Lecture 1. Intro, systems of linear equations, matrices

Aleks Avdiushenko

Neapolis University Paphos

May 10, 2023

First data scientist

First data scientist

Mathematician, mechanic, physicist, astronomer

First data scientist

Mathematician, mechanic, physicist, astronomer

At the age of 24 in 1801, he predicted where to look for the dwarf planet Ceres, hidden behind the Sun

Course program

Course program

Linear algebra

- **Systems of linear equations, matrices**, operations, block operations, reversibility and non-degeneracy
- **Determinants** (3 approaches), oriented volumes, explicit inverse matrix formulas, characteristic polynomial, polynomial calculus in matrices, spectrum, Hamilton-Cayley theorem
- **Vector spaces** and subspaces, dimensions, matrix ranks: row rank, column rank, factorial rank, tensor rank, minor rank. Properties of ranks and inequalities on ranks
- **Linear mappings** and their matrix description, change of coordinates. The image and the kernel, their geometric meaning, connection to the dimension. Linear operator invariants: trace, determinant, characteristic polynomial. Eigenvalues and vectors, connection with the spectrum A note about complex numbers. Diagonalizability and related matrix expansions
- **Bilinear forms**. Quadratic forms and symmetric bilinear forms Signature, its geometric meaning, methods for determining the signature. Relationship with LU-decomposition. Dot products, angles and distances. Orthogonalization and QR-decomposition. Linear manifolds and linear classifiers, margins
- **Operators in Euclidean spaces**. Motions and orthogonal matrices and their classification. Self-adjoint operators and symmetric matrices, their diagonalizability. Singular value decomposition (SVD). Finding SVD

Probability theory

- **Probability space**, random events, how to understand them. Probability (measure) and conditional probability, independence of events, geometric meaning. Bayes formulas and full probability
- **Random variables**, how to understand them. Distribution functions, probabilities (measures) on the line and how to set them. Classes of distributions, examples of discrete and continuous distributions. Joint distribution. Characteristics of random variables: mathematical expectation, variance, moments, median (in a good case). Normal or Gaussian distribution
- **Random vector** or multidimensional random variable, how to understand and set them. Classes of distributions, examples of discrete and continuous distributions. Recovery of distributions of coordinates. Mathematical expectation and covariance matrix. Independence of random variables. Properties of mathematical expectation and dispersion for independent random variables
- **Conditional mathematical expectations and probabilities**. Bayes formulas and full probability for the continuous case. Distribution of the sum of independent random variables and convolution of densities. Multivariate Gaussian distribution

Mathematical statistics

- **The basic model of mathematical statistics** (how to relate the formalism of probability theory to sample measurements). Estimates and their properties. Why convergence and limit theorems are needed. Types of convergences and the relationship between them. Laws of large numbers and Chebyshev's inequality. Sample mean and sample variance, sample covariance matrix, correlation coefficient. Maximum likelihood method. PCA and SVD. Central limit theorem and the Berry-Essen inequality
- **Generating a random sample**. Probability integral transformation. Direct and non-direct methods. Accept-Reject algorithm. MCMC algorithms, Metropolis Algorithm
- **Hypothesis testing**. Formal problem of hypothesis testing. Differences of parametric and non-parametric tests. Confidence intervals. Examples of hypotheses and tests
- **Parametric and non-parametric tests**. Type-I, Type-II errors, p-value. Choosing the methods. Student-T and U-Mann-Whitney. Normality tests, Shapiro-Wilk W-Test
- **ANOVA family**. Formal problem. ANOVA assumptions. Theoretical basis of one way ANOVA. Two-way ANOVA. N-way ANOVA, non-parametric ANOVA, ANCOVA
- **Bootstrap**. Theory and practice
- **Introduction in Bayesian Statistics**. Theory and practice

Homework Assignments and Grading Rules

- The course consists of 13 classes. There will be a homework assignment for each class. The deadline for each homework assignment is 10 days from the date of publication (the deadline will be indicated in the assignment).
- If an assignment is submitted after the deadline, the final grade is calculated using the formula $O_{\text{final}} = 0.7^t O_{\text{hw}}$, where t is the time after the deadline in days without rounding, O_{hw} is the grade for the homework assignment if it had been submitted on time, and O_{final} is the grade awarded for the late homework assignment.
- Homework assignments must be submitted in written form (handwritten or using LaTeX, virtual boards are also acceptable). The work should be submitted as a single multi-page PDF file (you can use the `notebloc` app on your mobile phone, as it does a good job of whitening the background and produces photos of acceptable size). All pages must be vertically oriented and in sequential order. Please make sure to adhere to this requirement. We would be very grateful for your cooperation.
- The rules for determining the final grade will be announced later, but by default, it is assumed that all 13 homework assignments must be completed. The threshold value for passing the course will be announced later.
- As a midterm there will be a offline written test with tasks similar to homework assignments.

Why do you need linear algebra?

Consider Kirchhoff's Circuit Laws for example.

- Two fundamental laws for electric circuits:
  1. Kirchhoff's Current Law (KCL)
  2. Kirchhoff's Voltage Law (KVL)

Why do you need linear algebra?

Consider Kirchhoff's Circuit Laws for example.

- Two fundamental laws for electric circuits:
  1. Kirchhoff's Current Law (KCL)
  2. Kirchhoff's Voltage Law (KVL)
- KCL: the sum of currents entering a node equals the sum of currents leaving the node

Why do you need linear algebra?

Consider Kirchhoff's Circuit Laws for example.

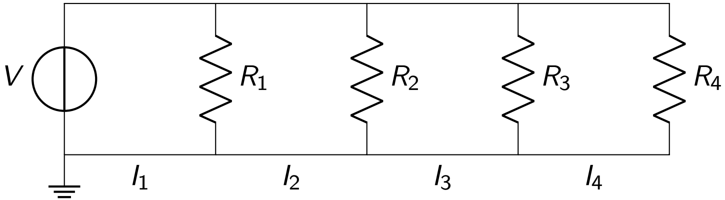
- Two fundamental laws for electric circuits:
  1. Kirchhoff's Current Law (KCL)
  2. Kirchhoff's Voltage Law (KVL)
- KCL: the sum of currents entering a node equals the sum of currents leaving the node
- KVL: the sum of the voltage differences (potential drops) around any closed loop or mesh in a network is zero

Why do you need linear algebra?

Consider Kirchhoff's Circuit Laws for example.

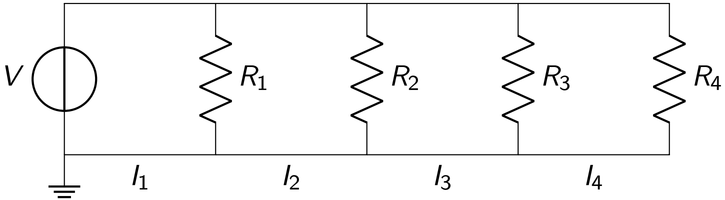
- Two fundamental laws for electric circuits:
  1. Kirchhoff's Current Law (KCL)
  2. Kirchhoff's Voltage Law (KVL)
- KCL: the sum of currents entering a node equals the sum of currents leaving the node
- KVL: the sum of the voltage differences (potential drops) around any closed loop or mesh in a network is zero
- Applying KCL and KVL to a circuit leads to a system of linear algebraic equations

Example Circuit



$I_1, \ldots, I_4 \in \mathbb{R}$ are variables, R_i — resistances

Example Circuit



$I_1, \ldots, I_4 \in \mathbb{R}$ are variables, R_i — resistances

$$
\begin{aligned}
R_1 I_1 - R_1 I_2 &= V \\
R_2 I_2 - R_2 I_3 &= V \\
R_3 I_3 - R_3 I_4 &= V \\
R_4 I_4 &= V
\end{aligned}
$$

System of linear equations in general form

$x_1, \ldots, x_n \in \mathbb{R}$ are variables

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = & b_2 \\ & & \vdots & & & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\ldots & +a_{nn}x_n & = & b_n \end{cases} \qquad a_{ij}, b_i \in \mathbb{R}$$

System of linear equations in general form

$x_1, \dots, x_n \in \mathbb{R}$ are variables

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = & b_2 \\ & & \vdots & & & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\dots & +a_{nn}x_n & = & b_n \end{cases} \qquad a_{ij}, b_i \in \mathbb{R}$$

We can easily check by substitution, if $(c_1, \dots, c_n)$ is solution.

System of linear equations in general form

$x_1, \dots, x_n \in \mathbb{R}$ are variables

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = & b_2 \\ & & \vdots & & & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\dots & +a_{nn}x_n & = & b_n \end{cases} \qquad a_{ij}, b_i \in \mathbb{R}$$

We can easily check by substitution, if $(c_1, \dots, c_n)$ is solution.

Question 1

But how to find all solutions effectively?

System of linear equations in general form

$x_1, \dots, x_n \in \mathbb{R}$ are variables

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = & b_2 \\ & & \vdots & & & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\dots & +a_{nn}x_n & = & b_n \end{cases} \qquad a_{ij}, b_i \in \mathbb{R}$$

We can easily check by substitution, if $(c_1, \dots, c_n)$ is solution.

Question 1

But how to find all solutions effectively?

We can **transform** the system to a form where all solutions become easy to find: $S_1 \to S_2 \to \dots \to S_k$

Elementary Row Transformations

Three kinds of elementary row transformations

Elementary Row Transformations

Three kinds of elementary row transformations

- Interchanging the rows: the entire row R_i is swapped with another row R_j

Three kinds of elementary row transformations

- Interchanging the rows: the entire row R_i is swapped with another row R_j
- Scaling the entire row with a non zero number: $R_i \to \lambda R_i$

Elementary Row Transformations

Three kinds of elementary row transformations

- Interchanging the rows: the entire row R_i is swapped with another row R_j
- Scaling the entire row with a non zero number: $R_i \to \lambda R_i$
- Add one row to another row multiplied by a non zero number: $R_i \to R_i + \lambda R_j$

A matrix can always be transformed into an **upper triangular matrix** using elementary row transformations.

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 2 & 1 & -3 & 5 \\ -4 & 7 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & 10 & 10 \end{bmatrix}$$

In genral rectangular case it is called row echelon form.

Demo Video

Youtube link

Principal and free variables

$$
\begin{array}{ccccc}
\boldsymbol{x}_1 & & \boldsymbol{x}_3 & & \\
\multicolumn{5}{c}{\left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 3 \\ 0 & 0 & 1 & 5 & 4 \end{array}\right]} \\
 & x_2 & & x_4 &
\end{array}
$$

Principal and free variables

$$
\begin{array}{ccccc}
\boldsymbol{x}_1 & & \boldsymbol{x}_3 & & \\
\multicolumn{5}{c}{\left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 3 \\ 0 & 0 & 1 & 5 & 4 \end{array}\right]} \\
 & x_2 & & x_4 &
\end{array}
$$

Question 2

How many principal variables is for different values x?

$$
\left[\begin{array}{cccc|c}
x & 1 & 1 & 1 & 0 \\
1 & x & \dots & 1 & 0 \\
1 & \dots & \ddots & 1 & 0 \\
1 & 1 & 1 & x & 0
\end{array}\right]
$$

Solution scheme and answers

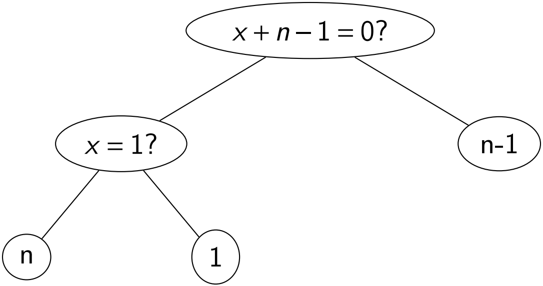
Hint!

Try to add all rows to the first.

Solution scheme and answers

Hint!

Try to add all rows to the first.



Homogeneous system

Right hand side equals zero:

$$\underbrace{\left[\begin{array}{c|c} R_1 & 0 \\ R_2 & 0 \\ \vdots & 0 \\ R_m & 0 \end{array}\right]}_{n+1}$$

Homogeneous system

Right hand side equals zero:

$$\underbrace{\left[\begin{array}{c|c} R_1 & 0 \\ R_2 & 0 \\ \vdots & 0 \\ R_m & 0 \end{array}\right]}_{n+1}$$

Fact

$n > m \Rightarrow \exists$ nontrivial solution

Operations on matrices

$M_{mn}(\mathbb{R}):$

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad a_{ij} \in \mathbb{R}$$

Operations

1. Addition and substraction — component wise
2. Multiplication by a number — component wise

3. Matrix multiplication

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. The product $C = AB$ of A and B has $m \times p$ shape and is given by:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{m \times n} \underbrace{\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}}_{n \times p} = \underbrace{\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{bmatrix}}_{m \times p}$$

where

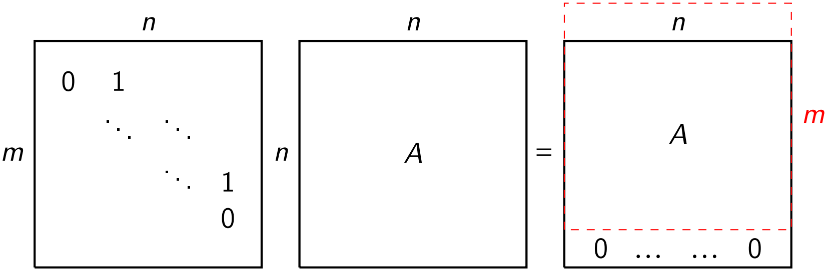
$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

for $i = 1, \dots, m$ and $j = 1, \dots, p$

First example

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \end{array} \begin{array}{cc} & n \\ n & \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \end{array} = \begin{array}{c} n \\ \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \end{array}$$

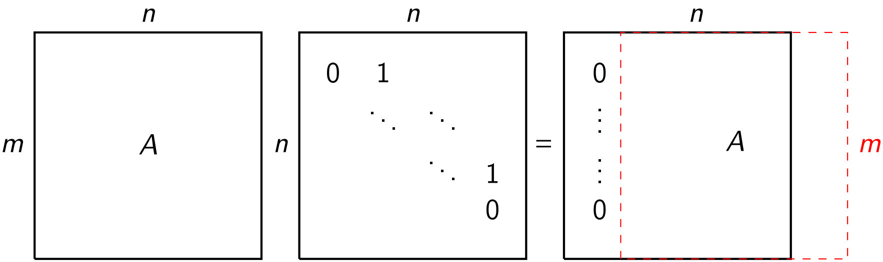
First example



Second example

$$\begin{array}{cc} & n \\ m & \boxed{\quad A \quad} \end{array} \quad \begin{array}{cc} & n \\ n & \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \end{array} = \begin{array}{c} n \\ \boxed{\qquad\quad} \end{array}$$

Second example



Mnemonic rule

Multiplication from the **right** side operates on **columns**, and on the **left** on **rows**.

Mnemonic rule

Multiplication from the **right** side operates on **columns**, and on the **left** on **rows**.

Question 3

What if we multiply on the diagonal matrix?

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}$$

Properties of matrix operations

1. $(AB)C = A(BC)$ — associativity
2. $A(B+C) = AB + AC$ — distributivity
3. $(A+B)C = AC + BC$

Properties of matrix operations

1. $(AB)C = A(BC)$ — associativity
2. $A(B + C) = AB + AC$ — distributivity
3. $(A + B)C = AC + BC$
4. $AB \neq BA$ (and even it is not possible sometimes)

Properties of matrix operations

1. $(AB)C = A(BC)$ — associativity
2. $A(B+C) = AB + AC$ — distributivity
3. $(A+B)C = AC + BC$
4. $AB \neq BA$ (and even it is not possible sometimes)
5. Zero divisors: $AB = 0$, when $A \neq 0$ and $B \neq 0$
6. Nilpotents: $A \neq 0$, but $A^k = 0$

Properties of matrix operations

1. $(AB)C = A(BC)$ — associativity
2. $A(B+C) = AB + AC$ — distributivity
3. $(A+B)C = AC + BC$
4. $AB \neq BA$ (and even it is not possible sometimes)
5. Zero divisors: $AB = 0$, when $A \neq 0$ and $B \neq 0$
6. Nilpotents: $A \neq 0$, but $A^k = 0$
7. $Ax = b$ — matrix form of system of linear equations

Properties of matrix operations

1. $(AB)C = A(BC)$ — associativity
2. $A(B+C) = AB + AC$ — distributivity
3. $(A+B)C = AC + BC$
4. $AB \neq BA$ (and even it is not possible sometimes)
5. Zero divisors: $AB = 0$, when $A \neq 0$ and $B \neq 0$
6. Nilpotents: $A \neq 0$, but $A^k = 0$
7. $Ax = b$ — matrix form of system of linear equations

Question 4

For which matrices A the equalty $A\Lambda = \Lambda A$ is true, where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_k \end{bmatrix}$$

Block formula

Blocks as «numbers»:

A	B
C	D

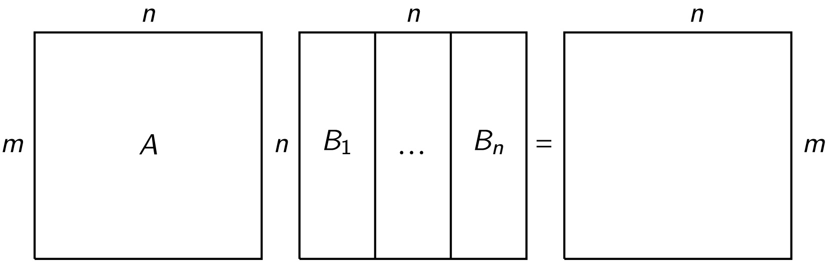
X	Y
Z	W

Block formula

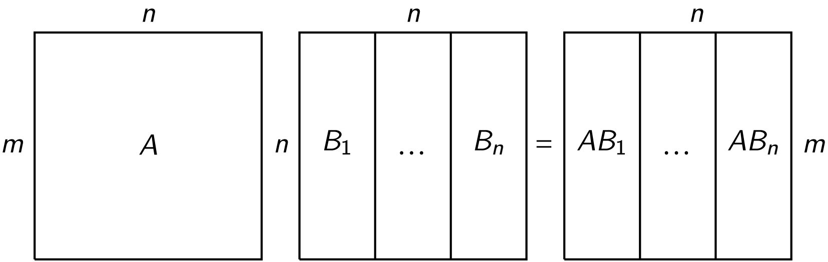
Blocks as «numbers»:

$$\begin{array}{|c|c|}\hline A & B \\ \hline C & D \\ \hline \end{array} \; \begin{array}{|c|c|}\hline X & Y \\ \hline Z & W \\ \hline \end{array} = \begin{array}{|c|c|}\hline AX + BZ & AY + BW \\ \hline CX + DZ & CY + DW \\ \hline \end{array}$$

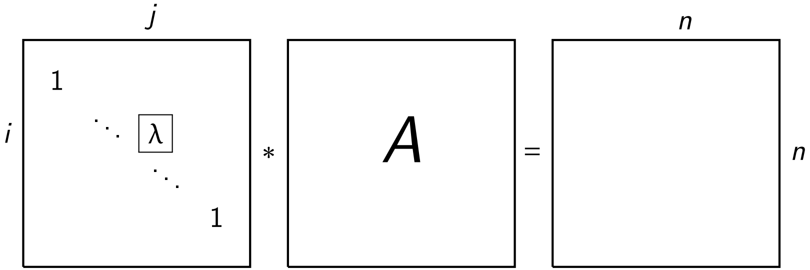
Block example



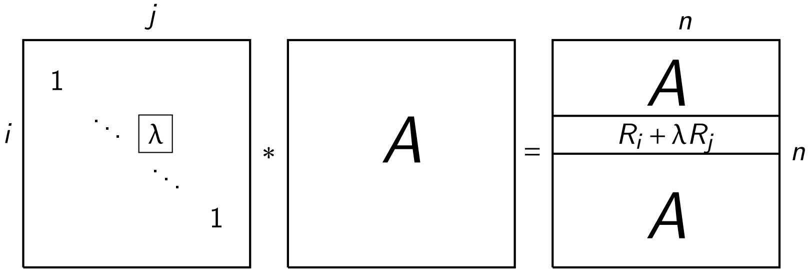
Block example



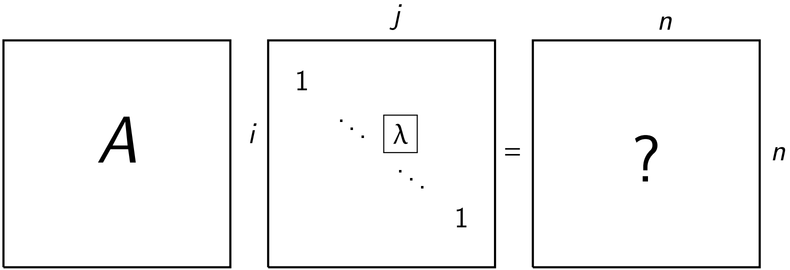
Elementary Row Transformations as Matrix Operations



Elementary Row Transformations as Matrix Operations



Elementary Row Transformations as Matrix Operations



Elementary Row Transformations as Matrix Operations

$$\boxed{\;A\;} \quad \begin{array}{cc} & \begin{array}{cccc} & & j & \end{array} \\ i & \boxed{\begin{array}{cccc} 1 & & & \\ & \ddots & \boxed{\lambda} & \\ & & \ddots & \\ & & & 1 \end{array}} \end{array} = \begin{array}{cc} n & \\ \boxed{\;?\;} & n \end{array}$$

Question 5

What if we transpose the elementary matrix?

Elementary swap of rows and columns

$$U_{ij} = \boxed{\begin{matrix} 1 & & & & \\ & 0 & & 1 & \\ & & \ddots & & \\ & 1 & & 0 & \\ & & & & 1 \end{matrix}}$$

Matrix division

Inverse matrix

$A \in M_n(\mathbb{R})$ is reversible $\iff \exists B \in M_n(\mathbb{R}) : AB = BA = E$

By definition $B = A^{-1}$

Matrix division

Inverse matrix

$A \in M_n(\mathbb{R})$ is reversible $\iff \exists B \in M_n(\mathbb{R}) : AB = BA = E$

By definition $B = A^{-1}$

Question 6

Prove that the inverse matrix is unique.

Matrix division

Inverse matrix

$A \in M_n(\mathbb{R})$ is reversible $\iff \exists B \in M_n(\mathbb{R}) : AB = BA = E$

By definition $B = A^{-1}$

Question 6

Prove that the inverse matrix is unique.

$$\begin{array}{cc} & j \\ i & \begin{bmatrix} 1 & & & \\ & \ddots & \boxed{\lambda} & \\ & & \ddots & \\ & & & 1 \end{bmatrix}^{-1} \end{array} = \begin{bmatrix} & & \\ & ? & \\ & & \end{bmatrix}$$

The following statements are equivalent

$A \in M_n(\mathbb{R})$

- $Ax = 0 \Rightarrow x = 0$
- $A^T y = 0 \Rightarrow y = 0$

The following statements are equivalent

$A \in M_n(\mathbb{R})$

- $Ax = 0 \Rightarrow x = 0$
- $A^T y = 0 \Rightarrow y = 0$
- $A = U_1 \dots U_k$ for elementary matrices U_i

The following statements are equivalent

$A \in M_n(\mathbb{R})$

- $Ax = 0 \Rightarrow x = 0$
- $A^T y = 0 \Rightarrow y = 0$
- $A = U_1 \dots U_k$ for elementary matrices U_i
- $\exists A^{-1}$
- $\exists L \in M_n(\mathbb{R}) : LA = E$
- $\exists R \in M_n(\mathbb{R}) : AR = E$

The following statements are equivalent

$A \in M_n(\mathbb{R})$

- $Ax = 0 \Rightarrow x = 0$
- $A^T y = 0 \Rightarrow y = 0$
- $A = U_1 \dots U_k$ for elementary matrices U_i
- $\exists A^{-1}$
- $\exists L \in M_n(\mathbb{R}) : LA = E$
- $\exists R \in M_n(\mathbb{R}) : AR = E$
- $det\, A \neq 0$

The following statements are equivalent

$A \in M_n(\mathbb{R})$

- $Ax = 0 \Rightarrow x = 0$
- $A^T y = 0 \Rightarrow y = 0$
- $A = U_1 \dots U_k$ for elementary matrices U_i
- $\exists A^{-1}$
- $\exists L \in M_n(\mathbb{R}) : LA = E$
- $\exists R \in M_n(\mathbb{R}) : AR = E$
- $det\, A \neq 0$

Consequence

$A, B \in M_n(\mathbb{R})$ A and B are invertible $\Leftrightarrow AB$ is invertible

Inverse Matrix with Gauss-Jordan

Demo video

What about rectangular matrices?

Lemma

If $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$ and $AB = E, BA = E$ then $m = n$.

What about rectangular matrices?

Lemma

If $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$ and $AB = E, BA = E$ then $m = n$.

Definition

The *trace* of a matrix A of order $m \times n$ is the sum of its diagonal elements. It is denoted as $\mathrm{tr}(A)$ or simply $\mathrm{tr}A$

Example

For a matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ where $m < n$, the trace is given by:

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}$$

Trace properties

1. $\mathrm{tr}(A + \lambda B) = \mathrm{tr}(A) + \lambda \mathrm{tr}(B)$
2. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

Lemma

Let $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$.
Prove that $E - AB$ is reversible $\Leftrightarrow$ $E - AB$ is reversible.

Lemma

Let $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$.
Prove that $E - AB$ is reversible $\Leftrightarrow$ $E - AB$ is reversible.

- $E - AB$ is reversible

Lemma

Let $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$.
Prove that $E - AB$ is reversible $\Leftrightarrow$ $E - AB$ is reversible.

- $E - AB$ is reversible
- $(E - AB)x = 0 \Rightarrow x = 0$

Lemma

Let $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$.
Prove that $E - AB$ is reversible $\Leftrightarrow$ $E - AB$ is reversible.

- $E - AB$ is reversible
- $(E - AB)x = 0 \Rightarrow x = 0$
- $\forall x: \; x = ABx \Rightarrow x = 0$

Lemma

Let $A \in M_{mn}(\mathbb{R}), B \in M_{nm}(\mathbb{R})$.
Prove that $E - AB$ is reversible $\Leftrightarrow$ $E - AB$ is reversible.

- $E - AB$ is reversible
- $(E - AB)x = 0 \Rightarrow x = 0$
- $\forall x: \; x = ABx \Rightarrow x = 0$
- From second $y = BAy \Rightarrow Ay = ABAy \underset{x=Ay}{\Rightarrow} x = ABx$

Example

$$E + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} -1 & -2 & -3 \end{bmatrix}$$

Polynomial matrix calculus

$A \in M_n(\mathbb{R})$
$p(x) = a_0 + a_1 x + \dots + a_k x^k$
$p(A) = a_0 E + a_1 A + \dots + a_k A^k$

Polynomial matrix calculus

$A \in M_n(\mathbb{R})$
$p(x) = a_0 + a_1 x + \cdots + a_k x^k$
$p(A) = a_0 E + a_1 A + \cdots + a_k A^k$

Vanishing polynomial

$\forall A \in M_n(\mathbb{R})\ \exists$ polynomial $p(x) \neq 0: \ p(A) = 0$

$\deg p \leq n^2$ as we get n^2 equations on $k+1$ variables

Example

$A \in M_n(\mathbb{R}),\ \ p(t) = t^4 + 5t - 2$

$p(A) = 0$

Example

$A \in M_n(\mathbb{R}),\ p(t) = t^4 + 5t - 2$

$p(A) = 0$

$$A^4 + 5A - 2E = 0$$

$$A^4 + 5A = 2E$$

$$\frac{1}{2}(A^3 + 5E)A = E$$

$$\boxed{A^{-1} = \frac{1}{2}(A^3 + 5E)}$$

Minimal vanishing polynomial

- $p_{min}(A) = 0$
- senior coefficient $= 1$
- deg is minimal

Minimal vanishing polynomial

- $p_{min}(A) = 0$
- senior coefficient $= 1$
- deg is minimal

Remark

p_{min} is unique

Minimal vanishing polynomial

- $p_{min}(A) = 0$
- senior coefficient $= 1$
- deg is minimal

Remark

p_{min} is unique

Property

$p(x)$ is polynomial, C is invertible: $p(CAC^{-1}) = Cp(A)C^{-1}$

Spectrum

$A \in M_n(\mathbb{R})$

Definition

$\mathrm{spec}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ is irreversible}\}$

Spectrum

$A \in M_n(\mathbb{R})$

Definition

$\mathrm{spec}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ is irreversible}\}$

Example

$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \Rightarrow \mathrm{spec}\, A = \{\lambda_1, \ldots, \lambda_n\}$

Statements

Spectrum

$A \in M_n(\mathbb{R})$

Definition

$\mathrm{spec}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ is irreversible}\}$

Example

$A = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow \mathrm{spec}\, A = \{\lambda_1, \dots, \lambda_n\}$

Statements

1. $\mathrm{spec}_{\mathbb{R}} A = \{\text{real roots of } p_{min}\}$

$A \in M_n(\mathbb{R})$

Definition

$\mathrm{spec}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ is irreversible}\}$

Example

$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \Rightarrow \mathrm{spec}\, A = \{\lambda_1, \ldots, \lambda_n\}$

Statements

1. $\mathrm{spec}_{\mathbb{R}} A = \{\text{real roots of } p_{min}\}$
2. for all vanishing polynomials $g(A) = 0 \Rightarrow p_{min} \mid g$

Spectrum

$A \in M_n(\mathbb{R})$

Definition

$\operatorname{spec}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ is irreversible}\}$

Example

$A = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow \operatorname{spec} A = \{\lambda_1, \dots, \lambda_n\}$

Statements

1. $\operatorname{spec}_{\mathbb{R}} A = \{\text{real roots of } p_{min}\}$
2. for all vanishing polynomials $g(A) = 0 \Rightarrow p_{min} \mid g$
3. $g(A) = 0 \Rightarrow \operatorname{spec}_{\mathbb{R}} A \subseteq \{\text{real roots of } g\}$

Example

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad A^2 = -E$$

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

is vanishing polynomial

$$\mathrm{spec}_{\mathbb{R}} A \subseteq \{\text{real roots of } x^2 + 1\}$$
$$\mathrm{spec}_{\mathbb{C}} A \subseteq \{\text{complex roots of } x^2 + 1\}$$

Task

$A \in M_{mn}(\mathbb{R}),\ B \in M_{nm}(\mathbb{R})$

$\mathrm{spec}(AB) \cup \{0\} = \mathrm{spec}(BA) \cup \{0\}$

If $m = n$ $\mathrm{spec}(AB) = \mathrm{spec}(BA)$