

24 Bilinear Forms

24.1 Review

Last week, we talked about the Sylow theorems, which are fundamental to the theory of finite groups.

24.2 Bilinear Forms

Throughout this class, we have been pivoting between group theory and linear algebra, and now we will return to some linear algebra.

Today, we will be discussing the notion of **bilinear forms**. Let's look at some examples first and then provide the general definition.

For now, we will be working with a vector space V over $F = \mathbb{R}$, and later on we will look at the case of $F = \mathbb{C}$, the complex numbers.

Let's consider three examples of bilinear forms on \mathbb{R}^3 .

Example 24.1

Consider these three different examples of mappings:

$$\begin{aligned} \mathbb{R}^3 \times \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &\xrightarrow{(1)} x_1y_1 + x_2y_2 + x_3y_3 \\ &\xrightarrow{(2)} x_1y_1 + 2x_2y_2 + 3x_2y_1 + 4x_2y_3 + 5x_3y_1 \\ &\xrightarrow{(3)} x_1y_1 + 2x_2y_1 + 2x_1y_2 + 3x_2y_2. \end{aligned}$$

These all take in pairs of vectors in \mathbb{R}^3 and return a real number. They all have the property that when keeping the y 's fixed, the mapping is "linear" in the x 's, and when keeping the x 's fixed, the mapping is "linear" in the y 's.⁸⁴ In particular, there are no constant terms or terms that are squared or higher order in x_i or y_i .

Definition 24.2

A **bilinear form** is a function

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \langle v, w \rangle^{85} \end{aligned}$$

such that

1. $\langle v, cw \rangle = c\langle v, w \rangle$
2. $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
3. $\langle cv, w \rangle = c\langle v, w \rangle$
4. $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$.

Requirements (1) and (2) are linearity in the second variable w , and requirements (3) and (4) are linearity in the first variable v .

⁸⁵The angle brackets $\langle \cdot, \cdot \rangle$ are how a bilinear form is usually denoted.

A bilinear form takes in *two* inputs and returns a real number in a way that is linear in either of its two inputs.⁸⁶ Intuitively, a bilinear form looks like Example 24.1.

⁸⁴The general idea of a bilinear form is that it is linear when varying in x (and keeping y fixed) and linear when varying in y (and keeping x fixed); hence, it is linear in two different variables, independently, so it is "bilinear." However, being bilinear is *not* the same as being linear; for example, if both x and y were doubled, the output would quadruple.

⁸⁶A "trilinear form" would also be possible.

Definition 24.3

A bilinear form is **symmetric** if

$$\langle v, w \rangle = \langle w, v \rangle$$

for all $v, w \in V$.

For instance, (1) and (3) in Example 24.1 are symmetric, but (2) is not, by looking at the coefficients.

Linear transformations from $\mathbb{R}^n \rightarrow \mathbb{R}^n$ can be written down explicitly using matrices. In a similar way, bilinear forms can also be described concretely using matrices. Consider the special case where $V = \mathbb{R}^n$. Then the **dot product** is a symmetric bilinear form. More generally, given any matrix $A \in \text{Mat}_{n \times n} R$, the mapping

$$\langle x, y \rangle := x^T A y \in \mathbb{R}$$

turns out to describe a bilinear form, satisfying the four properties.⁸⁷

Proposition 24.4

Given a symmetric matrix, the corresponding bilinear form is a symmetric bilinear form.

Proof. A matrix $A \in \text{Mat}_{n \times n} R$ is **symmetric** if $A^T = A$, and in this case,

$$\langle x, y \rangle = x^T A y$$

and

$$\langle y, x \rangle = y^T A x = (y^T A x)^T = x^T A^T y = x^T A y = \langle x, y \rangle.$$

□

For every matrix, there is an associated bilinear form, and for every symmetric matrix, there is an associated symmetric bilinear form. It turns out that *every* bilinear form arises in this manner.

Proposition 24.5

Every bilinear form $\langle \cdot, \cdot \rangle$ on \mathbb{R}^n arises from a matrix A . That is, there exists some A such that

$$\langle x, y \rangle = x^T A y.$$

Moreover, the form $\langle \cdot, \cdot \rangle$ is symmetric if and only if A is symmetric.

So there is a bijective correspondence between bilinear forms and $n \times n$ matrices.

In particular, for each example in 24.1, there is an associated matrix.

Example 24.6

The associated matrices come from the coefficients, and can be verified by simply carrying out the multiplication process.

$$1. A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$2. A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 4 \\ 5 & 0 & 0 \end{pmatrix}$$

$$3. A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

⁸⁷We won't verify this, but the properties follow from the way matrix multiplication works.

From this example, we see that the matrix entry A_{ij} is the coefficient of $x_i y_j$.

Proof of Proposition 24.5. Given a bilinear form, we want to produce the corresponding matrix. Let $V = \mathbb{R}^n$, and let the standard basis vectors be

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Any other column vector can be written as a linear combination of these basis vectors:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \vec{e}_1 + \dots + x_n \vec{e}_n.$$

In order to produce the matrix for the bilinear form, we look at the form evaluated on pairs of basis vectors. Let

$$a_{ij} = \langle \vec{e}_i, \vec{e}_j \rangle.$$

Now, placing these coefficients in a matrix, take

$$A = (a_{ij})_{i,j=1,\dots,n}.$$

To verify that this matrix actually produces the same result as the bilinear form on any two pairs of vectors, take $\vec{x}, \vec{y} \in \mathbb{R}^n$, and use bilinearity on both coordinates x and y .

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= \left\langle \sum_{i=1}^n x_i \vec{e}_i, \sum_{j=1}^n y_j \vec{e}_j \right\rangle \\ &= \sum_{i=1}^n x_i \langle \vec{e}_i, \vec{y} \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i \langle \vec{e}_i, \vec{e}_j \rangle y_j \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} y_j \\ &= \vec{x}^T A \vec{y}. \end{aligned}$$

In addition, if and only if $\langle \cdot, \cdot \rangle$ is symmetric, $a_{ij} = a_{ji}$, which is precisely the condition that A is symmetric. \square

From the bilinear hypothesis, the bilinear form on any two vectors can be written in terms of the form on a basis, which provides us the matrix. The upshot is that when $V = \mathbb{R}^n$, the information of a bilinear form can be encoded in a matrix.

Like when studying linear transformations, we do not have to restrict ourselves to only \mathbb{R}^n . More generally, for *any* vector space V along with a basis $\{v_1, \dots, v_n\}$ of V , a (symmetric) bilinear form on V corresponds with a (symmetric) matrix $A \in \text{Mat}_{n \times n}(\mathbb{R})$.⁸⁸

Guiding Question

What is the correspondence between a bilinear form on a vector space and the matrix, given a basis?

⁸⁸The way this form depends on the basis chosen will differ from the case of linear transformations, and will be discussed later in this lecture.

In some sense, a basis is simply a linear isomorphism

$$\mathcal{B} : \mathbb{R}^n \longrightarrow V.$$

A basis provides a dictionary between vectors in V and column vectors in \mathbb{R}^n . Given two vectors, the result of the bilinear form will be

$$\langle \vec{v}, \vec{w} \rangle = \vec{x}^T A \vec{y},$$

where

$$\mathcal{B}\vec{x} = \vec{v} \text{ and } \mathcal{B}\vec{y} = \vec{w}.$$

How can we find the entries of A ? We take $a_{ij} = \langle \vec{v}_i, \vec{v}_j \rangle$, and the same argument as in the proof of Proposition 24.5 holds, by using bilinearity.

24.3 Change of Basis

As always, we need to be careful about what basis we are working in and what effects the basis has.

Note 24.7 (Warning!)

A linear operator $T : V \longrightarrow V$ corresponds to an $n \times n$ matrix by picking a basis:

$$\text{linear operator } T : V \longrightarrow V \rightsquigarrow n \times n \text{ matrix}$$

Today, we saw that a bilinear form on V also corresponds to an $n \times n$ matrix by picking a matrix:

$$\text{bilinear form on } V \rightsquigarrow n \times n \text{ matrix}$$

But in fact, these two correspondences act extremely differently!

For a linear transformation, where the change of basis matrix is Q , the change of basis formula takes

$$P \longmapsto QPQ^{-1}.$$

Now, we can explore a change of basis for a bilinear form instead. Pick two bases

$$B : \mathbb{R}^n \longrightarrow V, B' : \mathbb{R}^n \longrightarrow V$$

for V , and consider a bilinear form $\langle \cdot, \cdot \rangle_V$ on V . The two bases are related by some invertible matrix P such that $B' = BP$ and $P \in GL_n(\mathbb{R})$.⁸⁹ Using B , there is one bilinear form $\langle \cdot, \cdot \rangle$ associated with some matrix A and using B' , there is another bilinear form $\langle \cdot, \cdot \rangle'$ associated with a matrix A' .

$$\begin{array}{ccc} \mathbb{R}^n & & \\ \uparrow P & \searrow B & \\ \mathbb{R}^n & \xrightarrow{B'} & V \end{array}$$

Given two column vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, the result in B' is

$$\langle \vec{x}, \vec{y} \rangle' = \langle B'\vec{x}, B'\vec{y} \rangle_V = \langle BP\vec{x}, BP\vec{y} \rangle_V.$$

This is the same as

$$\langle P\vec{x}, P\vec{y} \rangle = (P\vec{x})^T A (Py) = \vec{x}^T P^T A Py.$$

So the matrices are related by

$$\boxed{A' = P^T A P},$$

which is *not* $P^{-1}AP$. Changing basis for bilinear forms, unlike linear transformations, does *not* change the matrix by conjugation! If A is a symmetric matrix, then A' is also a symmetric matrix, which is expected. That's kind of alarming.

The same question for linear mappings can be asked in this situation.

⁸⁹The columns of P indicate how to write one basis in terms of the other.

Guiding Question

Given V and $\langle \cdot, \cdot \rangle_V$, can we pick a basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ of V such that A is as nice as possible?

For linear mappings, we ended up with the Jordan normal form. It turns out that the answer for bilinear forms is very nice! We will discuss this in the future.

24.4 Bilinear Forms over \mathbb{C}

The definitions provided so far generally work over any field. The **standard dot product**, which is a typical example of a bilinear form, has an additional property.

Definition 24.8

A **dot**, or **inner product** is a symmetric bilinear form such that

$$\langle x, x \rangle \geq 0,$$

and if $x \neq 0$, then $\langle x, x \rangle > 0$.^a

^aThis condition is called being positive definite.

We can use an inner product to measure distances and lengths in a vector space. In \mathbb{R}^n , $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{\vec{v} \cdot \vec{v}}$.

Guiding Question

Can we extend this to the complex numbers, when $F = \mathbb{C}$?

First, let's extend the notion of a dot product. We would like to do so in a way that captures our notion of distance. Naively setting the dot product in the same way as over \mathbb{R} results in a complex number, which does not measure distance in a way that we would prefer.

In \mathbb{C} , the length of a complex number z is $z\bar{z}$, which is the distance from the complex number z to the origin in the complex plane. The analogue of an inner product over \mathbb{C} will coincide with this definition of distance and use complex conjugation.

Definition 24.9

The standard Hermitian form on \mathbb{C}^n looks almost like the normal inner product, but with some complex conjugates thrown in. We have

$$\langle \vec{x}, \vec{y} \rangle = \bar{x}_1 y_1 + \bar{x}_2 y_2 + \dots + \bar{x}_n y_n \in \mathbb{C}.$$

In particular,

$$\langle \vec{x}, \vec{x} \rangle = \bar{\vec{x}}^T \vec{x} \in \mathbb{C}.$$

Once we do this, we get

$$\begin{aligned} \langle \vec{x}, \vec{x} \rangle &= \bar{x}_1 x_1 + \bar{x}_2 x_2 + \dots \\ &= |x_1|^2 + |x_2|^2 + \dots, \end{aligned}$$

which is actually a non-negative real number! So we prefer to use this Hermitian form over the complex numbers, as it can capture some notion of distance.

In the definition of the standard Hermitian form, we took the transpose and then the complex conjugate of every entry. This is a move we will do over and over.

Definition 24.10

For $M \in \text{Mat}_{m \times n}(\mathbb{C})$, the **adjoint** matrix is

$$M^* := \overline{M^T} \in \text{Mat}_{n \times m}(\mathbb{C}).$$

It behaves very much like taking the transpose does: $(AB)^* = B^* A^*$.

Then the equation from before becomes

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} \in CC.$$

Notice that for $\alpha \in \mathbb{C}$,

$$\langle \alpha \vec{x}, \vec{y} \rangle \neq \alpha \langle \vec{x}, \vec{y} \rangle,$$

so it is *not* bilinear in the first entry! We instead get

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

so it is linear in the second factor but nonlinear (it is only linear up to complex conjugation) in the first factor. This leads us to our last definition for today.

We can generalize the properties of the standard Hermitian form for a complex vector space.

Definition 24.11

For V a vector space over $F = \mathbb{C}$, then a Hermitian form is a function from

$$\begin{aligned} V \times V &\longrightarrow \mathbb{C} \\ (\vec{v}, \vec{w}) &\longmapsto \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

where

1. $\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle$
2. $\langle \vec{v}, \alpha \vec{w} \rangle = \alpha \langle \vec{v}, \vec{w} \rangle$
3. $\langle \vec{w}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{w} \rangle}$.

A Hermitian form is like a symmetric form, except instead of being symmetric, it is symmetric with a conjugation thrown in.

Notice that

$$\begin{aligned} \langle \alpha \vec{v}, \vec{w} \rangle &= \langle \vec{w}, \alpha \vec{v} \rangle \\ &= \overline{\alpha \langle \vec{w}, \vec{v} \rangle} \\ &= \overline{\alpha} \overline{\langle \vec{w}, \vec{v} \rangle} \\ &= \overline{\alpha} \langle \vec{v}, \vec{w} \rangle. \end{aligned}$$

So the Hermitian product of a vector with itself is in \mathbb{R} .