Further linear algebra. Chapter III. Revision of linear algebra.

Andrei Yafaev

As in the previous chapter, we consider a field k. Typically, $k = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p$.

A vector space over a field k (one also says a k-vector space) is a set V with two operations: addition and scalar multiplication by elements of k. Elements of V are called vectors, and elements of k are called scalars. The axioms are:

- $(V, +)$ is an abelian (commutative) group (in particular there is an identity for addition called zero 0).
- $(xy)v = x(yv)$ for $x, y \in k$, $v \in V$.
- $(x + y)(v + w) = xv + xw + yv + yw$ for $x, y \in k$, $v, w \in V$.
- $1v = v$.

A typical example of a vector space is the space k^n of n-tuples of elements in k. In particular k itself is a vector space over itself.

Another example is $k[X]$. The set of polynomials with coefficients in k is a vector space. Fix $n \geq 0$ and let $k[X]_n$ be the set of polynomials of degree less or equal to n. This is a vector space . If $n = 0$, then this vector space is just k.

The set of polynomials of degree exactly n is not a vector space. For example because 0 is not there.

Take $k = \mathbb{R}$ and let C be the set of all continuous functions from $[0, 1]$ to $\mathbb{R}$. Then C is an $\mathbb{R}$-vector space.

Similarly, take $k = \mathbb{C}$ and let H be the set of all holomorphic functions from the unit ball to $\mathbb{C}$. This is a $\mathbb{C}$ vector space. Of course it also an $\mathbb{R}$-vector space.

Another example. Let $a, b \in \mathbb{R}$ and consider the set of all twice differentiable functions f such that

$$\frac{d^2 f}{dx} + a\frac{df}{dx} + bf = 0$$

This is an $\mathbb{R}$ vector space (exercise).

Let k be a field, the set of matrices $M_n(k)$ with coefficients in k is a k-vector space.

- A linear combination of $\{v_1, \ldots, v_n\}$ is a vector of the form $x_1 v_1 + \ldots + x_n v_n$.

  For example, consider the vector space $k[x]_n$ as before. This vector space is in fact the set of all linear combinations of the elements $1, x, \ldots, x^n$.

- The span of a set of vectors is the set of linear combinations of those vectors.

  As above, $k[x]_n$ is the span of the set $\{1, x, \ldots, x^n\}$. We say that the vectors $\{1, x, \ldots, x^n\}$ span or generate this vector space.

  Consider k^2 and the vecotrs

  $$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

  and

  $$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

  Then the set of vectors $\{e_1, e_2\}$ spans k^2 for any field k.

- Let V be a k-vector space. A subset A of V is said to generate V is V is the span of A.

  In the examples above $\{1, x, \ldots, x^n\}$ generates $k[x]_n$ but 1, $\{1, x\}$, $\{1, x, x^2\}$, $\{1, x, x^2, \ldots, x^{n-1}\}$ do not generate V.

  The set $\{e_1, e_2\}$ certainly generates $\mathbb{R}^2$ while $\{e_1\}$ or $\{e_2\}$ do not.

- Let V be a k-vector space. A subset $W \subset V$ is called a subspace if any linear combination of elements in W is in W. In other words, a subspace is a subset which is a vector space with the same addition and scalar multiplication as V.

  Let V be a k-vector space and take $v \in V$. The set kv of all multiples of v by elements of k is a subspace. More generally, take any set $A \subset V$, then the set of linear combinations of elements of A is a vector subspace.

  As an (easy) exercise, prove that given any collection W_i, $i \in I$ (I is some set, finite or infinite) of subspaces of V, the intersection $\cap_{i \in I} V_i$ is a subspace. The union is not ! For example, let $V = k^2$ and $W_1 = ke_1$ and $e_2 = ke_2$. It is quite clear that $W_1 \cup W_2$ not a subspace, for example $e_1 + e_2$ is not in it.

  Let C be the set of continuous functions $[0,1] \longrightarrow \mathbb{R}$. We have seen that this is an $\mathbb{R}$-vector space. Let $W = \{f \in C : f(0) = 0\}$. This is a subspace (easy exercise).

  We have seen that $k[x]$ is a vector space. The space $k[x]_n$ is a vector subspace.

- Vectors $v_1, \ldots, v_n \in V$ are called linearly independent if whenever $\sum_{i=1}^n \lambda_i v_i = 0$ (for some $\lambda_i \in \mathbb{R}$), then $\lambda_i = 0$ for all i.

  For example, in k^2, vetors e_1, e_2 are linearly independent. Clearly e_1 and $2e_1$ are not linearly independent. If v is any vector, v and 0 are obviously never linearly independent.

  If $k = \mathbb{R}$ then the vectors e_1 and $2e_2$ are linearly independent. Not if $k = \mathbb{F}_2$, indeed in this case $2e_2 = 0$.

  In $k[x]$, the vectors $\{1, x, x^2, \ldots,$ are linearly independent.

- A set $\{v_1, \ldots, v_n\}$ of vectors is a basis for V if it is linearly independent and it's span is V (it generates V). Such a set is called a *basis*. If this is the case then every vector has a unique expression as a linear combination of $\{v_1, \ldots, v_n\}$.

  For example $\{e_1, e_2\}$ is a basis of k^2. The set $\{1, x, x^2, \ldots, x^n\}$ is basis of $k[x]_n$.

  The set $\{1, x, x^2, \ldots,$ is a basis of $k[x]$.

  A basis of the space $M_n(k)$ consists of the matrices $E_{i,j}$ which have 1 at the position (i, j) and zero elsewhere.

- Any vector space has a basis. If it has a basis consisting of finitely many elements, it is called finite dimensional.

  The dimension of a vector space is the number of vectors in a basis. This does not depend on the basis: any two bases have the same number of elements.

  For example k^n has dimension n, $k[x]_n$ has dimension $n+1$ while $k[x]$ is infinite dimensional and so is C.

  $\mathbb{R}$ viewed as vactor space over itself has dimension 1 but viewed as vector space over $\mathbb{Q}$ is infinite dimensional.

  $\mathbb{C}$ viewed as a vector space over itself has dimension 1 but as a vector space over $\mathbb{R}$ it has dimension 2 : a basis is $\{1, i\}$.

  The space $M_n(k)$ has dimension n^2.

- The fundamental properties are the following.

  Suppose that V is a finite dimensional vector space.

  1. Let $\{v_1, \cdots, v_n\}$ be a *linearly independent* set of vectors. Then one can find $v_{n+1}, \ldots, v_r\}$ such that $\{v_1, \ldots, v_r\}$ is a basis of V.

     In particular, if $W \subset V$ is a subspace, then any basis of W can be completed to a basis of V.

  2. Let $\{w_1 \ldots, w_r\}$ be a family of vectors spanning V, then one can extract a basis from $\{w_1 \ldots, w_r\}$

- Direct sums.

  Let V be a vector space and U, W two subspaces. The sum $U + W = \{u + w : u \in U, w \in W\}$ and the intersection $U \cap W$ are subspaces of V. The sum $U + W$ is called *direct* if $U \cap W = \{0\}$, the notation is $U \oplus W$. One has $\dim(U \oplus W) = \dim(U) + \dim(V)$.

  For example, take k^n with its standard basis, then k^n is the direct sum

  $$k^n = (ke_1) \oplus \cdots \oplus (ke_n)$$

  Consider in $V = k^2$ the vectors $v_1 = e_1 + e_2$ and $v_2 = e_1 - e_2$. If $k = \mathbb{R}$, then $k^2 = Span(v_1) \oplus Span(v_2)$.

Indeed, let $v = \alpha e_1 + \beta e_2 \in V$, then

$$v = \frac{\alpha}{2}(v_1 + v_2) + \frac{\beta}{2}(v_1 - v_2)$$

hence $Span(v_1) + Span(v_2) = V$.

Now suppose $v \in Span(v_1) \cap Span(v_2)$. Then $v = \lambda v_1 = \mu v_2$. As e_1 and e_2 are linearly independent, we get $\lambda = -\mu = \mu$ hence, because $k = \mathbb{R}$, $\mu = 0$, therefore $v = 0v_2 = 0$ hence $Span(v_1) \cap Span(v_2) = \{0\}$ and the sum is direct.

If $k = \mathbb{F}_2$, then it fails as in this case $v_1 = v_2$.

Let V, W be vector spaces. A function $T : V \to W$ is a linear map if

- $T(v + w) = T(v) + T(w)$,
- $T(xv) = xT(v)$.

or equivalently, $T(v + xw) = T(v) + xT(w)$. A bijective linear map is called an isomorphism of vector spaces.

For example the map $T \colon \mathbb{C} \longrightarrow \mathbb{C}$ that sends z to $\overline{z}$ is not a linear map of $\mathbb{C}$-vector spaces : $T(\lambda z) = \overline{\lambda} T(z)$! But it is a map of real vector spaces : if $\lambda \in \mathbb{R}$, then $\overline{\lambda} = \lambda$.

Similarly, the map $k \longrightarrow k$ sending $x \mapsto x^2$ is not linear when $k = \mathbb{R}$ or $k = \mathbb{C}$, but it is linear if $k = \mathbb{F}_2$.

The map $tr \colon M_n(k) \longrightarrow k$ sending a matrix M to $\sum_{i,j} M_{i,j}$ is linear.

If T is linear, we define its kernel and image:

$$\ker(T) = \{v \in V : T(v) = 0\},$$

$$im(T) = \{T(v) : v \in V\}.$$

The rank of T is the dimension of the image of T, and the nullity of T is the dimension of the kernel of T.

This implies the following:

Theorem 0.1 (Rank-Nullity Theorem) *Let $T : V \to W$ be a linear map. Then*

$$rank(T) + nullity(T) = \dim V.$$

Proof. Let $\{v_1, \ldots, v_r\}$ be a basis of $\ker(T)$ and $\{w_1, \ldots, w_s\}$ be a basis of $im(T)$. By definition of the image, there exist $\{u_1, \ldots, u_s\}$ vectors of V such that

$$T(u_i) = w_i$$

We claim that $\{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ form a basis of V which will conclude the proof.

First we show linear independence. Suppose that

$$a_1 v_1 + \cdots + a_s v_r + b_1 u_1 + \cdots + b_r u_s = 0$$

Apply T, we get

$$0 = T(0) = b_1 T(u_1) + \cdots + b_r T(u_s) = b_1 w_1 + \cdots + b_s w_s$$

(note that $a_1 v_1 + \cdots + a_s v_r = 0$ because the v_is are in the kernel of T). Now, as $\{w_1, \ldots, w_s\}$ is a basis of $im(T)$ (in particular it is linearly independent), we get that $b_i = 0$ for all i.

So we have $a_1 v_1 + \cdots + a_s v_r = 0$ and, because v_is for a basis of $\ker(T)$ (and in particular are linearly independent), we get that $a_i = 0$ for all i.

We have shown that a_is and b_is are all zero hence $\{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ is linearly independent.

It remains to show that $\{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ spans V.

Let $x \in V$. By the choice of $\{w_1, \ldots, w_s\}$ as a basis of the image of T, we have

$$T(x) = \sum_{i=1}^{s} a_i w_i = \sum_{i=1}^{s} a_i T(u_i) = T(\sum_{i=1}^{s} a_i u_i)$$

Therefore

$$T(x - \sum_{i=1}^{s} a_i u_i) = 0$$

and hence

$$x - \sum_{i=1}^{s} a_i u_i \in \ker(T)$$

and now, by the choice of $\{v_i\}$ as basis of $\ker(T)$, there exist b_is such that

$$x = \sum_{i=1}^{s} a_i u_i + \sum_{j=1}^{r} b_i v_i$$

which shows that $\{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ generates V.

This finishes the proof. $\square$

Here are some consequences of this theorem.

Definition 0.1 *A linear map $T\colon V \longrightarrow W$ is called* ***isomorphism*** *if there exists*

1. $T_1\colon W \longrightarrow V$ *such that* $TT_1 = I_V$ *(identity of* W*)*
2. $T_2\colon W \longrightarrow V$ *such that* $T_2T = I_W$ *(identity of* V*)*

In particular, a linear map $T^{-1}\colon V \longrightarrow V$ *is an isomorphism if there exists* T^{-1} *such that* $T^{-1}T$ *is the identity.*

It is easy (and left as exercise) to see that $T\colon V \longrightarrow W$ is an isomorphism if and only if T is both surjective and injective. (for the converse you will need to constrict T_1 and T_2 as maps and then show that they are linear.)

If $T\colon V \longrightarrow V$ is an isomorphism, one also says that T is invertible.

Corollary 0.2 *Let* $T\colon V \longrightarrow W$ *be a linear map* ***with*** $\dim V = \dim W$*. If* T *is injective, then* T *is an isomorphism. If* T *is surjective, then* T *is an isomorphism.*

Proof. If T is injective, then $\ker(T) = \{0\}$. By the above theorem, $\dim(im(T)) = \dim(V) = \dim(W)$ and hence $im(T) = W$ and T is surjective. Injective + Surjective = Isomorphism.

Similarly, if T is surjective, then $\dim(im(T)) = \dim(W) = \dim(V)$ and hence $\dim(\ker(T)) = 0$. It follows that T is injective. Injective + Surjective = Isomorphism. □

Corollary 0.3 *Let* V *and* W *be two vector spaces of same dimension. Then* V *is isomorphic to* W *(i.e there is an isomorphism between* V *and* W*).*

Proof. Let $r = \dim(V) = \dim(W)$ and let $\{v_1, \dots, v_r\}$ be a basis of V and $\{w_1, \dots, w_r\}$ be a basis of W. Define T by $T(v_i) = w_i$. By construction, T is surjective and by the theorem it's also injective hence an isomorphism. □

If $T\colon V \longrightarrow W$ and $T'\colon W \longrightarrow U$, then we denote by TT' the composition $T \circ T'\colon V \longrightarrow U$.

If $T\colon V \longrightarrow V$ and n is an integer, we write T^n for T composed with itself n times.

1 Matrix representation of linear maps.

Let V and W be two finite dimensional vector space over a field k. Suppose that V is of dimension r and W is of dimension t.

Let $B = \{b_1, \ldots, b_r\}$ be a basis for V and $B' = \{b'_1, \ldots, b'_s\}$ be a basis for W.

For any vector $v \in V$ we shall write $[v]_B$ (in the future, we will by abuse of notation simply call this column vector v when it is obvious which basis we are referring to) for the column vector of coefficients of v with respect to the basis B, i.e.

$$[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}, \quad v = x_1 b_1 + \ldots + x_r b_r.$$

Given a linear map $T : V \to W$ we have

$$T(v) = T(\sum_{i=1}^{r} x_i b_i) = \sum_{i=1}^{r} x_i T(b_i)$$

Now we have

$$T(b_i) = \sum_{j=1}^{s} a_{ji} b'_j$$

We get :

$$T(v) = \sum_{1 \le i \le r, 1 \le j \le s} x_i a_{ji} b'_j$$

In other words it is the $s \times r$ matrix , usually denoted by $[T]_{\mathcal{B},\mathcal{B}'}$, with entries

$$([T]_{\mathcal{B},\mathcal{B}'})_{i,j} = a_{ji}$$

In practice, to write a matrix of T with respect to given bases, decompose $T(b_i)$ in the basis of W and write column vectors, this gives the matrix $[T]$

The matrix $[T]_{\mathcal{B},\mathcal{B}'}$ is called the matrix of T with respect to bases $\mathcal{B}$ and $\mathcal{B}'$.

A LINEAR TRANSFORMATION IS THE MATRIX WITH RESPECT TO SPECIFIED BASES OF THE SOURCE AND THE TARGET SPACES.

Example. $V = W = \mathbb{R}^3$ with canonical bases $\{e_1, e_2, e_3\}$.

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

(notice that this is the projection onto the plane $z = 0$).
One finds

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

One can find the kernel and the image. In this case, clearly the image is the span of e_1 and e_2 hence $\dim(\mathrm{im}T) = 2$. By rank-nullity theorem, $\dim\ker(T) = 1$ and quite clearly it is generated by e_3.

Let us look at $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$T = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + 3z \\ 4x - 2y + 6z \\ -6x + 3y - 9z \end{pmatrix}$$

One finds :

$$\begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & -3 & -9 \end{pmatrix}$$

Quite clearly, the first column vector in this matrix is -2 times the second and the third is the first minus the second, therefore $\mathrm{im}(T)$ is one dimensional and spanned by $\begin{pmatrix} 2 \\ 4 \\ -6 \end{pmatrix}$ The rank-nullity theorem implies that the dimension of $\ker(T)$ is 2. To find $\ker(T)$ one needs to solve $[T]v = 0$. By elimination, one easily shows that the kernel has equation $2x - y + 3 = 0$, hence can be spanned by the vectors $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$

Another example : $k[x]_n \longrightarrow k[x]_{n-1}$ sending f to its derivative f'. Quite clearly it's a linear map. Find its matrix, image and kernel.

Same question with $k[x]_n \longrightarrow k[x]_n$ sending f to $f + f'$.

Let us consider the transformation $T\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$T = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \end{pmatrix}$$

and $B_1 = B_2 = \{v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_1 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}\}$.

One calculates $T(v_1) = v_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix} = -6v_1 - 2v_2$ and $T(v_2) = v_1 = \begin{pmatrix} -3 \\ 2 \end{pmatrix} = 17v_1 + 6v_2$.

The matrix of T with respect to these bases is

$$\begin{pmatrix} -6 & 17 \\ -2 & 6 \end{pmatrix}$$

In the canonical bases, of course the matrix is:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Ex. Let $T\colon M_2(k) \longrightarrow M_2(k)$ sending M to M^t. Write down the matrix of this linear map in the standard bases.

When one changes the bases, the matrix gets multiplied on the left and on the right by appropriate 'base change' matrices.

More precisely, let $T\colon V \longrightarrow W$ be a linear map. Let $B_1 = \{v_1, \ldots, v_r\}$ be a basis for V and let $B_1' = \{v_1', \ldots, v_r'\}$ be another basis for V.

Similarly let $B_2 = \{w_1, \ldots, w_s\}$ be a basis for W and let $B_2' = \{w_1', \ldots, w_s'\}$ be another basis for W.

Let $[T]_{B_1,B_2}$ be the matrix of T in the bases B_1 and B_2.

Let A be the matrix whose jth column consists of coordinates of v_j' in the basis B_1. Then A is called the transition matrix from B_1 to B_1'. The matrix A is invertible and A^{-1} is the transition matrix from B_1' to B_1.

Similarly, let B be the matrix whose jth column consists of coordinates of w_j' in the basis B_2.

Then
$$[T]_{B_1',B_2'} = B^{-1}[T]_{B_1,B_2}A$$

In the particular case where $V = W$ and $B_1 = B_2$ and $B_1' = B_2'$, we get

$$[T]_{B_1',B_1'} = A^{-1}[T]_{B_1,B_1}A$$

Example.

$$T = \begin{pmatrix} x - 2y \\ 2x + y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \end{pmatrix}$$

and $B_1 = B_2 = \{v_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, v_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}\}$.

One calculates $T(v_1) = v_1 = \begin{pmatrix} 5 \\ 0 \end{pmatrix} = \frac{5}{4}(v_1 + v_2)$

and $T(v_2) = v_1 = \begin{pmatrix} -1 \\ 8 \end{pmatrix} = -\frac{13}{4}v_1 + \frac{3}{4}v_2$.

The matrix of T with respect to these bases is

$$[T]_{B_1,B_2} = \begin{pmatrix} 5/4 & -13/4 \\ 5/4 & 3/4 \end{pmatrix}$$

In the canonical basis B the matrix is

$$[T]_{B,B} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

The transition matrix from B to B_1 is

$$A = \begin{pmatrix} 1 & 3 \\ -2 & 2 \end{pmatrix}$$

And the transition matrix from B_1 to B is

$$A^{-1} = \frac{1}{8}\begin{pmatrix} 2 & -3 \\ 2 & 1 \end{pmatrix}$$

One easily checks that

$$[T]_{B_1,B_2} = A^{-1}[T]_{B,B}A$$

We also have the following :

Proposition 1.1 *Let $T_1 \colon V \longrightarrow W$ and $T_2 \colon W \longrightarrow U$ be two linear maps and suppose we are given bases B, B_1, B_2 of the vector spaces V, W, U. Then for the composed map $T_2 \circ T_1$ (usually simply denoted by T_2T_1) the matrix is*

$$[T_2T_1]_{B,B_2} = [T_2]_{B_1,B_2}[T_1]_{B,B_1}$$

In particular if $T \colon V \longrightarrow V$ (such a map is called ***endomorphism****) and B is a basis for V, then*

$$[T^n]_B = [T]_B^n$$

and, if we suppose that T is an isomorphism

$$[T^{-1}]_B = [T]_B^{-1}$$

In particular, if B and B' are two bases of V, then the transition matrix of B to B' is the inverse of the transition matrix of B' to B.