Linear Algebra

Lecture 1 - Solving Linear Equations

Oskar Kędzierski

2 October 2017

Other Details

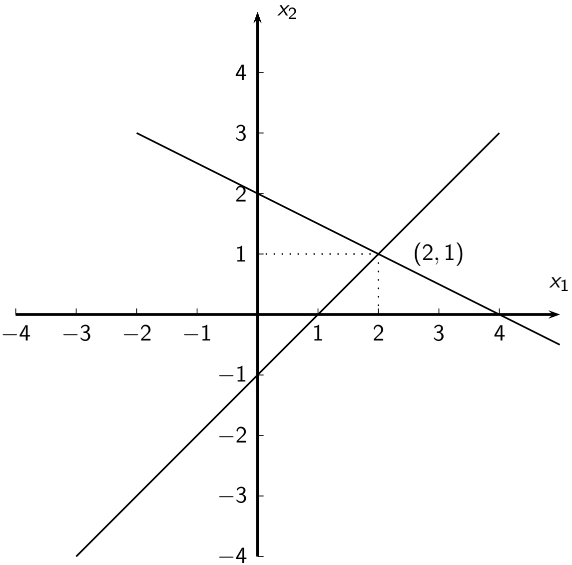
email: oskar@mimuw.edu.pl,
website:
`http://www.mimuw.edu.pl/~oskar/linear_algebra_wne_2017.html`
60 problems for classes:
`http://www.mimuw.edu.pl/~tkozn/ALWNEcw.pdf`
address: room 5310, the Faculty of Mathematics, Informatics, and Mechanics, ul. Banacha 2, Warsaw.

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*
ii) G. Strang, *Linear Algebra and its Applications*
iii) T. Koźniewski, *Wykłady z algebry liniowej* (in Polish)

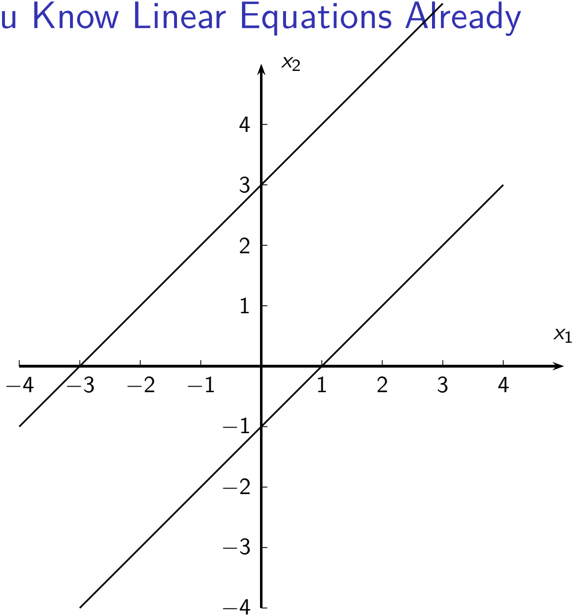
You Know Linear Equations Already



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

Exactly one solution $(2, 1)$

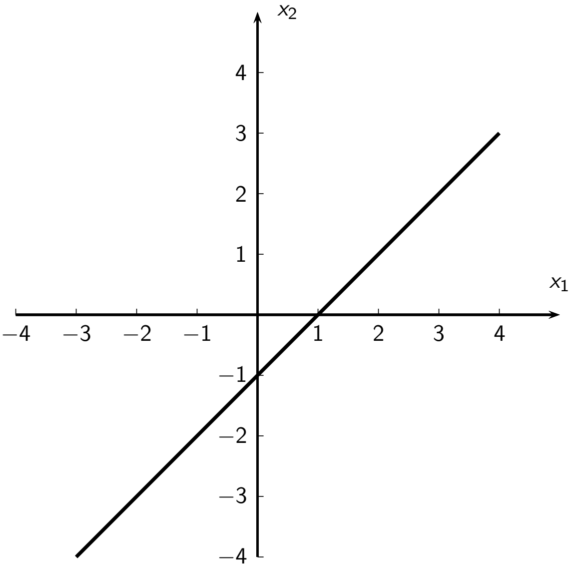
You Know Linear Equations Already



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

No solutions at all

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 $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ 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 $x_1, \ldots, x_n$

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

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. 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 $ca_1x_1 + ca_2x_2 + \ldots + ca_nx_n = cb$.

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):

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:

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

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 $d_{ij} \in \mathbb{R}$. 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 every system of linear equations

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

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

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

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:

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.

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 lower 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 a 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 a 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 $n > 1$. 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 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.

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

The Gaussian Elimination

Proof.

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

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

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

The Gaussian Elimination

How to solve a system of linear equations?

Bring a matrix of a system of linear equation to a 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 a 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 $\begin{cases} x_1 &=& 2x_2 &+& 4x_4 &+& 6 \\ x_3 &=& &-& 3x_4 &-& 4 \end{cases}$

Every solution of this linear system is of the form
$(2x_2 + 4x_4 + 6, x_2, -3x_4 - 4, x_4),\ x_2, x_4 \in \mathbb{R}$.