Calculus and Linear Algebra for Biomedical Engineering

Week 3: Matrices and systems of linear equations

H. Führ, Lehrstuhl A für Mathematik, RWTH Aachen, WiSe 07/08

Motivation

Sample problem: Suppose we are given two solutions of a certain chemical in water, one with a 2 % concentration, the other with a 10% concentration. Our aim is to mix the two solutions in such a way that we obtain 3 liters with a 3 % concentration.

Mathematical Formulation. Let x denote the quantity of 2 % concentration and y the quantity of 10 % concentration that we use for mixing.
The fact that we want three litres of the final product gives rise to the equation

$$x + y = 3 \tag{1}$$

Motivation cont'd

Moreover, the amount of substance contributed by quantity x with a 2 % concentration is $x \cdot 0.02$, whereas quantity y of a 10 % concentration contributes $y \cdot 0.10$. The concentration after mixing is obtained by dividing this by the total amount of solution, 3 litres, which results in the equation

$$x \cdot 0.02/3 + y \cdot 0.10/3 = 0.03 \tag{2}$$

We are thus looking for solutions x, y of the system of linear equations (1) and (2).

Solution. We solve (1) for x, getting $x = 3 - y$. Plugging this into (2) gives

$$(3 - y) \cdot 0.02/3 + y \cdot 0.10/3 = 0.03 \Leftrightarrow y = 0.375 \ .$$

Hence mixing 2.625 litres of the 2 % solution and 0.375 litres of the 10 % is the only way of achieving the desired quantity and concentration.

General systems of linear equations

A system of linear equations with m equations and n variables is a system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2 \\ \vdots \quad &\vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n &= y_m \end{aligned}$$

Here $a_{ij} \in \mathbb{R}$ are called the coefficients of the system, $y_1, \ldots, y_m$ is the right-hand side. Both the coefficients and the right-hand side are known.
By contrast, $x_1, \ldots, x_n$ are the variables. Solutions of the system are all possible vectors $(x_1, \ldots, x_n)^T$ such that all equations above equations are fulfilled simultaneously.

Central task: Finding all solutions

Given a system of linear equations as on the previous slide, we want to find the set of all solutions, given as

$$\mathbb{S} = \{(x_1, \ldots, x_n)^T : \text{ for all } l = 1, \ldots, m : a_{l1}x_1 + \ldots + a_{ln}x_n = y_l\}$$

Any such set will be either empty, contain a single point, or infinitely many of them. In the latter case, we want a parametrization of $\mathbb{S}$. This parametrization will usually depend on certain free variables.

There is no simple formula for these sets. We describe a systematically applicable method for the computation of $\mathbb{S}$.

Example: Solving a single linear equation

Consider the equation $x + y - z = 5$
How to get all solutions:

- Insert arbitrary real numbers for y, z
- Solving for x gives $x = 5 - y + z$

Hence, the set of all solutions is $\mathbb{S} = \{(5 + y - z, y, z) : y, z \in \mathbb{R}\}$.

Observations:

- The set of solutions is parameterized by two free variables $y, z \in \mathbb{R}$; i.e., it is a plane in $\mathbb{R}^3$.
- Solving for a different variable (e.g., y) results in the same set of solutions, only in a different parameterization.

Example: Solving a system of two linear equations

Consider the system

$$\begin{aligned} x + y - z &= 5 \\ x - y - 2z &= 3 \end{aligned}$$

Substituting the solution $x = 5 - y + z$ for the first into the second equation provides

$$5 - y + z - y - 2z = 3 \Leftrightarrow -2y - z = -2 \quad .$$

Here, we may choose any value for z, and obtain $y = 1 - \frac{z}{2}$. Plugging this into the equation for x gives $x = 5 - y + z = 4 + \frac{3z}{2}$. Thus, the set of all solutions is

$$\mathbb{S} = \left\{ \left(4 + \frac{3z}{2}, 1 - \frac{z}{2}, z \right)^T : z \in \mathbb{R} \right\} \quad .$$

Observations

- The set of solutions is parameterized by one free variable $z \in \mathbb{R}$; i.e., it is a line in $\mathbb{R}^3$.
- Solving for a different variable (e.g., y) results in the same set of solutions, only in a different parameterization.

$\Rightarrow$ Rules of thumb

- Each equation fixes one variable
- The solutions of a system of m equations with n variables is parameterized by $n - m$ free variables.

Exceptions to the rule

- The system

$$\begin{aligned} x + y - z &= 5 \\ 2x + 2y - 2z &= 10 \end{aligned}$$

  is redundant: Two equations, three variables, but two degrees of freedom.

- The system

$$\begin{aligned} x + y - z &= 5 \\ 2x + 2y - 2z &= 11 \end{aligned}$$

  is contradictory: It has no solution.

For larger numbers of variables, these cases are not easily recognized.

Matrix calculus

Effective, systematic notation for the treatment of linear equations.

- Matrix by vector multiplication
- Systems of linear equations and matrix-by-vector multiplication
- Simple matrices and their solutions
- Solving linear systems of equations via the Gauss algorithm

Definition of matrices

Definition. Let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$, and $m, n \in \mathbb{N}$. A $m \times n$-matrix in $\mathbb{K}$ is a mapping $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{K}$, denoted as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = (a_{ij}) \,.$$

In other words: An $m \times n$ matrix is a rectangular array of numbers.
m = number of lines in A = length of columns in A
n = number of columns in A = length of lines in A
The space of $m \times n$-matrices is denoted by $\mathbb{R}^{m \times n}$.

Multiplying a matrix with a column vector

Definition. Given $A \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$, the column vector $\mathbf{y} = A \cdot \mathbf{x}$ is defined by

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}$$

Rules for the matrix-vector product

The product of a matrix $A \in \mathbb{R}^{k \times m}$ with a column vector $\mathbf{x} \in \mathbb{R}^n$ is only defined if $m = n$, i.e., if the length of the rows in A equals the length of $\mathbf{x}$.

Main Property: Linearity.

Let $A \in \mathbb{R}^{m \times n}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and $s, t \in \mathbb{R}$. Then

$$A \cdot (s\mathbf{x} + t\mathbf{y}) = sA \cdot \mathbf{x} + tA \cdot \mathbf{y}$$

As before, the "$\cdot$" is sometimes omitted where no confusion can arise.

Linear systems of equations and matrices

Let $A \in \mathbb{R}^{m \times n}$ and $\mathbf{y} \in \mathbb{R}^m$. The matrix-vector equation

$$A \cdot \mathbf{x} = \mathbf{y}$$

is the short-hand form of the system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2 \\ \vdots \qquad\qquad\qquad &\vdots \;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= y_m \end{aligned}$$

Each column corresponds to an unknown, each row corresponds to an equation.
Alternatively, one represents the system by the $m \times (n+1)$-matrix $A' = (A|\mathbf{y})$ obtained by appending $\mathbf{y}$ as $n+1$st column to A.

An example

The linear system

$$\begin{aligned} x_1 + x_2 - x_3 &= 5 \\ x_1 - x_2 - 2x_3 &= 3 \end{aligned}$$

is equivalent to

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

alternatively represented by the matrix

$$A' = \left(\begin{array}{ccc|c} 1 & 1 & -1 & 5 \\ 1 & -1 & -2 & 3 \end{array}\right)$$

Homogeneous linear equations

Let $A \in \mathbb{R}^{m\times n}$, and $\mathbf{y} \in \mathbb{R}^m$. The equation

$$A \cdot \mathbf{x} = \mathbf{y}$$

is called homogeneous equation if $\mathbf{y} = \mathbf{0}$, and inhomogeneous otherwise.

Main property of homogeneous equations: If $\mathbf{x}, \mathbf{z}$ solve the homogeneous equation $A \cdot \mathbf{x} = \mathbf{0}$, the same is true for $s\mathbf{x} + t\mathbf{z}$, $s, t \in \mathbb{R}$ arbitrary: By linearity of the matrix-vector product,

$$A \cdot (s\mathbf{x} + t\mathbf{z}) = sA \cdot \mathbf{x} + tA \cdot \mathbf{z} = s\mathbf{0} + t\mathbf{0} = \mathbf{0} \quad .$$

Solving homogeneous and inhomogeneous equations

Theorem 1.
Let $\mathbf{y} \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$. Let

$$\begin{aligned} \mathbb{S}(A, \mathbf{y}) &= \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{y}\} \\ \mathbb{S}(A, \mathbf{0}) &= \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\} \end{aligned}$$

i.e., the sets of all solutions to the inhomogeneous and homogeneous system, respectively. Suppose that $\mathbf{z} \in \mathbb{S}(A, \mathbf{y})$. Then

$$\mathbb{S}(A, \mathbf{y}) = \{\mathbf{z} + \mathbf{x}_0 : \mathbf{x}_0 \in \mathbb{S}(A, \mathbf{0})\}$$

Hence: In order to find all solutions of the equation $A\mathbf{x} = \mathbf{y}$,

- find one such solution $\mathbf{z}$, and
- determine all solutions $\mathbf{x}_0$ of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

Gauss elimination: Motivation

A systematic solution of linear systems of equations relies on

- Simplification (Gaussian elimination)
- Substitution

Definition. Let A be an $m \times n$-matrix. A is called simple if there exist indices

$$1 \leq l \leq m \text{ and } 1 \leq i_1 < i_2 < \ldots < i_l \leq n$$

such that

$$\begin{aligned} a_{j,i_j} &\neq 0 && \text{for} && j = 1, \ldots, l \\ a_{j,i} &= 0 && \text{if } i < i_j \text{ or } j > l \end{aligned}$$

Informally: i_j is the index of the first non-zero entry in the jth line.

Illustration of simple matrices

General form:

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1,i_1} & a_{1,i_1+1} & \dots & \dots & \dots & \dots & a_{1,n} \\ 0 & \dots & \dots & \dots & \dots & 0 & a_{2,i_2} & a_{2,i_2+1} \dots & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_{li_l} & \dots & a_{l,n} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

A concrete example:

$$A = \begin{pmatrix} 1 & 3 & 0 & 0 & -1 & 7 \\ 0 & 0 & 2 & 1 & \pi^2 & 4 \\ 0 & 0 & 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Solving a simple system of equations

Consider the equation

$$A \cdot \mathbf{x} = \mathbf{y}$$

with A an $m \times n$ matrix and $\mathbf{y} \in \mathbb{R}^m$.

Consider the extended $m \times (n+1)$ matrix $A' = (A|\mathbf{y})$ of the system. Assume that it is simple, with indices

$$1 \leq l \leq m \text{ and } 1 \leq i_1 < i_2 < \ldots < i_l \leq n+1$$

of the first nonzero entries.

Two cases can arise for the last equation:

Looking at the last equation

First case: $i_l = n+1$. This means that the last nonzero equation reads

$$0x_1 + \ldots + 0x_n = b_l \quad ,$$

with $b_l \neq 0$. This does not have a solution, hence the system is not solvable.

Second case: Here, the last nonzero equation is

$$a_{l,i_l} x_{i_l} + \ldots a_{l,n} x_n = b_l$$

with $a_{l,i_l} \neq 0$. We can therefore divide by a_{l,i_l} and solve for x_{i_l}:

$$x_{i_l} = \frac{b_l}{a_{l,i_l}} - \frac{a_{l,i_l+1}}{a_{l,i_l}} - \ldots - \frac{a_{l,n}}{a_{l,i_l}} x_n \ .$$

Here $x_{i_l+1}, \ldots, x_n$ are free parameters of the solution.

Solving a simple system of equations, cont'd

Having solved the last nonzero equation, we substitute the expression for x_{i_l} and the free parameters $x_{i_l+1}, \ldots, x_n$ into the second-to-last equation

$$a_{l-1,i_{l-1}} x_{i_{l-1}} + \ldots + a_{l-1,n} x_n = b_{l-1} \quad .$$

We can now solve this equation for $x_{i_{l-1}}$, using

$$a_{l-1,i_{l-1}} \neq 0 \text{ and } i_{l-1} < i_l \quad .$$

Note that the latter inequality means that the variable $x_{i_{l-1}}$ did not occur in the equation we solved first.

Working our way up through all equations, we obtain all solutions of the system.

General result: Solvability of simple equation systems

Theorem 2.
Let A be an $m \times n$ matrix, $\mathbf{y} \in \mathbb{R}^m$. Denote by $A' = (A|\mathbf{y})$ the extended matrix. Assume that A' is simple, with l' nonzero rows.
Then A is simple, with l nonzero rows.

- If $l \neq l'$, the equation $A\mathbf{x} = \mathbf{y}$ has no solution $\mathbf{x} \in \mathbb{R}^n$.
- If $l = l'$, the equation $A\mathbf{x} = \mathbf{y}$ has a solution. The general solution of the equation has $n - l$ free parameters.

$\rightsquigarrow$ Informally: If we have a simple system A',

- solvability can be decided by counting nonzero rows of A, A';
- the rule of thumb, “one nonzero equation fixes one variable” is applicable.

Definition. The number l from the Theorem is called **rank** of the linear system.

Simplifying systems of linear equations

Theorem 3. Let A', B' be the extended matrices of systems of linear equations. Assume that B' is obtained from A' by one of the following operations

- Interchanging two lines of A'.
- Multiplying a line of A' by a nonzero scalar.
- Adding a scalar multiple of one line to another.
  I.e., if $b^1, \ldots, b^m$ denote the lines of B', and $a^1, \ldots, a^m$ the lines of A', then there exists $1 \le k, l \le m$, with $k \neq l$, and $s \in \mathbb{R}$ such that

$$b^i = a^i \text{ for } i \neq l \quad , \quad b^l = a^l + sa^k$$

Then every solution of the system A' is also a solution for B' and vice versa.

Solving arbitrary systems of equations

Remark. The operations from Theorem 3. are called basic transformations. They can be applied repeatedly without changing the set of solutions.

Theorem 4. For every matrix A there is a simple matrix B obtainable by finitely many basic transformations from A.

Corollary. Combining Theorems 2 and 3, the set of solutions of a linear system of equations $A' = (A|\mathbf{y})$ is computable in finitely many steps.

The Gauss algorithm is a method to systematically convert an arbitrary system of linear equations to a simple one, and thus to solve linear systems of equations.

The Gauss algorithm

Given a matrix A, perform the following steps:

1. If A is the zero matrix, we are done. Otherwise, go to step 2.
2. Locate the first nonzero column from the left. One line has a nonzero entry in this column. If necessary, swap this line with the first line.
3. After step 2, the first nonzero entry of the first line is in the first nonzero column. Using this entry, subtract suitable multiples of the first line from the lines below to eliminate all other entries in that column.
4. After step 3, the matrix B consisting of the lines below the first line has at least one more zero column than A. Continue with step 1, with the smaller matrix B instead of A.

Summary

- Matrices allow a compact notation for writing and solving linear equations.
- General procedure for solving systems of linear equations.
  - Write a linear system $A\mathbf{x} = \mathbf{y}$ in extended matrix form $A' = (A|\mathbf{y})$.
  - Using basic transformations, compute simple matrix B' having the same set of solutions as A' ($\rightsquigarrow$ Gauss algorithm)
  - Using Theorem 2, determine all solutions for the matrix B'.
- Important definitions: Matrix-by-vector product, simple matrices, the rank of a matrix, homogeneous equations, free variables in the solution of systems of linear equation, simple transformations, Gauss algorithm