D-INFK | Linear Algebra | HS 2023

Afonso Bandeira
Bernd Gärtner

Solution for Assignment 4

1. **a)** Since H is a hyperplane, there exists $\mathbf{d} \in \mathbb{R}^n$ such that $H = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{d} = 0\}$. In order to prove that H is a subspace of $\mathbb{R}^n$, we have to prove that H is non-empty and closed under vector addition and scalar multiplication. By definition, $\mathbf{0} \in H$ and hence H is non-empty. It remains to prove that, given arbitrary $\mathbf{v}, \mathbf{w} \in H$ and $c \in \mathbb{R}$, we also have $(\mathbf{v} + \mathbf{w}) \in H$ and $c\mathbf{v} \in H$. Indeed, we observe

$$(\mathbf{v} + \mathbf{w}) \cdot \mathbf{d} = \mathbf{v} \cdot \mathbf{d} + \mathbf{w} \cdot \mathbf{d} = 0 + 0 = 0$$

and hence $(\mathbf{v} + \mathbf{w})$ is in H. Similarly, we have

$$(c\mathbf{v}) \cdot \mathbf{d} = c(\mathbf{v} \cdot \mathbf{d}) = c0 = 0$$

and therefore $(c\mathbf{v}) \in H$. We conclude that H is a subspace of $\mathbb{R}^n$.

b) Note that precise notation is very important here: $\mathbf{f} \in V$ is a function and we write $\mathbf{f}(x) \in \mathbb{R}$ for the value of $\mathbf{f}$ evaluated at point $x \in [0,1]$. In particular, $\mathbf{f}$ and $\mathbf{f}(x)$ are very different types of objects. Note that the symbol $+$ is overloaded in the following sense: for two functions $\mathbf{f}, \mathbf{g} \in V$ and $x \in [0,1]$, the $+$ in the expression $\mathbf{f}(x) + \mathbf{g}(x)$ denotes the normal addition of real numbers while the $+$ in the expression $\mathbf{f} + \mathbf{g}$ is the addition of functions defined in this exercise.

First, note that U is non-empty since every constant function is in U. Thus, consider arbitrary functions $\mathbf{f}, \mathbf{g} \in U$ and scalar $c \in \mathbb{R}$. For any $x \in [0,1]$, we have

$$(\mathbf{f} + \mathbf{g})(x) \stackrel{\text{def}}{=} \mathbf{f}(x) + \mathbf{g}(x) \stackrel{\mathbf{f} \in U}{=} \mathbf{f}(1-x) + \mathbf{g}(x) \stackrel{\mathbf{g} \in U}{=} \mathbf{f}(1-x) + \mathbf{g}(1-x) \stackrel{\text{def}}{=} (\mathbf{f} + \mathbf{g})(1-x)$$

and therefore the function $\mathbf{f} + \mathbf{g}$ is in U. Similarly, we have

$$(c\mathbf{f})(x) \stackrel{\text{def}}{=} c\mathbf{f}(x) \stackrel{\mathbf{f} \in U}{=} c\mathbf{f}(1-x) \stackrel{\text{def}}{=} (c\mathbf{f})(1-x)$$

and hence $c\mathbf{f} \in U$. We conclude that U is indeed a subspace of V.

2. Note that A is already upper triangular. Hence, we can solve the three systems $A\mathbf{x}_1 = \mathbf{e}_1$, $A\mathbf{x}_2 = \mathbf{e}_2$, and $A\mathbf{x}_3 = \mathbf{e}_3$ to find the inverse

$$A^{-1} = \begin{bmatrix} | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ | & | & | \end{bmatrix}.$$

In the first system

$$A\mathbf{x}_1 = \begin{bmatrix} a & b & c \\ 0 & 1 & d \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{e}_1$$

we find

$$\mathbf{x}_1 = \begin{bmatrix} \frac{1}{a} \\ 0 \\ 0 \end{bmatrix}$$

by backwards substitution. In the second system

$$A\mathbf{x}_2 = \begin{bmatrix} a & b & c \\ 0 & 1 & d \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2$$

we find

$$\mathbf{x}_2 = \begin{bmatrix} \frac{-b}{a} \\ 1 \\ 0 \end{bmatrix}$$

by backwards substitution. Finally, we find

$$\mathbf{x}_3 = \begin{bmatrix} \frac{bd-c}{a} \\ -d \\ 1 \end{bmatrix}$$

in the third system

$$A\mathbf{x}_3 = \begin{bmatrix} a & b & c \\ 0 & 1 & d \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{e}_3$$

by backward substitution. The inverse of A is hence given by

$$A^{-1} = \begin{bmatrix} | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & -\frac{b}{a} & \frac{bd-c}{a} \\ 0 & 1 & -d \\ 0 & 0 & 1 \end{bmatrix}$$

whenever $a \neq 0$. In the case where $a = 0$, the matrix A is not invertible as its columns are not linearly independent (the first column is $\mathbf{0}$).

3. **a)** There are many possible solutions here. For example, we could choose

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

with

$$(AB)^\top = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^\top = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$A^\top B^\top = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

b) Yes, it is possible to find examples where both A and B are symmetric. For example, consider the matrices

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

We have

$$(AB)^\top = B^\top A^\top = BA = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}$$

but also

$$A^\top B^\top = AB = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}.$$

Note that any two symmetric 2×2 matrices that do not commute (i.e. $AB \neq BA$) work as an example here.

4. **a)** We proceed with the standard elimination procedure on matrix A. In a first step, we multiply A with

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

to get

$$E_{31}E_{21}A = \begin{bmatrix} 1 & 2 & -4 \\ 0 & 0 & -3 \\ 0 & -9 & 5 \end{bmatrix}.$$

Next, we have to switch rows 2 and 3 by multiplying with

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

to get

$$P_{23}E_{31}E_{21}A = \begin{bmatrix} 1 & 2 & -4 \\ 0 & -9 & 5 \\ 0 & 0 & -3 \end{bmatrix}.$$

As we have seen in the lecture, we want to think of doing the permutations first. In particular, we could now just forget E_{21} and E_{31} and start a new elimination on the matrix $P_{23}A$ to obtain the desired decomposition. In a first step, we would multiply $P_{23}A$ with

$$E'_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E'_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

to get

$$E'_{31}E'_{21}P_{23}A = \begin{bmatrix} 1 & 2 & -4 \\ 0 & -9 & 5 \\ 0 & 0 & -3 \end{bmatrix}.$$

Notice how the -2 and 4 have changed places when comparing E_{21}, E_{31} with E'_{21}, E'_{31}. This is not a coincidence. By applying P_{23} to A directly, we switched its second and third row. Hence, the coefficients for eliminating column 1 switched places as well. At this point we are done as the second column already contains the desired 0 in row 3. We conclude that $PA = LU$ where

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 2 & -4 \\ -4 & -8 & 13 \\ 2 & -5 & -3 \end{bmatrix} \\
P = P_{23} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\
L = (E'_{31}E'_{21})^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \\
U = E'_{31}E'_{21}P_{23}A &= \begin{bmatrix} 1 & 2 & -4 \\ 0 & -9 & 5 \\ 0 & 0 & -3 \end{bmatrix}.
\end{aligned}$$

***Hint**: In general, the LU decomposition is not unique and there are valid variations of it. However, if you strictly adhere to the procedure from the lecture to obtain it, you always get the same result (since the procedure is deterministic).*

b) We first solve $L\mathbf{y} = P\mathbf{b}$ for $\mathbf{y}$ and then $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. It then follows that

$$B\mathbf{x} = P^{-1}LU\mathbf{x} = P^{-1}L\mathbf{y} = P^{-1}P\mathbf{b} = \mathbf{b}.$$

In other words, $\mathbf{x}$ must be a solution to $B\mathbf{x} = \mathbf{b}$ as well. Concretely, from

$$L\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -1 & 1 \end{bmatrix} \mathbf{y} \stackrel{!}{=} \begin{bmatrix} -1 \\ 0 \\ -7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ -7 \end{bmatrix} = P\mathbf{b}$$

we derive thus $\mathbf{y} = \begin{bmatrix}-1 & -2 & -6\end{bmatrix}^\top$ by forward substitution. Next, we derive $\mathbf{x} = \begin{bmatrix}1 & -1 & 2\end{bmatrix}^\top$ by backward substitution in

$$U\mathbf{x} = \begin{bmatrix}3 & 2 & -1\\ 0 & 2 & 0\\ 0 & 0 & -3\end{bmatrix}\mathbf{x} \overset{!}{=} \begin{bmatrix}-1\\ -2\\ -6\end{bmatrix} = \mathbf{y}.$$

The final solution is hence $\mathbf{x} = \begin{bmatrix}1 & -1 & 2\end{bmatrix}^\top$.

5. **a)** Let $\mathbf{a}_1, \dots, \mathbf{a}_n$ be the rows of A and let $\mathbf{b}_1, \dots \mathbf{b}_n$ be the columns of B. By the definition of matrix multiplication, we have

$$AB = \begin{bmatrix}— & \mathbf{a}_1 & —\\ — & \mathbf{a}_2 & —\\ \vdots & \vdots & \vdots\\ — & \mathbf{a}_n & —\end{bmatrix}\begin{bmatrix}| & | & \cdots & |\\ \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n\\ | & | & \cdots & |\end{bmatrix} = \begin{bmatrix}\mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \dots & \mathbf{a}_1 \cdot \mathbf{b}_n\\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \dots & \mathbf{a}_2 \cdot \mathbf{b}_n\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{a}_n \cdot \mathbf{b}_1 & \mathbf{a}_n \cdot \mathbf{b}_2 & \dots & \mathbf{a}_n \cdot \mathbf{b}_n\end{bmatrix}.$$

Notice that by the triangular shapes of A, the last $n-i$ entries of $\mathbf{a}_i$ are zero for all $i \in [n]$. Similarly, the first $i-1$ entries of $\mathbf{b}_i$ are zero for all $i \in [n]$. In particular, for all $i, j \in [n]$ with $i < j$, we get $\mathbf{a}_i \cdot \mathbf{b}_j = 0$. Hence, AB is indeed lower triangular.

b) Let A and B be $n \times n$ upper triangular matrices. Then $A^\top$ and $B^\top$ are lower triangular. By subtask a), this implies that $B^\top A^\top$ is lower triangular and we know $B^\top A^\top = (AB)^\top$. Hence, AB is upper triangular.

6. **a)** Using the formula for 2×2 matrices, we get

$$L^{-1} = \begin{bmatrix}1 & 0\\ -a & 1\end{bmatrix}.$$

b) We prove both directions individually. First, assume that we are given an arbitrary $n \times n$ lower triangular matrix L with no zeroes on its diagonal. We want to prove that L is invertible. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be the columns of L. We claim that these vectors are linearly independent. Indeed, checking them from right to left, we can see that for any $i \in [n]$, there is no way of obtaining $\mathbf{v}_{n-i}$ from $\mathbf{v}_{n-i+1}, \dots, \mathbf{v}_n$: by assumption, the $(n-i)$-th entry of $\mathbf{v}_{n-i}$ is non-zero, but by the triangular shape of L, the $(n-i)$-th entries in all of $\mathbf{v}_{n-i+1}, \dots, \mathbf{v}_n$ are zero. This argument works for all i and we conclude that the columns of L are linearly independent. Hence, by the inverse theorem, L is invertible.

The reverse direction is a bit more difficult. We provide an indirect proof. For this, let L be an arbitrary lower triangular $n \times n$ matrix with a zero on its diagonal. We want to prove that L is not invertible. Let $i \in [n]$ be such that $\ell_{ii} = 0$. Without loss of generality, assume that this is the last zero on the diagonal, i.e. we have $\ell_{jj} \neq 0$ for all integers $i < j \leq n$. If $i = n$, the last column of L is $\mathbf{0}$ and hence the columns of L are dependent. By the inverse theorem, this implies that L is not invertible, as desired. Thus, assume from now on $i < n$.

Now consider the $(n-i) \times (n-i)$ submatrix L' in the bottom right corner of L, i.e. the entries of L' are $\ell'_{jk} := \ell_{(i+j),(i+k)}$. Observe that L' is lower triangular and all entries on its diagonal are non-zero (by our choice of i). But then, we know from above (the other direction of this subtask) that L' is invertible. By the inverse theorem, this implies that the linear system

$$L'\mathbf{x} = \begin{bmatrix}\ell_{(i+1),i}\\ \ell_{(i+2),i}\\ \vdots\\ \ell_{n,i}\end{bmatrix}$$

has a solution. By adding rows of zeroes to the system, we then get that

$$\begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ \ell_{(i+1),(i+1)} & \ell_{(i+1),(i+2)} & \cdots & \ell_{(i+1),n} \\ \ell_{(i+2),(i+1)} & \ell_{(i+2),(i+2)} & \cdots & \ell_{(i+2),n} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n,(i+1)} & \ell_{n,(i+2)} & \cdots & \ell_{n,n} \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 = \ell_{ii} \\ \ell_{(i+1),i} \\ \ell_{(i+2),i} \\ \vdots \\ \ell_{n,i} \end{bmatrix}$$

has a solution as well. In other words, the last $n-i+1$ columns of L are linearly dependent. Hence, L is not invertible by the inverse theorem.

c) Let L be an arbitrary $n \times n$ lower triangular matrix with inverse L^{-1}. We want to prove that L^{-1} is lower triangular. Assume for a contradiction, that there exists $(L^{-1})_{ij} \neq 0$ for some $i, j \in [n]$ with $i < j$. Moreover, assume without loss of generality that $(L^{-1})_{ij}$ is the rightmost non-zero entry in row i, i.e. j is maximized given i. Let $\mathbf{u} \in \mathbb{R}^n$ be the i-th row of L^{-1} and let $\mathbf{v} \in \mathbb{R}^n$ be the j-th column of L. Note that we must have $\mathbf{u} \cdot \mathbf{v} = (L^{-1}L)_{ij} = 0$ since $i \neq j$. But observe that by the triangular shape of L, $v_k = 0$ for all integers $1 \leq k < j$. Recall that $(L^{-1})_{ij}$ is the rightmost non-zero entry in row i. Hence, we also have $u_k = 0$ for all $j < k \leq n$. We conclude that

$$0 = (L^{-1}L)_{ij} = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_j v_j + \cdots + u_n v_n = u_j v_j \neq 0$$

since we have $u_j \neq 0$ by assumption and $v_j \neq 0$ from the previous subtask. This is a contradiction. Hence, L^{-1} must be lower triangular.

d) Yes, both statements also hold for upper triangular matrices. To see this, consider an arbitrary $n \times n$ upper triangular matrix U. Then $U^\top$ is lower triangular. The diagonal entries of U are non-zero if and only if the diagonal entries of $U^\top$ are non-zero. Moreover, U is invertible if and only if $U^\top$ is invertible, since $(U^{-1})^\top = (U^\top)^{-1}$ (so if either of the inverses exists, the other exists as well). Because $U^\top$ is lower triangular, we know that $(U^\top)^{-1}$, if it exists, must be lower triangular as well. In particular, the inverse of U, if it exists, must be upper triangular by $U^{-1} = ((U^{-1})^\top)^\top = ((U^\top)^{-1})^\top$.

7. Subspaces (★☆☆)

1. Let U_1, U_2 be subspaces of $\mathbb{R}^n$. Which of the following subsets of $\mathbb{R}^n$ are also subspaces of $\mathbb{R}^n$?

√ **(a)** $U_1 \cap U_2$

Explanations: For vectors $\mathbf{u}, \mathbf{v} \in U_1 \cap U_2$ and a scalar $c \in \mathbb{R}$ we need to prove that $\mathbf{u}+\mathbf{v} \in U_1 \cap U_2$ and $c\mathbf{v} \in U_1 \cap U_2$. By $\mathbf{u}, \mathbf{v} \in U_1 \cap U_2$, we also get $\mathbf{u}, \mathbf{v} \in U_1$ and $\mathbf{u}, \mathbf{v} \in U_2$. Since U_1 and U_2 are subspaces, this implies $\mathbf{u}+\mathbf{v} \in U_1$, $c\mathbf{v} \in U_1$, $\mathbf{u}+\mathbf{v} \in U_2$, and $c\mathbf{v} \in U_2$. Hence, we also have $\mathbf{u}+\mathbf{v} \in U_1 \cap U_2$, and $c\mathbf{v} \in U_1 \cap U_2$.

(b) $U_1 \cup U_2$

Explanations: The set $U_1 \cup U_2$ is in general not a subspace of $\mathbb{R}^n$. For example, U_1 and U_2 could be distinct hyperplanes of $\mathbb{R}^n$. Then, by exercise 1, both U_1 and U_2 are subspaces of $\mathbb{R}^n$ but adding a vector $\mathbf{u}_1 \in U_1$ with a vector $\mathbf{u}_2 \in U_2$ can take us outside of $U_1 \cup U_2$.

(c) $U_1 \setminus U_2 := \{\mathbf{u} \in U_1 : \mathbf{u} \notin U_2\}$

Explanations: The set $U_1 \setminus U_2$ can never be a subspace because the $\mathbf{0}$ is missing.

(d) $\varnothing$

Explanations: By definition, a subspace has to be nonempty.

√ **(e)** $\{0\}$

Explanations: Adding any two vectors from $\{\mathbf{0}\}$ gives us $\mathbf{0}$ again. Similarly, multiplying with a scalar always gives us $\mathbf{0}$ as well.

√ **(f)** $U_1 + U_2 := \{\mathbf{u}_1 + \mathbf{u}_2 : \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}$

Explanations: The set $U_1 + U_2$ is a subspace by design. Consider arbitrary vectors $\mathbf{u}, \mathbf{v} \in U_1 + U_2$ and scalar $c \in \mathbb{R}$. By definition of $U_1 + U_2$, we can write $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ with $\mathbf{u}_1 \in U_1$ and $\mathbf{u}_2 \in U_2$. Similarly, we can write $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ with $\mathbf{v}_1 \in U_1$ and $\mathbf{v}_2 \in U_2$. Then we have

$$\mathbf{u}+\mathbf{v} = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{v}_1 + \mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{v}_1) + (\mathbf{u}_2 + \mathbf{v}_2) \in U_1 + U_2$$

and

$$c\mathbf{v} = c(\mathbf{v}_1 + \mathbf{v}_2) = (c\mathbf{v}_1) + (c\mathbf{v}_2) \in U_1 + U_2.$$