Lecture 16

Andrei Antonenko

March 10, 2003

1 Image and kernel

Last lecture we studied image and kernel of a linear function. Now we will prove one of the properties of image and kernel. First let's consider kernel.

Let $f : V \to U$ be a linear function, and let its kernel be $\operatorname{Ker} f$ — set of all elements v from V which map to $\mathbf{0}$. Then we can state the following properties of it.

Existence of zero. The zero vector $\mathbf{0}$ belongs to kernel of f, since $f(\mathbf{0}) = \mathbf{0}$ — maps to $\mathbf{0}$, so $\mathbf{0}$ is in kernel.

Summation. Let vectors v and u belong to kernel, so, $f(v) = \mathbf{0}$ and $f(u) = \mathbf{0}$. Then

$$f(v + u) = f(v) + f(u) = \mathbf{0},$$

and thus $u + v$ belongs to $\operatorname{Ker} f$.

Multiplication by a scalar. Let vector v belongs to the kernel of f. Then we know that $f(v) = \mathbf{0}$. Now for any constant k we have:

$$f(kv) = kf(v) = k \cdot \mathbf{0} = \mathbf{0},$$

thus kv belongs to $\operatorname{Ker} f$.

So, we proved the following theorem:

Theorem 1.1. *The kernel of linear function $f : V \to U$ is a vector subspace in V.*

Example 1.2. *Consider the projection function $f(x, y, z) = (x, y, 0)$. It's kernel consists of vectors of the form $(0, 0, c)$ for any constant c. Geometrically speaking, this is a z-axis in the 3-dimensional space. This is a vector subspace.*

Now let's consider the image. Let $f : V \to U$ be a linear function, and it's image $\operatorname{Im} f$ is the set of all vectors from U where we can get by applying a function to vectors from V. We'll state some properties of it.

Existence of zero. The zero vector is in $\operatorname{Im} f$ since by taking $f(\mathbf{0})$ we can get to $\mathbf{0}$: $f(\mathbf{0}) = \mathbf{0}$.

Addition. Let u_1 and u_2 be elements from the image of f, so there exist v_1 and v_2 from V such that $f(v_1) = u_1$ and $f(v_2) = u_2$. Now we can consider the element $v_1 + v_2$ from V. We have:

$$f(v_1 + v_2) = f(v_1) + f(v_2) = u_1 + u_2,$$

and thus $u_1 + u_2$ belongs to $\operatorname{Im} f$.

Multiplication by a scalar. Let u be a vector from $\operatorname{Im} f$. Then there exists a vector v from V such that $f(v) = u$. So, let's consider an element kv for any constant k. We have:

$$f(kv) = kf(v) = ku,$$

thus ku belongs to $\operatorname{Im} f$.

As for the kernel, we proved the following theorem:

Theorem 1.3. *The image of a linear function $f : V \to U$ is a vector subspace in U.*

Example 1.4. *Consider the projection function $f(x, y, z) = (x, y, 0)$. It's image consists of vectors of the form $(x, y, 0)$ for all $x, y \in \mathbb{R}$. Geometrically speaking, this is an xy-plane in the 3-dimensional space. This is a vector subspace.*

In order to continue studies of image and kernel, we would like to know more about linear functions.

2 Matrix of a linear function

When we studied linear function for the first time we considered the following example. If A is an $m \times n$ matrix, then we can define a linear function $F_A : \mathbb{R}^n \to \mathbb{R}^m$ by the following formula: $F_A(x) = Ax$ for any vector $x \in \mathbb{R}^n$. In this part we can see, that it is one of the general cases of linear functions.

Let's consider any linear function $f : V \to W$. Let vectors $e_1, e_2, \dots, e_n$ form a basis in the space V. And let we know the values $f(e_1), f(e_2), \dots, f(e_n)$. Then we can compute the function f for any vector from V using only these given values. To show it let's note, that is e_i's form a basis, then any vector v from V can be represented as a linear combination of them:

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n.$$

Now let's show how to compute the value $f(v)$:

$$\begin{aligned} f(v) &= f(a_1e_1 + a_2e_2 + \cdots + a_ne_n) \\ &= f(a_1e_1) + f(a_2e_2) + \cdots + f(a_ne_n) \\ &= a_1f(e_1) + a_2f(e_2) + \cdots + a_nf(e_n). \end{aligned}$$

So as we stated, function f can be computed for any vector v is we know its values on basis vectors.

Example 2.1. *Let's consider* $\mathbb{R}^3$ *and a standard basis* $e_1 = (1,0,0), e_2 = (0,1,0)$*, and* $e_3 = (0,0,1)$*. For projection function* $f(x,y,z) = (x,y,0)$ *we have:*

$$\begin{aligned} f(1,0,0) &= (1,0,0) \\ f(0,1,0) &= (0,1,0) \\ f(0,0,1) &= (0,0,0) \end{aligned}$$

So, for any vector $v = (a,b,c) = a(1,0,0) + b(0,1,0) + c(0,0,1)$ *one can compute* f*:*

$$\begin{aligned} f(v) &= af(1,0,0) + bf(0,1,0) + cf(0,0,1) \\ &= a(1,0,0) + b(0,1,0) + c(0,0,0) \\ &= (a,b,0). \end{aligned}$$

Now we'll make a trick: we'll write the coordinates of $f(e_1), f(e_2), \ldots, f(e_n)$ as columns of matrix:

$$A_f = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ f(e_1) & f(e_2) & \ldots & f(e_n) \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

This matrix is called a **matrix of a linear function**. Now we can compute the value of $f(x)$ for any vector x by writing coordinates of x in column, and multiplying the matrix A_f by vector-column x.

Example 2.2. *For projection function* $f(x,y,z) = (x,y,0)$ *the matrix is*

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

For example, to compute $f(1,2,3)$ *we can multiply* A_f *by* $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$*:*

$$A_f x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

Example 2.3. *Now let's consider the function of taking a derivative in the space* $\mathbb{P}_2$*:* $D(at^2 + bt + c) = 2at + b$*. Let's take the standard basis in the space of polynomials* $\mathbb{P}_2$ *and compute values of function on basis vectors:*

- $e_1 = t^2$*, and* $D(e_1) = D(t^2) = 2t$*. The coordinates are* $(0, 2, 0)$ *since it is equal to* $0t^2 + 2t + 0$*.*
- $e_2 = t$*, and* $D(e_2) = D(t) = 1$*. The coordinates are* $(0, 0, 1)$ *since it is equal to* $0t^2 + 0t + 1$*.*
- $e_3 = 1$*, and* $D(e_3) = D(1) = 0$*. The coordinates are* $(0, 0, 0)$ *since it is equal to* $0t^2 + 0t + 0$*.*

So, the matrix is

$$A_D = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

For example, let's take a derivative of $3t^2 + 5t + 7$*. We'll write this polynomial as a column-vector* $\begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}$*, and multiply* A_D *by it:*

$$\begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 5 \end{pmatrix}$$

So, the derivative of this polynomial $6t + 5$*.*

2.1 Proof

Let's prove that if a matrix A_f is constructed using the method provided here, then

$$f(x) = A_f x.$$

Proof. Let's take any vector $x = (x_1, x_2, \ldots, x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$. Since we have that

$$f(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj})$$

— j-th column of the matrix A, then

$$\begin{aligned} f(x) &= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_n f(e_n) \\ &= x_1(a_{11}, a_{21}, \ldots, a_{m1}) + \cdots + x_n(a_{1n}, a_{2n}, \ldots, a_{mn}) \\ &= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n) \\ &= (\sum_j a_{1j}x_j, \sum_j a_{2j}x_j, \ldots, \sum_j a_{mj}x_j). \end{aligned}$$

Comparing this with the formal definition of matrix multiplication, we get that

$$f(x) = A_f x.$$

□