Linear Algebra II
Exercise Sheet no. 11

TECHNISCHE UNIVERSITÄT DARMSTADT

Summer term 2011
June 20, 2011

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

Exercise 1 (Warm-up)

(Exercise 3.1.1 in the notes, see also T7.2.) Show that the relation $\approx$ on $\mathbb{R}^{(n,n)}$ defined as $A \approx A'$ iff $A' = C^t AC$ for some $C \in \mathrm{GL}_n(\mathbb{R})$ is an equivalence relation. What are sufficient criteria for $A \not\approx A'$?

Solution:

- Reflexivity: $A \approx A$ since $A = E^t AE$ and $E \in \mathrm{GL}_n(\mathbb{R})$.
- Symmetry: if $A \approx B$, then $B = C^t AC$ for some $C \in \mathrm{GL}_n(\mathbb{R})$, and $(C^{-1})^t BC^{-1} = (C^{-1})^t C^t ACC^{-1} = A$.
- Transitivity: Assume $A \approx B$ and $B \approx C$. So $B = F^t AF$ for some $F \in \mathrm{GL}_n(\mathbb{R})$ and $C = G^t BG$ for some $G \in \mathrm{GL}_n(\mathbb{R})$. Then $C = (FG)^t AFG$, and $FG \in \mathrm{GL}_n(\mathbb{R})$, that is, $A \approx C$.

For instance, if two matrices A and A' have different ranks, $A \not\approx A'$. Another example, if A is symmetric and A' is not, $A \not\approx A'$.

Exercise 2 (Normal matrices)

Recall that a matrix A is called normal if $AA^+ = A^+A$. We have seen (cf Exercise T11.1) that unitary, hermitian, and skew-hermitian matrices are normal. (Similarly in the real case, orthogonal, symmetric, skew-symmetric matrices are normal.) In this exercise we will see that there are normal matrices that do not belong to any of these classes.

(a) Prove that every real 2×2 normal matrix is either symmetric or a scalar multiple of an orthogonal matrix.

(b) Find a sufficient (and also necessary) condition for a complex 2×2 matrix to be normal. Give an example of such a matrix which is neither hermitian, skew-hermitian, nor a scalar multiple of a unitary matrix.

(c) Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. Show that A is normal, but is neither symmetric, skew-symmetric, nor a scalar multiple of an orthogonal matrix.

Solution:

a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. So $AA^t = A^tA$ iff $b^2 = c^2$ and $(a-d)(c-b) = 0$. If $c = -b$ then $a = d$.

b) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. So $AA^+ = A^+A$ iff $b\bar{b} = c\bar{c}$ and $c(\bar{a} - \bar{d}) = \bar{b}(a-d)$. In the case where $a \neq d$, let $b = re^{i\beta}$ and $c = re^{i\gamma}$, so $AA^+ = A^+A$ iff $a - d = \rho e^{i\frac{\beta+\gamma}{2}}$ for some $\rho \in \mathbb{R}$. So the matrix $\begin{pmatrix} 2+i & i \\ 1 & 1 \end{pmatrix}$ is normal, but neither hermitian, skew-hermitian, nor a scalar multiple of a unitary matrix. (Why not?)

c) To see that A is not a scalar multiple of any orthogonal matrix, check that at least two of its column (or row) vectors are not orthogonal.

Exercise 3 (Canonical form of an orthogonal map)

Consider the endomorphism $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ represented in the standard basis by the following orthogonal matrix in $\mathbb{R}^{(3,3)}$:

$$A = \begin{pmatrix} -1/2 & 1/2 & -1/\sqrt{2} \\ 1/2 & -1/2 & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}.$$

(a) Regard A as a complex matrix via the inclusion $\mathbb{R}^{(3,3)} \subseteq \mathbb{C}^{(3,3)}$, and find its characteristic polynomial over $\mathbb{C}$.

(b) Find a basis of complex eigenvectors $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ of A.

(c) Use this information to find the invariant subspaces of φ regarded again as an endomorphism of $\mathbb{R}^3$. Find an orthonormal basis for $\mathbb{R}^3$ such that in this basis, φ is given by a rotation followed by a reflection.

Solution:

a) Over $\mathbb{C}$, we have $p_A = (X+i)(X-i)(X+1)$.

b) The vectors $\mathbf{v}_1 = \begin{pmatrix} i/\sqrt{2} \\ i/\sqrt{2} \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -i/\sqrt{2} \\ -i/\sqrt{2} \\ 1 \end{pmatrix}$, and $\mathbf{v}_3 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$ are eigenvectors of A in $\mathbb{C}^3$, with eigenvalues $i, -i, -1$, respectively.

c) Note that $\mathbf{v}_3$ actually lies in $\mathbb{R}^3$, and the vector space $V \subseteq \mathbb{C}^3$ spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$ is invariant under A. We seek a new basis $(\mathbf{u}_1, \mathbf{u}_2)$ for V consisting of vectors in $\mathbb{R}^3$, such that V is the complexification of the vector space $U \subseteq \mathbb{R}^3$ spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. In fact, we have already seen in Exercise (T5.3) how to find these vectors. Let $\mathbf{u}_1 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$, $\mathbf{u}_2 = \frac{1}{2i}(\mathbf{v}_1 - \mathbf{v}_2)$. As shown in that exercise, the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ have the desired properties. The invariant subspaces of φ regarded as an endomorphism of $\mathbb{R}^3$ are therefore U and the one-dimensional space spanned by $\mathbf{v}_3$.

Letting $\mathbf{u}_3 = \mathbf{v}_3$, we see that with respect to the basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for $\mathbb{R}^3$, φ is given by rotation through the angle π in the plane spanned by $\mathbf{u}_1, \mathbf{u}_2$, followed by reflection in this plane. In particular, letting S be the (orthogonal) matrix whose columns are $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$, we have

$$S^{-1}AS = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Exercise 4 (Dual maps)

Let $(V, \langle\cdot,\cdot\rangle^V)$ and $(W, \langle\cdot,\cdot\rangle^W)$ be finite-dimensional euclidean spaces. Recall from Exercise T8.4 that the scalar product of V induces a canonical (i.e., basis-independent) isomorphism $\rho^V : V \to V^*$, where $V^* = Hom(V, \mathbb{R})$ is the *dual space* of V.

$$\rho^V : \quad V \to V^*, \qquad \mathbf{v} \mapsto \langle \mathbf{v}, \cdot \rangle^V$$

where

$$\langle \mathbf{v}, \cdot \rangle^V : \quad V \to \mathbb{R}, \qquad \mathbf{u} \mapsto \langle \mathbf{v}, \mathbf{u} \rangle^V$$

Note that $\rho^W : W \to W^*$ is defined similarly.

(a) Let $\varphi \in Hom(V, W)$ be a linear map. We define the *dual* of φ to be the map $\varphi^* \in Hom(W^*, V^*)$ as follows:

$$\varphi^* : \quad W^* \to V^*, \qquad \eta \mapsto \eta \circ \varphi$$

Note that everything we have defined so far does not depend on a choice of basis. Now let $B_V = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be any basis for V. We define the *dual basis* $B_V^* = (\mathbf{b}_1^*, \dots, \mathbf{b}_n^*)$ for V^* by the condition $\mathbf{b}_j^*(\mathbf{b}_j) = 0$ for $i \neq j$ and $\mathbf{b}_j^*(\mathbf{b}_j) = 1$ for $i = j$. Similarly, fix a basis $B_W = (\hat{\mathbf{b}}_1, \dots, \hat{\mathbf{b}}_m)$ for W, with associated dual basis B_W^*. Show that the relationship between the matrix representations of φ and φ^* w.r.t. these bases is
$[\![\varphi^*]\!]_{B_V^*}^{B_W^*} = ([\![\varphi]\!]_{B_W}^{B_V})^t$.

(b) What is the status of the map $\varphi^+ := (\rho^V)^{-1} \circ \varphi^* \circ \rho^W$ w.r.t. $\langle \cdot, \cdot \rangle^W$ and $\langle \cdot, \cdot \rangle^V$? Discuss its matrix representations w.r.t. the orthonormal bases B_V and B_W.

(c) In the special case of $V = W = (V, \langle \cdot, \cdot \rangle)$, consider the map $\varphi^+ = (\rho^V)^{-1} \circ \varphi^* \circ \rho^W$ and try to interpret the adjoint of the endomorphism φ in terms of an isomorphic copy of the dual φ^* via canonical identifications of V with V^* via ρ^V.

Analyse the change of basis transformations w.r.t. changes from an onb $B_V (= B_W)$ to another onb $B'_V (= B'_W)$.

Solution:

a) By definition, $[\![\varphi]\!]_{B_W}^{B_V}$ is the matrix A whose entries are given by $\varphi(\mathbf{b}_i) = \sum_{j=1}^m a_{ji} \hat{\mathbf{b}}_j$. Hence $\hat{\mathbf{b}}_j^*(\varphi(\mathbf{b}_i)) = a_{ji}$.

Next, we calculate the matrix $[\![\varphi^*]\!]_{B_V^*}^{B_W^*}$. By definition, $\varphi^*(\hat{\mathbf{b}}_i^*) = \hat{\mathbf{b}}_i^* \circ \varphi$, so

$$\varphi^*(\hat{\mathbf{b}}_i^*)(\mathbf{b}_j) = \hat{\mathbf{b}}_i^*(\varphi(\mathbf{b}_j)) = \hat{\mathbf{b}}_i^*(\sum_{k=1}^m a_{kj} \hat{\mathbf{b}}_k) = a_{ij}.$$

The claim follows.

b) Let $B_V = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and $B_W = (\hat{\mathbf{b}}_1, \ldots, \hat{\mathbf{b}}_m)$ now be orthonormal bases for V and W, respectively. It follows from the definition of ρ^V and ρ^W that $\rho^V(\mathbf{b}_i) = \mathbf{b}_i^*$ and $\rho^W(\hat{\mathbf{b}}_i) = \hat{\mathbf{b}}_i^*$. Hence

$$\varphi^+(\hat{\mathbf{b}}_i) = (\rho^V)^{-1} \circ \varphi^* \circ \rho^W(\hat{\mathbf{b}}_i) = (\rho^V)^{-1} \circ \varphi^*(\hat{\mathbf{b}}_i^*) = (\rho^V)^{-1}(\sum_{j=1}^n a_{ij} \mathbf{b}_j^*) = \sum_{j=1}^n a_{ij} \mathbf{b}_j.$$

The equality $[\![\varphi^+]\!]_{B_V}^{B_W} = ([\![\varphi]\!]_{B_W}^{B_V})^t$ follows.

c) If $V = W$ and we identify V with V^* via ρ^V, it follows from (a) and (b) that φ^+ corresponds to φ^* under this identification.

Exercise 5 (Positive definiteness and compactness of the unit surface)

(a) Let σ_A be a bilinear form on $\mathbb{R}^n$, which in the standard basis is represented by a symmetric matrix A, whose ijth entry $a_{ij} = \sigma_A(\mathbf{e}_i, \mathbf{e}_j)$. Define the *unit surface*

$$S_A = \{\mathbf{v} \in \mathbb{R}^n : \sigma_A(\mathbf{v}, \mathbf{v}) = 1\}.$$

Suppose that S_A is non-empty. Prove that S_A is compact if and only if σ_A is positive definite.

(b) Let A and B be matrices representing scalar products $\langle \cdot, \cdot \rangle_A$ and $\langle \cdot, \cdot \rangle_B$ on $\mathbb{R}^n$. Show that the corresponding norms are equivalent in the sense that there exist positive real numbers m and M satisfying

$$m \langle \mathbf{v}, \mathbf{v} \rangle_A \leqslant \langle \mathbf{v}, \mathbf{v} \rangle_B \leqslant M \langle \mathbf{v}, \mathbf{v} \rangle_A$$

for all $\mathbf{v} \in \mathbb{R}^n$.

Solution:

a) Let $S = \{\mathbf{v} : \|\mathbf{v}\| = 1\}$ denote the unit sphere in $\mathbb{R}^n$ with respect to the standard inner product, and let $f : \mathbb{R}^n \to \mathbb{R}$ denote the quadratic function $f(\mathbf{v}) = \sigma_A(\mathbf{v}, \mathbf{v})$, which is clearly continuous. Since $S_A = f^{-1}(1)$ is closed by continuity, S_A is compact if and only if it is bounded.

Suppose first that σ_A is positive definite. Since S is compact, f achieves a minimum value m on S, and since σ_A is positive definite, we have $m > 0$. Let $\mathbf{v} \in S_A$. Since $\frac{\mathbf{v}}{\|\mathbf{v}\|} \in S$, we have

$$f(\frac{\mathbf{v}}{\|\mathbf{v}\|}) = \sigma_A(\frac{\mathbf{v}}{\|\mathbf{v}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|}) = \frac{1}{\|\mathbf{v}\|^2} \sigma_A(\mathbf{v}, \mathbf{v}) = \frac{1}{\|\mathbf{v}\|^2} \geqslant m.$$

Finally, since $m > 0$ we have $\|\mathbf{v}\| \leqslant \frac{1}{\sqrt{m}}$, so S_A is bounded.

Conversely, suppose that σ_A is not positive definite, so there exists $\mathbf{u} \in S$ such that $f(\mathbf{u}) \leqslant 0$. Since S_A is not empty, there exists $\mathbf{w} \in S$ for which $f(\mathbf{w}) > 0$. The two vectors $\mathbf{u}$ and $\mathbf{w}$ are therefore independent, so for all $\lambda \in [0, 1]$ the vector $\frac{\lambda \mathbf{u} + (1-\lambda)\mathbf{w}}{\|\lambda \mathbf{u} + (1-\lambda)\mathbf{w}\|}$ is well-defined, and in S. So for any $\varepsilon > 0$ there exists $\lambda \in [0, 1]$ such that $\mathbf{v} = \frac{\lambda \mathbf{u} + (1-\lambda)\mathbf{w}}{\|\lambda \mathbf{u} + (1-\lambda)\mathbf{w}\|}$ satisfies $0 < f(\mathbf{v}) < \varepsilon$, by the intermediate value theorem. It follows that $\mathbf{w} = \frac{\mathbf{v}}{\sqrt{|f(\mathbf{v})|}} \in S_A$ and $\|\mathbf{w}\| > \frac{1}{\sqrt{\varepsilon}}$. Since ε was arbitrary, it follows that S_A is not bounded.

b) Let m_A and M_A be the minimum and maximum values achieved by the function f_A defined by $f_A(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v} \rangle_A$ on the sphere S. Clearly m_A and M_A are positive real numbers. Similarly, let m_B and M_B be the minimum and maximum values achieved by the function f_B defined by $f_B(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v} \rangle_B$ on S. Define $M = \frac{M_B}{m_a}$ and $m = \frac{m_B}{M_A}$. The desired inequality

$$m\langle \mathbf{v}, \mathbf{v} \rangle_A \leqslant \langle \mathbf{v}, \mathbf{v} \rangle_B \leqslant M\langle \mathbf{v}, \mathbf{v} \rangle_A$$

follows for all $\mathbf{v} \in S$. The fact that this holds for all $\mathbf{v} \in \mathbb{R}^n$ follows.