LECTURE 1
FUNCTIONS OF AN OPERATOR ARGUMENT

A.A.KIRILLOV

1. Introduction

Our goal is to make sense to the expressions like

$$e^A, \qquad \sin A, \qquad \frac{A}{\sqrt{1+A^2}}, \qquad \text{or, in general,} \quad f(A)$$

where A is not a number but a linear operator in some vector space V. We shall see that it can be done only under some restrictions on the function f and the operator A in question.

2. Functions of a matrix argument

2.1. **Polynomial functions.** We start with functions of the simplest kind – polynomial functions. Let $P(x) = p_0x^n + \cdots + p_{n-1}x + p_n$ be a polynomial with real or complex coefficients and $A = ||a_{i,j}||$ be a matrix of size N. Then it is rather clear that the expression $P(A)$ must be understood as the matrix

$$P(A) = p_0 \cdot A^n + \cdots + p_{n-1} \cdot A + p_n \cdot \mathbf{1}$$

where $\mathbf{1}$ denotes the unit matrix of size N.

Remark 1. *One can ask, why we should write the last summand in this form. The most natural answer is that only under this agreement we ensure the map $P \mapsto P(A)$ to be a homomorphism of the polynomial algebra to the number field, i.e. the following equalities hold:*

$$(P_1 + P_2)(A) = P_1(A) + P_2(A), \quad (\lambda \cdot P)(A) = \lambda \cdot P(A),$$
$$(P_1 \cdot P_2)(A) = P_1(A) \cdot P_2(A)$$

2.2. **Rational functions.** A rational function is by definition a ratio of two polynomial functions: $R(x) = \frac{P(x)}{Q(x)}$. So, we can try to define the quantity $R(A)$ as $R(A) = \frac{P(A)}{Q(A)}$. But there is two delicate points. First, $Q(A)$ must be invertible matrix; second, the matrix multiplication is not commutative, so we must choose between $P(A) \cdot Q(A)^{-1}$ and $Q(A)^{-1} \cdot P(A)$.

Actually, the second obstacle is inessential, because for a given matrix A all matrices of the form $P(A)$ pairwise commute. So, $P(A) \cdot Q(A) = Q(A) \cdot P(A)$, hence, $Q(A)^{-1} \cdot P(A) = P(A) \cdot Q(A)^{-1}$.

Date: Sept 2007.

Consider the first point. For a given A we can define $R(A)$ only for those $R = \frac{P(x)}{Q(x)}$, for which $Q(A)$ is invertible. Recall that a number $\lambda \in \mathbb{C}$ is called an **eigenvalue** of a matrix A if $\det(A - \lambda \cdot \mathbf{1}) = 0$, i.e. when $(A - \lambda \cdot \mathbf{1})$ is not invertible. Collection of all eigenvalues of A is called **spectrum** of A. We denote it by $\operatorname{Spec} A$

Proposition 1. *The matrix $Q(A)$ is invertible if and only if $Q(\lambda) \neq 0$ for every eigenvalue $\lambda \in \operatorname{Spec} A$.*

Indeed, the polynomial Q can be written as a product of linear factors:

$$Q(x) = c(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_N) \tag{1}$$

where λ_1, λ_2, λ_N are roots of Q (taken with multiplicities). Therefore,

$$Q(A) = c(A - \lambda_1)(A - \lambda_2) \cdots (A - \lambda_N) \quad \text{and}$$
$$Q(A)^{-1} = c^{-1}(A - \lambda_1 \cdot \mathbf{1})^{-1}(A - \lambda_2 \cdot \mathbf{1})^{-1} \cdots (A - \lambda_N \cdot \mathbf{1})^{-1}.$$

So, $Q(A)$ is invertible when all matrices $(A - \lambda_i \cdot \mathbf{1})$ are. But this is the case when no roots of Q belong to $\operatorname{Spec} A$.

2.3. **General functions.** The statement of proposition 1 suggests that properties of the matrix $f(A)$ depend on behavior of the function f on the spectrum of A. It is not completely true, but becomes true if we replace the set $\operatorname{Spec} A$ by its **infinitesimal neighborhood**. To explain this, we start with two simple examples.

Example 1. *Assume that our matrix A is diagonal, or, more generally, can be reduced to the diagonal form by the transformation $A \mapsto SAS^{-1}$ with some invertible matrix S.*[1] *Then for any polynomial function F (hence, for any admissible rational function F) we have*

$$F\left(\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_N \end{pmatrix}\right) = \begin{pmatrix} F(a_1) & 0 & \dots & 0 \\ 0 & F(a_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F(a_N) \end{pmatrix} \tag{2}$$

This equality confirms the suggestion above and can be easily extended to all functions F of a complex variable z.

But the life is not so simple.

Example 2. *Suppose that our matrix A can not be reduced to the diagonal form. The simplest example of such matrix is the so-called Jordan block*

[1]The "raison d'être" of the notion of a matrix is that matrices can be used as an algebraic counterpart of the geometric notion of a linear operator. But to associate a matrix to a given operator we have to choose a basis. If we change the basis, the matrix also changes according to the rule $A \mapsto SAS^{-1}$ where S is the matrix of transition from one basis to another. Thus, the similar matrices A and SAS^{-1} are just two "portraits" of the same operator and their properties are completely parallel.

$J_N(\lambda)$ of size N with an eigenvalue λ given by

$$J_N(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \tag{3}$$

The direct computation[2] shows that the answer has the following beautiful form:

$$F\big(J_N(\lambda)\big) = \begin{pmatrix} F(\lambda) & F'(\lambda) & \frac{1}{2}F''(\lambda) & \dots & \frac{1}{(N-1)!}F^{(N-1)}(\lambda) \\ 0 & F(\lambda) & F'(\lambda) & \dots & \frac{1}{(N-2)!}F^{(N-2)}(\lambda) \\ 0 & 0 & F(\lambda) & \dots & \frac{1}{(N-3)!}F^{(N-3)}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & F'(\lambda) \\ 0 & 0 & 0 & \dots & F(\lambda) \end{pmatrix} \tag{4}$$

We see, that the result depends not only on the values of F on the spectrum of A but also on the values of the first $N-1$ derivatives of F at the points of $\operatorname{Spec} A$. This is the exact meaning of the expression: "$F(A)$ depends on the values of F on $U_{N-1}(\operatorname{Spec} A)$ – the infinitesimal neighborhood of $\operatorname{Spec} A$ of order $N-1$".

Introduce the notation

$$F_1 \overset{N}{\sim} F_2 \quad \text{on the set X} \tag{5}$$

which means that the difference $F_1 - F_2$ vanishes at all points $x \in X$ with multiplicity $\geq N$. We can also express this fact, saying that F_1 and F_2 coincide on $U_N(X)$.

Note, that for a polynomial F the matrix $F\big(J_N(\lambda)\big)$ can be non-zero even when λ is a root of F. To have a zero value at $J_N(\lambda)$, F must have λ as a root of multiplicity at least N. In this case we say that F vanishes on the infinitesimal neighborhood $U_{N-1}(\lambda)$.

Now we can discuss the case of a general matrix A. It is known that any matrix is similar to the direct sum of Jordan blocks of arbitrary sizes and eigenvalues. In other words, A is similar to a block-diagonal matrices with K blocks of the form

$$J_{N_k}(\lambda_k), \quad k = 1, 2, \dots, K, \quad \text{where} \quad \sum_{k=1}^{K} N_k = N. \tag{6}$$

---

[2]There are several ways to make this computation in a "smart" way, practically with no computations at all. But it is very instructive to make it directly at least for a momomial $F(x) = x^n$.

Then for any polynomial function P the value $P(A)$ is similar to the direct sum of K blocks of the form

$$P\left(J_{N_k}(\lambda_k)\right), \quad k = 1, 2, \ldots, K. \tag{7}$$

These matrices depend only on the values of P on the infinitesimal neighborhood of Spec$A$ of order $\widetilde{N}-1$, where

$$\widetilde{N} := \max_{1\leq k\leq K} N_k. \tag{8}$$

Exercise 1. *Show that for any finite set $X \subset \mathbb{R}$, any $N \in \mathbb{N}$ and any smooth function F on $\mathbb{R}$ there exist a polynomial P such that $P \overset{N}{\sim} F$ on X.*

Let now F be any smooth function on $\mathbb{R}$. Choose a polynomial function P satisfying $P \overset{\widetilde{N}-1}{\sim} F$ on $\operatorname{Spec} A$. Then the value $P(A)$ is determined uniquely and does not depend on the choice of P.

So, we put by definition

$$F(A) := P(A) \quad \text{for any polynomial } P \text{ satisfying } P \overset{\widetilde{N}-1}{\sim} F \text{ on } \operatorname{Spec} A; \tag{9}$$

Exercise 2. *Let* $A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$. *Compute the following matrices*

a) $\sin A$; *b)* $|A|$; *c)* e^A; *d)* $\log A$ *where*

$$\log(re^{i\theta}) := \log r + i\theta \qquad \text{for } r > 0 \text{ and } -\pi < \theta < \pi;$$

e) $\log A$ *where*

$$\log(re^{i\theta}) := \log r + i\theta \qquad \text{for } r > 0 \text{ and } 0 < \theta < 2\pi.$$

Department of Mathematics, The University of Pennsylvania, Philadelphia, PA 19104-6395, USA *E-mail addresses: kirillov@math.upenn.edu*