LECTURE : BILINEAR FORMS

Today we will focus on extra structure we can put on a vector space that render certain constructions more natural.

**Definition.** A bilinear form \(<-,-\>\) on \(V\) is a function \(<-,-\>: V \times V \to F\) that is bilinear:
\[
\langle a\vec{v} + b\vec{u}, \vec{w} \rangle = a\langle \vec{v}, \vec{w} \rangle + b\langle \vec{u}, \vec{w} \rangle
\]
and
\[
\langle \vec{v}, a\vec{u} + b\vec{w} \rangle = a\langle \vec{v}, \vec{u} \rangle + b\langle \vec{v}, \vec{w} \rangle.
\]

It is important to note that \(<-,-\>\) is a function on \(V \times V\) as a set and not a linear map on the direct sum \(V \oplus V\). This makes this definition a little unsettling: we consider structure on vector spaces that does not stay in the category of vector spaces.

We have a few types of special bilinear forms.

**Definition.**
1. \(<-,-\>\) is symmetric if \(\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle\) for all \(\vec{v}, \vec{w} \in V\).
2. \(<-,-\>\) is skew-symmetric if \(\langle \vec{v}, \vec{w} \rangle = -\langle \vec{w}, \vec{v} \rangle\) for all \(\vec{v}, \vec{w} \in V\).
3. \(<-,-\>\) is alternating if \(\langle \vec{v}, \vec{v} \rangle = 0\) for all \(\vec{v} \in V\).

These are not unrelated concepts, and for the first time, specific properties of the field \(F\) appear.

**Proposition.** Alternating always implies skew-symmetric. If \(\text{char}(F) = 2\), then skew-symmetric is the same as symmetric. If \(\text{char}(F) \neq 2\), then skew-symmetric implies alternating.

**Proof.** If \(\langle -,-\rangle\) is alternating, then we know
\[
0 = \langle \vec{v} + \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle.
\]
Thus \(\langle -,-\rangle\) is skew-symmetric. If the characteristic is 2, then 1 = -1, so symmetric is clearly the same as skew-symmetric. Assume that the characteristic is not 2 and \(\langle -,-\rangle\) is skew-symmetric. Then
\[
\langle \vec{v}, \vec{v} \rangle = -\langle \vec{v}, \vec{v} \rangle,
\]
and so we conclude that \(\langle \vec{v}, \vec{v} \rangle = 0\).

We should think of \(\langle -,-\rangle\) as a generalization of the dot product on \(\mathbb{R}^n\). This provides a fantastic example.

**Example.**
1. The dot product on \(\mathbb{R}^n\) is a bilinear form.
2. The generalization of the dot product on \(\mathbb{F}^n\) is a bilinear form.
3. If \(a_1, \ldots, a_n\) are elements of \(\mathbb{R}\) (or \(\mathbb{F}\)), then \(\langle \vec{v}, \vec{w} \rangle = a_1v_1w_1 + \cdots + a_nv_nw_n\) is a bilinear form.

The last part is very helpful, and it helps us tie bilinear forms to matrices.
Proposition. Choose a basis $\mathcal{B} v_1, \ldots, v_n$ for $V$, and let $[\vec{v}]_\mathcal{B} \in \mathbb{F}^n$ denote the column vector that represents $\vec{v}$ with respect to $\mathcal{B}$. Let $B$ denote the matrix with $B_{i,j} = \langle v_i, v_j \rangle$.

Then $[\vec{v}]_\mathcal{B}^t B [\vec{w}]_\mathcal{B} = \langle \vec{v}, \vec{w} \rangle$.

Proof. This is an exercise in bilinearity. Both sides are clearly bilinear (the left one is so since multiplication distributes over addition). By linearity in the first factor, to know $\langle \vec{v}, \vec{w} \rangle$ (or the column vector form), it suffices to know $\langle \vec{v}_i, \vec{w} \rangle$ for all $i$. Similarly, by linearity in the second factor, to know $\langle \vec{v}, \vec{w} \rangle$, it suffices to know $\langle \vec{v}, \vec{w}_j \rangle$ for all $j$. Thus we are reduced to showing $[\vec{v}_i]_\mathcal{B}^t B [\vec{v}_j]_\mathcal{B} = \langle \vec{v}_i, \vec{v}_j \rangle$.

However, by definition, $[\vec{v}_i]_\mathcal{B} = e_i$. Matrix multiplication tells us that $[\vec{v}_i]_\mathcal{B}^t B [\vec{v}_j]_\mathcal{B} = B_{i,j} = \langle \vec{v}_i, \vec{v}_j \rangle$, as required. \qed

Thus when we have a basis $\mathcal{B}$, we have a connection between matrices and bilinear forms.

Proposition. A choice of basis $\mathcal{B}$ of $V$ provides a 1−1 correspondence between bilinear forms and the associated matrices.

What happens then if we change basis? Let $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_n\}$ be another basis, and let $P = c P_B$ be the change-of-basis matrix. Then we know $[\vec{v}]_\mathcal{C} = c [\vec{v}]_\mathcal{B}$.

Let $C$ be the matrix associated to $\langle -, - \rangle$ in the $\mathcal{C}$-basis, and similarly for $B$. Then we have $[\vec{v}]_\mathcal{C}^t C [\vec{w}]_\mathcal{C} = \langle \vec{v}, \vec{w} \rangle = [\vec{v}]_\mathcal{B}^t B [\vec{w}]_\mathcal{B}$.

Combining these two equations, we see that for all $\vec{v}$ and $\vec{w}$, we have $[\vec{v}]_\mathcal{C}^t C [\vec{w}]_\mathcal{C} = (P[\vec{v}]_\mathcal{B})^t C (P[\vec{w}]_\mathcal{B}) = [\vec{v}]_\mathcal{B}^t (P^t C P)[\vec{w}]_\mathcal{B} = [\vec{v}]_\mathcal{B}^t B [\vec{w}]_\mathcal{B}$.

We’ve therefore shown the following.

Proposition. If $P$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$, and if $B$ (resp $C$) represents $\langle -, - \rangle$ in the $\mathcal{B}$-basis (resp $\mathcal{C}$-basis), then $P^t C P = B$.

Thus we’ll say that two matrices are transpose equivalent if there is an invertible matrix $P$ such that $P^t C P = B$. This is not standard.

Row vectors showed up in another context: dual spaces. This is one of the most powerful aspects of bilinear forms: they provide a natural way to connect a vector space and its dual.

Theorem. Let $\vec{v} \in V$, and let $R_{\vec{v}}$ denote $\langle \vec{v}, - \rangle$. Then

1. For all $\vec{v}$, $R_{\vec{v}}$ is a linear transformation $V \to \mathbb{F}$.
2. The assignment $R : V \to V^*$ given by $\vec{v} \mapsto R_{\vec{v}}$ is linear.
Proof. Both of these are given by linearity in their respective factors. The bilinear form \(<-,-\>\) is linear in the second factor, so for all \(\vec{v}, \langle \vec{u}, - \rangle\) is a linear map. Thus we have the first part, and we learn that \(R\) is a function from \(V\) to \(V^*\).

For the second, we need to know that for all \(\vec{v} \in V\),

\[
R_{a\vec{v} + b\vec{w}}(\vec{u}) = aR_{\vec{v}}(\vec{u}) + bR_{\vec{w}}(\vec{u}).
\]

This will show that \(V \to V^*\) is a linear map. So we check:

\[
R_{a\vec{v} + b\vec{w}}(\vec{u}) = \langle a\vec{v} + b\vec{w}, \vec{u} \rangle = a\langle \vec{v}, \vec{u} \rangle + b\langle \vec{w}, \vec{u} \rangle = aR_{\vec{v}}(\vec{u}) + bR_{\vec{w}}(\vec{u}),
\]

since \(<-,-\>\) is linear in the first factor.

To fully understand the map \(R\), we have to determine \(\ker(R)\) and the image of \(R\).

**Definition.** The **radical** of \(V\), \(\text{Rad}(V)\) is defined by

\[
\text{Rad}(V) = \{\vec{v} \in V | \forall \vec{w} \in V, \langle \vec{v}, \vec{w} \rangle = 0\}.
\]

**Proposition.** We have \(\ker(R) = \text{Rad}(V)\).

**Proof.** This is immediate. The zero functional is the one that assigns the value 0 to all \(\vec{w} \in V\). So \(R(\vec{v}) = 0\) iff for all \(\vec{w} \in V\), \(\langle \vec{v}, \vec{w} \rangle = 0\). This is the same thing as \(\vec{v} \in \text{Rad}(V)\). \(\square\)

To make this more useful, we tether this to the matrix form. If \(\vec{v} \in V\) is in the radical, then we have

\[
[\vec{v}]_B^t B [\vec{w}]_B = 0
\]

for all \(\vec{w} \in V\). This is the same condition as \([\vec{v}]_B^t B = \vec{0}\), or \([\vec{v}]_B\) is in the null-space of \(B^t\).

**Definition.** A bilinear form \(<-,-\>\) is singular if \(B\) is singular.

A bilinear form \(<-,-\>\) is non-singular if \(B\) is invertible.

**Corollary** (Reisz Representation Theorem). If \(<-,-\>\) is non-singular, then \(R: V \to V^*\) is an injection. If \(V\) is finite dimensional, then \(R\) is an isomorphism.

Thus in the finite dimensional case, for any \(f \in V^*\), we have a vector \(\vec{v}_f \in V\) such that \(f(\vec{w}) = \langle \vec{v}_f, \vec{w} \rangle\) for all \(\vec{w} \in V\). This is the Reisz vector for \(f\).

Now if \(S\) is a subspace of \(V\), then \(S\) inherits a bilinear form by restriction. It can be the case that \(S\) is singular even if \(V\) is not. If \(V\) is non-singular and \(S\) is a singular subspace, then the Reisz vector for \(f \in S^*\) need not be a vector in \(S\). In fact, it could be just a vector in \(V\). \(S\) is non-singular exactly when the Reisz vector is back in \(S\).