Technische Universität München
Fakultät für Informatik
PD Dr. Tobias Lasser
Rüdiger Göbl
Anca Stefanoiu
Walter Simson

WiSe 2017
Solution for Exercise Sheet 1
October 24, 2017

Exercises in Basic Mathematical Tools

Assignment 1 Norm – Equality (Exemplary Solution)

The *Euclidean norm* of a vector $\mathbf{x} \in \mathbb{R}^3$ is defined as:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

Further, we know that the *scalar product* is positive definite:

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \text{ and } \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ we have $\|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$, and:

$$\begin{aligned} & \|\mathbf{x} - \mathbf{y}\| = 0 \\ \Leftrightarrow \quad & \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle} = 0 \\ \Leftrightarrow \quad & \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0 \\ \Leftrightarrow \quad & \mathbf{x} - \mathbf{y} = 0 \\ \Leftrightarrow \quad & \mathbf{x} = \mathbf{y} \end{aligned}$$

The third step uses that the root function satisfies $\sqrt{x} = 0 \Leftrightarrow x = 0$. The forth step uses the positive definiteness of the scalar product.

Assignment 2 Cauchy-Schwarz Inequality (Exemplary Solution)

Reminder *Scalar Product*

For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define the scalar product

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} x_i y_i$$

with the following properties $\forall \mathbf{x}, \mathbf{x}', \mathbf{y} \in \mathbb{R}^n$, $\forall \lambda \in \mathbb{R}$:

- Linearity
  $\langle \mathbf{x} + \mathbf{x}', \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}', \mathbf{y} \rangle$
  $\langle \mathbf{x}\lambda, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \lambda$
- Commutativity
  $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- Positive definiteness
  $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$
  $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \in \mathbb{R}^n$

Proof *Cauchy-Schwarz Inequality*

$\forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n$ and $\forall \lambda \in \mathbb{R}$, it holds:

$$0 \leq \langle \mathbf{x}+\lambda\mathbf{y}, \mathbf{x}+\lambda\mathbf{y}\rangle = \langle \mathbf{x},\mathbf{x}\rangle + 2\lambda\langle \mathbf{x},\mathbf{y}\rangle + \lambda^2\langle \mathbf{y},\mathbf{y}\rangle \tag{1}$$

Consider this equation in the special case $\langle \mathbf{y},\mathbf{y}\rangle = 0$, then:

$$\mathbf{y} = \mathbf{0}$$

and thus

$$0 \leq \langle \mathbf{x},\mathbf{x}\rangle \quad \forall \mathbf{x} \in \mathbb{R}^n$$

$\hookrightarrow$ The Cauchy-Schwarz Inequality holds in this special case!

Therefore, let now w.l.o.g.[1] $\langle \mathbf{y},\mathbf{y}\rangle > 0$ and select

$$\lambda = -\frac{\langle \mathbf{x},\mathbf{y}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle} \in \mathbb{R}.$$

Insert this into equation (1):

$$\begin{aligned}
0 &\leq \langle \mathbf{x},\mathbf{x}\rangle - 2\frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\langle \mathbf{y},\mathbf{y}\rangle} + \frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\langle \mathbf{y},\mathbf{y}\rangle^2}\langle \mathbf{y},\mathbf{y}\rangle \quad | \cdot \langle \mathbf{y},\mathbf{y}\rangle (> 0!) \\
0 &\leq \langle \mathbf{x},\mathbf{x}\rangle\langle \mathbf{y},\mathbf{y}\rangle - \langle \mathbf{x},\mathbf{y}\rangle^2 \\
\langle \mathbf{x},\mathbf{y}\rangle^2 &\leq \langle \mathbf{x},\mathbf{x}\rangle\langle \mathbf{y},\mathbf{y}\rangle
\end{aligned}$$

$\hookrightarrow$ This is the Cauchy-Schwarz Inequality in the general case! ■

Assignment 3 Function Space (Exemplary Solution)

a) As discussed in the lecture, the function space $X = \mathbb{R}^{\mathbb{R}}$ is a vector space when defining the necessary operations (addition and scalar multiplication) in a point-wise manner as

$$\begin{aligned}
(f+g)(x) &:= f(x)+g(x) \\
(\lambda f)(x) &:= \lambda f(x)
\end{aligned}$$

for arbitrary "vectors" (mappings) $f,g \in X$ and parameters $x \in \mathbb{R}$.

A mathematical proof of this fact requires checking all the vector space properties for the function space X and is a rather lengthy procedure (and omitted here).

b) In order to show that a certain subset $S \subset X$ is a subspace, we need to show that all subspace properties hold for that set. The contrary can, of course, be shown with a single counter-example.

Reminder *Subspace*

Let V be a linear space. A nonempty set $S \subset V$ is called subspace of V if

- $\mathbf{x},\mathbf{y} \in S \Rightarrow \mathbf{x}-\mathbf{y} \in S$
- $\mathbf{x} \in S, \lambda \in \mathbb{R} \Rightarrow \lambda\mathbf{x} \in S$

Note that the first property is an elegant way of requiring both completeness with respect to addition and existence of the inverse element (subtraction does not work otherwise). The requirement of the zero element forming part of the subspace is implicitly contained in the second property, for $\lambda = 0$.

[1] "without loss of generality"; general and well-used remark that a restriction can be made as the other cases have already been dealt with

(i) $S := \{f \in X | f(0) = 1 + f(1)\} \subset X$ is **no** subspace.

Define

$$f(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{else} \end{cases},$$

then $f \in S$, but $2f \notin S$ as $(2f)(0) = 2f(0) = 2 \neq 1$.

(ii) $S := \{f \in X | 2f(-1) = f(1)\} \subset X$ is a **subspace**.

$(f - g) \in S$:

$$(f-g)(1) = f(1) - g(1) = 2f(-1) - 2g(-1) = 2(f-g)(-1)$$

$(\lambda f) \in S$:

$$(\lambda f)(1) = \lambda f(1) = 2\lambda f(-1) = 2(\lambda f)(-1)$$

(iii) $S := \{f \in X | f(-1) = 0\} \subset X$ is a **subspace**.

$(f - g) \in S$:

$$(f-g)(-1) = f(-1) - g(-1) = 0$$

$(\lambda f) \in S$:

$$(\lambda f)(-1) = \lambda f(-1) = 0$$

(iv) $S := \{f \in X | f(x) = f(1-x) \quad \forall x \in \mathbb{R}\} \subset X$ is a **subspace**.

$(f - g) \in S$:

$$(f-g)(x) = f(x) - g(x) = f(1-x) - g(1-x) = (f-g)(1-x)$$

$(\lambda f) \in S$:

$$(\lambda f)(x) = \lambda f(x) = \lambda f(1-x) = (\lambda f)(1-x)$$

Note that these equalities are directly derived from the definitions of the addition and scalar multiplication operations.

(v) $S := \{f \in X | f(x^3) = f(x)^5 \quad \forall x \in \mathbb{R}\} \subset X$ is **no** subspace.

Define the constant function

$$f(x) = 1,$$

then $f \in S$, but $2f \notin S$ as $(2f)(x^3) = 2f(x^3) = 2 \neq (2f)(x)^5 = (2f(x))^5 = 2^5$.

(vi) $S := \{f \in X | f \text{ is continuous}\} \subset X$ is a **subspace**.

Both, addition and scalar multiplication, are continuous functions themselves, and applying them to further continuous functions does not change this property.
Two mappings $f, g \in S$ are continuous by definition of the set, and therefore $(f - g) \in S$ and $(\lambda f) \in S$.
$\hookrightarrow$ These concepts will be discussed in detail in the lecture's calculus part!

<u>Assignment 4</u> Linear Dependence (Exemplary Solution)

In general, the question (for whether a linear combination exists) yields a system of linear equations[2]:

$$v_i = \lambda x_i + \mu y_i + \nu z_i \quad i = 1,2,3$$

For the first set $\mathbf{x} = (1,1,0)^\top$, $\mathbf{y} = (1,0,1)^\top$, $\mathbf{z} = (0,1,0)^\top$, one gets a particularly simple system:

$$\begin{aligned} 3 &= \lambda + \mu \\ 0 &= \lambda + \nu \\ 2 &= \mu \end{aligned}$$

From these equations, it is simple to solve for λ and ν:

$$\begin{aligned} \lambda &= 3 - \mu = 1 \\ \nu &= -\lambda = -1 \end{aligned}$$

It is thus possible to express $\mathbf{v}$ as linear combination of these $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$.

It is advisable to change (even standardize) the notation to matrix-vector-form. Then, the general system is:

$$\underbrace{\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}}_{=:\,A} \underbrace{\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}}_{=:\,\mathbf{x}} = \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}}_{=:\,\mathbf{b}}$$

The names A, $\mathbf{x}$ (“unknowns”), and $\mathbf{b}$ (“right-hand side”) are commonly used to refer to the three components of the system, and there are many general (and specialized) approaches to solving such problems. Some are going to be discussed in the lecture.

For the first set, the matrix form is:

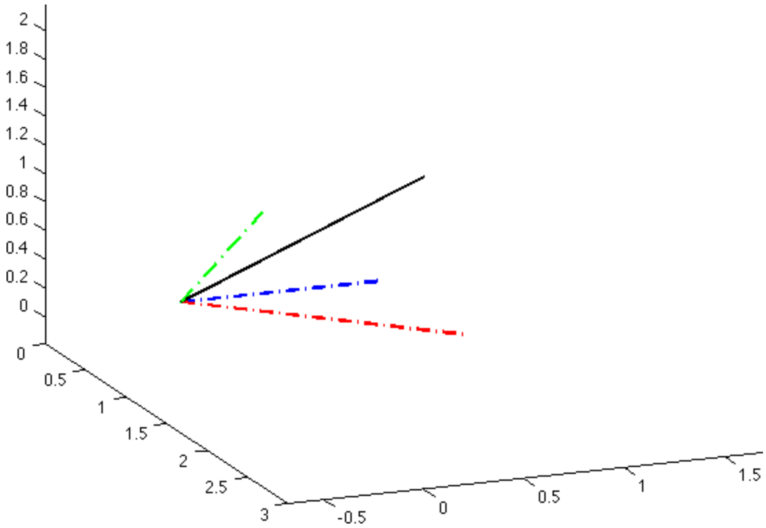
$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right)$$

The best-known solution algorithm is probably *Gaussian Elimination* which attempts to bring the system into triangular or even diagonal shape by adding up multiples of single rows, commuting rows, etc. Once that simple shape is reached, it is easy to get the solution.

a) $$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right) \overset{I-III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right) \overset{II-I}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 2 \end{array}\right) \overset{II\leftrightarrow III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

The three vectors $\mathbf{x}$ (red), $\mathbf{y}$ (green) and $\mathbf{z}$ (blue) are not linearly dependent themselves (i.e. in general configuration) and form a **basis** of $\mathbb{R}^3$, however, neither an orthogonal nor an orthonormal one.

[2] also called “linear system”

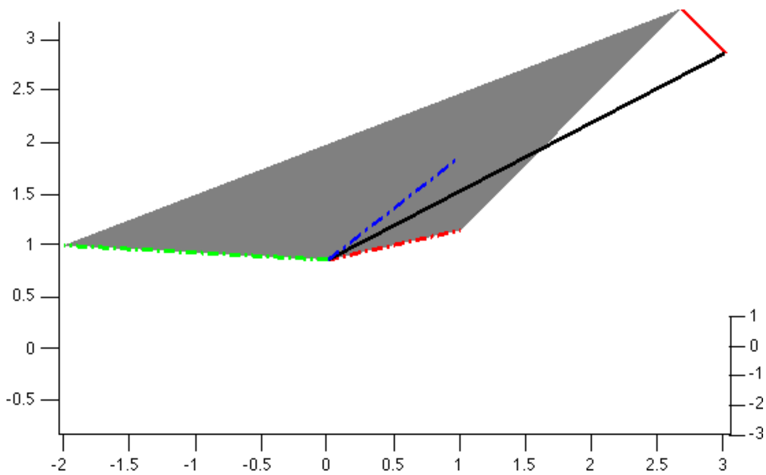


b) $$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 1 & -3 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right) \overset{II-I}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 0 & -1 & -1 & -3 \\ 0 & 1 & 1 & 2 \end{array}\right) \overset{II+III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 1 & 2 \end{array}\right)$$

In the second row of the system, a problematic situation has occurred. That part is logically equivalent to the statement $0 = -1$ which is always wrong. Consequently, the system does not have a solution.

The three vectors $\mathbf{x}$ (red), $\mathbf{y}$ (green) and $\mathbf{z}$ (blue) turn out to be linearly dependent. They are **no basis** and do not generate $\mathbb{R}^3$, but just a two-dimensional plane within the three-dimensional space. Such matrices are called *singular* and said to be *rank-deficient*.

The right-hand-side (black) points to a location outside of that plane, a solution is therefore impossible to exist. The three vectors just miss the information of the error vector (red, solid).



c) $$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right)$$

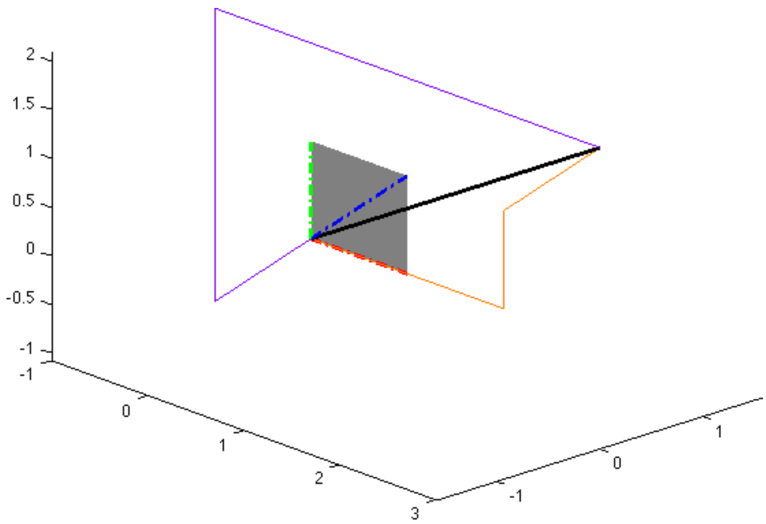
This time, the matrix is **also singular**, but the right-hand side vector (black) happens to point into the plane generated by the selection of vectors $\mathbf{x}$ (red), $\mathbf{y}$ (green) and $\mathbf{z}$ (blue).

It is thus possible to express the right-hand side as linear combination of the three vectors, however, the solution is not unique. The reason for this phenomenon is that the third vector contains information already available through the first two, and is thus redundant.

One can consequently find an infinite set of solutions, for instance $(2,1,1)^\top$ (orange) or $(4,3,-1)^\top$ (violet). If you want to determine the set of solutions, you can expand the linear system by a parameter $p \in R$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right) \overset{\text{reduces to}}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 2 \end{array}\right) \overset{p\in\mathbb{R}}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & p \end{array}\right) \overset{I-III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3-p \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & p \end{array}\right)$$
$$\overset{II-III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3-p \\ 0 & 1 & 0 & 2-p \\ 0 & 0 & 1 & p \end{array}\right)$$

Consequently, every vector $\begin{pmatrix} 3-p \\ 2-p \\ p \end{pmatrix}$ with $p \in \mathbb{R}$ is a solution.



Assignment 5 Properties of the Matrix Product (Exemplary Solution)

a) *No commutativity:* Choose for example:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

Now compute AB and BA:

$$AB = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \neq \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix} = BA$$

Thus the matrix-multiplication is not commutative.

b) Using the hint, we find:

$$AB = (B^\top A^\top)^\top \underbrace{=}_{\text{sym}} (BA)^\top$$

Thus, with $(BA)^\top = BA$ commutativity holds.

c) *Distributivity:* Consider three matrices $A, B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times k}$. Then, for $D = (A+B)C$ $D \in \mathbb{R}^{n\times k}$ holds. For the elements of the matrix D we find:

$$d_{ij} = \sum_{k=1}^{m} (a_{ik} + b_{ik}) * c_{kj} \qquad = \sum_{k=1}^{m} a_{ik} * c_{kj} + b_{ik} * c_{kj} = \sum_{k=1}^{m} a_{ik} * c_{kj} + \sum_{k=1}^{m} b_{ik} * c_{kj}$$

The last line states exactly the matrix-matrix product $AC + BC$. Thus, the distributive property holds.