Lecture 13

Hypervectors, the Change of Basis Matrix and Gram-Schmidt

This lecture does not come from the text. I is helpful for understanding the change of basis formula and also the relation of the determinant function $D(v_1, v_2, \ldots, v_n)$ on an n-dimensional vector space V to determinants of matrices. See Chapter 5. You will find hypervectors to be a very useful computational tool.

Definition

Let V be a vector space

bold

is an ordered k-tuple $\mathbf{v} = (v_1, v_2, \ldots, v_k)$ of vectors from V. We denote the set of k-hypervectors by V^k

$$V^k = \underbrace{V \times V \times \cdots \times V}_{k}$$

A Matrix Interpretation of Hypervectors

Choose a basis $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ for V. Then every vector v has coordinates $[v]_{\mathcal{B}}$ relative to $\mathcal{B}$, which we will think of as a column vector. Now let $\vec{v} = (v_1, \ldots, v_k)$ bold be a k-hypervector. We associate to bold $\vec{v}$ the n by k matrix $M(\vec{v})$ bold whose first column is the coordinates of v_1, the second the coordinates of v_2 etc.

__Example__ $V = \mathrm{Pol}_2(\mathbb{R})$ and bold $\vec{v} = (v_1, v_2)$ with $v_1 = 1 + 3x + 2x^2$ and $v_2 = 5 + x$. Take $\mathcal{B} = (1, x, x^2)$. Then

bold
$$M(\vec{v}) = \begin{pmatrix} 1 & 5 \\ 3 & 1 \\ 2 & 0 \end{pmatrix}$$

Operations on Hypervectors

Since V^k is a vector space you add hypervectors and multiply them by scalars

But we have new operations

I Matrices Operate on V^k from the Right

I will give an example of how this works. Consider the matrix $A = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}$

Then we get a linear map

$R_A : V^3 \longrightarrow V^2$ as follows

$$R_A((v_1, v_2, v_3)) = (v_1, v_2, v_3)\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}$$

$$= (av_1 + bv_2 + cv_3,\ bv_1 + dv_2 + fv_3)$$

So pretend that (v_1, v_2, v_3) is a row vector of numbers.

More generally if A is a k by ℓ matrix we define

$$R_A : V^k \longrightarrow V^{\ell}$$ by

$$R_A((v_1, v_2, \ldots, v_k)) = (v_1, v_2, \ldots, v_k)A$$

$$= \left(\sum_{i=1}^{k} a_{i1} v_i , \sum_{i=1}^{k} a_{i2} v_i , \ldots , \sum_{i=1}^{k} a_{i\ell} v_i \right) \quad (*)$$

\ell ll

__Remark__

If you rewrite $\sum_{i=1}^{k} a_{i1} v_i$ as $\sum_{i=1}^{k} v_i a_{i1}$ then the two i's are adjacent. This is the way we should write it because we are multiplying by A on the right but we can't because we write the product of a vector v and a scalar c as $c \cdot v$ not $v \cdot c$.

Examples

1. It is important to note that a basis of an n-dimensional vector space V is an n-hypervector an $\vec{b}$
$\mathcal{B} = \vec{b} = (b_1, b_2, \ldots, b_n) \in V^n$.
The notation $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ is incorrect because the elements of a set are not ordered. By definition $\{x_1, x_2\} = \{x_2, x_1\}$.
But everybody writes bases as $\{b_1, b_2, \ldots, b_n\}$.

2. The coordinates $(x_1, x_2, \ldots, x_n)$ of vector $v \in V$ relative to a basis $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ are a special case of the operation of a matrix A on the n-hypervector given by the basis. Put $A = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

so A is an n by 1 matrix

so $R_V : V^n \longrightarrow V$

$$R_A((b_1, b_2, \dots, b_n)) = (b_1, b_2, \dots, b_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \sum_{i=1}^{n} x_i b_i = v$$

Linear Transformations Operate on Hypervectors from the Left

Now let $T \in L(V, V)$. Then for any k we define left multiplication L_T of T on V^k by

$$L_T((v_1, v_2, \dots, v_k)) = (Tv_1, Tv_2, \dots, Tv_k)$$

So $L_T : V^k \longrightarrow V^k$

Proposition

Let A be a k by k matrix. Then L_T and R_A both take V^k to V^k. Then

$$\underbrace{L_T \circ R_A = R_A \circ L_T}_{\text{the two operations commute}} \quad (**)$$

There is an important principle behind this lemma that occurs in many places in mathematics - operations on the left commute with operations on the right. To see how this works we give a matrix interpretation of (**)

Choose a basis $\mathcal{B} = (b_1, b_2, \ldots, b_n)$

Let $\mathbf{v} = (v_1, v_2, \ldots, v_k)$ and $M(\mathbf{v})$ (bold)

be the n by k matrix of $\mathbf{v}$

Let A be as above.

Then we have

<u>Lemma</u>

The matrix of the k-hypervector $R_A(\mathbf{v})$ is the n by k

$$\underbrace{M(\mathbf{v}) \cdot A}_{\text{right matrix multiplication}}$$

I won't prove this.

Now let $M(T) = [T]_{\mathcal{B}}$ be the matrix of T relative to the basis $\mathcal{B}$

Lemma

The matrix of the k-hypervector $L_T((v_1, \ldots, v_k))$ is

$$\underbrace{M(T) \cdot M(v)}_{\text{left matrix multiplication}}$$

I won't prove this either.

Now the Proposition follows because left matrix multiplication commutes with right matrix multiplication. Why?

Given X an n by k matrix
First multiply on the right by A
to get XA. Then multiply on the left by B to get $B(XA) = BXA$.

Now reverse the order. First multiply on the left by B
to get BX, then multiply on the right by A to get $(BX)A = BXA$

Now we can better understand the definitions of the change of basis matrix and the matrix of a linear transformation and the Gram-Schmidt orthogonalization process

The Change of Basis Matrix from $\mathcal{B}$ to $\mathcal{C}$ (see Lecture 7)

Let $\mathcal{B} = \mathbf{b} = (b_1, b_2, \ldots, b_n)$ and $\mathcal{C} = \mathbf{c} = (c_1, c_2, \ldots, c_n)$ be bases for a vector space V. Then

bold / bold

the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$ satisfies the equation of n-hypervectors

$$(b_1, b_2, \ldots, b_n) = (c_1, c_2, \ldots, c_n) \cdot P_{\mathcal{C} \leftarrow \mathcal{B}}$$

The Matrix of a Linear Transformation

Let $T \in L(V, V)$ and

$\mathcal{B} = \overset{\text{bold}}{b} = (b_1, b_2, \ldots, b_n)$ be a basis for V.

Then the matrix $M(T) = [T]_{\mathcal{B}}$ satisfies the equation of n-hypervectors

$$(Tb_1, Tb_2, \ldots, Tb_n) = (b_1, b_2, \ldots, b_n) \cdot M(T)$$

Remark

It is because we are multiplying on the right by the basis (hypervector) that the a matrix in these two cases (not matched) indices are strange

$$b_j = \sum_{i=1}^{n} x_{ij} c_i \qquad \left(\text{so } P_{\mathcal{C} \leftarrow \mathcal{B}} = (x_{ij})\right)$$

$$T(b_j) = \sum_{i=1}^{n} y_{ij} b_i \qquad \left(\text{so } [T]_{\mathcal{B}} = (y_{ij})\right)$$

Gram-Schmidt Orthogonalization

We start with a two-dimensional space and a basis $\mathcal{B} = (v_1, v_2)$. We want to convert to an orthonormal basis $\mathcal{U} = (u_1, u_2)$. The Gram-Schmid orthogonalization process does this in three steps.

(1) Make v_1 have unit length so put $u_1 = \dfrac{v_1}{\|v_1\|}$

so now we have (u_1, v_2)

(2) Make v_2 orthogonal to u_1 so put

$$v_2' = v_2 - (u_1, v_2) u$$

$$= v_2 - \frac{(v_1, v_2) v_1}{\|v_1\|^2}$$

(3) Make v_2' have unit length so

$$u_2 = \frac{v_2 - \dfrac{(v_1, v_2)}{\|v_1\|^2} v_1}{\left\| v_2 - \dfrac{(v_1, v_2)}{\|v_1\|^2} v_1 \right\|}$$

Hence

$$(u_1, u_2) = \left(\frac{v_1}{\|v_1\|}, \frac{v_2 - \frac{(v_1, v_2)}{\|v_1\|^2} v_1}{\left\| v_2 - \frac{(v_1, v_2)}{\|v_1\|^2} v_1 \right\|} \right)$$

We can write this as

$$(u_1, u_2) = (v_1, v_2) \begin{pmatrix} \frac{1}{\|v_1\|} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{(v_1, v_2)}{\|v_1\|^2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\left\| v_2 - \frac{(v_1, v_2)}{\|v_1\|^2} \right\|} \end{pmatrix}$$

So multiplying the three matrices we get

$$(u_1, u_2) = (v_1, v_2) \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

The key point is that the matrix is upper triangular. This reflects the fact that the Gram-Schmidt process operating on $(v_1, v_2, \ldots, v_i, \ldots v_n)$ changes v_i by a function of v_i and vectors to the left of v_i.

So Gram-Schmidt moves from left to right.

In general we have

Theorem

Let $(v_1, v_2, \ldots, v_n)$ be a basis for a vector space V. Then there is an upper triangular matrix T whose entries are a function of the inner products (v_i, v_j) such

$$(u_1, u_2, \ldots, u_n) = (v_1, v_2, \ldots, v_n)\, T$$

is an orthonormal basis for V.