

ALGEBRAIC PRELIMINARIES FOR THE STUDY OF MULTI-DIMENSIONAL MATRICES

EUGENE TYRTYSHNIKOV

Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow
eugene.tyrtysnikov@gmail.com

1. What do we need it for. A matrix is introduced as a 2-dimensional table with rows and columns. A 3-dimensional matrix can be easily imagined as a 3-dimensional table, in other words a function of 3 indices

$$a_{ijk} = a(i, j, k), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n, \quad 1 \leq k \leq q,$$

with natural ordering of the indices. Main big questions about matrices are the decompositions, e.g. the *skeleton decomposition*:

$$a(i, j) = \sum_{\alpha=1}^r u_{\alpha}(i)v_{\alpha}(j) \quad \Leftrightarrow \quad A = \sum_{\alpha=1}^r u_{\alpha}v_{\alpha}^T = UV^T,$$

$$U = [u_1, \dots, u_r], \quad V = [v_1, \dots, v_r].$$

If we exclude zero summands, then it expresses the matrix as the sum of matrices of rank 1. Each nonzero matrix of the form $u_{\alpha}(i)v_{\alpha}(j)$ is called a nonzero *skeleton*. Similarly, a 3-dimensional matrix is called a *skeleton* if it is of the form $u_{\alpha}(i)v_{\alpha}(j)w_{\alpha}(k)$. Any decomposition of the form

$$a(i, j, k) = \sum_{\alpha=1}^r u_{\alpha}(i)v_{\alpha}(j)w_{\alpha}(k)$$

is called *canonical tensor decomposition* of $a(i, j, k)$. A canonical decomposition with minimal number of nonzero skeletons is called *minimal decomposition*, the corresponding number of skeletons is called the *tensor rank* of $a(i, j, k)$. In this course we always assume that the entries of all multi-dimensional matrices and their decompositions are *complex numbers*. Multi-dimensional matrices are often called *tensors*.

We want to answer the following questions.

1. Is it possible to compute the tensor rank in finitely many arithmetic operations?
2. Is there a single natural number $r = \text{grank}(m, n, q)$ such that the set of 3-dimensional tensors of size $m \times n \times q$ and of rank bounded by r is dense in the space of all tensors of size $m \times n \times q$? Such a number r is called *generic rank* for tensors of size $m \times n \times q$.
3. Let L be any subspace of dimension $(n-1)^2 + 1$ in the linear space of all $n \times n$ matrices. Does it contain a matrix of rank 1?

The answer to each of the questions is positive. However, what do we need from algebra to be able to prove this? The corresponding algebraic background turns out

to be strikingly rich, deep, certainly beautiful but not of the early stage of the study of mathematics. The purpose of this course is to pave the way to it as easy as possible. In the end we will be able to have complete rigorous proofs for the above questions.

2. Algebraic dependence and independence (Lecture 1). Field extensions. Adjunction of an element. Primitive elements. Degree of extension. Algebraically dependent elements. Algebraically independent elements. Transcendence degree. Transcendence base. Differentiations on a field. Linear space of differentiations.

EXERCIZES

1. Find all intermediate fields in between of \mathbb{R} and \mathbb{C} .
2. Calculate the extension degree (dimension) of $\mathbb{Q}(\sqrt{2}+\sqrt{3})$ over \mathbb{Q} .
3. Elements a and b are algebraic over \mathbb{Q} . Prove that $a+b$ and ab are algebraic over \mathbb{Q} .
4. Let $p_1 < \dots < p_n$ be prime numbers and $L = \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_n})$. Prove that the extension degree (dimension) of L over \mathbb{Q} is equal to 2^n .
5. Prove that there exist infinitely many differentiations on the field \mathbb{R} .

3. Ideals and bases (Lecture 2). We consider only commutative rings and their subrings. A subring I is called an *ideal* in R if it is closed under multiplications by an arbitrary element of R , i.e. possesses the *absorption property*: $f \in I$ and $g \in R$ imply $fg \in I$.

Our favorite ring is $R = \mathbb{C}[x_1, \dots, x_n]$. The set of polynomials

$$I = \langle f_1, \dots, f_s \rangle := \{f_1g_1 + \dots + f_sg_s : g_1, \dots, g_s \in R\}$$

is obviously an ideal. We say that I is generated by f_1, \dots, f_s and the system f_1, \dots, f_s is called a *basis* of I . Such ideals are referred to as *finitely generated*.

THEOREM 3.1. (HILBERT THEOREM ON BASES) *Any ideal in $\mathbb{C}[x_1, \dots, x_n]$ is finitely generated.*

An ideal $I \subsetneq R = \mathbb{C}[x_1, \dots, x_n]$ is called *maximal* if any ideal strictly larger than I must coincide with R .

COROLLARY 3.2. *Any proper ideal in $\mathbb{C}[x_1, \dots, x_n]$ is contained in a maximal ideal.*

Given a polynomial $f \in R$, how can we check that it lies in the ideal I ? We can do this in finitely many operations if we have a special basis of I , the so called *Groebner basis*.

In order to introduce the Groebner bases we need to choose and fix an ordering of monomials in R . Let it be the *lex-order*:

$$x_1^{\alpha_1} \dots x_n^{\alpha_n} > x_1^{\beta_1} \dots x_n^{\beta_n} \quad \text{iff}$$

$$\alpha_1 = \beta_1, \dots, \alpha_{t-1} = \beta_{t-1}, \alpha_t > \beta_t \quad \text{for some } 1 \leq t \leq n.$$

Then, each polynomial f has the *leading term* with respect to the lex-order of monomials. A basis f_1, \dots, f_s of I is called a Groebner basis of I if the leading term of each polynomial of I is divisible by the leading term of at least one of polynomials from the basis. The leading term of f will be denoted by $\text{LT}(f)$.

To check if $f \in I$, all we need to do is to find the remainder of the division of f by f_1, \dots, f_s . The process of division of f by a system of nonzero polynomials is defined as follows.

ALGORITHM 3.1. *On input: f, f_1, \dots, f_s . On output: r (the remainder).*

1. $g := f, r := 0$.
2. If $g = 0$, then quit.
3. $\mathcal{M} := \{j : \text{LT}(g) \dot{\vdash} \text{LT}(f_j), 1 \leq j \leq s\}$.
4. If $\mathcal{M} \neq \emptyset$, then $i := \min_{j \in \mathcal{M}} j$, $g := g - f_i \text{LT}(g)/\text{LT}(f_i)$, go to 2.
5. $r := r + \text{LT}(g)$, $g := g - \text{LT}(g)$, go to 2.

THEOREM 3.3. *Assume that f_1, \dots, f_s is a Groebner basis of the ideal $I = \langle f_1, \dots, f_s \rangle$. Then $f \in I$ iff the remainder of the division of f by f_1, \dots, f_s is equal to zero.*

The existence of the Groebner basis easily follows from Dixon's result about monomial ideals. An ideal is called *monomial* if it is generated by a possibly infinite set of monomials.

THEOREM 3.4. (DIXON LEMMA) *Any monomial ideal has a finite basis consisting of some of the monomials that generate this ideal.*

How to check if a basis of an ideal is its Groebner basis? It is easy to do using the so called *syzygy polynomials*: for given polynomials $f(x_1, \dots, x_n)$ and $g(x_1, \dots, x_n)$, this name is reserved for the new polynomial

$$S(f, g) := \frac{h}{\text{LT}(f)}f - \frac{h}{\text{LT}(g)}g,$$

where $h := x_1^{\gamma_1} \dots x_n^{\gamma_n}$ is the lowest-degree monomial divisible by both $\text{LT}(f)$ and $\text{LT}(g)$. We emphasize that the definition depends upon the order of the monomials.

THEOREM 3.5. (BUCHBERGER) *A given basis g_1, \dots, g_s of an ideal is its Groebner basis iff the remainder in the division of any syzygy polynomial $S(g_i, g_j)$ by this basis is equal to zero.*

Making more use of the syzygy polynomials, Buchberger suggested an algorithm that finds a Groebner basis in finitely many operations.

ALGORITHM 3.2. *On input: $F = \{f_1, \dots, f_s\}$ (the basis of an ideal). On output: $G = \{g_1, \dots, g_t\}$ (a Groebner basis of this ideal).*

1. $G := F$.
2. $H := G$

3. For all polynomials $p, q \in H$ compute the remainder r of the division of $S(p, q)$ by G , if $r \neq 0$ then $H := H \cup \{r\}$.
4. If $H = G$ then quit, otherwise go to 2.

Why does this algorithm stop after finitely many steps? Assume that G is strictly less than H before we set $H := G$. Then one can prove that the ideal $\text{LT}(G)$ of the leading terms of polynomials from G is strictly less than the corresponding ideal $\text{LT}(H)$. Indeed, if $r \neq 0$ then r is not divisible by any of the leading terms of polynomials from G , and hence $r \notin \text{LT}(G)$. As a corollary of the Hilbert theorem on bases, an infinite chain of strictly increasing polynomial ideals does not exist. Consequently, at some step we obtain $H = G$.

EXERCIZES

1. Show that the remainder in the division by a system f_1, \dots, f_s may depend of the ordering of the system polynomials.
2. Prove that the remainder in the division by the Groebner basis does not depend on the ordering of the basis polynomials.
3. A Groebner basis f_1, \dots, f_s is called *reduced* if no term of any of f_i is not divisible by the leading term of any of f_j with $j \neq i$ and each of the leading coefficients is equal to 1. Prove that the reduced Groebner basis exists and is unique up to the ordering of the basis polynomials.
4. Assume that $x > y > z$ in the lex-order. Find Groebner bases of the ideals $I = \langle x - y, y - z \rangle$ and $J = \langle x + y, x - z \rangle$.

4. Varieties and Nullstellensatz (Lecture 3). Let S be an arbitrary subset of points in \mathbb{C}^n . Then $\mathbb{I}(S)$ denotes the set of all polynomials $f(x_1, \dots, x_n) \in \mathbb{C}[x_1, \dots, x_n]$ that are zeroed on every point from S . By definition, $\mathbb{I}(\emptyset) = \mathbb{C}[x_1, \dots, x_n]$. It is trivial to prove that $\mathbb{I}(S)$ is an ideal.

On the other hand, for any ideal I we define $\mathbb{V}(I)$ as the set of all common zeroes of all polynomials from I . Any set of this form is called an *algebraic set* or *affine variety* or *algebraic variety*. In this text we will say simply *variety*. According to the Hilbert theorem on bases, any variety is the set of *all common zeroes* of a system of finitely many polynomial equations.

THEOREM 4.1. (HILBERT THEOREM ON ZEROES IN THE WEAK FORM) *Let I be an ideal in $\mathbb{C}[x_1, \dots, x_n]$. Then $\mathbb{V}(I) \neq \emptyset$ iff $I \neq \mathbb{C}[x_1, \dots, x_n]$, or equivalently, iff $1 \notin I$.*

COROLLARY 4.2. *For any given tensor, its tensor rank can be computed in finitely many arithmetic operations.*

THEOREM 4.3. (NULLSTELLENSATZ) *Let I be an ideal in $\mathbb{C}[x_1, \dots, x_n]$ and $V = \mathbb{V}(I)$. Then $f \in \mathbb{I}(V)$ iff $f^s \in I$ for some natural number s depending on f .*

The set of all polynomials f such that $f^s \in I$ for some natural number $s = s(f)$ is called the *radical* of the ideal I and denoted by \sqrt{I} . Prove that \sqrt{I} is itself an ideal. An ideal I is called *radical* if $\sqrt{I} = I$. An equivalent form of the Nullstellensatz is the assertion that $\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}$. Prove this. As a corollary, we have the one-to-one correspondence $I \leftrightarrow V$ between radical ideals and varieties, it is provided by the

equations $I = \mathbb{I}(V)$ and $V = \mathbb{V}(I)$.

A variety is called *irreducible* if it is not a union of two proper varieties. An ideal I is called *prime* if the inclusion $fg \in I$ implies that $f \in I$ or $g \in I$.

THEOREM 4.4. *Any variety is the union of finitely many irreducible varieties. The latter varieties are unique provided that none of them is part of another one.*

EXERCIZES

1. Prove that the union and intersection of two varieties is a variety.
2. Show that the sum of two radical ideals is an ideal that is not necessarily radical.
3. Prove that any prime ideal is radical.
4. Prove that an ideal I is prime iff the corresponding variety $V = \mathbb{V}(I)$ is irreducible.
5. Prove that the radical of an ideal I is the intersection of all maximal ideals containing I and equal to the intersection of all prime ideals containing I .

5. Varieties and projections (Lecture 4). A protagonist in the study of 3-dimensional matrices of size $m \times n \times q$ is certainly the set of those of them whose rank is bounded by r . Denote this set by S_r . How does it look like? Anyway, S_r is evidently the image of a polynomial mapping from \mathbb{C}^N to \mathbb{C}^{mnq} , where $N = (m+n+q)r$. This simple observation is paramount as it allows one to recognize that S_r is a projection of some algebraic variety. This variety is to be considered in \mathbb{C}^{mnq+N} and defined by the equations

$$y_i - f_i(x_1, \dots, x_N) = 0, \quad 1 \leq i \leq mnq,$$

where f_1, \dots, f_{mnq} are polynomials. Set $y_i = a_i \in \mathbb{C}$ and $x_j = b_j \in \mathbb{C}$. Then the point $(a_1, \dots, a_{mnq}, b_1, \dots, b_N)$ satisfies each of the equations iff $(a_1, \dots, a_{mnq}) \in S_r$.

Consider the rings $\mathbb{K}_l := \mathbb{C}[x_1, \dots, x_{n-l}]$ for different $0 \leq l \leq n$. For an ideal $I \subseteq \mathbb{K}_0$, by $I_l = I \cap \mathbb{K}_l$ we denote the subset of those polynomials from I that do not include x_j with $j > n-l$.

Denote by π_l^i a projection from \mathbb{C}^{n-i} to \mathbb{C}^{n-l} according to the rule

$$\pi_l^i : (a_1, \dots, a_{n-i}) \rightarrow (a_1, \dots, a_{n-l}).$$

Of course, we assume that $i \leq l$. The notation for π_l^0 can be shortened to π_l .

Let V be a variety in \mathbb{C}^n . Obviously, the set $\pi_l(V)$ is contained in $V_l = \mathbb{V}(I_l)$.

LEMMA 5.1. *Assume that V is an irreducible variety in \mathbb{C}^n . For any proper subvariety $W_0 \subsetneq V$, there exists a proper subvariety $W_1 \subsetneq V_1$ such that*

$$V_1 \setminus W_1 \subseteq \pi_1(V \setminus W_0).$$

THEOREM 5.2. *For an arbitrary variety $V \subseteq \mathbb{C}^n$ and any proper subvariety $W_0 \subsetneq V$, there exists a proper subvariety $W_l \subsetneq V_l$ such that*

$$V_l \setminus W_l \subseteq \pi_l(V \setminus W_0).$$

Using the Hilbert theorem on bases, it is then not very difficult to find that

$$\pi_l(V) = \bigcup_{1 \leq i \leq s} A_i \setminus B_i,$$

where A_i, B_i are subvarieties in V_l . Sets of such a form are called *constructive sets*.

For a set $S \subseteq \mathbb{C}^n$, denote by \bar{S} the smallest variety containing S . Prove that $\bar{S} = \mathbb{V}(\mathbb{I}(S))$.

Recall that $S_l \subseteq \mathbb{C}^{m \times n \times q}$ is the set of all 3-dimensional matrices of rank bounded from above by l and consider the chain

$$\bar{S}_1 \subsetneq \bar{S}_2 \subsetneq \dots \subsetneq \bar{S}_r = \bar{S}_{r+1}.$$

Prove that \bar{S}_r coincides with the space of all tensors of size $m \times n \times q$.

THEOREM 5.3. *If W is a proper subvariety of an irreducible variety V , then the set $V \setminus W$ is dense in V in the standard topology.*

COROLLARY 5.4. *Let r be minimal such that $\bar{S}_r = \bar{S}_{r+1}$. Then $r = \text{grank}(m, n, q)$.*

EXERCIZES

1. Prove that $S_1 \subseteq \mathbb{C}^{m \times n \times q}$ is a variety.
2. Prove that S_2 is not a variety, provided that $m, n, q \geq 2$.
3. Prove that each variety \bar{S}_l is irreducible.
4. Prove that $\text{grank}(2, 2, 2) = 2$.
5. Prove that the maximal possible value of tensor rank for 3-dimensional matrices from $\mathbb{C}^{2 \times 2 \times 2}$ is equal to 3.

6. Dimension of varieties (Lecture 5). Since a variety is defined implicitly as solutions to a finite system of polynomial equations, it seems very pertinent to remember the classical *implicit function theorem*. Suppose that a mapping $f: (x_1, \dots, x_n) \rightarrow (y_1, \dots, y_m)$ is defined by polynomial equations

$$\begin{aligned} y_1 &= f_1(x_1, \dots, x_n), \\ &\dots \\ y_m &= f_m(x_1, \dots, x_n). \end{aligned}$$

A point $a = (a_1, \dots, a_n)$ is called *regular* if the rank of Jacoby matrix

$$J_f = \left[\frac{\partial f_i}{\partial x_j} \right]_{m \times n}$$

is constant in a domain around a . By a domain, in this text we assume an open polydisc, i.e. a set of all points $x = (x_1, \dots, x_n)$ of the form

$$D = \{|x_1 - a_1| < \varepsilon_1\} \times \dots \times \{|x_n - a_n| < \varepsilon_n\}$$

with some positive radii $\varepsilon_1, \dots, \varepsilon_n$.

If $f(a) = 0$ and the point a is regular with $r = \text{rank } J_f(a)$ and the first r columns are linearly independent, then the implicit function theorem states that there exists a domain D around a in which the set $V = \{x : f(x) = 0\}$ is a graph of some analytic function

$$\phi : (x_{r+1}, \dots, x_n) \rightarrow (x_1, \dots, x_r).$$

One may consider this function to be defined at each point of

$$D \cap \{x_1 = a_1, \dots, x_r = a_r\}.$$

LEMMA 6.1. *Any point of a variety is a limit of its regular points with the same value of rank.*

In this approach, one should still remark that we deal, as a matter of fact, with regular points for a system of polynomials that define a variety rather than variety itself, and the regularity may (and does!) depend on the choice of this system. This is by no means satisfactory. However, it can be proved that *if a point is regular with respect to one system of polynomials then it remains regular for any larger system defining the same variety*. We can consider the Jacoby matrix (with infinite number of rows) for a set of all polynomials from the ideal $I = \mathbb{I}(V)$. Then, a point of V is called regular if the rank of this matrix is constant in a domain around this point.

The *dimension of variety at a point* is defined as n minus minimal value of rank for all sequences of regular points converging to this point. The *dimension of variety* is the maximal dimension of its points.

The key algebraic object for the study of dimension is a *coordinate ring*. It is juxtaposed to any irreducible variety V . The coordinate ring R for V is defined as the *quotient ring* $R = \mathbb{C}[X_1, \dots, X_n]/I$ (also known by the names *factor ring* and *residue-class ring*). Here we follow a custom of many texts on algebraic geometry where the independent variables associated with the coordinates are denoted by capital letters. The reason is that the lowercase letters x_i may be reserved for the classes $[X_i] = X_i + I$ of polynomials equal modulo I . Another useful point of view is that the elements of R are polynomial functions restricted to the points of V . Note that different polynomials may be equal as functions considered only on V , the difference of such polynomials evidently belongs to I .

As we know, the ideal $I = \mathbb{I}(V)$ of an irreducible variety V is prime. Irreducibility of V is important for the assertion that R is an *integral domain*, i.e. R is free from divisors of zero. Consequently, R can be embedded into the *field of rational functions modulo I* . Denote this field by \tilde{R} .

THEOREM 6.2. *If V is an irreducible variety with the coordinate ring R , then there is a proper subvariety $W \subsetneq V$ such that every point in $V \setminus W$ is regular with the rank equal to the transcendence degree of the field \tilde{R} over \mathbb{C} .*

To taste the flavor and style of characteristic proofs we present next a proof of one of the previously formulated theorems. It is useful and instructive for the reader to check each of the steps and be assured that each particular claim is entirely understandable.

PROOF OF THEOREM 5.3. From the contrary, if $V \setminus W$ is not dense in V , then there is a domain D with the property $D \cap W = D \cap V$. Let $a \in D \cap W$ be a regular point with the rank r . We do not lose generality assuming that $D \cap W$ is a graph of

an analytic function from X_1, \dots, X_d , where $d = n - r$. The transcendence degree, denote it by t , is then less than or equal to d . Otherwise, the points of V cannot lie on a graph of a function of X_1, \dots, X_d in a vicinity of a , because there should be infinitely many points of V above the points (b_1, \dots, b_d) near (a_1, \dots, a_d) . If $t < d$ then there could be only finitely many points of V above the most of points (b_1, \dots, b_t) , which is not our case.

Consequently, $t = d$ and the elements

$$x_1 = [X_1], \dots, x_d = [X_d]$$

are algebraically independent over \mathbb{C} . Then, each element of the coordinate ring R is algebraic over the field $K := \mathbb{C}(x_1, \dots, x_d)$. Consider a polynomial f that is equal to zero on W . Since $W \neq V$, we can choose this polynomial to be nonzero at least at one point of V . Thus, $[f] \neq 0$. As any element of R , $[f]$ is algebraic over K and, therefore, is a root of its minimal polynomial over K :

$$p_0 + p_1 f + \dots + p_s f^s \in I = \mathbb{I}(V),$$

where p_0, \dots, p_s are polynomials in X_1, \dots, X_d and $p_0 \neq 0$ due to minimality. However, $p_0(b_1, \dots, b_d) = 0$ for all points (b_1, \dots, b_d) in a domain around the point (a_1, \dots, a_d) . Hence, p_0 must be zero polynomial and we conclude that $f \in \mathbb{I}(V)$. Thus, any $f \in \mathbb{I}(W)$ belongs to $\mathbb{I}(V)$. Consequently, $W = V$, that is, W cannot be a proper subvariety of V \square

THEOREM 6.3. *Assume that varieties V and W are irreducible and $V \cap D = W \cap D$ in a domain D . Then $V = W$.*

THEOREM 6.4. *Let W be a proper subvariety of an irreducible variety V . Then $\dim W < \dim V$.*

THEOREM 6.5. *Suppose that varieties V and W have at least one common point. Then*

$$\dim(V \cap W) \geq \dim V + \dim W - n.$$

EXERCIZES

1. For any varieties A and B , prove that $A \times B$ is a variety of dimension $\dim A + \dim B$.
2. Prove that any variety defined by a single equation $f = 0$ for a nonzero polynomial f is irreducible iff the polynomial f is irreducible. Prove that the dimension of this variety is equal to $n - 1$.
3. Prove that $S_1 \subseteq \mathbb{C}^{m \times n \times q}$ is a variety of dimension $m + n + q - 2$.
4. Prove that the set of $n \times n$ matrices with rank bounded by r is an irreducible variety of dimension $2r(n - r)$.
5. Let V be an irreducible variety defined by an ideal

$$I = \mathbb{I}(V) \subseteq \mathbb{C}[X_1, \dots, X_n],$$

and assume that $a \in \mathbb{C}^n$ is a regular point of V with the rank r and the columns $d + 1, \dots, n$ of the Jacoby matrix for I at a are

linearly independent, where $d = n - r$. Prove that the elements

$$[X_1] = X_1 + I, \dots, [X_d] = X_d + I$$

of the coordinate ring $\mathbb{C}[X_1, \dots, X_n]/I$ are algebraically independent over \mathbb{C} .

6. Prove that any variety of dimension $n - 1$ in \mathbb{C}^n can be defined by a single polynomial equation.
7. Prove that a variety consists of finitely many points iff its dimension is equal to zero.
8. Using Theorem 6.5, prove that the answer to Question 3 from the introduction is positive.

REFERENCES

- [1] D. Cox, J. Little, D. O' Shea, Idelas, varieties, algorithms, 3rd edition, Springer, 2007.
- [2] K. Kendig, Elementary algebraic geometry, Springer-Verlag, New York, 1977.
- [3] S. Lang, Algebra, Springer, 2002.
- [4] D. Mumford, Algebraic geometry I. Complex projective varieties. Springer-Verlag, Berlin, Heidelberg, New York, 1976.
- [5] V. Strassen, Rank and optimal computation of generic tensors, Linear Algebra Appl., 52/53 (1983), pp. 645–685.