9 Computing with bases

We spent a lot of time looking how to deal with vectors in $\mathbf{R}^n$, or, more generally, in $\mathbf{F}^n$. What about other finite dimensional vector spaces? It turns out that even if we are working in some abstract finite dimensional vector space V over $\mathbf{F}$, it is still possible to do all the computations with vectors in $\mathbf{F}^n$. I will say more about it later, but now I will just show how it works.

Let V be a finite dimensional vector space over the field $\mathbf{F}$, and assume that $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis in V. This is the first time, when the order of these vectors is actually important and hence in this lecture I will be talking about *ordered collections* of vectors from V. Recall that having a basis means exactly that each vector $v \in V$ has a *unique* representation

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

The scalars α_i are called the *coordinates* of v with respect to basis $\mathcal{B}$. If for $u \in V$

$$u = \beta_1 v_1 + \ldots + \beta_n v_n,$$

then, clearly,

$$u + v = (\alpha_1 + \beta_1)v_1 + \ldots + (\alpha_n + \beta_n)v_n,$$

so that the i-th coordinate of $u + v$ is $\alpha_i + \beta_i$. Similarly, the i-th coordinate of γu is $\gamma \alpha_i$.

The key fact is that I can consider the coordinates of u as an elements of $\mathbf{F}^n$, and, in the opposite direction, any vector $\boldsymbol{x} \in \mathbf{F}^n$ gives coordinates of some vector $w \in V$. To summarize, each ordered basis for V determines one-to-one correspondence

$$v \in V \mapsto (\alpha_1, \ldots, \alpha_n) \in \mathbf{F}^n$$

between the set of all vectors in V and the set of all n-tuples in $\mathbf{F}^n$. Note that the sum $u + v$ corresponds to the sum of two vectors in $\mathbf{F}^n$ and γv corresponds to the product by a scalar in $\mathbf{F}^n$. This means that *all the calculations can be performed in* $\mathbf{F}^n$, and any finite dimensional vector space V of dimension n actually can be represented as $\mathbf{F}^n$ (we say, mathematically, that these two vector spaces are *isomorphic*).

Now, since it is possible to have many different bases, I want to see how the coordinates of my vector change with respect to the given basis. From now on, my vectors $\boldsymbol{x} \in \mathbf{F}^n$ are *column-vectors*, and I will denote $\boldsymbol{x} = [v]_{\mathcal{B}}$ to emphasize the dependence of $\boldsymbol{x}$ on the basis. Let $\mathcal{B}' = (v_1', \ldots, v_n')$ be another basis in V, hence I will have a different vector $\boldsymbol{x}' = [v]_{\mathcal{B}'}$. How the vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$ are related?

I have that

$$v_j' = \sum_{i=1}^{n} p_{ij} v_i,$$

Math 329: *Intermediate Linear Algebra* by Artem Novozhilov©
e-mail: artem.novozhilov@ndsu.edu. Spring 2017

for some constants p_{ij}. Now,

$$\begin{aligned} v &= \alpha'_1 v'_1 + \ldots + \alpha'_n v'_n \\ &= \sum_{j=1}^n \alpha'_j v'_j \\ &= \sum_{j=1}^n \alpha'_j \sum_{i=1}^n p_{ij} v_i \\ &= \sum_{j=1}^n \sum_{i=1}^n (p_{ij}\alpha'_j) v_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij}\alpha'_j \right) v_i. \end{aligned}$$

This implies, due to the uniqueness of the coordinates with respect to a given basis, that

$$\alpha_i = \sum_{j=1}^n p_{ij}\alpha'_j, \quad i = 1, \ldots, n,$$

or, denoting $\boldsymbol{P} = [p_{ij}]_{n \times n}$,

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}'.$$

Since both bases are linearly independent, it implies that $\boldsymbol{x} = 0$ if and only if $\boldsymbol{x}' = 0$, therefore, the corresponding system of homogeneous equations has only the trivial solution, hence $\boldsymbol{P}$ must be invertible, and hence

$$\boldsymbol{x}' = \boldsymbol{P}^{-1}\boldsymbol{x}.$$

Using the introduced notation, the same connection can be written

$$\begin{aligned} [v]_{\mathcal{B}} &= \boldsymbol{P}[v]_{\mathcal{B}'}, \\ [v]_{\mathcal{B}'} &= \boldsymbol{P}^{-1}[v]_{\mathcal{B}}. \end{aligned}$$

Thus I have proved

Theorem 9.1. *Let V be n-dimensional vector space over $\mathbf{F}$ and let $\mathcal{B}$, $\mathcal{B}'$ be two bases of V. Then there is a unique, necessarily invertible, $n \times n$ matrix $\boldsymbol{P}$ with entries in $\mathbf{F}$ such that*

$$\begin{aligned} [v]_{\mathcal{B}} &= \boldsymbol{P}[v]_{\mathcal{B}'}, \\ [v]_{\mathcal{B}'} &= \boldsymbol{P}^{-1}[v]_{\mathcal{B}}, \end{aligned}$$

for every vector $v \in V$. The columns of $\boldsymbol{P}$ are given by

$$\boldsymbol{p}_j = [v'_j]_{\mathcal{B}}, \quad j = 1, \ldots, n.$$

Remark 9.2. Actually, a converse is true: For any invertible matrix $\boldsymbol{P}$ and fixed basis $\mathcal{B}$ in V there is a unique basis $\mathcal{B}'$ with the aforementioned properties. I will leave proving this fact as an exercise.

Example 9.3. Recall that in the previous lectures I showed that the vectors

$$\begin{aligned}\boldsymbol{v}_1 &= (1,2,2,1),\\ \boldsymbol{v}_2 &= (0,2,0,1),\\ \boldsymbol{v}_3 &= (-2,0,-4,3).\end{aligned}$$

are linearly independent and hence form a basis on the subspace $W = \operatorname{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$. I also shower that an arbitrary vector $\boldsymbol{x} \in \mathbf{R}^4$ is also in W if and only if $x_3 = 2x_1$. Now consider another three vectors:

$$\begin{aligned}\boldsymbol{v}_1' &= (1,0,2,0),\\ \boldsymbol{v}_2' &= (0,2,0,1),\\ \boldsymbol{v}_3' &= (0,0,0,3).\end{aligned}$$

They are clearly in W. Also they are linearly independent (check this carefully) and hence form another basis $\mathcal{B}'$ of W. Consider a vector $\boldsymbol{z} \in \mathbf{R}^4$ whose coordinates relative to the $\mathcal{B}$ and $\mathcal{B}'$ bases are $\boldsymbol{x}$ and $\boldsymbol{x}'$ respectively. I want to find matrix $\boldsymbol{P}$, such that $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}'$.

I know that to find $\boldsymbol{P}$, I need to find the coordinates of $\boldsymbol{v}_i'$ with respect to $\mathcal{B}$. In coordinates, for example, for the first vector, I must have

$$\begin{bmatrix}1\\0\\2\\0\end{bmatrix} = \begin{bmatrix}1\\2\\2\\1\end{bmatrix} p_{11} + \begin{bmatrix}0\\2\\0\\1\end{bmatrix} p_{21} + \begin{bmatrix}-2\\0\\-4\\3\end{bmatrix} p_{31}.$$

Since I know that my third coordinate must be doubled first coordinate, I can disregard the third row and hence end up with the system of linear equations with a square matrix, which I can invert to find

$$\begin{bmatrix}1 & 0 & -2\\ 2 & 2 & 0\\ 1 & 1 & 3\end{bmatrix}^{-1} = \begin{bmatrix}1 & -1/3 & 2/3\\ -1 & 5/6 & -2/3\\ 0 & -1/6 & 1/3\end{bmatrix}.$$

Therefore, the first column of $\boldsymbol{P}$ is given as

$$\begin{bmatrix}1 & 0 & -2\\ 2 & 2 & 0\\ 1 & 1 & 3\end{bmatrix}^{-1} \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\-1\\0\end{bmatrix}$$

Similarly, I find the rest of coordinates and obtain

$$\boldsymbol{P} = \begin{bmatrix}1 & 0 & 2\\ -1 & 1 & -2\\ 0 & 0 & 1\end{bmatrix}.$$

Hence, every time when we are dealing with finite dimensional vector spaces and performing some computations with coordinates, it should be clear that the computations are performed with respect to some fixed basis. Why then we never mention this when doing computations in Math 129 with the usual vectors and matrices in $\mathbf{R}^n$? The answer is because the general expression "vector $\boldsymbol{x} \in \mathbf{R}^3$ has coordinates $[x_1 \ x_2 \ x_3]$" is an abbreviation of the correct "vector $\boldsymbol{x} \in \mathbf{R}^3$ has coordinates $[x_1 \ x_2 \ x_3]$ with respect to the standard ordered basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$." Keep this in mind when you perform such computations.