MATH 211: Introduction to Linear Algebra - Fall 2017

FORMULAS ABOUT BASE CHANGE AND MATRIX REPRESENTATION OF A LINEAR FUNCTION

Base change is considered to be one of the most difficult topics in linear algebra. In order to help you to understand this concept and solve the exercises, I prepared these notes in which I summarize the most important formulas about base change. You should compare or supplement these notes with what we covered in the class of Nov. 15th.

1. Base change

Denote by V a vector space of dimension n and fix a basis $\mathcal{B} = (v_1, \dots v_n)$ of V. Then every element v of V can be uniquely written as a linear combination of the elements forming the basis $\mathcal{B}$:

$$v \;=\; c_1 v_1 + \ldots + c_n v_n \qquad \text{for a unique choice of scalars} \qquad c_1, \ldots, c_n.$$

The scalars $c_1, \ldots, c_n$ are the $\mathcal{B}$-coordinates of v and we write

$$[v]_{\mathcal{B}} \;=\; \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

In this way we can represent any element v of V as a column vector with n entries.

Now suppose that we have another basis $\mathcal{C} = (w_1, \ldots, w_n)$ of V. (You should think that $\mathcal{B}$ is the old basis, while $\mathcal{C}$ is a new basis.) Then every element v of V has two different sets of coordinates, the $\mathcal{B}$-coordinates and the $\mathcal{C}$-coordinates. We can pass from one type of coordinates to another thanks to the *matrix of change of basis* $S_{\mathcal{C}\to\mathcal{B}}$. This matrix transforms the $\mathcal{C}$-coordinates of every element v of V, into the $\mathcal{B}$-coordinates of v. Hence for any element v in V we have

$$\boxed{[v]_{\mathcal{B}} \;=\; S_{\mathcal{C}\to\mathcal{B}}\,[v]_{\mathcal{C}}.} \tag{1}$$

The matrix of change of basis $S_{\mathcal{C}\to\mathcal{B}}$ is defined as:

$$\boxed{S_{\mathcal{C}\to\mathcal{B}} \;=\; \begin{pmatrix} | & | & \cdots & | \\ [w_1]_{\mathcal{B}} & [w_2]_{\mathcal{B}} & \cdots & [w_n]_{\mathcal{B}} \\ | & | & \cdots & | \end{pmatrix}.}$$

In other words the columns of $S_{\mathcal{C}\to\mathcal{B}}$ are the $\mathcal{B}$-coordinates of the elements forming the basis $\mathcal{C}$.

Proof of equation (1). Say that $[v]_{\mathcal{C}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ are the $\mathcal{C}$-coordinates of v with respect to the basis $\mathcal{C}$. This means that

$$v = c_1 w_1 + \ldots + c_n w_n.$$

Therefore starting from the right hand side of (1) we find that:

$$S_{\mathcal{C}\to\mathcal{B}}\,[v]_{\mathcal{C}} \;=\; \begin{pmatrix} | & | & \dots & | \\ [w_1]_{\mathcal{B}} & [w_2]_{\mathcal{B}} & \dots & [w_n]_{\mathcal{B}} \\ | & | & \dots & | \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \;=\; c_1[w_1]_{\mathcal{B}} + \ldots + c_n[w_n]_{\mathcal{B}} \;=$$

$$= \; [c_1 w_1 + \ldots + c_n w_n]_{\mathcal{B}} \;=\; [v]_{\mathcal{B}}.$$

□

We can also pass from the $\mathcal{B}$-coordinates of an element v, into the $\mathcal{C}$-coordinates of v. This is achieved by the inverse $S^{-1}_{\mathcal{C}\to\mathcal{B}}$ of $S_{\mathcal{C}\to\mathcal{B}}$. Hence for any element v in V we have

$$\boxed{[v]_{\mathcal{C}} \;=\; S^{-1}_{\mathcal{C}\to\mathcal{B}}\,[v]_{\mathcal{B}}.}$$

2. The matrix associated to a linear function (after having fixed a basis)

After having chosen a basis $\mathcal{B}$ of V, we can represent a linear function $T\colon V \to V$ by a matrix, So let's say that V is a vector space of dimension n, that $T\colon V \to V$ is a linear function, and let's fix a basis $\mathcal{B} = (v_1, \ldots, v_n)$ of V. Then there exists a matrix $C_{\mathcal{B}}$ (depending on the basis $\mathcal{B}$) which transforms the $\mathcal{B}$-coordinates of an element v of V, into the $\mathcal{B}$-coordinates of its image $T(v)$ under T. Hence for any v in V the following relation holds:

$$\boxed{\Big[T(v)\Big]_{\mathcal{B}} \;=\; C_{\mathcal{B}}[v]_{\mathcal{B}}.} \tag{2}$$

The matrix $C_{\mathcal{B}}$ is defined as:

$$\boxed{C_{\mathcal{B}} \;=\; \begin{pmatrix} | & | & \dots & | \\ \Big[T(v_1)\Big]_{\mathcal{B}} & \Big[T(v_2)\Big]_{\mathcal{B}} & \dots & \Big[T(v_n)\Big]_{\mathcal{B}} \\ | & | & \dots & | \end{pmatrix}.}$$

Proof of equation (2). Say that $[v]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ are the $\mathcal{B}$-coordinates of v. This means that $v = c_1 v_1 + \ldots + c_n v_n$. Starting from the right hand side of (2) we find that:

$$C_{\mathcal{B}}[v]_{\mathcal{B}} \;=\; \begin{pmatrix} | & | & \dots & | \\ \Big[T(v_1)\Big]_{\mathcal{B}} & \Big[T(v_2)\Big]_{\mathcal{B}} & \dots & \Big[T(v_n)\Big]_{\mathcal{B}} \\ | & | & \dots & | \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \;=\; c_1\Big[T(v_1)\Big]_{\mathcal{B}} + \ldots + c_n\Big[T(v_n)\Big]_{\mathcal{B}} \;=$$

$$=\; \Big[c_1 T(v_1) + \ldots + c_n T(v_n)\Big]_{\mathcal{B}} \;=\; \Big[T\big(c_1 v_1 + \ldots + c_n v_n\big)\Big]_{\mathcal{B}} \;=\; [T(v)]_{\mathcal{B}}.$$

□

Once we worked out the matrix $C_{\mathcal{B}}$ of a linear function T, it becomes very simple to check whether T is an isomorphism. In fact T is an isomorphism if and only if $\det(C_{\mathcal{B}}) \neq 0$. In addition we can find bases of the kernel and the image of T simply by computing the bases of the kernel and the image of $C_{\mathcal{B}}$. One issue in this context (that we will solve in the next section) is the following: Since we may have many different matrix-representations of a linear function (in fact for every choice of a basis $\mathcal{B}$ we obtain a different matrix $C_{\mathcal{B}}$), why are the information that we get from one representation equal to the information we obtain from a different representation? The point is that even though the matrices that represent a linear function with respect to different basis are not equal, they are however similar matrices. As such they share several important properties. For instance, the dimensions of the kernels and images of two similar matrices are equal, and they have the same determinant. From this we deduce that in order to check if T is an isomorphism, it does not matter which matrix-representation we work with.

3. From one representation of a linear function to one other

Suppose that we have two matrices $C_{\mathcal{B}}$ and $C_{\mathcal{C}}$ that represent a linear function $T\colon V \to V$ with respect to two bases $\mathcal{B}$ and $\mathcal{C}$ of V, respectively. What is the relation between $C_{\mathcal{B}}$ and $C_{\mathcal{C}}$? How much do they differ? Because we have two bases $\mathcal{B}$ (the old one) and $\mathcal{C}$ (the new one) of V, then we also have the matrix of change of basis $S_{\mathcal{C}\to\mathcal{B}}$ from the basis $\mathcal{C}$ to the basis $\mathcal{B}$. The relation between $C_{\mathcal{B}}$ and $C_{\mathcal{C}}$ is the following:

$$\boxed{C_{\mathcal{C}} \;=\; S_{\mathcal{C}\to\mathcal{B}}^{-1}\, C_{\mathcal{B}}\, S_{\mathcal{C}\to\mathcal{B}}.} \tag{3}$$

In other words $C_{\mathcal{B}}$ and $C_{\mathcal{C}}$ are similar matrices, because the matrix $S_{\mathcal{C}\to\mathcal{B}}$ is invertible and the previous equation can be rewritten as

$$S_{\mathcal{C}\to\mathcal{B}}\, C_{\mathcal{C}} \;=\; C_{\mathcal{B}}\, S_{\mathcal{C}\to\mathcal{B}}.$$

The equation (3) can also be rewritten as:

$$\boxed{C_{\mathcal{B}} \;=\; S_{\mathcal{C}\to\mathcal{B}}\, C_{\mathcal{C}}\, S_{\mathcal{C}\to\mathcal{B}}^{-1}.}$$

In conclusion we can calculate a matrix-representation of a linear function starting from any other matrix-representation. Moreover we conclude that:

the matrices that represent a linear function with respect to two different bases are similar.

What properties have similar matrices in common? Here is a list of them.

Theorem 3.1. *Suppose that A and B are similar matrices, so that there exists an invertible matrix S such that $SB = AS$. The the following statements hold.*

(i). $\dim \operatorname{Ker}(A) = \dim \operatorname{Ker}(B)$ *and* $\dim \operatorname{Im}(A) = \dim \operatorname{Im}(B)$.

(ii). $\det(A) = \det(B)$ *and* $\operatorname{tr}(A) = \operatorname{tr}(B)$.

(iii). A and B have the same eigenvalues with the same algebraic and geometric multiplicities.

(iv). A and B have the same characteristic polynomial.

Proof. We prove first the first point. Because A and B are similar matrices we have that $SB = AS$. We define a linear function $T_S\colon \operatorname{Ker}(B) \to \operatorname{Ker}(A)$ as $T(\vec{x}) = S\vec{x}$. This function is well defined because if if $\vec{x}$ belongs to $\operatorname{Ker}(B)$, then $S\vec{x}$ belongs to $\operatorname{Ker}(A)$. In fact

$$\vec{0} = S\vec{0} = S(B\vec{x}) = (SB)\vec{x} = (AS)\vec{x} = A(S\vec{x}).$$

The function T_S is an isomorphism because it admits an inverse, which is defined as $T_S^{-1}(\vec{y}) = S^{-1}\vec{y}$ (you should check that this is the inverse of T_S). Hence $\operatorname{Ker}(A)$ and $\operatorname{Ker}(B)$ are isomorphism vector spaces and in particular they have the same dimension. By the Rank-Nullity Theorem we also deduce that $\dim \operatorname{Im}(A) = \dim \operatorname{Im}(B)$.

Now we prove the fourth point. We have that $B = S^{-1}AS$ as A and B are similar. Recall that Binet's theorem says that $\det(EF) = \det(E)\det(F)$ for any two square matrices E and F of the same size. Hence

$$p_B(\lambda) = \det(B - \lambda I) = \det(S^{-1}AS - \lambda I) = \det(S^{-1}AS - \lambda S^{-1}S) =$$
$$= \det\left(S^{-1}(A - \lambda I)S\right) = \det(S^{-1})\det(A - \lambda I)\det(S) =$$
$$= \det(S)^{-1}\det(A - \lambda I)\det(S) = \det(A - \lambda I) = p_A(\lambda).$$

Points two and three follow directly from the last point because the eigenvalues, the determinant, and the trace of a matrix can be recovered from its characteristic polynomial.

□

4. Examples

Example 1. Suppose that $V = \mathbf{R}^{2\times 2}$ and fix the basis

$$\mathcal{B} = \left(\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right).$$

What are the $\mathcal{B}$-coordinates of $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$? In order to answer to this question, we need to write A as a linear combination of the matrices forming the basis $\mathcal{B}$:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = a\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -a & a+c \\ b+d & -b+c \end{pmatrix}.$$

Therefore $a = -1, b = -1, c = 3, d = 4$ and the $\mathcal{B}$-cooridnates are

$$\left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} -1 \\ -1 \\ 3 \\ 4 \end{pmatrix}.$$

On the other hand, if a matrix B has $\mathcal{B}$-coordinates $\begin{pmatrix} -2 \\ 1 \\ 2 \\ 3 \end{pmatrix}$, then

$$B = -2\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} + 1\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} + 2\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 4 & 1 \end{pmatrix}.$$