Topics in Analysis and Linear Algebra

Le Chen
le.chen@emory.edu

Emory University

Atlanta GA

Last updated on

July 22, 2021

Summer Bootcamp for
Emory Biostatistics and Bioinformatics
PhD Program

July 22 - 28, 2021

https://github.com/chenle02/2021_Summer_Bio_bootcamp_slides/

Chapter 4. Topics in linear algebra

Chapter 4. Topics in linear algebra

What is a vector space?

1. $\mathbb{R}^n$
2. Polynomials of order at most n:

$$\{a_0 + a_1 x + \cdots + a_n x^n | a_i \in \mathbb{R},\ i = 1, \cdots, n\}$$

3. The set of $m \times n$ matrices.
4. The set of continuous functions on $[0, 1]$, i.e., $C([0, 1])$.
5. The set of functions on $[0, 1]$ having nth continuous derivatives, i.e., $C^n([0, 1])$.

$\vdots \qquad \vdots$

Def. Let V be a nonempty set of objects with two operations: vector addition and scalar multiplication.

Then V is called a *vector space* if it satisfies the following:

Axioms of Addition
and
Axioms of Scalar Multiplication.

The elements of V are called *vectors*.

Axioms of addition

A1. V is closed under addition.
$\vec{v}, \vec{w} \in V \implies \vec{u} + \vec{v} \in V$

A2. Addition is commutative.
$\vec{u} + \vec{v} = \vec{v} + \vec{u}$ for all $\vec{u}, \vec{v} \in V$.

A3. Addition is associative.
$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ for all $\vec{u}, \vec{v}, \vec{w} \in V$.

A4. Existence of an additive identity.
There exists an element $\vec{0}$ in V so that $\vec{u} + \vec{0} = \vec{u}$ for all $\vec{u} \in V$.

A5. Existence of an additive inverse.
For each $\vec{u} \in V$ there exists an element $-\vec{u} \in V$ so that $\vec{u} + (-\vec{u}) = 0$.

Axioms of scalar multiplication

S1. V is closed under scalar multiplication.
$\vec{v} \in V$ and $k \in \mathbb{R}, \implies k\vec{v} \in V$.

S2. Scalar multiplication distributes over vector addition.
$a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ for all $a \in \mathbb{R}$ and $\vec{u}, \vec{v} \in V$.

S3. Scalar multiplication distributes over scalar addition.
$(a + b)\vec{u} = a\vec{u} + b\vec{u}$ for all $a, b \in \mathbb{R}$ and $\vec{u} \in V$.

S4. Scalar multiplication is associative.
$a(b\vec{u}) = (ab)\vec{u}$ for all $a, b \in \mathbb{R}$ and $\vec{u} \in V$.

S5. Existence of a multiplicative identity for scalar multiplication.
$1\vec{u} = \vec{u}$ for all $\vec{u} \in V$.

Def. Let V be a vector space and $\vec{u}, \vec{v} \in V$. The *difference* of $\vec{u}$ and $\vec{v}$ is defined as

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$$

(where $-\vec{v}$ is the additive inverse of $\vec{v}$).

Thm Let V be a vector space, $\vec{u}, \vec{v}, \vec{w} \in V$, and $a \in \mathbb{R}$.

1. If $\vec{u} + \vec{v} = \vec{u} + \vec{w}$, then $\vec{v} = \vec{w}$.
2. The equation $\vec{x} + \vec{v} = \vec{u}$, has a unique solution $\vec{x} \in V$ given by $\vec{x} = \vec{u} - \vec{v}$.
3. $a\vec{v} = \vec{0}$ if and only if $a = 0$ or $\vec{v} = \vec{0}$.
4. $(-1)\vec{v} = -\vec{v}$.
5. $(-a)\vec{v} = -(a\vec{v}) = a(-\vec{v})$.

Examples

E.g.1 Let $V = \{(x, y) \mid x, y \in \mathbb{R}\}$, with addition $\oplus$ and scalar multiplication $\odot$ defined as follows:

For $(x_1, y_1), (x_2, y_2) \in V$, and $a, b \in \mathbb{R}$:

Addition: $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 + y_2 + 1)$.

Scalar multiplication: $a \odot (x_1, y_1) = (ax_1, ay_1 + a - 1)$.

Show that V, with addition and scalar multiplication as defined, is a vector space.

Sol. It is clear that V is closed under $\oplus$ and $\odot$, since both operations produce ordered pairs of real numbers.

1. It is routine to verify that $\oplus$ is commutative and associative.
2. What is the additive identity?
3. What is the additive inverse of $(x, y) \in V$?
4. Verify that $(a + b) \odot (x_1, y_1) = (a \odot (x_1, y_1)) \oplus (b \odot (x_1, y_1))$.
5. Verify that $a \odot ((x_1, y_1) \oplus (x_2, y_2)) = (a \odot (x_1, y_1)) \oplus (a \odot (x_2, y_2))$.
6. Verify that $a \odot (b \odot (x_1, y_1)) = (ab) \odot (x_1, y_1)$.
7. Verify that $1 \odot (x, y) = (x, y)$. $\square$

E.g.2 Let $\mathbb{R}_+$ be the set of positive reals.

Let the addition $\oplus$ and the scalar multiplication $\odot$ defined as follows:

For $x, y \in \mathbb{R}_+$, and $a \in \mathbb{R}$:

Addition: $x \oplus y = xy$.

Scalar multiplication: $a \odot x = x^a$.

Prove that $\mathbb{R}_+$ equipped with $\oplus$ and $\odot$ is a vector space.

E.g.3 Let $C([0,1])$ be the set of continuous functions defined on $[0,1]$ equipped with usual addition and scalar multiplication. Prove that $C([0,1])$ is a vector space.

E.g.4 Let $C^n([0,1])$ be the set of functions that have continuous nth derivatives ($n \geq 0$) defined on $[0,1]$, equipped with usual addition and scalar multiplication. Prove that $C^n([0,1])$ is a vector space.

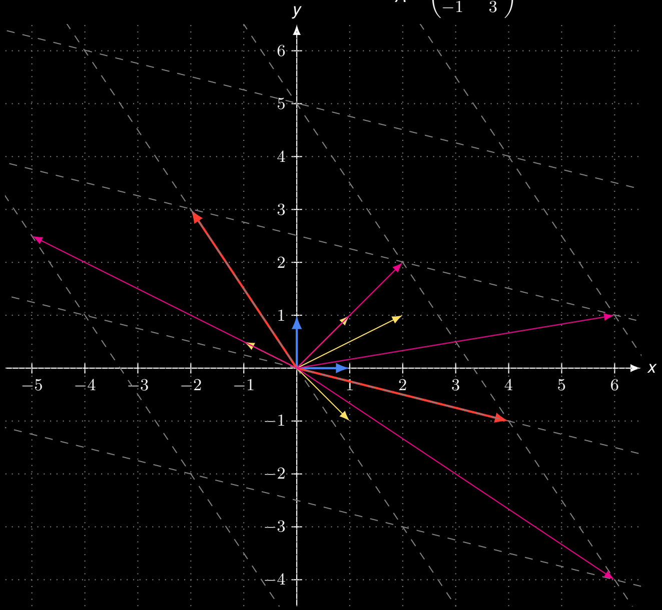
E.g.5 The set of $m \times n$ matrices M_{mn}.

E.g.6 Polynomials of degree n.

E.g. $A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$ has two eigenvalues: $\lambda_1 = 2$ and $\lambda_2 = 5$ with corresponding eigenvectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}$$

$$A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$$



Def. Let A be a 2×2 matrix and L a line in $\mathbb{R}^2$ through the origin. Then L is said to be A-invariant if the vector $A\vec{x}$ lies in L whenever $\vec{x}$ lies in L,

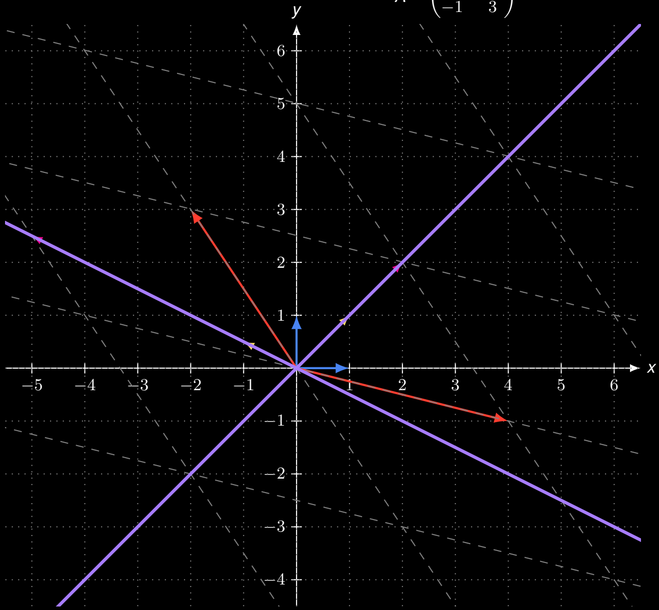
i.e., $A\vec{x}$ is a scalar multiple of $\vec{x}$,
i.e., $A\vec{x} = \lambda\vec{x}$ for some scalar $\lambda \in \mathbb{R}$,
i.e., $\vec{x}$ is an eigenvector of A.

Def. Let $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ be a nonzero vector in $\mathbb{R}^2$. Then $L_{\vec{v}}$ is the set of all scalar multiples of $\vec{v}$, i.e.,

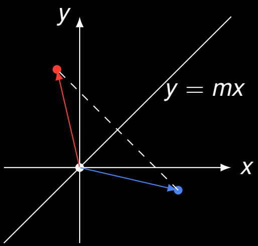
$$L_{\vec{v}} = \mathbb{R}\vec{v} = \left\{ t\vec{v} \mid t \in \mathbb{R} \right\}.$$

Thm Let A be a 2×2 matrix and let $\vec{v} \neq 0$ be a vector in $\mathbb{R}^2$. Then $L_{\vec{v}}$ is A-invariant if and only if $\vec{v}$ is an eigenvector of A.

$$A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$$



E.g. Let $m \in \mathbb{R}$ and consider the linear transformation $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$, i.e., reflection in the line $y = mx$.

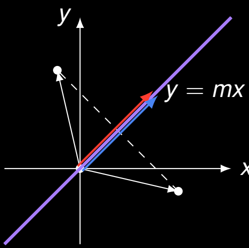


Recall that this is a matrix transformation induced by

$$A = \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.$$

Find the lines that pass through origin and are A-invariant. Determine corresponding eigenvalues.

Sol.

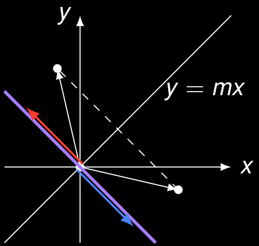


Let $\vec{x}_1 = \begin{bmatrix} 1 \\ m \end{bmatrix}$. Then $L_{\vec{x}_1}$ is A-invariant, that is, $\vec{x}_1$ is an eigenvector.

Since the vector won't change, its eigenvalue should be 1. Indeed, one can verify that

$$A\vec{x}_1 = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{pmatrix} 1 \\ m \end{pmatrix} = \dots = \begin{pmatrix} 1 \\ m \end{pmatrix} = \vec{x}_1.$$

Sol. (Continued)



Let $\vec{x}_2 = \begin{bmatrix} -m \\ 1 \end{bmatrix}$. Then $L_{\vec{x}_2}$ is A-invariant, that is, $\vec{x}_2$ is an eigenvector.

Since the vector won't change the size, only flip the direction, its eigenvalue should be -1. Indeed, one can verify that

$$A\vec{x}_2 = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{pmatrix} -m \\ 1 \end{pmatrix} = \cdots = \begin{pmatrix} m \\ -1 \end{pmatrix} = -\vec{x}_2.$$

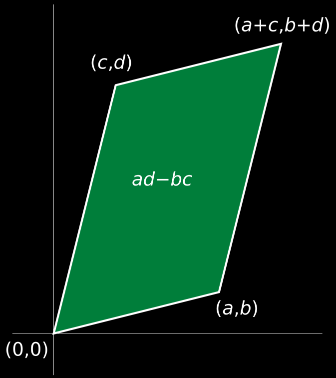
$\square$

E.g. Let θ be a real number, and $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ rotation through an angle of θ, induced by the matrix

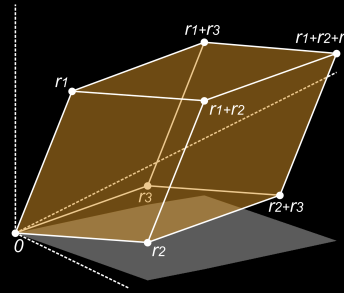
$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Claim A has no real eigenvalues unless θ is an integer multiple of π, i.e., $\pm\pi, \pm 2\pi, \pm 3\pi$, etc.

Sol. a line L in $\mathbb{R}^2$ is A invariant if and only if θ is an integer multiple of π.



$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \text{signed area of parallelogram}$$



$$\det\begin{pmatrix}\vec{r}_1 & \vec{r}_2 & \vec{r}_3\end{pmatrix} = \text{signed volume of the parallelepipe}$$

Cofactor and cofactor expansion

Def. Let $A = [a_{ij}]$ be an $n \times n$ matrix.

– The sign of the (i, j) position is $(-1)^{i+j}$.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \Rightarrow \begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

– Let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j. The (i,j)-cofactor of A is

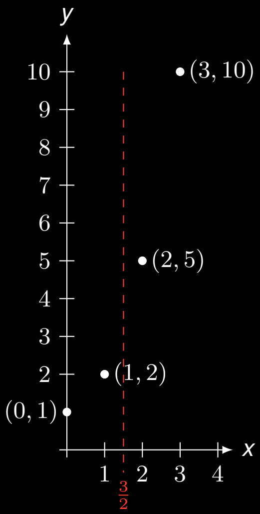
$$c_{ij}(A) = (-1)^{i+j} \det(A_{ij}).$$

– The determinant of A is defined as

$$\det A = a_{11}c_{11}(A) + a_{12}c_{12}(A) + a_{13}c_{13}(A) + \cdots + a_{1n}c_{1n}(A)$$

and is called the cofactor expansion of $\det A$ along row 1.

Problem Given data points $(0,1)$, $(1,2)$, $(2,5)$ and $(3,10)$, find an interpolating polynomial $p(x)$ of degree at most three, and then estimate the value of y corresponding to $x = 3/2$.



Sol. We want to find the coefficients r_0, r_1, r_2 and r_3 of

$$p(x) = r_0 + r_1 x + r_2 x^2 + r_3 x^3$$

so that $p(0) = 1$, $p(1) = 2$, $p(2) = 5$, and $p(3) = 10$.

$$\begin{aligned}
p(0) &= r_0 = 1 \\
p(1) &= r_0 + r_1 + r_2 + r_3 = 2 \\
p(2) &= r_0 + 2r_1 + 4r_2 + 8r_3 = 5 \\
p(3) &= r_0 + 3r_1 + 9r_2 + 27r_3 = 10
\end{aligned}$$

Solve this system of four equations in the four variables r_0, r_1, r_2 and r_3.

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 4 & 8 & 5 \\ 1 & 3 & 9 & 27 & 10 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

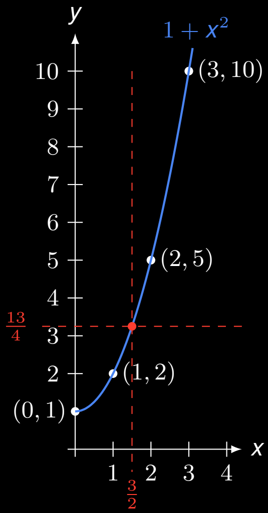
Therefore, $r_0 = 1$, $r_1 = 0$, $r_2 = 1$, $r_3 = 0$, and so

$$p(x) = 1 + x^2.$$

Finally, the estimate is

$$y = p\left(\frac{3}{2}\right) = 1 + \left(\frac{3}{2}\right)^2 = \frac{13}{4}.$$

$\square$



y
$1 + x^2$
10
9
8
7
6
5
4
$\frac{13}{4}$
3
2
$(3, 10)$
$(2, 5)$
$(1, 2)$
$(0, 1)$
x
1
2
3
4
$\frac{3}{2}$

Thm (Polynomial Interpolation)

Given n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ with the x_i distinct, there is a unique polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}$$

such that $p(x_i) = y_i$ for $i = 1, 2, \ldots, n$.

The polynomial $p(x)$ is called the *interpolating polynomial* for the data.

To find $p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}$, set up a system of n linear equations in the n variables $r_0, r_1, r_2, \ldots, r_{n-1}$.

$$\begin{aligned}
r_0 + r_1 x_1 + r_2 x_1^2 + \cdots + r_{n-1} x_1^{n-1} &= y_1 \\
r_0 + r_1 x_2 + r_2 x_2^2 + \cdots + r_{n-1} x_2^{n-1} &= y_2 \\
r_0 + r_1 x_3 + r_2 x_3^2 + \cdots + r_{n-1} x_3^{n-1} &= y_3 \\
\vdots \qquad\qquad &\vdots \quad \vdots \\
r_0 + r_1 x_n + r_2 x_n^2 + \cdots + r_{n-1} x_n^{n-1} &= y_n
\end{aligned}$$

The coefficient matrix for this system is

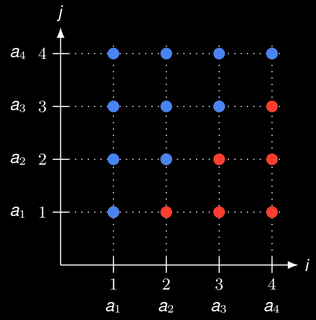
$$\begin{bmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\
1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x_n & x_n^2 & \cdots & x_n^{n-1}
\end{bmatrix}$$

- Such matrix is called Vandermonde matrix.
- Its determinant is called Vandermonde determinant.

Thm (Vandermonde Determinant)

Let $a_1, a_2, \ldots, a_n$ be real numbers, $n \geq 2$. The corresponding Vandermonde determinant is

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \prod_{1 \leq j < i \leq n} (a_i - a_j).$$



Proof. We will prove this by induction. It is clear that when $n = 2$,

$$\det \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \end{pmatrix} = a_2 - a_1 = \prod_{1 \le j < i \le 2} (a_i - a_j).$$

Assume that it is true for $n - 1$. Now let's consider the case n. Denote

$$p(x) := \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \\ 1 & x & x^2 & \cdots & x^{n-1} \end{bmatrix}.$$

Because $p(a_1) = \cdots = p(a_{n-1}) = 0$ (why?), $p(x)$ has to take the following form:

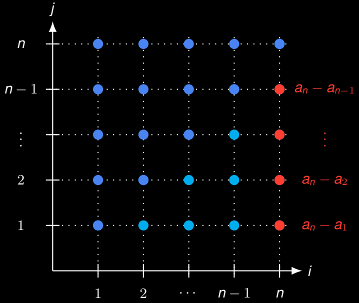
$$p(x) = c(x - a_1)(x - a_2)\cdots(x - a_{n-1}).$$

To identify the constant c, notice that c is the coefficient for x^{n-1}. By cofactor expansion of the determinant along the last row,

$$\begin{aligned} c &= (-1)^{n+n} \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \end{bmatrix} \\ &= \prod_{1 \le j < i \le n-1} (a_i - a_j). \end{aligned}$$

Hence,

$$p(a_n) = \left(\prod_{1 \le j < i \le n-1} (a_i - a_j) \right) \times (a_n - a_1)(a_n - a_2) \cdots (a_n - a_{n-1})$$



$$p(a_n) = \prod_{1 \le j < i \le n} (a_i - a_j).$$

□

E.g. In our earlier example with the data points $(0,1)$, $(1,2)$, $(2,5)$ and $(3,10)$, we have

$$a_1 = 0, \quad a_2 = 1, \quad a_3 = 2, \quad a_4 = 3$$

giving us the *Vandermonde* determinant

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{vmatrix}$$

According to the previous theorem, this determinant is equal to

$$\begin{aligned} &(a_2 - a_1)(a_3 - a_1)(a_3 - a_2)(a_4 - a_1)(a_4 - a_2)(a_4 - a_3) \\ =&(1-0)(2-0)(2-1)(3-0)(3-1)(3-2) \\ =&2 \times 3 \times 2 \\ =&12. \end{aligned}$$

Corr. The Vandermonde determinant is nonzero if $a_1, a_2, \ldots, a_n$ are distinct.

This means that given n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ with distinct x_i, then there is a unique interpolating polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}.$$