Linear Algebra II
Exercise Sheet no. 7

TECHNISCHE UNIVERSITÄT DARMSTADT

Summer term 2011
May 18, 2011

Prof. Dr. Otto
Dr. Le Roux
Dr. Linshaw

Exercise 1 (Warm up: the trace)

Recall Exercise E4.3 about the trace.

Let $V := \mathbb{R}^{(n,n)}$ be the $\mathbb{R}$-vector space of all real $n \times n$ matrices and let $S \subseteq V$ be the subspace consisting of all symmetric matrices (i.e., all matrices A with $A^t = A$). For $A, B \in V$, we define

$$\langle A, B \rangle := \mathrm{Tr}(AB)\,,$$

where the *trace* $\mathrm{Tr}(A)$ of a matrix $A = (a_{ij})$ is defined as

$$\mathrm{Tr}(A) := \sum_{i=1}^{n} a_{ii}\,.$$

(a) Show that $\langle .\,, . \rangle$ is bilinear.

(b) Show that $\langle .\,, . \rangle$ is a scalar product on S.

Solution:

a) Let $A = (a_{ij})$, $B = (b_{ij})$, and $C = (c_{ij})$ be matrices and $\lambda \in R$. Since

$$\langle A, B \rangle = \sum_{i,k=1}^{n} a_{ik} b_{ki}$$

it follows that

$$\langle A + C, B \rangle = \sum_{i,k=1}^{n} (a_{ik} + c_{ik}) b_{ki} = \sum_{i,k=1}^{n} a_{ik} b_{ki} + \sum_{i,k=1}^{n} c_{ik} b_{ki} = \langle A, B \rangle + \langle C, B \rangle\,,$$

$$\langle \lambda A, B \rangle = \sum_{i,k=1}^{n} \lambda a_{ik} b_{ki} = \lambda \sum_{i,k=1}^{n} a_{ik} b_{ki} = \lambda \langle A, B \rangle\,.$$

In the same way, we show that $\langle A, B + C \rangle = \langle A, B \rangle + \langle A, C \rangle$ and $\langle A, \lambda B \rangle = \lambda \langle A, B \rangle$.

b) We have

$$\langle A, A \rangle = \sum_{i,k=1}^{n} a_{ik} a_{ki} = \sum_{i,k=1}^{n} (a_{ik})^2 \geqslant 0\,.$$

Furthermore, it follows that we have $\langle A, A \rangle = 0$ if and only if $A = 0$.

Exercise 2 (Cauchy-Schwarz and triangle inequalities)

(a) (Exercise 2.1.4 on page 60 of the notes)
Let $(V, \langle .\,, . \rangle)$ be a euclidean or unitary vector space. Show that equality holds in the Cauchy-Schwarz inequality, i.e., we have $\|\langle \mathbf{v}, \mathbf{w} \rangle\| = \|\mathbf{v}\| \cdot \|\mathbf{w}\|$, if, and only if, $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent.

(b) (Exercise 2.1.5 on page 60 of the notes)
Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be pairwise distinct vectors in a euclidean or unitary vector space $(V, \langle . , . \rangle)$, and write $\mathbf{a} := \mathbf{v} - \mathbf{u}$, $\mathbf{b} := \mathbf{w} - \mathbf{v}$. Show that equality holds in the triangle inequality

$$d(\mathbf{u}, \mathbf{w}) = d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w}), \text{ or, equivalently, } \|\mathbf{a} + \mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\|,$$

if, and only if, $\mathbf{a}$ and $\mathbf{b}$ are *positive real* scalar multiples of each other (geometrically: $\mathbf{v} = \mathbf{u} + s(\mathbf{w} - \mathbf{u})$ for some $s \in (0, 1) \subseteq \mathbb{R}$).

Solution:

a) Without loss of generality we may assume that $\mathbf{v}, \mathbf{w} \neq \mathbf{0}$.

When $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent, then $\mathbf{w} = \lambda \mathbf{v}$ for some λ. It follows that

$$\|\langle \mathbf{v}, \mathbf{w} \rangle\| = \|\langle \mathbf{v}, \lambda \mathbf{v} \rangle\| = \|\lambda\| \, \|\langle \mathbf{v}, \mathbf{v} \rangle\| = \|\lambda\| \|\mathbf{v}\|^2 = \|\mathbf{v}\| \|\lambda \mathbf{v}\| = \|\mathbf{v}\| \|\mathbf{w}\|.$$

Conversely, suppose that

$$\|\langle \mathbf{v}, \mathbf{w} \rangle\| = \|\mathbf{v}\| \cdot \|\mathbf{w}\|$$

and write $\lambda = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$. Then it follows, as in the proof of Proposition 2.1.10 (the Cauchy-Schwarz inequality) on page 59 of the notes, that

$$\langle \mathbf{w} - \lambda \mathbf{v}, \mathbf{w} - \lambda \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle - \frac{\langle \mathbf{w}, \mathbf{v} \rangle \langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \|\mathbf{w}\|^2 - \frac{\|\mathbf{v}\|^2 \|\mathbf{w}\|^2}{\|\mathbf{v}\|^2} = 0.$$

So by positive definiteness of the scalar product, $\mathbf{w} = \lambda \mathbf{v}$.

b) Note that $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$, because $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are pairwise distinct.

When $\mathbf{b} = \lambda \mathbf{a}$ with $0 < \lambda \in \mathbb{R}$, then

$$\|\mathbf{a} + \mathbf{b}\| = \|\mathbf{a} + \lambda \mathbf{a}\| = (1 + \lambda)\|\mathbf{a}\| = \|\mathbf{a}\| + \lambda \|\mathbf{a}\| = \|\mathbf{a}\| + \|\lambda \mathbf{a}\| = \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Conversely, when $\|\mathbf{a} + \mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\|$, also $(\|\mathbf{a} + \mathbf{b}\|)^2 = (\|\mathbf{a}\| + \|\mathbf{b}\|)^2$. But

$$\begin{aligned}
(\|\mathbf{a} + \mathbf{b}\|)^2 &= \langle \mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b} \rangle = \langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{b}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle \\
&\leqslant \|\mathbf{a}\|^2 + 2\|\langle \mathbf{a}, \mathbf{b} \rangle\| + \|\mathbf{b}\|^2, \text{ and} \\
(\|\mathbf{a}\| + \|\mathbf{b}\|)^2 &= \|\mathbf{a}\|^2 + 2\|\mathbf{a}\|\|\mathbf{b}\| + \|\mathbf{b}\|^2.
\end{aligned}$$

Therefore $\|\mathbf{a}\|\|\mathbf{b}\| \leqslant \|\langle \mathbf{a}, \mathbf{b} \rangle\|$, and $\|\mathbf{a}\|\|\mathbf{b}\| = \|\langle \mathbf{a}, \mathbf{b} \rangle\|$ by Cauchy-Schwarz. So we know that $\mathbf{b} = \lambda \mathbf{a}$ for some $\lambda \in \mathbb{C}$ by Exercise (E3.2). From $\|\mathbf{a} + \mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\|$, we deduce that $\|1 + \lambda\| = 1 + \|\lambda\|$, which implies that λ is a positive real.

Exercise 3 (Orthogonal matrices)
We consider real $n \times n$ matrices. Set

$$\mathrm{O}(n) := \{A \in \mathbb{R}^{(n,n)} \mid A^t A = E_n\}.$$

Show that $\mathrm{O}(n)$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$.

Solution:
We have to show that $E_n \in \mathrm{O}(n)$ and that $\mathrm{O}(n)$ is closed under multiplication and inverses.
Since $E_n^t E_n = E_n$, we have $E_n \in \mathrm{O}(n)$. Furthermore, for $A, B \in \mathrm{O}(n)$, we have

$$(AB)^t AB = B^t A^t AB = B^t E_n B = B^t B = E_n.$$

Hence, $AB \in \mathrm{O}(n)$. Similarly, one can show that $A^{-1} \in \mathrm{O}(n)$. For the inverse, we first note that $A^t A = E_n$ implies $A^t = A^{-1}$. Therefore, we have

$$(A^{-1})^t A^{-1} = (A^t)^t A^t = AA^t = (A^t A)^t = E_n^t = E_n.$$

Exercise 4 (Orthogonal vectors)

Let V be a euclidean or unitary space and $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a set of non-null pairwise orthogonal vectors.

(a) Show that S is linearly independent.

(b) Let $\mathbf{u} \in V$. Show that the vector

$$\mathbf{w} := \mathbf{u} - \sum_{i=1}^{n} \frac{\langle \mathbf{v}_i, \mathbf{u} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v}_i$$

is orthogonal to S. Note that $\sum_{i=1}^{n} \frac{\langle \mathbf{v}_i, \mathbf{u} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v}_i$ is the orthogonal projection of $\mathbf{w}$ on $\operatorname{span}(S)$.

(c) **[Parseval's identity]** Suppose that V is finite dimensional and that S is an othornormal basis of V. Show that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v}_i \rangle \langle \mathbf{v}_i, \mathbf{w} \rangle \quad \text{for all } \mathbf{v}, \mathbf{w} \in V .$$

(d) **[Bessel's inequality]** Suppose that V is euclidean and S is orthonormal. Show that

$$\sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle^2 \leq \|\mathbf{u}\|^2 \quad \text{for all } \mathbf{u} \in V .$$

Solution:

a) Suppose $\mathbf{v}_1, \dots, \mathbf{v}_n$ satisfy $\sum_{i=1}^{n} \lambda_i \mathbf{v}_i = \mathbf{0}$. We need to show that each λ_i is zero. For each $j = 1, \dots, n$, we get that

$$0 = \langle \mathbf{v}_j, \mathbf{0} \rangle = \left\langle \mathbf{v}_j, \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \right\rangle = \sum_{i=1}^{n} \lambda_i \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \lambda_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle ,$$

since $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ whenever $j \neq i$. But $\mathbf{v}_j \neq \mathbf{0}$. So $\langle \mathbf{v}_j, \mathbf{v}_j \rangle \neq 0$ since the scalar product is positive definite. Hence, $\lambda_j = 0$ for each $j = 1, \dots, n$. Therefore S is linearly independent.

b) For each $j = 1, \dots, n$, we have that

$$\begin{aligned} \langle \mathbf{v}_j, \mathbf{w} \rangle = \left\langle \mathbf{v}_j, \mathbf{u} - \sum_{i=1}^{n} \frac{\langle \mathbf{v}_i, \mathbf{u} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v} \right\rangle &= \langle \mathbf{v}_j, \mathbf{u} \rangle - \sum_{i=1}^{n} \frac{\langle \mathbf{v}_j, \mathbf{u} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \langle \mathbf{v}_j, \mathbf{v}_i \rangle \\ &= \langle \mathbf{v}_j, \mathbf{u} \rangle - \frac{\langle \mathbf{v}_j, \mathbf{u} \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle} \langle \mathbf{v}_j, \mathbf{v}_j \rangle \\ &= 0 . \end{aligned}$$

c) By Lemma 2.3.2, we have that $\mathbf{w} = \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{w} \rangle \mathbf{v}_i$. Applying the operation $\langle \mathbf{v}, . \rangle$ on both sides, we obtain the result.

d) Setting $\mathbf{w} := \mathbf{u} - \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle \mathbf{v}_i$ we have

$$\begin{aligned} \|\mathbf{w}\|^2 = \langle \mathbf{w}, \mathbf{w} \rangle &= \left\langle \mathbf{u} - \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle \mathbf{v}_i,\ \mathbf{u} - \sum_{j=1}^{n} \langle \mathbf{v}_j, \mathbf{u} \rangle \mathbf{v}_j \right\rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - 2 \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle^2 + \sum_{i,j=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle \langle \mathbf{v}_j, \mathbf{u} \rangle \langle \mathbf{v}_i, \mathbf{v}_j \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle^2 . \end{aligned}$$

Since $\|\mathbf{w}\|^2 \geq 0$, the inequality follows.

Exercise 5 (Jordan normal form for describing processes)

Suppose that we use vectors $\mathbf{s}_n \in \mathbb{R}^3$ to describe the state of a 3-dimensional system at step $n \in \mathbb{N}$ (for example, the position of a particle in space). The evolution of the system from stage n to $n+1$ is given by

$$\mathbf{s}_{n+1} = A\mathbf{s}_n\,, \quad \text{where} \quad A = \begin{pmatrix} -4 & 2 & -1 \\ -4 & 3 & 0 \\ 14 & -5 & 5 \end{pmatrix}.$$

(a) Use a transformation of the given A into Jordan normal form in order to get a feasible formula for $\mathbf{s}_n$, as a function of the index n and the initial state $\mathbf{s}_0$.

(b) Compute $\mathbf{s}_{100}$ for $\mathbf{s}_0 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$.

Solution:

(a) The characteristic polynomial of A is $p_A = (1-X)^2(2-X)$, so $\lambda_1 = 1$ and $\lambda_2 = 2$ are the eigenvalues of A. The corresponding eigenspaces are 1-dimensional, with generators

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \quad \text{for } V_{\lambda_1} \quad \text{and} \quad \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \quad \text{for } V_{\lambda_2}\,.$$

So the Jordan normal form of A has two blocks, one of size 2 and one of size 1. As

$$(A - E_3)^2 = \begin{pmatrix} 3 & -1 & 1 \\ 12 & -4 & 4 \\ 6 & -2 & 2 \end{pmatrix},$$

$\dim(\ker(A-E_3)^2) = 2$. Hence, the Jordan block of size 2 has entries 1 on the diagonal. Therefore the Jordan normal form of A is

$$J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

To find a matrix S such that $A = SJS^{-1}$, we take as third column $\mathbf{u}_3 = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$, an eigenvector with eigenvalue 2, and as second column an element of $\ker(A-E_3)^2 \setminus \ker(A-E_3)$, for example $\mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. The first column will then be

$$\mathbf{u}_1 = (A - E_3)\mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}. \quad \text{Hence,} \quad S = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 4 \\ -1 & 1 & 2 \end{pmatrix}.$$

We have $\mathbf{s}_n = A^n\mathbf{s}_0 = SJ^nS^{-1}\mathbf{s}_0$. Furthermore

$$J^n = \begin{pmatrix} 1 & n & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{pmatrix}.$$

(b) For $\mathbf{s}_0 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, we have

$$\mathbf{s}_n = 2^n \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - n \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}. \quad \text{Hence,} \quad \mathbf{s}_{100} = 2^{100} \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 100 \\ 201 \\ -99 \end{pmatrix}.$$