcal{F}$.

Definition

A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

A vector space equipped with an inner product is called an inner product space.

Note that (a) means
$\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$ and
$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
for all $x, x_1, x_2, y \in V$ and $\alpha \in \mathcal{F}$.

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Definition

A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

Consider the complex vector space $\mathbb{C}$ and the product map $\langle \cdot, \cdot \rangle \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ given by

$$\langle \alpha, \beta \rangle = \alpha\beta \qquad \text{for } \alpha, \beta \in \mathbb{C}.$$

Question: which of the following are true for the map $\langle \cdot, \cdot \rangle$ (multiple choice)?

(1) (a) holds

(2) (b) holds

(3) (c) holds

(4) None of (a), (b), (c) hold.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Definition

A map $\langle \cdot, \cdot \rangle : V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

Consider the complex vector space $\mathbb{C}$ and the product map $\langle \cdot, \cdot \rangle \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ given by $\langle \alpha, \beta \rangle = \alpha\beta$ for $\alpha, \beta \in \mathbb{C}$.
(a): is $\alpha \mapsto \alpha\beta$ linear? We have
$\langle \alpha_1 + \alpha_2, \beta \rangle = (\alpha_1 + \alpha_2)\beta = \alpha_1\beta + \alpha_2\beta = \langle \alpha_1, \beta \rangle + \langle \alpha_2, \beta \rangle$.
So $\alpha \mapsto \alpha\beta$ is additive. Similarly it preserves scalar multiplication. So (a) holds.
(b): Is $\beta\alpha = \overline{\alpha\beta}$ for all $\alpha, \beta \in \mathbb{C}$? We know $\alpha\beta = \beta\alpha$, so the question is: is $\alpha\beta = \overline{\alpha\beta}$ for all $\alpha, \beta \in \mathbb{C}$?
Counter example: If $\alpha = 1$ and $\beta = i$, then $\alpha\beta = i$ and $\overline{\alpha\beta} = \overline{i} = -i$. So (b) does not hold.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Definition

A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

Consider the complex vector space $\mathbb{C}$ and the product map $\langle \cdot, \cdot \rangle \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ given by $\langle \alpha, \beta \rangle = \alpha\beta$ for $\alpha, \beta \in \mathbb{C}$.
(c): Is $\alpha\alpha = \alpha^2 > 0$ for every non-zero $\alpha \in \mathbb{C}$?
Counter example: Take $\alpha = i$. Then

$$\alpha^2 = i^2 = -1$$

which is not > 0. Hence (c) does not hold.

Definition

A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

What could we do instead?

We could have instead defined $\langle \alpha, \beta \rangle = \alpha\overline{\beta}$ on $\mathbb{C}$.
Then (a) holds just as before.
(b): $\langle \beta, \alpha \rangle = \beta\overline{\alpha} = \overline{\alpha\overline{\beta}} = \overline{\langle \alpha, \beta \rangle}$
(c): $\langle \alpha, \alpha \rangle = \alpha\overline{\alpha} = |\alpha|^2 > 0$ whenever $\alpha \neq 0$ in $\mathbb{C}$.

Definition

A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathcal{F}$ is an inner product on V if

(a) $x \mapsto \langle x, y \rangle$ is a linear map $V \to \mathcal{F}$ for every $y \in \mathcal{F}$;

(b) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$;

(c) $\langle x, x \rangle > 0$ for every non-zero $x \in V$.

More generally: Consider $V = \mathbb{C}^n$. There is a standard inner product $\langle \cdot, \cdot \rangle \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ given by

$$\left\langle \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i} = \alpha_1 \overline{\beta}_1 + \alpha_2 \overline{\beta}_2 + \cdots + \alpha_n \overline{\beta}_n.$$

This is the usual "dot product" or "scalar product".

The same can be done with the real vector space $\mathbb{R}^n$, but here we do not need complex conjugation.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

If $\langle\cdot,\cdot\rangle\colon V \times V \to \mathcal{F}$ is an inner product, then $y \mapsto \langle x, y\rangle$ is anti-linear, meaning that

$$\langle x, y_1 + y_2\rangle = \langle x, y_1\rangle + \langle x, y_2\rangle \qquad \text{additive}$$

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

This can easily be proved using (a) and (b) in the definition of inner products.

Let V be an inner product space (so we have a fixed inner product on V).

(a) The length (or norm) of $x \in V$ is $\|x\| := \sqrt{\langle x, x\rangle}$.

(b) Two vectors $x, y \in V$ are orthogonal if $\langle x, y\rangle = 0$. We write $x \perp y$.

(c) If $X \subseteq V$ is a subset then $X^\perp := \{y \in V \mid \forall x \in X : x \perp y\}$.

Note:
$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\langle x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x\rangle} = |\alpha|\|x\|$.
And:

$$x \perp y \quad \Leftrightarrow \quad \langle x, y\rangle = 0 \quad \Leftrightarrow \quad \underbrace{\overline{\langle x, y\rangle}}_{\langle y,x\rangle} = 0 \quad \Leftrightarrow \quad y \perp x.$$

So orthogonality is symmetric.

Lemma (8.3 (Pythogoras' identity))

If $x \perp y$ *then* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Proof.

For all $x, y \in V$ we have

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \underbrace{\langle x,x\rangle}_{\|x\|^2} + \langle x,y\rangle + \langle y,x\rangle + \underbrace{\langle y,y\rangle}_{\|y\|^2}.$$

Hence if $x \perp y$ then the identity follows. $\square$

Lemma (The parallelogram law)

For all $x, y \in V$ *we have*

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

Proof.

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

Adding this to $\|x+y\|^2$ from the previous proof, and we get the identity. $\square$

SDU

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Theorem (8.4 (Cauchy–Schwartz inequality))

Let $x, y \in V$. *Then* $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Proof.

If $\langle x, y\rangle = 0$ then the inequality is trivial, so we may assume WLOG $\langle x, y\rangle \neq 0$.

Also, $|\langle x, y\rangle| \leq \|x\|\|y\|$ iff $|\langle \frac{1}{\|x\|}x, \frac{1}{\|y\|}y\rangle| \leq 1$. Since $\|\frac{1}{\|x\|}x\| = \|\frac{1}{\|y\|}y\| = 1$ we may assume WLOG that $\|x\| = \|y\| = 1$. $\square$

Theorem (8.4 (Cauchy–Schwartz inequality))

Let $x, y \in V$. *Then* $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Proof.

We have shown: WLOG $\langle x, y\rangle \neq 0$ and $\|x\| = \|y\| = 1$.
Let $\alpha = \frac{\overline{\langle x,y\rangle}}{|\langle x,y\rangle|}$. Then $|\alpha| = \frac{|\overline{\langle x,y\rangle}|}{|\langle x,y\rangle|} = \frac{|\langle x,y\rangle|}{|\langle x,y\rangle|} = 1$, hence $\|\alpha x\| = |\alpha|\|x\| = 1$. Also

$$\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = \frac{\overline{\langle x, y\rangle}}{|\langle x, y\rangle|}\langle x, y\rangle = \frac{|\langle x, y\rangle|^2}{|\langle x, y\rangle|} = |\langle x, y\rangle|.$$

$$0 \leq \|\alpha x - y\|^2 = \underbrace{\|\alpha x\|^2}_{1} + \underbrace{\|y\|^2}_{1} - \underbrace{\langle \alpha x, y\rangle}_{|\langle x,y\rangle|} - \underbrace{\langle y, \alpha x\rangle}_{\overline{\langle \alpha x,y\rangle} = |\langle x,y\rangle|}$$

So $0 \leq 2 - 2|\langle x, y\rangle|$ or equivalently
$|\langle x, y\rangle| \leq 1 = \|x\|\|y\|$.

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Theorem (8.4 (Cauchy–Schwartz inequality))

Let $x, y \in V$. Then $|\langle x, y\rangle| \leq \|x\|\|y\|$.

We saw in the proof that
$0 \leq \|\alpha x - y\|^2 = 2(\|x\|\|y\| - |\langle x, y\rangle|)$. Hence
$|\langle x, y\rangle| = \|x\|\|y\|$ implies that $\|\alpha x - y\| = 0$ and thus $\alpha x = y$.
In conclusion: If one has equality in the Cauchy-Schwarts inequality, then x and y are proportionate (i.e. one is a scalar multiple of the other).
This is actually "if and only if".

Theorem (8.4 (Cauchy–Schwartz inequality))

Let $x, y \in V$. Then $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Consider the real vector space $C([0,1], \mathbb{R})$ of continuous functions $f\colon [0,1] \to \mathbb{R}$.

There is an inner product given by

$$\langle f, g\rangle := \int_0^1 f(t)g(t)\mathrm{d}t.$$

Question: What does the Cauchy–Schwartz inequality say in this case?

(1) For a positive function f, $\int_0^1 f(t)\mathrm{d}t > 0$;

(2) $|\int_0^1 f(t)\mathrm{d}t| \leq \max_{t\in[0,1]} f(t)$;

(3) $|\int_0^1 f(t)g(t)\mathrm{d}t| \leq \left(\int_0^1 f(t)^2\mathrm{d}t\right)^{1/2} \left(\int_0^1 g(t)^2\mathrm{d}t\right)^{1/2}$.

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Theorem (8.4 (Cauchy–Schwartz inequality))

Let $x, y \in V$. Then $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Consider the real vector space $C([0,1],\mathbb{R})$ of continuous functions $f\colon [0,1] \to \mathbb{R}$.

There is an inner product given by

$$\langle f, g\rangle := \int_0^1 f(t)g(t)\mathrm{d}t.$$

Note that

$\|f\| = \langle f, f\rangle^{1/2} = (\int_0^1 f(t)f(t)\mathrm{d}t)^{1/2} = (\int_0^1 f(t)^2\mathrm{d}t)^{1/2}$.

Hence Cauchy–Schwartz states:

$$\underbrace{|\int_0^1 f(t)g(t)\mathrm{d}t|}_{|\langle f,g\rangle|} \leq \|f\|\|g\| = \left(\int_0^1 f(t)^2\mathrm{d}t\right)^{1/2}\left(\int_0^1 g(t)^2\mathrm{d}t\right)^{1/2}.$$

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Theorem (8.5 (the triangle inequality))

Let $x, y \in V$. Then $\|x+y\| \leq \|x\| + \|y\|$.

Proof.

$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle$.

Hence

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

Taking square roots we get $\|x+y\| \leq \|x\| + \|y\|$. $\square$

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

A vector $x \in V$ is called normal if $\|x\| = 1$.
A subset $X \subseteq V$ is orthonormal if all its vectors are normal and orthogonal to each other.

Lemma (8.6)

Let $X \subseteq V$ be orthonormal.

(i) *If $z = \sum_{i=1}^{n} \alpha_i x_i$ with distinct vectors x_i from X, then*
$$\alpha_j = \langle z, x_j \rangle, \qquad j = 1, \dots, n;$$

(ii) *X is linearly independent;*

(iii) *If X is finite then $y - \sum_{x \in X} \langle y, x \rangle x \in X^{\perp}$ for all $y \in V$.*

Proof.

(i): Since $\langle x_i, x_j \rangle = \begin{cases} \|x_i\|^2 = 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$ we get

$$\langle z, x_j \rangle = \langle \sum_{i=1}^{n} \alpha_i x_i, x_j \rangle = \sum_{i=1}^{n} \alpha_i \langle x_i, x_j \rangle = \alpha_j.$$

Lemma (8.6)

Let $X \subseteq V$ be orthonormal.

(i) *If $z = \sum_{i=1}^{n} \alpha_i x_i$ with distinct vectors x_i from X, then*
$$\alpha_j = \langle z, x_j \rangle, \qquad j = 1, \dots, n;$$

(ii) *X is linearly independent;*

(iii) *If X is finite then $y - \sum_{x \in X} \langle y, x \rangle x \in X^{\perp}$ for all $y \in V$.*

Proof.

(ii): If $z = 0$ then $\alpha_j = 0$ for all j by (i). Hence X is linearly independent.

(iii): Let $y \in V$ and $z := \sum_{x \in X} \langle y, x \rangle x$. By (i) we have $\langle y, x \rangle = \langle z, x \rangle$ for all $x \in X$. Hence $\langle y - z, x \rangle = 0$ for all $x \in X$, so $y - z \in X^{\perp}$. $\square$

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Gram–Schmidt orthonomalisation: a process to take linearly independent vectors $y_1, \ldots, y_n$ and obtain an orthonormal set $\{x_1, \ldots, x_n\}$ s.t. $\mathrm{Span}\{y_1, \ldots, y_k\} = \mathrm{Span}\{x_1, \ldots, x_k\}$ for $k = 1, \ldots, n$.

We do this by recursion: let $x_1 = \frac{1}{\|y_1\|} y_1$ (this is called to normalise y_1).

Suppose we have found orthonormal $x_1, \ldots, x_k$ such that $\mathrm{Span}\{x_1, \ldots, x_k\} = \mathrm{Span}\{y_1, \ldots, y_k\}$. Then $x'_{k+1} := y_{k+1} - \sum_{i=1}^{k} \langle y_{k+1}, x_i \rangle x_i \in \{x_1, \ldots, x_k\}^{\perp}$ by Lemma 8.6(iii). As $y_{k+1} \notin \mathrm{Span}\{y_1, \ldots, y_k\} = \mathrm{Span}\{x_1, \ldots, x_k\}$, we have $x'_{k+1} \neq 0$. Let $x_{k+1} = \frac{1}{\|x'_{k+1}\|} x'_{k+1}$. It is easy to check that $\mathrm{Span}\{y_1, \ldots, y_{k+1}\} = \mathrm{Span}\{x_1, \ldots, x_{k+1}\}$.

Theorem (8.7)

Every finite-dimensional inner product space V contains an orthonormal basis.

Proof.

Apply Gram–Schmidt to any basis for V. $\square$

Definition (8.8)

An orthogonal direct sum in an inner product space V is a direct sum $U \oplus W$ of subspaces for which $U \perp W$ i.e. $u \perp w$ for all $u \in U$ and $w \in W$.
When $V = U \oplus W$ is orthogonal, we say that W is an orthogonal complement of U.

Lemma (8.9)

If $V = U \oplus W$ is an orthogonal direct sum, then $W = U^{\perp}$.

Proof.

By assumption $W \subseteq U^{\perp}$. Let $y \in U^{\perp}$. Write $y = u + w \in U + W$. Since $w \in U^{\perp}$ we have $u = y - w \in U \cap U^{\perp} = \{0\}$. Hence $u = 0$ and thus $y = w \in W$. □

Warning: in general V is not equal to $U + U^{\perp}$!

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Recall: if $V = U \oplus W$, then there is a unique linear $E \in \mathrm{End}(V)$ called the projection onto U along W, given by $E(v) = u$ where $v = u + w$ is the unique decomposition. These were characterised abstractly as idempotents: $E \in \mathrm{End}(V)$ satisfying $E = E^2$. In this case, $V = R(E) \oplus N(E)$ and E is the projection onto $R(E)$ along $N(E)$.

Definition

An orthogonal projection is an idempotent $E \in \mathrm{End}(V)$ for which

$$R(E) \perp N(E)$$

or equivalently for which $V = R(E) \oplus N(E)$ is an orthogonal direct sum.

E is called the orthogonal projection onto $R(E)$.

Note: There is no need for the "along $N(E)$" since $N(E) = R(E)^{\perp}$ by Lemma 8.9.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Theorem (8.11)

Let $U \subseteq V$ be a subspace. Let $v \in V$ and $u \in U$. Then $v - u \in U^\perp$ if and only if

$$\|v - u\| = \min_{x \in U} \|v - x\|.$$

Proof.

"$\Rightarrow$": Assume $v - u \in U^\perp$ and let $x \in U$. Then $v - u \perp u - x$. Hence by Pythagoras

$$\|v-x\|^2 = \|(v-u)+(u-x)\|^2 = \|v-u\|^2+\|u-x\|^2 \geq \|v-u\|^2.$$

So $\|v - u\| \leq \min_{x \in U} \|v - x\|$.
As $u \in U$, $\|v - u\| \geq \min_{x \in U} \|v - x\|$. $\square$

Theorem (8.11)

Let $U \subseteq V$ be a subspace. Let $v \in V$ and $u \in U$. Then $v - u \in U^{\perp}$ if and only if

$$\|v-u\| = \min_{x \in U} \|v-x\|.$$

Proof.

"$\Leftarrow$": Assume $\|v-u\| = \min_{x\in U} \|v-x\|$. In particular, $\|v-u\| \leq \|v-x\|$ for all $x \in U$.

We want $\langle v-u, x\rangle = 0$ for all $x \in U$. It suffices to check this for normal $x \in U$. Let $z := \langle v-u, x\rangle x \in U$. Then $\{x\}$ is orthonormal, so by Lemma 8.6(iii), $(v-u) - z \perp x$. In particular, $(v-u) - z \perp z$. By Pythagoras

$$\|\underbrace{z + ((v-u) - z)}_{v-u}\|^2 \overset{Pyt}{=} \|z\|^2 + \|v - \underbrace{(u+z)}_{\in U}\|^2 \geq \|z\|^2 + \|v-u\|^2.$$

This forces $\|z\| = 0$ so $z = 0$. Hence $\langle v-u, x\rangle = 0$.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Recall that in general V is not equal to $U \oplus U^\perp$. However

Theorem (8.12)

Assume U is a finite dimensional subspace of V. Then $V = U \oplus U^\perp$.

Proof.

U is itself an inner product space, so by Theorem 8.7 it has an orthonormal basis $\{x_1, \ldots, x_r\}$.
For $y \in V$ let $u := \sum_{i=1}^{r} \langle y, x_i \rangle x_i \in U$. By Lemma 8.6(iii), $y - u \in U^\perp$.
Hence $y = u + (y - u) \in U + U^\perp$. Consequently $V = U + U^\perp$ and $U \cap U^\perp = \{0\}$, so $V = U \oplus U^\perp$. $\square$