Aalto University

Linear algebra

Exercise sheet 10 / Model solutions

1. (a) Let $\mathcal{V} \subset \mathbb{R}^n$ be a subspace and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ an *orthonormal basis* for $\mathcal{V}$ with respect to the inner product $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Show that for any $\boldsymbol{x} \in \mathbb{R}^n$ it holds that

$$\boldsymbol{x}^\perp := \boldsymbol{x} - \sum_{i=1}^k \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \boldsymbol{v}_i \ \in \mathcal{V}^\perp,$$

where $\mathcal{V}^\perp \subset \mathbb{R}^n$ is the orthogonal complement of $\mathcal{V}$ with respect to the inner product $\langle \cdot, \cdot \rangle$, see Definition 2.3 on page 45 of the lecture notes. (This proves that equation (75) there does define a direct sum $\mathbb{R}^n = \mathcal{V} \oplus \mathcal{V}^\perp$.)

(b) Prove that any projection matrix $P \in \mathbb{R}^{n\times n}$ is an orthogonal projection in the Euclidian inner product, if $P^T = P$.

Hints: (a) Write an arbitrary vector of the subspace $\mathcal{V}$ in the form $\boldsymbol{v} = \sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ and show by direct calculation that $\langle \boldsymbol{x}^\perp, \boldsymbol{v} \rangle = 0$. (b) With a projection P we may write the direct sum $\mathbb{R}^n = R(P) \oplus R(I - P)$; check equation (72) from the lecture notes. You have to show that $R(I-P) = R(P)^\perp$. Since moreover $\mathbb{R}^n = R(P) \oplus R(P)^\perp$, it's actually enough to prove that $R(I-P) \subseteq R(P)^\perp$.

Solution:

(a) We may write as $\boldsymbol{v} = \sum_{j=1}^k \alpha_j \boldsymbol{v}_j$ an abitrary element of the subspace $\mathcal{V} \subset \mathbb{R}^n$. Then

$$\begin{aligned}
\langle \boldsymbol{x}^\perp, \boldsymbol{v} \rangle &= \langle \boldsymbol{x} - \sum_{i=1}^k \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \boldsymbol{v}_i, \sum_{j=1}^k \alpha_j \boldsymbol{v}_j \rangle \\
&= \langle \boldsymbol{x}, \sum_{j=1}^k \alpha_j \boldsymbol{v}_j \rangle - \langle \sum_{i=1}^k \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \boldsymbol{v}_i, \sum_{j=1}^k \alpha_j \boldsymbol{v}_j \rangle \\
&= \sum_{j=1}^k \alpha_j \langle \boldsymbol{x}, \boldsymbol{v}_j \rangle - \sum_{i=1}^k \sum_{j=1}^k \alpha_j \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle \\
&= \sum_{j=1}^k \alpha_j \langle \boldsymbol{x}, \boldsymbol{v}_j \rangle - \sum_{j=1}^k \alpha_j \langle \boldsymbol{x}, \boldsymbol{v}_j \rangle = 0,
\end{aligned}$$

where we used for the fourth equality that

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Hence $\boldsymbol{x}^\perp$ is orthogonal to every element of $\mathcal{V}$, that is, $\boldsymbol{x}^\perp \in \mathcal{V}^\perp$.

(b) As suggested in the hint, let us show the inclusion $R(I-P) \subset R(P)^{\perp}$. Take $\boldsymbol{x} \in R(I-P)$, meaning that $\boldsymbol{x} = (I-P)\boldsymbol{z}$ for some $\boldsymbol{z} \in \mathbb{R}^n$. Let us show that $\boldsymbol{x}$ is orthogonal to any element $\boldsymbol{y} = P\boldsymbol{w}$ of $R(P)$, where $\boldsymbol{w} \in \mathbb{R}^n$:

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{y}\rangle_2 &= \boldsymbol{x}^T\boldsymbol{y}\\ &= ((I-P)\boldsymbol{z})^T(P\boldsymbol{w})\\ &= \boldsymbol{z}^T(I-P)^TP\boldsymbol{w}\\ &= \boldsymbol{z}^T(I^T-P^T)P\boldsymbol{w}\\ &= \boldsymbol{z}^T(I-P)P\boldsymbol{w}\\ &= \boldsymbol{z}^T(P-P^2)\boldsymbol{w} = 0,\end{aligned}$$

where in the fifth step we used that $P = P^T$ by assumption and in the last step that $P^2 = P$ because P is a projection. So we have that $\boldsymbol{x} \in R(P)^{\perp}$, that is, $R(I-P) \subseteq R(P)^{\perp}$.

Why is this enough? Namely, why must it be that $R(I-P) = R(P)^{\perp}$? Let $\boldsymbol{v} \in R(P)^{\perp}$. Since $\mathbb{R}^n = R(P) \oplus R(I-P)$ (see formula (72) on page 45 of the lecture notes) there exist uniquely determined vectors $\boldsymbol{w}_1 \in \mathbb{R}(P)$ and $\boldsymbol{w}_2 \in R(I-P)$ such that

$$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2.$$

Since $\boldsymbol{v} \in R(P)^{\perp}$, by taking the inner product with $\boldsymbol{w}_1$we get

$$0 = \langle \boldsymbol{v}, \boldsymbol{w}_1\rangle = \langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle + \langle \boldsymbol{w}_2, \boldsymbol{w}_1\rangle = \|\boldsymbol{w}_1\|^2,$$

where $\langle \boldsymbol{w}_2, \boldsymbol{w}_1\rangle = 0$ by the inclusion proved in the exercise. So then $\boldsymbol{w}_1 = 0$ and $\boldsymbol{v} = \boldsymbol{w}_2 \in R(I-P)$.

2. Let

$$\mathcal{V} = \text{span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}.$$

(a) Find an orthogonal projection P such that $R(P) = \mathcal{V}$.

(b) Choose some other basis for the subspace $\mathcal{V}$. Compute P again in this new basis.

Solution: Let us consider the situation in general. We want to find a matrix $P \in \mathbb{R}^{n\times n}$ such that P is a projection matrix for the orthogonal projection to the subspace $R(V)$, where $V \in \mathbb{R}^{n\times m}$ is the matrix whose columns generate the subspace $\mathcal{V}$.

Let $\boldsymbol{x} \in \mathbb{R}^n$ be a vector we want to project. It can be written as

$$\boldsymbol{x} = V\boldsymbol{a} + W\boldsymbol{b}, \tag{1}$$

where $\boldsymbol{a} \in \mathbb{R}^m$, $\boldsymbol{b} \in \mathbb{R}^k$, and $W \in \mathbb{R}^{n\times k}$. Furthermore let us choose W so that $R([V \;\; W]) = \mathbb{R}^n$ (that is, the columns of the matrices V and W together generate $\mathbb{R}^n$) and $V^TW = 0$, hence the projection would be orthogonal. (We can make this choice, since for $\mathbb{R}^n$ there exists an orthogonal basis such that for some $k \in \mathbb{N}$ some k basis vectors generate the subspace $R(V)$. Now the rest of the basis vectors go to the matrix W and they are orthogonal to $R(V)$, as they are orthogonal to the columns of V.)

So now the task is to find a P such that $P\boldsymbol{x} = V\boldsymbol{a}$. Multiplying equation (1) by V^T on the left, we get

$$V^T\boldsymbol{x} = V^TV\boldsymbol{a} + V^TW\beta = V^TV\boldsymbol{a}.$$

The columns of V are chosen to be linearly independent, so that $N(V) = \{\mathbf{0}\}$, from which follows by some previous homework that we also have $N(V^TV) = \{\mathbf{0}\}$, meaning that V^TV is invertible. We then get that $\boldsymbol{a} = (V^TV)^{-1}V^T\boldsymbol{x}$, hence $P\boldsymbol{a} = V(V^TV)^{-1}V^T\boldsymbol{x}$. And therefore

$$P = V(V^TV)^{-1}V^T. \tag{2}$$

(a) So now when we pick

$$V = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

as given in the assignment, and pluggin it in (2) we get

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

And to conclude we also have $R(P) = \mathcal{V}$, because clearly every point of $\mathcal{V}$ we can express as a linear combination of the columns of P, and conversely P is a projection to the subspace $\mathcal{V}$.

(b) This time let us choose as basis $\{(1,1,0),(0,0,1)\}$ for $\mathcal{V}$. We get, again by using (2),

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This was expected, as the projection depends on the space on which we project, not on its basis.

3. Let

$$\mathcal{V} = \text{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \mathcal{W} = \text{span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Denote

$$V = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

You may assume known that $\mathbb{R}^4 = \mathcal{V} \oplus \mathcal{W}$.

(a) Using the matrix V give a formula for the orthogonal projection matrix $Q \in \mathbb{R}^{4\times4}$ such that $R(Q) = \mathcal{V}$.

(b) Using the matrices V and W give a formula for the projection matrix $P \in \mathbb{R}^{4\times4}$ such that $R(P) = \mathcal{V}$ and $N(P) = \mathcal{W}$.

(c) Compute the projection matrices in (a) and (b) using Matlab. Check that $Q^2 = Q$, $P^2 = P$, $Q^T(I - Q) = 0$, $PV = V$ and $PW = 0$ (within floating point accuracy).

Hints: (b) If P is constructed so that it projects the direct sum $\mathbb{R}^4 = \mathcal{V} \oplus \mathcal{W}$ to the first component, what is the null space of P?

(a) We can use equation (2), in which now we have $P = Q$ and

$$V = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

(b) Let us divide an arbitrary vector $\boldsymbol{x}$ in components so that the columns of V and W ($V \in \mathbb{R}^{n\times m}$, $W \in \mathbb{R}^{n\times k}$) together form a basis for $\mathbb{R}^n$ (from which follows that $[V\ W]$ is invertible):

$$\boldsymbol{x} = V\boldsymbol{a} + W\boldsymbol{b}.$$

Next we find the matrix P for the projection $\boldsymbol{x} \mapsto V\boldsymbol{a}$: from

$$\boldsymbol{x} = V\boldsymbol{a} + W\boldsymbol{b} = \begin{bmatrix} V & W \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix},$$

it follows that

$$\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} V & W \end{bmatrix}^{-1} \boldsymbol{x}$$

which we can use in order to compute

$$P\boldsymbol{x} = V\boldsymbol{a} = V \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0}_{m\times k} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} = V \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0}_{m\times k} \end{bmatrix} \begin{bmatrix} V & W \end{bmatrix}^{-1} \boldsymbol{x}.$$

And this means that

$$P = V \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0}_{m\times k} \end{bmatrix} \begin{bmatrix} V & W \end{bmatrix}^{-1} = \begin{bmatrix} V & \boldsymbol{0}_{m\times k} \end{bmatrix} \begin{bmatrix} V & W \end{bmatrix}^{-1}. \tag{3}$$

Let us now choose W so that $N(P) = \mathcal{W}$. First of all, we know that $N(P) = R(I-P)$, and here is a proof: ($\subseteq$) If $\boldsymbol{x} \in N(P)$, then $P\boldsymbol{x} = 0$, so that $\boldsymbol{x} = (I - P)\boldsymbol{x}$, and so $\boldsymbol{x} \in R(I - P)$. ($\supseteq$) If $\boldsymbol{x} \in R(I - P)$, then $\boldsymbol{x} = (I - P)\boldsymbol{y}$ for some $\boldsymbol{y} \in \mathbb{R}^n$, and since $(I - P)$ is a projection matrix (see the first part of problem 1), we get

$$\boldsymbol{x} = (I - P)\boldsymbol{y} = (I - P)(I - P)\boldsymbol{y} = (I - P)\boldsymbol{x},$$

hence $P\boldsymbol{x} = 0$, so that $\boldsymbol{x} \in N(P)$. So that's why $N(P) = R(I - P)$. Now, since $W\boldsymbol{b} = (I - P)\boldsymbol{x}$, as $R(W) = R(I - P)$, we have

$$\mathcal{W} = N(P) = R(I - P) = R(W),$$

that is, we can choose

$$W = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Since we want to project to the subspace $\mathcal{V}$, we choose as in part (a)

$$V = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Now we then get the desired projection P in (3).

(c) Matlab code:

```
V = [1 1; 2 1; 1 2; 0 1]
W = [2 1; 1 1; 1 1; 0 1]
Q = V*inv(V'*V)*V'
P = V*[eye(size(V)(2)) zeros(size(V)(2))]*inv([V W])
Q^2                                % =Q? i.e.  is it a proj.  matrix?
P^2                                % =P?
Q'*(eye(size(Q))-Q)                % =0?  orthogonality
```

We get the expected results. As projection matrices we get

$$Q = \begin{bmatrix} 0.176471 & 0.294118 & 0.235294 & 0.058824 \\ 0.294118 & 0.823529 & 0.058824 & -0.235294 \\ 0.235294 & 0.058824 & 0.647059 & 0.411765 \\ 0.058824 & -0.235294 & 0.411765 & 0.352941 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 3 & 5 & 4 & 1 \\ 5 & 14 & 1 & -4 \\ 4 & 1 & 11 & 7 \\ 1 & -4 & 7 & 6 \end{bmatrix},$$

$$P = \begin{bmatrix} -6.6667e-01 & 3.3333e-01 & 1.0000e+00 & -6.6667e-01 \\ -1.0000e+00 & 1.0000e+00 & 1.0000e+00 & -1.0000e+00 \\ -1.0000e+00 & 1.1102e-16 & 2.0000e+00 & -1.0000e+00 \\ -3.3333e-01 & -3.3333e-01 & 1.0000e+00 & -3.3333e-01 \end{bmatrix}$$
$$= \frac{1}{3}\begin{bmatrix} -2 & 1 & 3 & -2 \\ -3 & 3 & 3 & -3 \\ -3 & 0 & 6 & -3 \\ -1 & -1 & 3 & -1 \end{bmatrix}.$$

4. Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$ be such that $\boldsymbol{a}^T\boldsymbol{b} = 0$ and define the matrix $P \in \mathbb{R}^{3\times 3}$ as
$$P = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\|\boldsymbol{a}\|_2^2} + \frac{\boldsymbol{b}\boldsymbol{b}^T}{\|\boldsymbol{b}\|_2^2}.$$
   (a) Is P a projection matrix? How about $I - P$, where I is the identity matrix?
   (b) Is P an orthogonal projection?
   (c) Let $\boldsymbol{a} = [1, 1, 1]^T$ and $\boldsymbol{b} = [1, -2, 1]^T$. Find $\boldsymbol{c} \in \mathbb{R}^3$ such that
$$I - P = \frac{\boldsymbol{c}\boldsymbol{c}^T}{\|\boldsymbol{c}\|_2^2}.$$

Hints: (a) See Definition 2.1 on page 45 of the lecture notes. (b) Use problem 1(b). (c) The vector $\boldsymbol{c}$ has to be orthogonal to $\boldsymbol{a}$ and $\boldsymbol{b}$, since $R(I - P) = R(P)^{\perp}$.

Solution:

(a) By definition a matrix A is a projection matrix if $A^2 = A$ (that is, once projected a point doesn't move anymore). For P we get
$$\begin{aligned}
P^2 &= \frac{(\boldsymbol{a}\boldsymbol{a}^T)(\boldsymbol{a}\boldsymbol{a}^T)}{\|\boldsymbol{a}\|_2^4} + \frac{(\boldsymbol{b}\boldsymbol{b}^T)(\boldsymbol{b}\boldsymbol{b}^T)}{\|\boldsymbol{b}\|_2^4} + \frac{(\boldsymbol{a}\boldsymbol{a}^T)(\boldsymbol{b}\boldsymbol{b}^T)}{\|\boldsymbol{a}\|_2^2\|\boldsymbol{b}\|_2^2} + \frac{(\boldsymbol{b}\boldsymbol{b}^T)(\boldsymbol{a}\boldsymbol{a}^T)}{\|\boldsymbol{b}\|_2^2\|\boldsymbol{a}\|_2^2} \\
&= \frac{\boldsymbol{a}(\boldsymbol{a}^T\boldsymbol{a})\boldsymbol{a}^T}{\|\boldsymbol{a}\|_2^4} + \frac{\boldsymbol{b}(\boldsymbol{b}^T\boldsymbol{b})\boldsymbol{b}^T}{\|\boldsymbol{b}\|_2^4} + \frac{\boldsymbol{a}(\boldsymbol{a}^T\boldsymbol{b})\boldsymbol{b}^T}{\|\boldsymbol{a}\|_2^2\|\boldsymbol{b}\|_2^2} + \frac{\boldsymbol{b}(\boldsymbol{b}^T\boldsymbol{a})\boldsymbol{a}^T}{\|\boldsymbol{b}\|_2^2\|\boldsymbol{a}\|_2^2} \\
&= \frac{\boldsymbol{a}\boldsymbol{a}^T}{\|\boldsymbol{a}\|_2^2} + \frac{\boldsymbol{b}\boldsymbol{b}^T}{\|\boldsymbol{b}\|_2^2} = P,
\end{aligned}$$
where in the third step we use that $\boldsymbol{a}^T\boldsymbol{a} = \|\boldsymbol{a}\|_2^2$ and $\boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{b}^T\boldsymbol{a} = 0$. So P is a projection matrix.
For an arbitrary projection matrix, it holds that $I - P$ is also a projection matrix, because
$$(I - P)^2 = I^2 - 2P + P^2 = I - 2P + P = I - P.$$

(b) The projection P splits any vector $\boldsymbol{x}$ in two components:
$$\boldsymbol{x} = P\boldsymbol{x} + (I - P)\boldsymbol{x}.$$
The projection P is orthogonal if these components are orthogonal to each other. So P is orthogonal if $(P\boldsymbol{x})^T((I - P)\boldsymbol{x}) = 0$. We get
$$\begin{aligned}
(P\boldsymbol{x})^T((I - P)\boldsymbol{x}) &= (P\boldsymbol{x})^T\boldsymbol{x} - (P\boldsymbol{x})^T P\boldsymbol{x} \\
&= \boldsymbol{x}^T P^T \boldsymbol{x} - \boldsymbol{x}^T P^T P\boldsymbol{x} \\
&= \boldsymbol{x}^T P\boldsymbol{x} - \boldsymbol{x}^T P\boldsymbol{x} \\
&= 0,
\end{aligned}$$
where we use that P is symmetric and $P^2 = P$. Then P is an orthogonal projection. The same we could show by checking that $P^T(I - P) = \boldsymbol{0}$, because then the subspaces $R(P)$ and $R(I - P)$ are orthogonal, from which follows that the associated components are orthogonal to each other.

(c) The projection matrix of the form $A = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\|\boldsymbol{a}\|^2}$ is an orthogonal projection to the subspace span$\{\boldsymbol{a}\}$ (see the lecture notes). Then P is a map to the subspacespan$\{\boldsymbol{a}, \boldsymbol{b}\}$, also an orthogonal projection, which we proved in the previous part.

For an orthogonal projection P the components $P\boldsymbol{x}$ and $(I - P)\boldsymbol{x}$ are always orthogonal, and $(I - P)\boldsymbol{x}$ is now the projection to the subspace span$\{\boldsymbol{c}\}$, the vector $\boldsymbol{c}$ must be orthogonal to the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. So then we get $\boldsymbol{c}$ by cross product:

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}.$$

We can still rescale the vector $\boldsymbol{c}$, for instance we can choose $\boldsymbol{c} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$.