Lectures 2, September 15

LAGRANGE INTERPOLATION POLYNOMIALS

AND LINEAR ALGEBRA

Let T be a monic polynomial of degree n over an arbitrary field K. Suppose that the polynomial T has exactly n different roots

$$\lambda_1, \ldots, \lambda_n.$$

With every root λ_i, we associate the polynomial

$$T_i(t) = \frac{T(t)}{T'(\lambda_i)(t - \lambda_i)}.$$

Lemma 1. *The polynomial T_i is the unique polynomial of degree at most $n-1$ that is equal to one at the root λ_i and to zero at all other roots of the polynomial T.*

Proof. Since λ_i is a root of T the polynomial T is divisible by $t - \lambda_i$, and the value of the ratio at the point λ_i by definition is equal to $T'(\lambda_i)$. The ratio obviously vanishes at any other root λ_j

If there is two polynomials of degree $< n$ which coincide in n points then they are identically equal. Indeed their difference is a polynomial of degree $< n$ which has n roots, thus it is identically equal to zero. □

Let $c_1, \ldots, c_n$ be any collection of elements of the field K.

Definition 1. *The polynomial $L(t)$ of degree $< n$ is called the Lagrange interpolating polynomial with the interpolation points λ_1, $\ldots$, λ_n and the interpolation data c_1, $\ldots$, c_n if the following identities hold:*

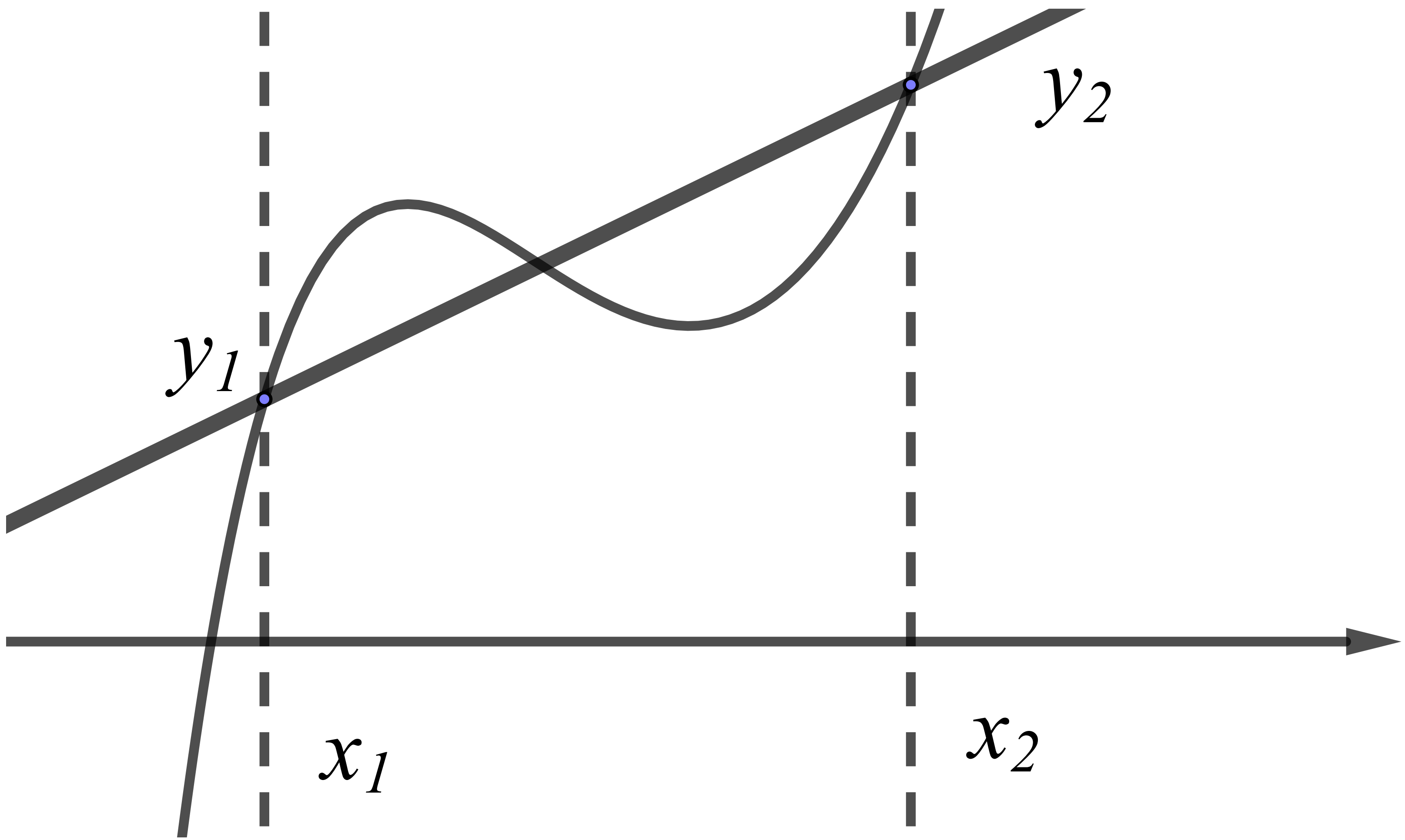
$$L(\lambda_1 = c_1, \ldots, L(\lambda_n) = c_n.$$

Lemma 2. *The Lagrange interpolating polynomia with the interpolation points $\lambda_1, \ldots, \lambda_n$ and the interpolation data $c_1, \ldots, c_n$ is given by the following formula:*

$$L(t) = \sum c_i T_i(t).$$

Proof. By Lemma 1 the above polynomial satisfies the needed identities. Such polynomial is unique since any polynomial of degree $< n$ which has at least n roots is identically equal to zero.

This is the unique polynomial of degree at most n that takes the value c_i at every point λ_i, $i = 1, \ldots, n$. $\square$


y_1
y_2
x_1
x_2

Consider a vector space V (possibly, infinite dimensional) over the field K and a linear operator

$$A : V \to V.$$

Suppose that the operator A satisfies a polynomial equation

$$T(A) = A^n + a_1 A^{n-1} + \cdots + a_{n-1} A + a_n E = 0,$$

where $a_i \in K$, and E is the identity operator.

Assume that the polynomial

$$T(t) = t^n + a_1 t^{n-1} + \cdots + a_n$$

has n different roots

$$\lambda_1, \ldots, \lambda_n$$

in the field K.

Definition 2. *The operator $L_i = T_i(A)$, where*

$$T_i(t) = \frac{T(t)}{T'(\lambda_i)(t - \lambda_i)},$$

will be called the generalized Lagrange resolvent *of the operator A corresponding to the root λ_i.*

Definition 3. *For every vector $x \in V$, the vector $x_i = L_i x$ will be called the generalized Lagrange resolvent (corresponding to the root λ_i) of the polynomial T.*

Proposition 1. *The following statements hold.*

1. *Generalized Lagrange resolvents L_i of the operator A satisfy the following relations:*

   $L_1 + \cdots + L_n = E$,

   $L_i L_j = 0$ *for* $i \neq j$,

   $L_i^2 = L_i$,

   $AL_i = \lambda_i L_i$.

2. *Every vector $x \in V$ is representable as the sum of its generalized Lagrange resolvents, i.e.* $x = x_1 + \cdots + x_n$.

   Moreover, the nonzero resolvents x_i of the vector x are linearly independent and are equal to eigenvectors of the operator A with the corresponding eigenvalues λ_i.

Proof. 1. Let

$$\Lambda = \{\lambda_i\}$$

be the set of all roots of the polynomial T. By definition, the polynomial T_i is equal to one at the point λ_i and is equal to zero at all other points of this set.

It is obvious that the following polynomials vanish on the set Λ:

$T_1 + \cdots + T_n - 1$,

$T_i T_j$ for $i \neq j$,

$T_i^2 - T_i$,

$tT_i - \lambda_i T_i$.

Therefore, each of the polynomials indicated above is divisible by the polynomial T, which has simple roots at the points of the set Λ.

Since the polynomial T annihilates the operator A, i.e. $T(A) = 0$, this implies the relations

$$L_1 + \cdots + L_n = E,$$

$$L_i L_j = 0 \text{ for } i \neq j,$$

$$L_i^2 = L_i,$$

$$A L_i = \lambda_i L_i.$$

2. The second part of the statement is a formal consequence of the first.

Indeed, since $E = L_1 + \cdots + L_n$, every vector x satisfies

$$x = L_1 x + \cdots + L_n x = x_1 + \cdots + x_n.$$

Assume that the vector x is nonzero, and that some linear combination

$$\sum \mu_j x_j$$

of the vectors $x_1, \ldots x_n$ vanishes. Then

$$0 = L_i \sum \mu_j L_j x = \sum L_i L_j \mu_j x = \mu_i x_i,$$

i.e. every nonzero vector x_i enters this linear combination with coefficient zero: $\mu_i = 0$.

The identity $AL_i = \lambda_i L_i$ implies that $AL_i x = \lambda_i L_i x$, i.e. either the vector $x_i = L_i x$ is an eigenvector of L_i with the eigenvalue λ_i, or $x_i = 0$. $\square$ $\square$

CASE OF SEVERAL COMMUTING OPERATORS

Let us discuss the case of two commuting operators in more detail.

Suppose that, along with the linear operator A on the space V, we are given another linear operator $B : V \to V$ that commutes with A and satisfies a polynomial relation of the form

$$Q(B) = B^k + b_1 B^{k-1} + \dots b_k E = 0,$$

where $b_i \in K$.

Assume that the polynomial

$$Q(t) = t^k + b_1 t^{k-1} + \dots b_k$$

has k distinct roots

$$\mu_1, \dots, \mu_k$$

in the field K.

With a root μ_j, we associate the polynomial

$$Q_j(t) = Q(t)/Q'(\mu_j)(t - \mu_j)$$

and the operator $Q_j(B)$, i.e. the generalized Lagrange resolvent of the operator B corresponding to the root μ_j.

Definition 4. *We call the operator*

$$L_{i,j} = T_i(A)Q_j(B)$$

the generalized Lagrange resolvent of the operators A and B corresponding to the pair of roots λ_i, μ_j. The vector

$$x_{i,j} = L_{i,j}x$$

will be called the generalized Lagrange resolvent of the vector $x \in V$ (corresponding to the pair of roots λ_i and μ_j) with respect to the operators A and B.

Proposition 2. *The following statements hold.*

1. *Generalized Lagrange resolvents* $L_{i,j}$ *of commuting operators* A *and* B *satisfy the following relations:*
   $\sum L_{i,j} = E$,
   $L_{i_1,j_1} L_{i_2,j_2} = 0$ *for* $(i_1, j_1) \neq (i_2, j_2)$,
   $L_{i,j}^2 = L_{i,j}$,
   $AL_{i,j} = \lambda_i L_{i,j}$, $BL_{i,j} = \mu_j L_{i,j}$.

2. *Every vector* $x \in V$ *is representable as the sum of its generalized Lagrange resolvents, i.e.* $x = \sum x_{i,j}$.
   Moreover, nonzero resolvents $x_{i,j}$ *of the vector* x *are linearly independent, and are equal to eigenvectors of the operators* A *and* B *with the eigenvalues* λ_i *and* μ_j*, respectively.*

To prove the first part of the proposition, it suffices to multiply the corresponding identities for the generalized resolvents of the operators A and B.

The second part of the proposition is a formal consequence of the first part.

CLASSICAL LAGRANGE RESOLVENT

Consider an operator A of order n, i.e. such operator that

$$A^n = E.$$

Generalized Lagrange resolvents for such operators are particularly important for solving equations by radicals. These are the resolvents that Lagrange has discovered, and we call them the *Lagrange resolvents* (omitting the word "generalized").

Suppose that the field K contains n roots of unity $\xi_1, \ldots, \xi_n$ of degree n, $\xi^n = 1$.

By our assumption, $T(A) = 0$, where

$$T(t) = t^n - 1.$$

Let us now compute the Lagrange resolvent corresponding to the root $\xi_i = \xi$. We have

$$T_i(t) = \frac{t^n - \xi^n}{n\xi^{n-1}(t-\xi)} = \frac{1}{n\xi^{n-1}}(t^{n-1} + \cdots + \xi^{n-1}) = \frac{1}{n}((\xi^{-1}t)^{n-1} + \cdots + 1).$$

The Lagrange resolvent $T_i(A)$ of the operator A corresponding to a root $\xi_i = \xi$ will be denoted by $R_\xi(A)$. We obtain

$$R_\xi(A) = \frac{1}{n} \sum_{0 \le k < n} \xi^{-k} A^k.$$

Corollary 3. *Consider a vector space V (maybe infinite dimensional) over a field K containing all roots of unity of degree n. Suppose that an operator A satisfies the relation*

$$A^n = E.$$

Then, for every vector $x \in V$, either the Lagrange resolvent $R_\xi(A)(x)$ is zero, or it is equal to an eigenvector of the operator A with the eigenvalue ξ.

The vector x is the sum of all its Lagrange resolvents.

Problem 1. *Verify the previous Corollary directly, without any reference to preceding results.*

REPRESENTATION OF FINITE ABILIAN GROUPS

Theorem 4 (On representations of finite abilian groups)**.** *Let G be a finite group of linear operators on a vector space V over the field K. Let n denote the order of the group G. Suppose that the field K contains all roots of unity of degree n. Then the space V is a direct sum of subspaces that are eigenspaces simultaneously for all operators from the group G.*

Proof. Every finitely generated abilian group is a direct sum of cyclic groups. Suppose that the group G is the direct sum of k cyclic groups of orders $m_1, \ldots, m_k$. Suppose that the operators $A_i \in G, \ldots, A_k \in G$ generate these cyclic subgroups. In particular,

$$A_1^{m_1} = \cdots = A_k^{m_k} = E.$$

For every collection $\lambda = \lambda_1, \dots, \lambda_k$ of roots of unity of degrees $m_1, \dots, m_k$, consider the *joint Lagrange resolvent*

$$L_\lambda = L_{\lambda_1}(A_1) \dots L_{\lambda_k}(A_k)$$

of all generators $A_1, \dots, A_k$ of the group G. $\square$

Corollary 5. *Every vector $x \in V$ is representable in the form*

$$x = \sum L_\lambda x.$$

Each of the vectors $L_\lambda x$ is either zero or a common eigenvector of the operators $A_1, \dots, A_k$ with the respective eigenvalues $\lambda_1, \dots, \lambda_k$.

Problem 2. *The theorem about simultaneous diagonalization of a linear operator belonging to a of a finite abelian group does not hold if the characteristic of the ground field divides the order of* *the group. Provide an example of non diaganizable matrix A with $A^2 = E$over a field of characteristic two.*

APPENDIX

LAGRANGE POLYNOMIALS AND LINEAR ALBGEBRA

Let me recall a classical Cayley–Hamilton theorem.

Theorem 6 (Cayley–Hamilton Theorem)**.** *Let A be a $(k\times k)$-matrix with entries belonging to a field K. Let $P(t) = \det(A - tE)$ be the characteristic polynomial of A (whose degree equals to k). Then*

$$P(A) = 0.$$

Example 1. *Let A be a (2×2)-matrix*

$$A = \begin{matrix} a & b \\ c & d \end{matrix}.$$

Then A satisfies the following relation

$$A^2 - (a+d)A + (ad - bc)E = 0.$$

This Theorem allows to evaluate a wide class of function of a $n \times n$-matrix A.

Theorem 7. *Assume that A satisfies a polynomial equation $T(A) = 0$ where T is a degree n polynomial. Let Q be any polynomial (maybe of a very big degree). Let us divide Q by T with a remainder R, i.e. let*

$$Q = Q_1 T + R$$

where Q_1, R *are polynomials and* $\deg R < n$. *Then*

$$Q(A) = R(A).$$

Proof. Indeed $Q(A) = Q_1(A)T(A) + R(A) = R(A)$. $\square$

Definition 5. *Let* $y : K \to K$ *be a function on a field* k *whose values belong to* k. *The Lagrange interpolation polynomial of* y *with the interpolations points* $x_1, \ldots, x_n$ *is the unique polynomial* T *whose degree is smaller than* n *which coincides with* y *at the points* $X_1, \ldots, x_n$, *i.e. the following identities hold:*

$$P(x_1) = y(x_1), \ldots, P(x_n) = y(x_n).$$

Example 2. *The interpolation polynomial of a function* $y(x)$ *with two interpolation points* x_1, x_2 *is the following function* $T(x)$*:*

$$T(x) = \frac{y(x_2) - y(x_1)}{x_2 - x_1} x + y(x_1) - \frac{y(x_2) - y(x_1)}{x_2 - x_1} x_1.$$

Theorem 8. *Assume that a degree* n *polynomial* t *has* n *different roots* $x_1, \ldots, x_n \in K$*, Then for any polynomial* Q *its remainder* R *of its division by* T *is equal to the interpolation polynomial of* Q *with the interpolation points* $x_1, \ldots, x_n$*.*

Proof. Indeed R is a degree $< n$ polynomial which coincides with Q at each root x_i of T since

$$Q(x_i) = Q_1(x_i)T(x_i) + R(x_i) = R(x_i).$$

$\square$

Theorem 9. *Let A be a $(n \times n)$-matrix with entries in the field k having n different eigenvalues $x_1, \ldots, x_n$ and let Q be any polynomial over K. Then*

$$Q(A) = R(A),$$

where R is the Lagrange interpolation polynomial of q with the interpolation points $x_1, \ldots, x_n$.

Proof. Theorem is an automatic corollary from the previous results. □

Problem 3. *Let A be a $(n \times n)$-matrix with entries in the field k having n different eigenvalues $x_1, \ldots, x_n$ and let Q be any polynomial over K. Using explicit formula for the Lagrange interpolation polynomial ang using values of Q at the interpolation points*

$x_1, \ldots, x_n$ represent $Q(A)$ as an explicit degree $(n-1)$ polynomial in Q.

Let $f(z)$ be an entire analytic function of complex variable z. Such function can be represented by converging power series

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n + \cdots >$$

In any bounded domain $U \subset \mathbb{C}$ the function can be approximate by Taylor polynomials of f centered at the origin:

$$Q_k = a_0 + a_1 z + \cdots + a_k z^k.$$

For any $(n \times n)$-matrix A over $\mathbb{C}$ one can defined $f(A)$ as the limit of polynomials

$$f_k(A) = a_0 E + a_1 A + \cdots + a_k A^k$$

when $k \to \infty$.

Corollary 10. *Let A be a $(n \times n)$-matrix with complex entries having n different eigenvalues $x_1, \ldots, x_n$. Let f be any complex entire function. Then*

$$f(A) = R(A),$$

where R is the Lagrange interpolation polynomial of f with the interpolation points $x_1, \ldots, x_n$.

Example 3. *For a complex $(n \times n)$–matric A its exponent $\exp A$ is defined as*

$$\exp A = E + A + \cdots + \frac{1}{n!}A^n + \ldots .$$

Problem 4. *Let A be a complex $(n \times n)$-matrix having n different eigenvalues $x_1, \ldots, x_n$. Using explicit formula for the Lagrange* *interpolation polynomial and using the numbers* $\exp x_1, \ldots, \exp x_n$ *represent* $\exp(A)$ *as an explicit degree* $(n-1)$ *polynomial in* A.

Problem 5. *Consider a system of linear differential equations*

$$y' = Ay$$

with initial data $y(x_0) = y_0$, *where* y *is unknown function of complex variable* x, A *is a constant* $(n \times n$ *complex matrix and* y_0 *is a given complex vector.*

1) Check that the vector

$$y(x) = \exp(x - x_0)Ay_0.$$

satisfies the system and the initial data. 2) Assume that A *has* n *different eigenvalues* $x_1, \ldots, x_n$. *Using explicit formula for the Lagrange interpolation polynomial represent the above solution as an explicit degree* $(n-1)$ *polynomial in* A *whose coefficients are explicit* *linear combinations of the functions:* $\exp(x - x_0)x_1, \ldots, \exp(x - x_0)x_n$.