Linear Algebra

Lecture 1 - Solving Linear Equations

Oskar Kędzierski

2 October 2023

Other Details

email: oskar@mimuw.edu.pl,

Other Details

email: oskar@mimuw.edu.pl,
website:
`https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html`

Other Details

email: oskar@mimuw.edu.pl,
website:
`https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html`
60 problems for classes:
`http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf`

Other Details

email: oskar@mimuw.edu.pl,
website:
https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html
60 problems for classes:
http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf
address: room 5230, the Faculty of Mathematics, Informatics, and Mechanics, ul. Banacha 2, Warsaw.

Other Details

email: oskar@mimuw.edu.pl,
website:
https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html
60 problems for classes:
http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf
address: room 5230, the Faculty of Mathematics, Informatics, and Mechanics, ul. Banacha 2, Warsaw.

By $\mathbb{R}$ we will denote the real numbers, for example $-1, 0, \sqrt{2}, 3, \pi \in \mathbb{R}$.

Other Details

email: oskar@mimuw.edu.pl,
website:
`https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html`
60 problems for classes:
`http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf`
address: room 5230, the Faculty of Mathematics, Informatics, and Mechanics, ul. Banacha 2, Warsaw.

By $\mathbb{R}$ we will denote the real numbers, for example $-1, 0, \sqrt{2}, 3, \pi \in \mathbb{R}$.
By $\mathbb{R}^n$ we will denote the n-tuples of real numbers.

Other Details

email: oskar@mimuw.edu.pl,
website:
`https://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2023.html`
60 problems for classes:
`http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf`
address: room 5230, the Faculty of Mathematics, Informatics, and Mechanics, ul. Banacha 2, Warsaw.

By $\mathbb{R}$ we will denote the real numbers, for example $-1, 0, \sqrt{2}, 3, \pi \in \mathbb{R}$.
By $\mathbb{R}^n$ we will denote the n-tuples of real numbers. For example, the 3-tuple, $(1, -2, 4) \in \mathbb{R}^3$.

Bibliography

i) K. Hoffman, R. Kunze, *Linear Algebra*

Bibliography

i) K. Hoffman, R. Kunze, *Linear Algebra*

ii) G. Strang, *Linear Algebra and its Applications*

Bibliography

i) K. Hoffman, R. Kunze, *Linear Algebra*

ii) G. Strang, *Linear Algebra and its Applications*

iii) M. Artin, *Algebra*

Bibliography

i) K. Hoffman, R. Kunze, *Linear Algebra*

ii) G. Strang, *Linear Algebra and its Applications*

iii) M. Artin, *Algebra*

iv) T. Koźniewski, *Wykłady z algebry liniowej* (in Polish)

Bibliography

i) K. Hoffman, R. Kunze, *Linear Algebra*

ii) G. Strang, *Linear Algebra and its Applications*

iii) M. Artin, *Algebra*

iv) T. Koźniewski, *Wykłady z algebry liniowej* (in Polish)

v) P. D. Lax, *Linear Algebra and Its Applications*

Bibliography

Simple introduction to Linear Algebra

i) C. W. Curtis, *Linear Algebra: An Introductory Approach*, Springer

Bibliography

Simple introduction to Linear Algebra

i) C. W. Curtis, *Linear Algebra: An Introductory Approach*, Springer

ii) G. Schay, *A Concise Introduction to Linear Algebra*, Springer

Bibliography

Simple introduction to Linear Algebra

i) C. W. Curtis, *Linear Algebra: An Introductory Approach*, Springer

ii) G. Schay, *A Concise Introduction to Linear Algebra*, Springer

Linear programming

Bibliography

Simple introduction to Linear Algebra

i) C. W. Curtis, *Linear Algebra: An Introductory Approach*, Springer

ii) G. Schay, *A Concise Introduction to Linear Algebra*, Springer

Linear programming

i) J. Matousek, B. Gärtner, *Understanding and Using Linear Programming*, Springer

Bibliography

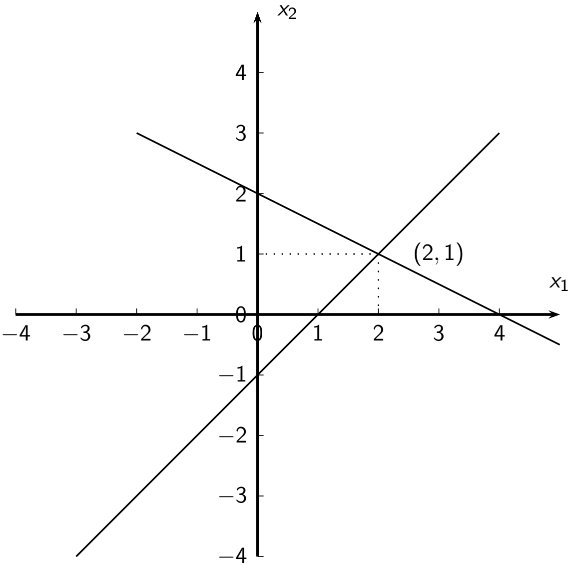
Simple introduction to Linear Algebra

i) C. W. Curtis, *Linear Algebra: An Introductory Approach*, Springer
ii) G. Schay, *A Concise Introduction to Linear Algebra*, Springer

Linear programming

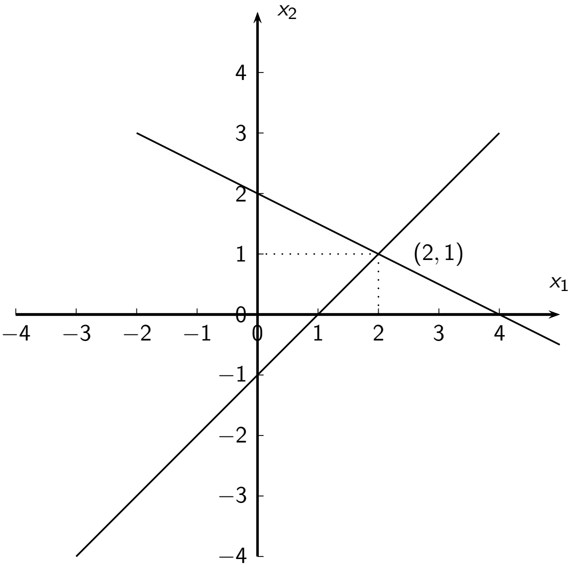
i) J. Matousek, B. Gärtner, *Understanding and Using Linear Programming*, Springer
ii) R. J. Vanderbei, *Linear Programming: Foundations and Extensions*, Springer

You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + 2x_2 = 4 \end{cases}$$

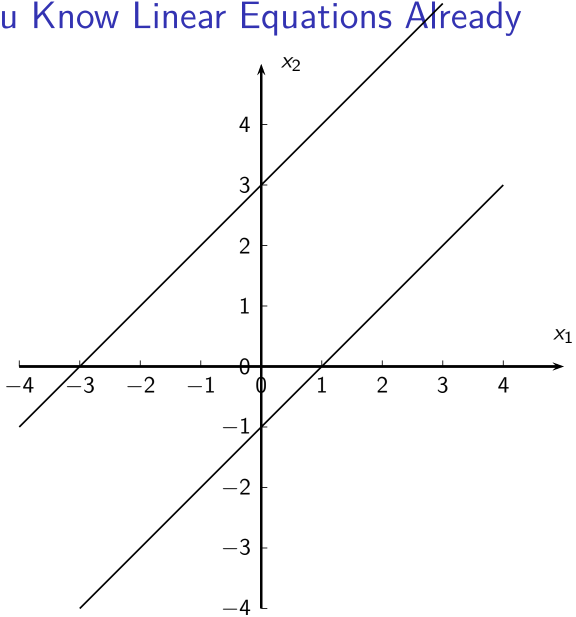
You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + 2x_2 = 4 \end{cases}$$

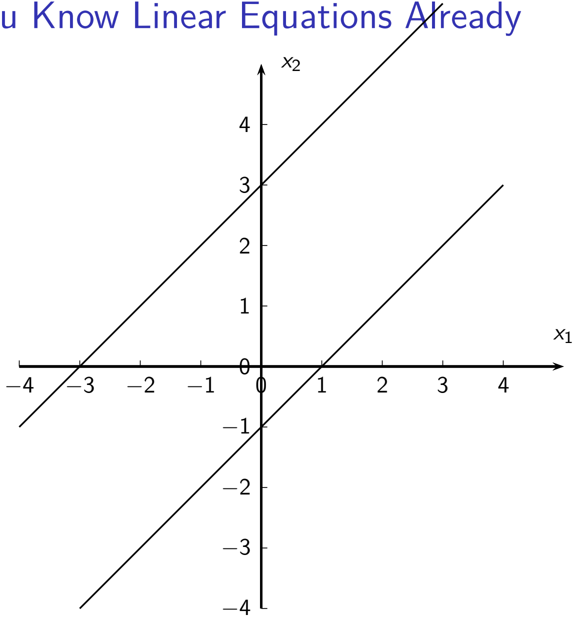
Exactly one solution $(2, 1)$

You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ x_1 - x_2 = -3 \end{cases}$$

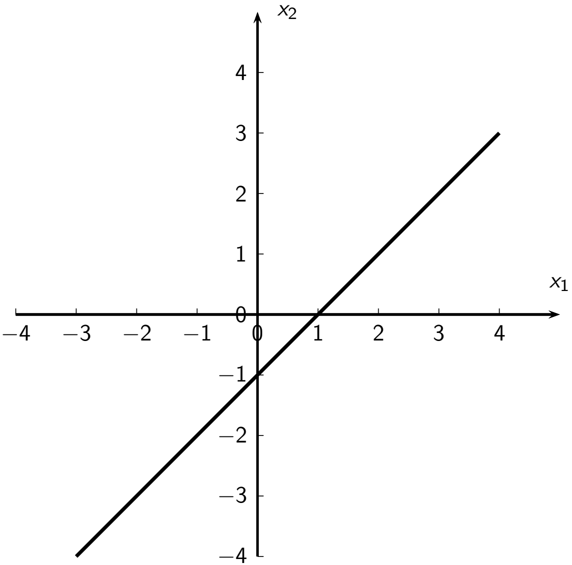
You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ x_1 - x_2 = -3 \end{cases}$$

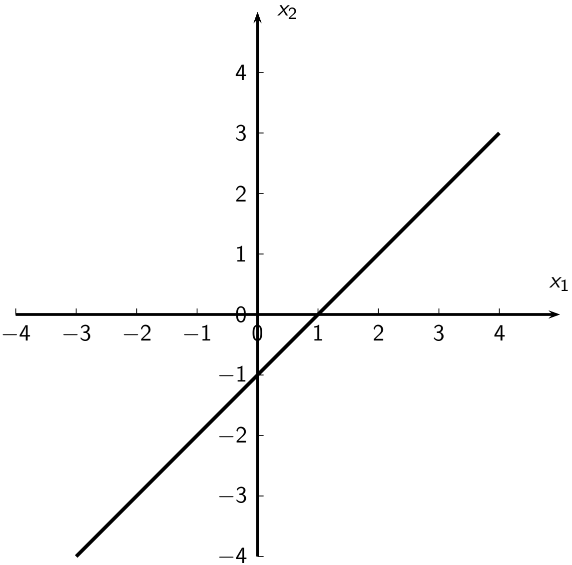
No solutions at all

You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ 2x_1 - 2x_2 = 2 \end{cases}$$

You Know Linear Equations Already



$$\begin{cases} x_1 - x_2 = 1 \\ 2x_1 - 2x_2 = 2 \end{cases}$$

Infinitely many solutions of the form $(x_2 + 1, x_2),\ x_2 \in \mathbb{R}$

Linear Equations

Linear equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$

Linear Equations

Linear equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ in n unknowns $x_1, \ldots, x_n$

Linear Equations

Linear equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ in n unknowns $x_1, \ldots, x_n$ with the **coefficients** $a_1, \ldots, a_n \in \mathbb{R}$

Linear Equations

Linear equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ in n unknowns $x_1, \ldots, x_n$ with the **coefficients** $a_1, \ldots, a_n \in \mathbb{R}$ and the **constant term** $b \in \mathbb{R}$.

System of Linear Equations

A system of m linear equations in n unknowns $x_1, \dots, x_n$

$$U : \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = b_m \end{cases}$$

System of Linear Equations

A system of m linear equations in n unknowns $x_1, \ldots, x_n$

$$U : \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = b_m \end{cases}$$

with coefficients $a_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n$ and constant terms $b_i \in \mathbb{R}$ for $i = 1, \ldots, m$.

System of Linear Equations

A system of m linear equations in n unknowns $x_1, \ldots, x_n$

$$U : \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = b_m \end{cases}$$

with coefficients a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ and constant terms $b_i \in \mathbb{R}$ for $i = 1, \ldots, m$. If $b_1 = b_2 = \ldots = b_m = 0$ we call the system **homogeneous.**

Solutions of Systems of Linear Equations

Any n-tuple $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that all equations in U are satisfied is called a **solution** of the system U.

Solutions of Systems of Linear Equations

Any n-tuple $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that all equations in U are satisfied is called a **solution** of the system U. For example, the n-tuple $(0, \ldots, 0)$ is a solution of a homogenous system of linear equations.

Solutions of Systems of Linear Equations

Any n-tuple $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that all equations in U are satisfied is called a **solution** of the system U. For example, the n-tuple $(0, \ldots, 0)$ is a solution of a homogenous system of linear equations.

A system with no solutions is called **inconsistent.**

Solutions of Systems of Linear Equations

Any n-tuple $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that all equations in U are satisfied is called a **solution** of the system U. For example, the n-tuple $(0, \ldots, 0)$ is a solution of a homogenous system of linear equations.

A system with no solutions is called **inconsistent.** Two systems of linear equations are called **equivalent** if they have the same sets of solutions.

Operations on Equations

Any equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ can be **multiplied** by a non-zero constant $c \in \mathbb{R} - \{0\}$ in order to get the equation $ca_1x_1 + ca_2x_2 + \ldots + ca_nx_n = cb$.

Operations on Equations

Any equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ can be **multiplied** by a non-zero constant $c \in \mathbb{R} - \{0\}$ in order to get the equation $ca_1x_1 + ca_2x_2 + \ldots + ca_nx_n = cb$.

One can add any two equations $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,\ a'_1x_1 + a'_2x_2 + \ldots + a'_nx_n = b'$ and get the equation $(a_1 + a'_1)x_1 + (a_2 + a'_2)x_2 + \ldots + (a_n + a'_n)x_n = b + b'$.

Equivalent System of Linear Equations

Theorem

The following operations on a system of linear equations do not change the set of its solutions (i.e. they lead to an equivalent system):

Equivalent System of Linear Equations

Theorem

The following operations on a system of linear equations do not change the set of its solutions (i.e. they lead to an equivalent system):

i) *swapping the order of any two equations,*

Equivalent System of Linear Equations

Theorem

The following operations on a system of linear equations do not change the set of its solutions (i.e. they lead to an equivalent system):

i) *swapping the order of any two equations,*

ii) *multiplying any equation by a non-zero constant,*

Equivalent System of Linear Equations

Theorem

The following operations on a system of linear equations do not change the set of its solutions (i.e. they lead to an equivalent system):

i) *swapping the order of any two equations,*
ii) *multiplying any equation by a non-zero constant,*
iii) *adding an equation to the other.*

Equivalent System of Linear Equations

Theorem

The following operations on a system of linear equations do not change the set of its solutions (i.e. they lead to an equivalent system):

i) *swapping the order of any two equations,*

ii) *multiplying any equation by a non-zero constant,*

iii) *adding an equation to the other.*

Proof.

Any solution of the original system is a solution of the new system. All above operations are reversible. $\square$

A General Solution

A **general solution** of the system of linear equation U is an equivalent linear system U' of the form:

$$U' : \begin{cases} x_{j_1} = c_{11}x_1 & + & c_{12}x_2 & + & \dots & + & c_{1n}x_n & +d_1 \\ x_{j_2} = c_{21}x_1 & + & c_{22}x_2 & + & \dots & + & c_{2n}x_n & +d_2 \\ \quad\vdots & & \vdots & & \ddots & & \vdots & \vdots \\ x_{j_k} = c_{k1}x_1 & + & c_{k2}x_2 & + & \dots & + & c_{kn}x_n & +d_k \end{cases}$$

A General Solution

A **general solution** of the system of linear equation U is an equivalent linear system U' of the form:

$$U' : \begin{cases} x_{j_1} = c_{11}x_1 & + & c_{12}x_2 & + & \dots & + & c_{1n}x_n & +d_1 \\ x_{j_2} = c_{21}x_1 & + & c_{22}x_2 & + & \dots & + & c_{2n}x_n & +d_2 \\ \qquad \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ x_{j_k} = c_{k1}x_1 & + & c_{k2}x_2 & + & \dots & + & c_{kn}x_n & +d_k \end{cases}$$

where $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$, $j_1 < j_2 < \dots < j_k$ and $c_{ij} = 0$ for any $i = 1, \dots, k$ and $j = j_1, \dots, j_k$. That is, the unknowns $x_{j_1}, \dots, x_{k_k}$ appear only on the **left hand-side** of each equation exactly once.

A General Solution

A **general solution** of the system of linear equation U is an equivalent linear system U' of the form:

$$U' : \begin{cases} x_{j_1} = c_{11}x_1 & + & c_{12}x_2 & + & \dots & + & c_{1n}x_n & +d_1 \\ x_{j_2} = c_{21}x_1 & + & c_{22}x_2 & + & \dots & + & c_{2n}x_n & +d_2 \\ \qquad \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ x_{j_k} = c_{k1}x_1 & + & c_{k2}x_2 & + & \dots & + & c_{kn}x_n & +d_k \end{cases}$$

where $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$, $j_1 < j_2 < \dots < j_k$ and $c_{ij} = 0$ for any $i = 1, \dots, k$ and $j = j_1, \dots, j_k$. That is, the unknowns $x_{j_1}, \dots, x_{k_k}$ appear only on the **left hand-side** of each equation exactly once.

The unknowns $x_{j_1}, \dots, x_{j_k}$ are called **basic** (or **dependent**) **variables**. The other unknowns are called **free variables** or **parameters**.

Matrices

A $m \times n$ **matrix** D with **entries** in $\mathbb{R}$ is a rectangular array of real numbers arranged in m rows and n columns, i.e.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

where $d_{ij} \in \mathbb{R}$.

Matrices

A $m \times n$ **matrix** D with **entries** in $\mathbb{R}$ is a rectangular array of real numbers arranged in m rows and n columns, i.e.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

where $d_{ij} \in \mathbb{R}$. Sometimes we write $D = [d_{ij}]$ for $i = 1, \dots, m$, $j = 1, \dots, n$.

Matrices

A $m \times n$ **matrix** D with **entries** in $\mathbb{R}$ is a rectangular array of real numbers arranged in m rows and n columns, i.e.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

where $d_{ij} \in \mathbb{R}$. Sometimes we write $D = [d_{ij}]$ for $i = 1, \dots, m$, $j = 1, \dots, n$. The set of all **m-by-n** matrices with entries in $\mathbb{R}$ will be denoted $M(m \times n; \mathbb{R})$.

Matrix of a System of Linear Equations

To each system of linear equations

$$U : \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = b_m \end{cases}$$

we associate its $m \times (n+1)$ matrix

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

Matrix of a System of Linear Equations

The **submatrix**

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

is called the **matrix of coefficients**. The last column

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

consists of constant terms.

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

i) swapping any two rows of the matrix,

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

i) swapping any two rows of the matrix,

ii) multiplying any row by a non-zero constant c, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \dots a_{in}]$ with the row $[ca_{i1}\ ca_{i2} \dots ca_{in}]$,

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

i) swapping any two rows of the matrix,

ii) multiplying any row by a non-zero constant c, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \ldots a_{in}]$ with the row $[ca_{i1}\ ca_{i2} \ldots ca_{in}]$,

iii) adding any row to the other, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \ldots a_{in}]$ with the row $[a_{i1} + a_{j1}\ a_{i2} + a_{j2} \ldots a_{in} + a_{jn}]$.

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

i) swapping any two rows of the matrix,

ii) multiplying any row by a non-zero constant c, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \dots a_{in}]$ with the row $[ca_{i1}\ ca_{i2} \dots ca_{in}]$,

iii) adding any row to the other, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \dots a_{in}]$ with the row $[a_{i1} + a_{j1}\ a_{i2} + a_{j2} \dots a_{in} + a_{jn}]$.

By the Theorem the elementary row operations lead to a matrix of an equivalent linear system.

Elementary Row Operations

There are three elementary row operations on a matrix of a linear system:

i) swapping any two rows of the matrix,

ii) multiplying any row by a non-zero constant c, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \ldots a_{in}]$ with the row $[ca_{i1}\ ca_{i2} \ldots ca_{in}]$,

iii) adding any row to the other, i.e. replacing the i-th row $[a_{i1}\ a_{i2} \ldots a_{in}]$ with the row $[a_{i1} + a_{j1}\ a_{i2} + a_{j2} \ldots a_{in} + a_{jn}]$.

By the Theorem the elementary row operations lead to a matrix of an equivalent linear system. The algorithm using the three elementary row operations, leading to a general solution is called the **Gaussian elimination**.

The (Reduced) Echelon Form

The **leading coefficient** (or **pivot**) of a non-zero row is the leftmost non-zero entry of the row.

The (Reduced) Echelon Form

The **leading coefficient** (or **pivot**) of a non-zero row is the leftmost non-zero entry of the row.

A matrix is in an echelon form if:

i) all non-zero rows are above all zero rows,

The (Reduced) Echelon Form

The **leading coefficient** (or **pivot**) of a non-zero row is the leftmost non-zero entry of the row.

A matrix is in an echelon form if:

i) all non-zero rows are above all zero rows,

ii) the leading coefficient of any row lies strictly to the right of the leading coefficient of any upper row.

The (Reduced) Echelon Form

The **leading coefficient** (or **pivot**) of a non-zero row is the leftmost non-zero entry of the row.

A matrix is in an echelon form if:

i) all non-zero rows are above all zero rows,

ii) the leading coefficient of any row lies strictly to the right of the leading coefficient of any upper row.

A matrix is in a **reduced echelon form** if it is in an echelon form, all leading coefficients are equal to 1 and every leading coefficient is the only non-zero element in its column.

Example

The following matrix is in an echelon form. The leading coefficients are marked with circles.

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It is not in the reduced echelon form because in columns 3 and 5 there are leading coefficients and other non-zero entries.

The Gaussian Elimination

Theorem

Any matrix can be brought into the reduced echelon form using elementary row operations.

The Gaussian Elimination

Theorem

Any matrix can be brought into the reduced echelon form using elementary row operations.

Proof.

Use induction on the number of columns to prove that every matrix can be brought into an echelon form.

The Gaussian Elimination

Theorem

Any matrix can be brought into the reduced echelon form using elementary row operations.

Proof.

Use induction on the number of columns to prove that every matrix can be brought into an echelon form. Let $A = [a_{ij}] \in M(m \times 1; \mathbb{R})$.

If $A \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ and, for example, $a_{11} \neq 0$ then

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} \xrightarrow{\substack{r_2 - \frac{a_{21}}{a_{11}} r_1 \\ \vdots \\ r_m - \frac{a_{m1}}{a_{11}} r_1}} \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The Gaussian Elimination

Proof.

Let $A = [a_{ij}] \in M(m \times n; \mathbb{R})$ and let $n > 1$.

The Gaussian Elimination

Proof.

Let $A = [a_{ij}] \in M(m \times n; \mathbb{R})$ and let $n > 1$. Let $k \in \mathbb{N}$ be the number of first non-zero column, changing the order of rows one can assume that $a_{1k} \neq 0$.

The Gaussian Elimination

Proof.

Let $A = [a_{ij}] \in M(m \times n; \mathbb{R})$ and let $n > 1$. Let $k \in \mathbb{N}$ be the number of first non-zero column, changing the order of rows one can assume that $a_{1k} \neq 0$. Then

$$\left[\begin{array}{cccc|ccc} 0 & \cdots & 0 & a_{1k} & a_{1(k+1)} & \cdots & a_{1n} \\ \hline 0 & \cdots & 0 & a_{2k} & a_{2(k+1)} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{mk} & a_{m(k+1)} & \cdots & a_{mn} \end{array}\right] \xrightarrow{\substack{r_2 - \frac{a_{2k}}{a_{1k}} r_1 \\ \vdots \\ r_m - \frac{a_{mk}}{a_{1k}} r_1}}$$

$$\left[\begin{array}{cccc|ccc} 0 & \cdots & 0 & a_{1k} & a_{1(k+1)} & \cdots & a_{1n} \\ \hline 0 & \cdots & 0 & 0 & b_{2(k+1)} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & b_{m(k+1)} & \cdots & b_{mn} \end{array}\right]$$

for some $b_{ij} \in \mathbb{R}$.

The Gaussian Elimination

Proof.
By the inductive assumption the matrix in the lower right corner, i.e.

$$\begin{bmatrix} b_{2(k+1)} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m(k+1)} & \cdots & b_{mn} \end{bmatrix}$$

can be brought to an echelon form by elementary operations.

The Gaussian Elimination

Proof.

By the inductive assumption the matrix in the lower right corner, i.e.

$$\begin{bmatrix} b_{2(k+1)} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m(k+1)} & \cdots & b_{mn} \end{bmatrix}$$

can be brought to an echelon form by elementary operations. The same operations will bring matrix

$$\left[\begin{array}{cccc|ccc} 0 & \cdots & 0 & a_{1k} & a_{1(k+1)} & \cdots & a_{1n} \\ \hline 0 & \cdots & 0 & 0 & b_{2(k+1)} & \cdots & b_{2n} \\ 0 & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & b_{m(k+1)} & \cdots & b_{mn} \end{array}\right]$$

to an echelon form.

The Gaussian Elimination

Proof.

Assume that matrix $A = [a_{ij}] \in M(m \times n; \mathbb{R})$ is in echelon form and the leading coefficients are $a_{1j_1}, a_{2j_2}, \ldots, a_{m'j_{m'}}$ where $j_1 < j_2 < \ldots < j_{m'}$ and $m' \leq m$, i.e. rows $m'+1, m'+2, \ldots, m$ are zero.

The Gaussian Elimination

Proof.

Assume that matrix $A = [a_{ij}] \in M(m \times n; \mathbb{R})$ is in echelon form and the leading coefficients are $a_{1j_1}, a_{2j_2}, \dots, a_{m'j_{m'}}$ where $j_1 < j_2 < \dots < j_{m'}$ and $m' \leq m$, i.e. rows $m'+1, m'+2, \dots, m$ are zero.

$$\begin{bmatrix}
0 & a_{1j_1} & * & * & * & \cdots & * & * & * & * & * \\
0 & 0 & a_{2j_2} & * & * & \cdots & * & * & * & * & * \\
0 & 0 & 0 & 0 & a_{3j_3} & \cdots & * & * & * & \dots & * \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_{m'j_{m'}} & * & \dots & * \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

The Gaussian Elimination

Proof.

The following elementary operations will bring the matrix A in an echelon form into the reduced echelon form

$$r_k - \frac{a_{kj_i}}{a_{ij_i}} r_i \text{ for } i = 2, \ldots, m', k = 1, \ldots, i - 1,$$

$$r_i / a_{ij_i} \text{ for } i = 1, \ldots, m'.$$

The Gaussian Elimination

Proof.

The following elementary operations will bring the matrix A in an echelon form into the reduced echelon form

$$r_k - \frac{a_{kj_i}}{a_{ij_i}} r_i \text{ for } i = 2, \ldots, m', k = 1, \ldots, i-1,$$

$$r_i / a_{ij_i} \text{ for } i = 1, \ldots, m'.$$

In short, in each of the column $j_1, j_2, \ldots, j_{m'}$ we use the leading coefficient to make the entries above it zero and then we divide the corresponding row to make the leading coefficient equal to 1.

The Gaussian Elimination

Proof.

The following elementary operations will bring the matrix A in an echelon form into the reduced echelon form

$$r_k - \frac{a_{kj_i}}{a_{ij_i}} r_i \text{ for } i = 2, \dots, m', k = 1, \dots, i-1,$$

$$r_i / a_{ij_i} \text{ for } i = 1, \dots, m'.$$

In short, in each of the column $j_1, j_2, \dots, j_{m'}$ we use the leading coefficient to make the entries above it zero and then we divide the corresponding row to make the leading coefficient equal to 1.

$$\begin{bmatrix} 0 & a_{1j_1} & * & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & 0 & 0 & a_{3j_3} & \cdots & * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_{m'j_{m'}} & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1/a_{1j_1}} \begin{bmatrix} 0 & 1 & * & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & 0 & 0 & a_{3j_3} & \cdots & * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_{m'j_{m'}} & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1 - \frac{a_{1j_2}}{a_{2j_2}} r_2}$$

The Gaussian Elimination

Proof.

$$\begin{bmatrix} 0 & 1 & 0 & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & 0 & 0 & a_{3j_3} & \cdots & * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_{m'j_{m'}} & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_2/a_{2j_2}} \begin{bmatrix} 0 & 1 & 0 & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & 1 & * & * & \cdots & * & * & * & * & * \\ 0 & 0 & 0 & 0 & a_{3j_3} & \cdots & * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_{m'j_{m'}} & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[r_2 - \frac{a_{2j_3}}{a_{3j_3}} r_3]{r_1 - \frac{a_{1j_3}}{a_{3j_3}} r_3} \cdots$$

$$\cdots \longrightarrow \begin{bmatrix} 0 & 1 & 0 & * & 0 & \cdots & * & 0 & * & * & * \\ 0 & 0 & 1 & * & 0 & \cdots & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 1 & \cdots & * & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

How to solve a system of linear equations?

Bring a matrix of a system of linear equation to the reduced echelon form. If there is a pivot in the column of constant terms the system if inconsistent. Otherwise, the general solution can be read from the echelon form by choosing the basic variables as those corresponding to columns with a pivot.

Example

Let's solve the system $\begin{cases} x_1 - 2x_2 + x_3 - x_4 = 2 \\ 2x_1 - 4x_2 + 3x_3 + x_4 = 0 \end{cases}$

The matrix of this system is $\left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 2 \\ 2 & -4 & 3 & 1 & 0 \end{array}\right]$

By the elementary row operation $r_2 - 2r_1$ we put the matrix in an echelon form, i.e.

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 2 \\ 0 & 0 & 1 & 3 & -4 \end{array}\right]$$

The elementary operation $r_1 - r_2$ puts matrix in the reduced echelon form, that is

Example (continued)

$$\left[\begin{array}{cccc|c} ① & -2 & 0 & -4 & 6 \\ 0 & 0 & ① & 3 & -4 \end{array}\right]$$

There is no leading coefficient in the constant term columnn so it has solutions. The basic variables are x_1, x_3 and the free variables are x_2, x_4.

The general solution is $\begin{cases} x_1 & = & 2x_2 & + & 4x_4 & + & 6 \\ x_3 & = & & - & 3x_4 & - & 4 \end{cases}, \; x_2, x_4 \in \mathbb{R}.$

Every solution of this linear system is of the form

$$(2x_2 + 4x_4 + 6, x_2, -3x_4 - 4, x_4), \; x_2, x_4 \in \mathbb{R}.$$

The Uniqueness of the Reduced Echelon Form

Proposition

Let $A \in M(m \times n; \mathbb{R})$ be a matrix. If matrices $B, C \in M(m \times n; \mathbb{R})$ were obtained from A by a series of elementary row operations and they are in the reduced echelon form then $B = C$.

The Uniqueness of the Reduced Echelon Form

Proposition

Let $A \in M(m \times n; \mathbb{R})$ be a matrix. If matrices $B, C \in M(m \times n; \mathbb{R})$ were obtained from A by a series of elementary row operations and they are in the reduced echelon form then $B = C$.

Proof.

Let j be the number of the leftmost column where the matrices B and C differ. Let

$$1 \leq j_1 < j_2 < \ldots < j_k < j,$$

be the numbers of the columns with pivots in B and C smaller than j. Let B' and C' be submatrices of matrices B and C, respectively, consisting of columns $j_1, \ldots, j_k, j$. Let U_B, U_C be systems of linear equations which matrices are equal to B' and C', respectively (the last column consists of constant terms).

The Uniqueness of the Reduced Echelon Form (continued)

Proof.

$$B' = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_{1j} \\ 0 & 1 & & \vdots & b_{2j} \\ \vdots & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & b_{kj} \\ \hline 0 & 0 & \cdots & 0 & b_{(k+1)j} \\ \vdots & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 0 & b_{mj} \end{array}\right], \quad C' = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_{1j} \\ 0 & 1 & & \vdots & c_{2j} \\ \vdots & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & c_{kj} \\ \hline 0 & 0 & \cdots & 0 & c_{(k+1)j} \\ \vdots & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 0 & c_{mj} \end{array}\right]$$

By the assumption, the systems U_B and U_C are equivalent, as their matrices were obtained by a series of elementary row transformations from the same submatrix of matrix A. The following may happen for B and C: the $(k+1)$-th pivot is in the j-th column, behind the j-th column or it does not exist. Say, if for matrix B the $(k+1)$-th pivot is behind the j-th column or it does not exist then $b_{ij} = 0$ for $i \geq k+1$. Analogously, if the same happens for matrix C then $c_{ij} = 0$ for $i \geq k+1$.

$\square$

The Uniqueness of the Reduced Echelon Form (continued)

Proof.

It is impossible that the $(k+1)$-th pivot is in the j-th column simultaneously in matrix B and in matrix C as this would mean the j-th columns of B and C are the same. If one of the matrices B, C has the $(k+1)$-th pivot is in the j-th column and the other one has the $(k+1)$-th pivot behind the j-th column or it does not exist then one of the systems U_B, U_C is inconsistent and the other is consistent. This leads to a contradiction. If both matrices B, C have the $(k+1)$-th pivot behind the j-th column or it does not exist then $b_{ij} = c_{ij} = 0$ for $i \geq k+1$. Therefore the system U_B has a unique solution $(b_{1j}, b_{2j}, \ldots, b_{kj})$ and the system U_C has a unique solution $(c_{1j}, c_{2j}, \ldots, c_{kj})$, which again leads to a contradiction. $\square$

The Uniqueness of the Reduced Echelon Form (continued)

Remark

Obviously, echelon form is not unique.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

The Uniqueness of the Reduced Echelon Form (continued)

Remark

Obviously, echelon form is not unique.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

The reduced echelon form of matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}$ *is* $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

The Uniqueness of the Reduced Echelon Form (continued)

Remark

One can read a general solution from a matrix which after a permutation (i.e. change of the order) of columns is in the reduced echelon form (by choosing basic variables as those corresponding to columns which after the permutation contain a pivot).

The Uniqueness of the Reduced Echelon Form (continued)

Remark

One can read a general solution from a matrix which after a permutation (i.e. change of the order) of columns is in the reduced echelon form (by choosing basic variables as those corresponding to columns which after the permutation contain a pivot).

A general solution of the system

$$\left[\begin{array}{rrrr|r} 2 & 1 & 3 & 0 & 2 \\ -1 & 0 & 5 & 1 & -7 \end{array}\right]$$

is

$$\begin{cases} x_2 = -2x_1 - 3x_3 + 2 \\ x_4 = x_1 - 5x_3 - 7 \end{cases}, \quad x_1, x_3 \in \mathbb{R}.$$

Reduced Echelon Form of a Square Matrix

Proposition

Let $A \in M(n \times n; \mathbb{R})$ *be a square matrix (i.e. it has the* n *rows and* n *columns). Then the reduced echelon form of* A *either has a zero row or it is equal to* $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in M(n \times n; \mathbb{R})$.

Reduced Echelon Form of a Square Matrix

Proposition

Let $A \in M(n \times n; \mathbb{R})$ be a square matrix (i.e. it has the n rows and n columns). Then the reduced echelon form of A either has a zero row or it is equal to $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in M(n \times n; \mathbb{R})$.

Proof.

By definition, the numbers of columns with pivots form a strictly increasing sequence

$$1 \leq j_1 < j_2 < \ldots < j_k \leq n.$$

Therefore $k \leq n$. If $k < n$ then there are $n - k$ zero rows (only k rows contain a pivot).

Reduced Echelon Form of a Square Matrix

Proposition

Let $A \in M(n \times n; \mathbb{R})$ be a square matrix (i.e. it has the n rows and n columns). Then the reduced echelon form of A either has a zero row or it is equal to $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in M(n \times n; \mathbb{R})$.

Proof.

By definition, the numbers of columns with pivots form a strictly increasing sequence

$$1 \leq j_1 < j_2 < \ldots < j_k \leq n.$$

Therefore $k \leq n$. If $k < n$ then there are $n - k$ zero rows (only k rows contain a pivot). The case $k = n$ is possible only if $j_1 = 1, j_2 = 2, \ldots, j_n = n$, i.e. the reduced form of A is equal to I_n. □

Generalized Inverse

Definition

Let $A \in M(m \times n; \mathbb{R})$ be a matrix. Matrix $A^g \in M(n \times m; \mathbb{R})$ is called **a generalized inverse** *of matrix A if*

$$A = AA^g A.$$

A generalized inverse always exists (the Moore-Penrose pseudoinverse A^+ is a generalized inverse) but it is not unique. For example any matrix is a generalized inverse of a zero matrix.

Proposition

Let $A \in M(m \times n; \mathbb{R})$, $b \in M(m \times 1; \mathbb{R})$. If $A^g b$ is a solution of the system of linear equations $Ax = b$ then all solutions of that system are given by the formula

$$x = A^g b + (I - A^g A)y,$$

where $y \in M(n \times 1)$ is any vector.

Generalized Inverse (continued)

Proof.

Let $x = A^g b + (I - A^g A)y$, where y is an arbitrary vector. Then

$$Ax = AA^g b + Ay - AA^g Ay = AA^g b = b.$$

Assume that $Ax = b$. Then

$$x = A^g b + (I - A^g A)x,$$

i.e., it is enough to take $y = x$. $\square$