D-INFK
Bernd Gärtner
Robert Weismantel

Linear Algebra

HS 2024

Solution for Assignment 10

1. **a)** Consider the matrix $M := AB$. We claim that it has $rank(M) = n$. To see this, observe that $\text{rank}(B) = n$ implies $\mathbf{C}(B) = \mathbb{R}^n$ because n is also the number of rows of B. Hence, we get $\mathbf{C}(M) = \mathbf{C}(A)$ (and therefore $\text{rank}(M) = \text{rank}(A) = n$). Finally, we can use Proposition 5.5.9 to get $(AB)^\dagger = M^\dagger = B^\dagger A^\dagger$.

b) Let $A = CR$ be the CR decomposition of A with $C \in \mathbb{R}^{m\times r}$ and $R \in \mathbb{R}^{r\times n}$ where $r = \text{rank}(A)$. Observe that C has full column rank and that R has full row rank. Using the definition of the pseudoinverse, we compute

$$A^\dagger A A^\dagger = (CR)^\dagger CR(CR)^\dagger = R^\dagger(C^\dagger C)(RR^\dagger)C^\dagger = R^\dagger C^\dagger = A^\dagger$$

where we used that $R^\dagger$ is a right inverse of R and $C^\dagger$ a left inverse of C.

c) Assume first that A has full column rank $n = \text{rank}(A)$. In this case, we have $A^\dagger = (A^\top A)^{-1}A^\top$ by definition of the pseudoinverse for matrices with full column rank. Moreover, notice that $A^\top$ has full row rank and hence we also get $(A^\top)^\dagger = A(A^\top A)^{-1}$ by definition of the pseudoinverse for matrices with full row rank. Hence, we get

$$(A^\dagger)^\top = ((A^\top A)^{-1}A^\top)^\top = A((A^\top A)^{-1})^\top = A((A^\top A)^\top)^{-1} = A(A^\top A)^{-1} = (A^\top)^\dagger.$$

We conclude that the statements holds for all matrices with full column rank.

Analogously, we can prove that the statement holds if A has full row rank $m = \text{rank}(A)$. In that case, we have $A^\dagger = A^\top(AA^\top)^{-1}$ and $(A^\top)^\dagger = (AA^\top)^{-1}A$. Hence, we indeed get

$$(A^\dagger)^\top = (A^\top(AA^\top)^{-1})^\top = ((AA^\top)^{-1})^\top A = ((AA^\top)^\top)^{-1}A = (AA^\top)^{-1}A = (A^\top)^\dagger.$$

We conclude that the statement holds for all matrices with full row rank.

It remains to prove the general case, i.e. we do not assume anymore that A has full row rank or full column rank. Then by definition, we have $A^\dagger = R^\dagger C^\dagger$ where $A = CR$ is a CR decomposition of A. In particular, we have $C \in \mathbb{R}^{m\times r}$ and $R \in \mathbb{R}^{r\times n}$ where $r = \text{rank}(A)$. Now observe that we also have $A^\top = R^\top C^\top$ with $R^\top \in \mathbb{R}^{n\times r}$ and $C^\top \in \mathbb{R}^{r\times m}$ and of course, $r = \text{rank}(A) = \text{rank}(A^\top)$. Hence, we can use Proposition 5.5.9 to get $(A^\top)^\dagger = (C^\top)^\dagger(R^\top)^\dagger$. We conclude that

$$(A^\top)^\dagger = (C^\top)^\dagger(R^\top)^\dagger = (C^\dagger)^\top(R^\dagger)^\top = (R^\dagger C^\dagger)^\top = (A^\dagger)^\top$$

by using that C has full column rank and R has full row rank and hence $(C^\top)^\dagger = (C^\dagger)^\top$ and $(R^\top)^\dagger = (R^\dagger)^\top$.

d) Let $A = CR$ be a CR decomposition of A with $C \in \mathbb{R}^{m\times r}$ and $R \in \mathbb{R}^{r\times n}$ where $r = \text{rank}(A)$. We can rewrite

$$A^\dagger A = (CR)^\dagger CR = R^\dagger C^\dagger CR \overset{\text{Prop. 5.5.2}}{=} R^\dagger IR = R^\top(RR^\top)^{-1}R$$

and hence we conclude symmetry of $AA^\dagger$ since

$$(A^\dagger A)^\top = (R^\top(RR^\top)^{-1}R)^\top = R^\top((RR^\top)^{-1})^\top R = R^\top((RR^\top)^\top)^{-1}R = R^\top(RR^\top)^{-1}R = A^\dagger A.$$

By Theorem 5.2.6, the matrix $R^\top(RR^\top)^{-1}R = A^\dagger A$ is exactly the projection matrix onto the subspace $\mathbf{C}(R^\top) = \mathbf{R}(R) = \mathbf{R}(A) = \mathbf{C}(A^\top)$ (the equality $\mathbf{R}(R) = \mathbf{R}(A)$ is due to the observation that R can be obtained from A through row operations and deleting 0-rows, and by recalling that row operations preserve the row space).

2. We provide two solutions.

- In this first solution, we solve this by using our knowledge on pseudoinverses. Consider the function $f^{-1} : \mathbf{C}(A) \to \mathbf{C}(A^\top)$ given by $f^{-1}(\mathbf{x}) = A^\dagger \mathbf{x}$ for all $\mathbf{x} \in \mathbf{C}(A)$. Observe that the composition $f^{-1} \circ f$ is the identity: we know from Exercise 1 that $A^\dagger A$ is the projection matrix that projects vectors onto the subspace $\mathbf{C}(A^\top)$, and hence we have
$$f^{-1}(f(\mathbf{x})) = A^\dagger A\mathbf{x} = \mathbf{x}$$
for all $\mathbf{x} \in \mathbf{C}(A^\top)$. This already implies that f is injective. Observe that with an analogous argument we get
$$f(f^{-1}(\mathbf{x})) = AA^\dagger \mathbf{x} = \mathbf{x}$$
for all $\mathbf{x} \in \mathbf{C}(A)$. Hence, f^{-1} is injective as well which implies that both f and f^{-1} are bijective.
Note that the matrix $A^\dagger A$ is in general not the identity matrix. It is crucial that the function f is only defined on $\mathbf{C}(A^\top)$ and not on all of $\mathbb{R}^n$.
- In this second solution, we start by proving injectivity. For this, let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{C}(A^\top)$ be arbitrary and assume that $f(\mathbf{x}_1) = f(\mathbf{x}_2)$. We want to argue that this implies $\mathbf{x}_1 = \mathbf{x}_2$. Observe that we have
$$0 = f(\mathbf{x}_1) - f(\mathbf{x}_2) = A(\mathbf{x}_1 - \mathbf{x}_2)$$
and therefore $\mathbf{x}_1 - \mathbf{x}_2 \in \mathbf{N}(A)$. Together with $\mathbf{C}(A^\top) \cap \mathbf{N}(A) = \{0\}$ and $\mathbf{x}_1 - \mathbf{x}_2 \in \mathbf{C}(A^\top)$, we conclude $\mathbf{x}_1 - \mathbf{x}_2 = 0$ and hence $\mathbf{x}_1 = \mathbf{x}_2$.
It remains to prove surjectivity. Let $\mathbf{y} \in \mathbf{C}(A)$ be arbitrary. By Theorem 5.1.10, we know that there exists $\mathbf{x} \in \mathbf{C}(A^\top)$ such that $\{\mathbf{z} \in \mathbb{R}^d : A\mathbf{z} = \mathbf{y}\} = \mathbf{x} + \mathbf{N}(A)$ (the theorem applies because the set is non-empty since $\mathbf{y} \in \mathbf{C}(A)$). In particular, we have $f(\mathbf{x}) = A\mathbf{x} = \mathbf{y}$, as desired.

3. For every $k \in \{1, \dots, n-1\}$, we define $S_{n-k} = \{1, \dots, n-k\}$.

Our strategy is as follows: We first prove inductively that $\mathsf{proj}_{S_{n-j}}(P)$ is a polyhedron for all $1 \leq j < n$. This will then allow us to generalize the proof of Lemma 5.6.4 accordingly.

- For the base case $j = 1$, observe that $\mathsf{proj}_{S_{n-j}}(P)$ by Theorem 5.6.3.
- Thus, fix now an arbitrary $n > j > 1$ and assume that $\mathsf{proj}_{S'_{n-j}}(P)$ is a polyhedron for $j' = j - 1$ (induction hypothesis).
- Under the above assumption, we want to prove that $\mathsf{proj}_{S_{n-j}}(P)$ is a polyhedron. Indeed, we know that from the induction hypothesis that $Q := \mathsf{proj}_{S'_{n-j}}(P)$ is a polyhedron. Using Theorem 5.6.3 on Q, we hence conclude that $\mathsf{proj}_{S_{n-j}}(P)$ is a polyhedron as well.

It remains to prove
$$\mathsf{proj}_{S_{n-j}}(P) = \mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P)).$$
for all indices $1 \leq k < j < n$. Let j, k be arbitrary such indices. Observe first that the expression $\mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P))$ is now valid because we proved that $\mathsf{proj}_{S_{n-k}}(P)$ is a polyhedron (and projections are defined on polyhedra).

Consider first an arbitrary $\mathbf{z} \in \mathsf{proj}_{S_{n-j}}(P)$. By definition of $\mathsf{proj}_{S_{n-j}}(P)$, there exist $x_{n-j+1}, \dots, x_n \in \mathbb{R}$ such that the vector
$$\begin{bmatrix} z_1 & \dots & z_{n-j} & x_{n-j+1} & \dots & x_n \end{bmatrix}^\top \in \mathbb{R}^n$$
is in P. By definition of $\mathsf{proj}_{S_{n-k}}(P)$, this directly implies that the vector
$$\begin{bmatrix} z_1 & \dots & z_{n-j} & x_{n-j+1} & \dots & x_{n-k} \end{bmatrix}^\top \in \mathbb{R}^{n-k}$$

is in $\mathsf{proj}_{S_{n-k}}(P)$. Using the definition of $\mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P))$, we conclude that $\mathbf{z} \in \mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P))$.

For the other direction, consider now an arbitrary vector $\mathbf{z} \in \mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P))$. By definition of $\mathsf{proj}_{S_{n-j}}(\mathsf{proj}_{S_{n-k}}(P))$, there exist $x_{n-j+1}, \dots, x_{n-k} \in \mathbb{R}$ such that the vector

$$\begin{bmatrix} z_1 & \dots & z_{n-j} & x_{n-j+1} & \dots & x_{n-k} \end{bmatrix}^\top \in \mathbb{R}^{n-k}$$

is in $\mathsf{proj}_{S_{n-k}}(P)$. Now using the definition of $\mathsf{proj}_{S_{n-k}}(P)$, there must exist $x_{n-k+1}, \dots, x_n \in \mathbb{R}$ such that the vector

$$\begin{bmatrix} z_1 & \dots & z_{n-j} & x_{n-j+1} & \dots & x_n \end{bmatrix}^\top \in \mathbb{R}^n$$

is in P. We conclude that $\mathbf{z} \in \mathsf{proj}_{S_{n-j}}(P)$ by the definition of $\mathsf{proj}_{S_{n-j}}(P)$.

4. In the lecture, it was already proven that, given an arbitrary polyhedron $P \subseteq \mathbb{R}^n$ for some n, we have $\mathsf{proj}_{S_{n-i}}(P) \subseteq P^{(i)}$ for all $i \in [n]$ and that $P^{(1)} \subseteq \mathsf{proj}_{S_{n-1}}(P)$. We can use this as base case and proceed by induction over i.

Fix an arbitrary $i > 1$ and assume as induction hypothesis that we have $P^{(i-1)} = \mathsf{proj}_{S_{n-i+1}}(P)$ for all polyhedra $P \subseteq \mathbb{R}^n$ where $n \geq n - i + 1$.

In the induction step, we want to prove that we also have $P^{(i)} = \mathsf{proj}_{S_{n-i}}(P)$ for all polyhedra $P \subseteq \mathbb{R}^n$ with some $n \geq i$. Thus, let P be an arbitrary such polyhedron and consider the polyhedron $P' = \mathsf{proj}_{S_{n-i+1}}(P) \subseteq \mathbb{R}^{n-i+1}$. Applying the base case for P' yields $\mathsf{proj}_{S_{n-i}}(P') = P'^{(1)}$. Further, we know from Lemma 5.6.4 that $\mathsf{proj}_{S_{n-i}}(P') = \mathsf{proj}_{S_{n-i}}(\mathsf{proj}_{S_{n-i+1}}(P)) = \mathsf{proj}_{S_{n-i}}(P)$. It remains to prove $P^{(i)} \subseteq P'^{(1)}$. Using the induction hypothesis, we indeed observe that $P'^{(1)} = (P^{(i-1)})^{(1)} = P^{(i)}$, where the equality $(P^{(i-1)})^{(1)} = P^{(i)}$ follows from Definition 5.6.5.

5. Assume that $P_1 = \{\mathbf{x} \in \mathbb{R}^2 \,:\, A_1\mathbf{x} \leq \mathbf{b}_1\}$ and $P_2 = \{\mathbf{x} \in \mathbb{R}^2 \,:\, A_2\mathbf{x} \leq \mathbf{b}_2\}$ for some $A_1, A_2 \in \mathbb{Q}^{m\times 2}$ and $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{Q}^m$ and natural number m (without loss of generality we can achieve that the two polyhedra have the same number of constraints by just repeating some constraints). Observe that the system

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$$

has no solution by our assumption $P_1 \cap P_2 = \emptyset$. Hence, Farkas lemma implies existence of a vector $\mathbf{y} \in \mathbb{R}^{2m}$ with $\mathbf{y} \geq \mathbf{0}$, $\mathbf{y}^\top \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \mathbf{0}$, and $\mathbf{y}^\top \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} < 0$. Let $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^m$ be such that $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$. Observe that $\mathbf{y}^\top \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \mathbf{0}$ can be rewritten as $\mathbf{y}_1^\top A_1 + \mathbf{y}_2^\top A_2 = \mathbf{0}$ and hence $\mathbf{y}_1^\top A_1 = -\mathbf{y}_2^\top A_2$. Similarly, we get $\mathbf{y}_1^\top \mathbf{b}_1 < -\mathbf{y}_2^\top \mathbf{b}_2$.

Now define $\mathbf{v} := \mathbf{y}_1^\top A_1 \in \mathbb{R}^2$ and $w := \mathbf{y}_1^\top \mathbf{b}_1 \in \mathbb{R}$. We claim that $P_1 \subseteq \{\mathbf{x} \in \mathbb{R}^2 \,:\, \mathbf{x} \cdot \mathbf{v} \leq w\}$ and $P_2 \subseteq \{\mathbf{x} \in \mathbb{R}^2 \,:\, \mathbf{x} \cdot \mathbf{v} > w\}$. To prove this, let first $\mathbf{x} \in P_1$ be arbitrary. Then $A_1\mathbf{x} \leq \mathbf{b}_1$. Using $\mathbf{y}_1 \geq \mathbf{0}$, we hence get $\mathbf{y}_1^\top A_1\mathbf{x} \leq \mathbf{y}_1^\top \mathbf{b}_1$ and thus $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^2 \,:\, \mathbf{x} \cdot \mathbf{v} \leq w\}$, as desired. Thus, let now $\mathbf{x} \in P_2$ be arbitrary. Then $A_2\mathbf{x} \leq \mathbf{b}_2$. By $\mathbf{y}_2 \geq \mathbf{0}$ we again get $\mathbf{y}_2^\top A_2\mathbf{x} \leq \mathbf{y}_2^\top \mathbf{b}_2$. Using our previous observations, we can rewrite this as $-\mathbf{y}_1^\top A_1\mathbf{x} \leq \mathbf{y}_2^\top \mathbf{b}_2 < -\mathbf{y}_1^\top \mathbf{b}_1$. Multiplying both sides with -1 yields $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^2 \,:\, \mathbf{x} \cdot \mathbf{v} > w\}$, as desired. This concludes the proof.