1111: Linear Algebra I

Dr. Vladimir Dotsenko (Vlad)

Lecture 18

Change of coordinates

Let V be a vector space of dimension n, and let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be two different bases of V. Then we can compute coordinates of each vector v with respect to either of those bases, so that

$$v = x_1 e_1 + \cdots + x_n e_n$$

and

$$v = y_1 f_1 + \cdots + y_n f_n.$$

Our goal now is to figure out how these are related. For that, we shall need the notion of a transition matrix.

Definition 1. Let us express the vectors $f_1, \ldots, f_n$ as linear combinations of $e_1, \ldots, e_n$:

$$\begin{aligned} f_1 &= a_{11}e_1 + a_{21}e_2 + \cdots + a_{m1}e_m, \\ f_2 &= a_{12}e_1 + a_{22}e_2 + \cdots + a_{m2}e_m, \\ &\cdots \\ f_n &= a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{mn}e_m. \end{aligned}$$

The matrix (a_{ij}) is called *the transition matrix* from the basis $e_1, \ldots, e_n$ to the basis $f_1, \ldots, f_n$. Its k-th column is the column of coordinates of the vector f_k relative to the basis $e_1, \ldots, e_n$.

Lemma 1. *In the notation above, we have*

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \ldots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

In plain words, if we call $e_1, \ldots, e_n$ the "old basis" and $f_1, \ldots, f_n$ the "new basis", then this system tells us that the product of the transition matrix with the columns of new coordinates of a vector is equal to the column of old coordinates.

Proof. The proof is fairly straightforward: we take the formula

$$v = y_1 f_1 + \cdots + y_n f_n,$$

and substitute instead of f_i's their expressions in terms of e_j's:

$$\begin{aligned} f_1 &= a_{11}e_1 + a_{21}e_2 + \cdots + a_{m1}e_m, \\ f_2 &= a_{12}e_1 + a_{22}e_2 + \cdots + a_{m2}e_m, \\ &\cdots \\ f_n &= a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{mn}e_m. \end{aligned}$$

What we get is

$$y_1(a_{11}e_1+a_{21}e_2+\dots+a_{n1}e_n)+y_2(a_{12}e_1+a_{22}e_2+\dots+a_{n2}e_n)+\ldots+y_n(a_{1n}e_1+a_{2n}e_2+\dots+a_{nn}e_n) = \\ = (a_{11}y_1+a_{12}y_2+\dots+a_{1n}y_n)e_1+\dots+(a_{n1}y_1+a_{n2}y_2+\dots+a_{nn}y_n)e_n.$$

Since we know that coordinates are uniquely defined, we conclude that

$$a_{11}y_1+a_{12}y_2+\dots+a_{1n}y_n = x_1,$$
$$\dots$$
$$a_{n1}y_1+a_{n2}y_2+\dots+a_{nn}y_n = x_n,$$

which is what we want to prove. □

If we denote, for a vector v, the column of coordinates of v with respect to the basis $e_1, \dots, e_n$ by $v_{\mathbf{e}}$, and also denote the transition matrix from the basis $e_1, \dots, e_n$ to the basis $f_1, \dots, f_n$ by $M_{\mathbf{e},\mathbf{f}}$, then the previous result can be written as

$$v_{\mathbf{e}} = M_{\mathbf{e},\mathbf{f}} v_{\mathbf{f}}.$$

Lemma 2. *We have*

$$M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{g}} = M_{\mathbf{e},\mathbf{g}}$$

and

$$M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{e}} = I_n$$

if $\dim(V) = n$.

Proof. Applying the formula above twice, we have

$$v_{\mathbf{e}} = M_{\mathbf{e},\mathbf{f}} v_{\mathbf{f}} = M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{g}} v_{\mathbf{g}}.$$

But we also have

$$v_{\mathbf{e}} = M_{\mathbf{e},\mathbf{g}} v_{\mathbf{g}}.$$

Therefore

$$M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{g}} v_{\mathbf{g}} = M_{\mathbf{e},\mathbf{g}} v_{\mathbf{g}}$$

for every $v_{\mathbf{g}}$. From our previous classes we know that knowing $A\mathbf{v}$ for all vectors $\mathbf{v}$ completely determines the matrix A, so $M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{g}} = M_{\mathbf{e},\mathbf{g}}$ as required. Since manifestly we have $M_{\mathbf{e},\mathbf{e}} = I_n$, we conclude by letting $g_k = e_k$, $k = 1, \dots, n$, that $M_{\mathbf{e},\mathbf{f}} M_{\mathbf{f},\mathbf{e}} = I_n$. □

Linear maps and transformations

Definition 2. Suppose that V and W are two vector spaces. A function $f\colon V \to W$ is said to be a *linear map*, or a *linear operator*, if

- for $v_1, v_2 \in V$, we have $f(v_1+v_2) = f(v_1)+f(v_2)$,
- for $c \in \mathbb{R}$, $v \in V$, we have $f(c \cdot v) = c \cdot f(v)$.

A linear map from a vector space V to the same vector space is said to be a *linear transformation* of V.

Example 1. As we know, every linear map $\varphi\colon \mathbb{R}^n \to \mathbb{R}^k$ is given by a $k \times n$-matrix A, so that $\varphi(x) = Ax$.

Example 2. Let V be the vector space of all polynomials in one variable x. Consider the function $X: V \to V$ that maps every polynomial $f(x)$ to $xf(x)$. This is a linear transformation of V:

$$x(f_1(x) + f_2(x)) = xf_1(x) + xf_2(x),$$
$$x(cf(x)) = c(xf(x)).$$

Let P_n be the vector space of all polynomials in one variable x of degree at most n. Then the rule X as above defines a linear map $: P_n \to P_{n+1}$. (Note that the target of φ has to be different, since multiplying by x increases degrees).

Example 3. Let V be the vector space of all polynomials in one variable x. Consider the function $D: V \to V$ that maps every polynomial $f(x)$ to $f'(x)$. This is a linear map:

$$(f_1(x) + f_2(x))' = f_1'(x) + f_2'(x),$$
$$(cf(x))' = cf'(x).$$

The function D define both a linear map $: P_n \to P_{n-1}$, and a linear transformation of P_n (since the degree of the derivative of a polynomial of degree *at most* n is *at most* $n-1$).