Week 6 Review session Notes Part 1

Christoph von Pezold

October 12, 2025

1 Nondegenerate bilinear forms and orthogonality

Before we get to inner products, it's worth making a brief final point about general bilinear forms:

Definition.

For a (in practice basically always symmetric or skew-symmetric) bilinear form $B : V^2 \to k$ and a subspace $S \subset V$, the *orthogonal complement* of S is:

$$S^\perp = \{w \in V \mid B(v, w) = 0 \ \forall v \in V\}$$

The key result here is this:

Proposition.

We have the following inequality for any bilinear form B and subspace S:

$$\dim S^\perp \geq \dim V - \dim S$$

When B is nondegenerate we have equality.

Proof Consider the induced map $\varphi_B : V \to V^*$ given by $v \mapsto B(v, \cdot)$. Composing with a map $V^* \to S^*$ given by restriction, $r : \ell \mapsto \ell_{|S}$ we get a map $V \to S^*$, and the orthogonal is exactly the kernel of this map. So by rank nullity $\dim S^\perp = \dim V - \operatorname{rank}(r \circ \varphi_B)$. Clearly $\operatorname{rank}(r \circ \varphi) \leq \dim S^* = \dim S$ since it is a map into that space, so the inequality follows. When B is nondegenerate we have equality since r is surjective and φ_B is an isomorphism, giving equality in that case.

We do not in general have a direct sum decomposition $S \oplus S^\perp$ though! This is an important thing to be careful about. Although often we think about nondegenerate bilinear forms as being inspired by things like the dot product, but this is not always the case; we can consider a nondegenerate bilinear form on $\mathbb{R}^2$ given by:

$$B\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$$

This bilinear form has the property that the orthogonal of a one dimensional subspace *is that subspace*, since the determinant is 0 exactly when the two are colinear! Trying to avoid such unfriendly scenarios is part of what motivates the move to inner products.

2 Inner Products

Definition.

An *inner product* $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ on an $\mathbb{R}$ vector space V is a symmetric, positive definite bilinear form on V. The pair $(V, \langle \cdot, \cdot \rangle)$ is then called an *inner product space*.

Reminder of definitions of terms:

1. $\langle \cdot, \cdot \rangle$ being symmetric means $\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in V$.
2. $\langle \cdot, \cdot \rangle$ being positive-definite means that we have for all $v \in V$ that $\langle v, v \rangle \geq 0$, with the only case where equality holds being the trivial one $\langle 0, 0 \rangle = 0$, trivial because $\langle \cdot, \cdot \rangle$ must be bilinear.

Note that, although it is not in the definition, inner products are always nondegenerate:

Proposition.

All inner products are nondegenerate.

Proof For any $v \in V \setminus \{0\}$, we have by positive-definiteness that $\langle v, v \rangle \neq 0$, so there exists an element the inner product of v with which is nonzero, namely v itself. Phrased in the language of the induced map $V \to V^*$, this says that the map $\varphi_{\langle \cdot, \cdot \rangle} : V \to V^*$ given by $v \mapsto \langle v, \cdot \rangle$ is an isomorphism, since it has trivial kernel. □

What does this definition mean? There are two main ways of looking at this. The first is that inner products give us a way of defining certain geometric properties that cannot be defined for general vector spaces, but can if we allow us to help ourselves to machinery inspired by the standard dot product on $\mathbb{R}^n$ (in fact it turns out that every finite dimensional inner product space is isomorphic to $(\mathbb{R}^n, \bullet)$). In particular, they allow us to define such concepts as *length* of a vector and *angle* between vectors, in particular they give us a notion of orthogonality that behaves in the way we want it to. To this end:

Definition.

1. The *norm* of a vector[a] is denoted by $\|v\|$ and is given by:
$$\|v\| = \sqrt{\langle v, v \rangle}$$
2. The *angle* between two vectors $\measuredangle(v, w)$ is given by:
$$\measuredangle(v, w) = \arccos\left(\frac{\langle v, w \rangle}{\|v\|\|w\|}\right)$$
3. In particular (this is review of more general concepts from bilinear forms but is worth reemphasising), two vectors are *orthogonal* – which can be thought of as a generalisation of the idea of being perpendicular – if $\langle v, w \rangle = 0$; notice that this lines up with our notion of angle since orthogonal vectors have $\measuredangle(v, w) = \frac{\pi}{2}$.

[a] this is what we think of as the length

There are a couple nice results from geometry that continue to be true for inner product spaces, for example the pythagorean theorem (and more generally the law of cosines):

Proposition.

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos \measuredangle(v, w)$$

Equivalently:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle$$

Proof The equivalence comes from just substituting the definition of the angle between two vectors. To see the second claim note:

$$\begin{aligned}\|v - w\|^2 &= \langle v - w, v - w\rangle \\ &= \langle v, v\rangle - \langle w, v\rangle - \langle v, w\rangle + \langle w, w\rangle \\ &= \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle\end{aligned}$$

□

Notice that in order for our geometric fact to continue to be true and line up with our intuition, we needed *all* of the properties of the inner product: positive-definiteness so that the length of a vector is nonzero if the vector is, and symmetry so that we could simplify down to $2\langle v, w\rangle$. This fact should thus give you some confidence that the inner product captures the essence of some of the properties we wish of a space with some nice notion of geometry. Another important fact backing up this intutition is:

Proposition.

If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space and $S \subset V$ is a subspace, then $V = S \oplus S^\perp$.

Proof We have from earlier results about orthogonals of general nondegenerate bilinear forms that $\dim S^\perp = \dim V - \dim S$, so it suffices to show independence. Suppose for the sake of contradiction that there exists nonzero $v \in S \cap S^\perp$. Then since $v \in S$ it is orthogonal to any element of $S^\perp$, in particular to itself; but this contradicts positive-definiteness. □

This allows us exactly to avoid the pathological case described earlier; positive-definiteness is exactly what we needed to avoid it! It is what allows us to cleanly break a space into a subspace and its orthogonal complement, a process that is often very useful when you want to induct on the dimension of V, for example.

Inner products do have uses beyond geometry though; to see this, consider what is perhaps the most important early theorem about them, the *Cauchy Schwarz Inequality*:

Theorem.

Cauchy-Schwarz Inequality

For vectors $u, v \in V$, we have that:

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

Proof.

If we replace u with λu the λ's cancel by bilinearity of the inner product[a] so we can without loss of generality assume that $\|u\| = 1$ and we seek to show that:

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

Since we know that $V = \text{span}(u) \oplus \text{span}(u)^\perp$, we can decompose v into its components in each of these subspaces. In particular, we can write it as $v = \langle u, v\rangle u + v - \langle u, v\rangle u$, since clearly $\langle u, v\rangle u \in \text{span}(u)$ and we have since $\|u\| = 1$ that:

$$\langle u, v - \langle u, v\rangle u\rangle = \langle u, v\rangle - \langle u, \langle u, v\rangle u\rangle = \langle u, v\rangle - \langle u, v\rangle \|u\|^2 = \langle u, v\rangle - \langle u, v\rangle = 0$$

Let $v_1 = \langle u, v\rangle u$, $v_2 = v - \langle u, v\rangle u$. Since we have just show $v_1 \perp v_2$:

$$\begin{aligned}
\|v\|^2 &= \langle v_1 + v_2, v_1 + v_2\rangle \\
&= \langle v_1, v_1\rangle + 2\langle v_1, v_2\rangle + \langle v_2, v_2\rangle \\
&= \|v_1\|^2 + \|v_2\|^2 \\
&\geq \|v_1\|^2 \\
&= \langle \langle u, v\rangle u, \langle u, v\rangle u\rangle \\
&= \langle u, v\rangle^2 \|u\|^2 \\
\|v\| &\geq \langle u, v\rangle
\end{aligned}$$

□

[a] $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ and $\|\lambda u\| = \sqrt{\langle \lambda u, \lambda u\rangle} = \sqrt{\lambda^2\langle u, u\rangle} = \lambda\|u\|$

Within the realm of geometry, this inequality is fairly obvious – what it is telling us is that the inner product, which is given by $\|v\|\|w\| \cos \measuredangle(v, w)$, is maximised when the two vectors are colinear. What happens though when we go beyond the finite dimensional? One inner product you're having to consider on your problem sets is the one on $\mathbb{R}[x]$, and in fact this works for any continuous function $[0, 1] \to \mathbb{R}$, given by:

$$\langle f, g\rangle := \int_0^1 f(x)g(x)dx$$

This is clearly symmetric and positive definite (integrating the square of a nonzero continuous function yields a positive number), and so it's an inner product to which the Cauchy-Schwarz inequality applies! This allows us to immediately prove, almost without doing any analysis, a rather substantive fact about continuous functions on $[0, 1]$:

Theorem.

For any continuous $f, g : [0, 1] \to \mathbb{R}$, we have:

$$\int_0^1 f(x)g(x)dx \leq \sqrt{\int_0^1 f(x)^2 dx}\sqrt{\int_0^1 g(x)^2 dx}$$

This hints one of the great utilities of inner products – they allow us to scalpel out the truly essential properties of geometry and apply them to a more general setting. One more of these crucially useful properties is the ability to project onto a coordinate axes. Thinking as we often do of coordinate axes as the span of basis elements, this might suggest we can take the inner product with a basis element to project onto its span. In order for this to exhibit the behaviour we want though, in particular that we can isolate the component of a vector in a particular direction via the inner product, we need a stricter kind of set than just a basis:

Definition.

An *orthonormal basis* of an inner product space $(V, \langle \cdot, \cdot \rangle)$ is a basis (e_i) of the vector space such that:

$$\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

This kind of basis is very convenient; for example, suppose we have some expansion of a vector in this basis $v = \sum a_i e_i$. Then we can very easily isolate its components:

$$\left\langle e_k, \sum_{i=1}^{n} a_i e_i \right\rangle = \sum_{i=1}^{n} a_i \langle e_k, e_i \rangle = a_k$$

The standard basis of $\mathbb{R}^n$ is one such basis, which is part of why it's so standard. Be careful though; as you may know if you've looked at the problem set, many otherwise nice bases like the basis $(1, x, x^2, \dots)$ of the inner product space $\left(\mathbb{R}[x], \int_0^1 fg dx\right)$ are *not* orthonormal. We might hope though that every inner product space admits some such basis. Good news!

Proposition.

Every finite-dimensional inner product space admits an orthonormal basis.

Proof There are two standard proofs of this fact. One of these uses the Gram-Schmidt process, which is covered in lecture and is very useful to understand if you want to do practical things with linear algebra, so look there if that's something you're interested. Thankfully for my computation-averse self though, there is a much less hands-on version, by induction on the dimension of V. Any norm 1 vector of a 1-dimensional inner product space[a] is trivially an orthonormal basis of that space. For the induction step, take any norm 1 vector $v_1, \|v_1\| = 1$ in the space and decompose $V = \text{span}(v) \oplus \text{span}(v)^\perp$. Inductively take an orthonormal basis $(v_2, ..., v_n)$ of the orthogonal $\text{span}(v)^\perp$, $n = \dim V$. Then the only pairwise products we need to check to check that $(v_1, ..., v_n)$ is orthonormal are those between v_1 and the others, since the rest are already given to us by induction hypothesis; but since $(v_2, \dots v_n)$ is in the orthogonal of the span of v_1 these must all be 0, and we are done. □

[a] this is just the field with its standard multiplication

Practically the Gram-Schmidt proof actually gives you an explicit basis, but there is a reason I prefer the one given here of the two given in lecture: this reflects a fundamental principle of how to prove problems about inner product spaces using induction: we can always split a space into a direct sum of a subspace and the orthogonal complement, splitting it into manageable components that interact with each other as minimally as possible. Never underestimate how many problems can be broken down by splitting into orthogonal complements like this. Another quick thing to note:

Note.

In an orthonormal basis, the matrix representing an inner product is just the identity matrix I, which makes the inner product in this basis just the sum of pointwise multiplications of components, that is the standard dot product. This means that the universal existence of orthonormal bases for finite dimensional inner product spaces gives an isomorphism respecting the inner product of any inner product space with $(\mathbb{R}^n, \bullet)$ as promised.

3 Adjoint and Self-Adjoint Operators

Before going over what adjoint and self-adjoint operators are, I'd like to go over that second perspective on what inner products are and why they're important. Recall that although we have an isomorphism $V \to V^*$ this isomorphism is in general not canonical; it requires an arbitrary choice of basis. The second perspective is that inner products are a canonical way of identifying a vector space with its dual. The utility of this is that there are many concepts that are more naturally introduced in the dual context that have phenomenal utility when transferred over to the vector space itself in this way. You've actually seen one of these already! Consider:

Proposition.

Let $\varphi : V \xrightarrow{\sim} V^*$ be the isomorphism induced by the inner product, $v \mapsto \langle v, \cdot \rangle$. Then for any subspace $S \subset V$:

$$\mathrm{Ann}(S) = \varphi(S^\perp)$$

Proof That $\varphi(S^\perp) \subset \mathrm{Ann}(S)$ is immediate from the definitions. Moreover, since we know φ is an isomorphism, any $\ell \in \mathrm{Ann}(S)$ can be written as $\langle w, \cdot \rangle$ for some $w \in V$; that this annihilates S is precisely the statement that $w \in S^\perp$. □

So a more concise definition of $S^\perp$ is that $S^\perp = \varphi^{-1}(\mathrm{Ann}(S))$. Positive-definiteness is essentially a condition to ensure that this does not overlap with S at all.

This perspective then leads to a very powerful question: what happens to the transpose under this identification? We know that if we have a linear operator $T : V \to V$ we have a naturally induced map, the *transpose* $T^t : V^* \to V^*$ given by $\ell \mapsto \ell \circ T$. Whenever we have an isomorphism ψ between spaces, the natural way to turn a map from one of the spaces to itself into a map from the other to itself is by conjugation, that is sending $T \mapsto \psi \circ T \circ \psi^{-1}$. This is because conjugation by an isomorphism as a general construction in algebra means "changing our perspective" by first shifting to a different domain, performing an operation, and then returning to the original space. So what is the transpose from the perspective of the identification of V^* with V via an inner product? Since our isomorphism is $V \to V^*$ and the transpose is a map $V^* \to V^*$ to get a corresponding map $V \to V$ we do this in reverse, and we can write $T^* = \varphi^{-1} \circ T^t \circ \varphi$. What is this? We can write it as a composition:

$$V \xrightarrow{\varphi} V^* \xrightarrow{T^t} V^* \xrightarrow{\varphi^{-1}} V$$

$$v \longmapsto \langle v, \cdot \rangle \longmapsto \langle v, T(\cdot) \rangle \longmapsto T^*(v)$$

So the transpose is, under this identification, just given by the linear map such that $\langle v, T(\cdot) \rangle = \langle T^*(v), \cdot \rangle$, which must exist and be unique because $\langle v, T(\cdot) \rangle \in V^*$ and φ is an isomorphism. This is precisely the *adjoint* of T:

Definition.

The *adjoint* of a linear operator on an inner product space $T : V \to V$ is the unique linear operator $T^* : V \to V$ such that $\langle v, Tw \rangle = \langle T^* v, w \rangle$ for all $v, w \in V$.

Now recall we have some properties relating the annihilator and the transpose:

Proposition.

1. $\ker T^t = \mathrm{Ann}(\mathrm{Im} T)$

2. $\mathrm{Im}T^t = \mathrm{Ann}(\ker T)$ assuming finite-dimensionality

Proof

1. $\ell \in \mathrm{Ann}(\mathrm{Im}T) \iff \forall v \in V : \ell(T(v)) = 0 \iff \ell \circ T = 0 \iff \ell \in \ker T^t$
2. $\ell' \in \mathrm{Im}T^t \iff \exists \ell \in V^* : \ell' = \ell \circ T \implies \ell'_{|\ker T} = 0 \implies \ell' \in \mathrm{Ann}(\ker T)$, other direction follows from (1) and rank-dimension formula

Passing through our isomorphism gives corresponding properties for the adjoint; in fact, these properties are stronger for adjoints because of the special properties of inner products, not requiring finite-dimensionality anymore:

Proposition.

1. $\ker T^* = (\mathrm{Im}T)^\perp$
2. $\mathrm{Im}T^* = (\ker T)^\perp$

Proof For (1), note that:

$$T^*u = 0 \iff \forall v \in V : \langle T^*u, v\rangle = 0 \iff \forall v \in V : \langle u, Tv\rangle = 0 \iff u \in (\mathrm{Im}T)^\perp$$

The other follows immediately from the facts that $(S^\perp)^\perp = S$ and that $T^{**} = T$, since indeed T is the unique linear operator such that $\langle v, T^*(\cdot)\rangle = \langle T(v), \cdot\rangle$ by symmetry. □

These properties are *incredibly* useful. They naturally encode a lot of the information about what makes adjoints useful, and often if you're trying to prove something about adjoints odds are you will be using these facts at some point. If you're ever stuck on a problem using adjoints and the spectral theorem hasn't made your problems vanish yet, throw these two facts at the problem for a bit, it will often make the problem go away. They are also suggestive of something however; ever since you've seen rank-nullity you've probably felt that the image and the kernel are complementary to each other, and now we have a very nice notion of complement to try this with. These properties almost get you there, if it were only the case that $T = T^*$. It turns out linear operators with that property are extraordinarily useful, precisely because they play very nicely with decompositions like this. So we make a definition:

Definition.

An operator is *self-adjoint* if $T^* = T$

Note.

Since the isomorphism $\varphi : V \to V^*$ sends an orthonormal basis (e_i) to its standard dual basis (e_i^*), in which the matrix of the transpose is just the matrix of the original map but with rows and columns swapped, the matrix of a self-adjoint linear operator in an orthonormal basis is always symmetric, $\mathcal{M}(T)_{ij} = \mathcal{M}(T)_{ji}$.