

## Solution for Assignment 4

1. Consider the three constraints  $p(-1) = 0$ ,  $p(0) = 2$  and  $p(1) = 2$  that we have on  $p$ . Each of these constraints gives us an equation involving the unknowns  $a$ ,  $b$  and  $c$ . In particular, we get the three equations

$$\begin{array}{ll} a - b + c = 0 & \text{from } p(-1) = 0 \\ c = 2 & \text{from } p(0) = 2 \\ a + b + c = 2 & \text{from } p(1) = 2 \end{array}$$

that we can also write down in matrix form

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

In order to solve this system, let us use the elimination method from the lecture. For this, let us define

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

We already highlighted the first pivot in  $A$ . After one step of elimination, we get

$$E_{21}A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } E_{21}b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

with  $E_{21} = I$  as we already had  $a_{21} = 0$ . In the second step, we obtain

$$E_{31}E_{21}A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \text{ and } E_{31}E_{21}b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

with  $E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ . Next, we need to permute rows 2 and 3 with the matrix  $P_{23} =$

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$  in order to get our next pivot. We obtain

$$P_{23}E_{31}E_{21}A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } P_{23}E_{31}E_{21}b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

In the last elimination step we again find that we do not need to do anything. In other words, we have  $E_{32} = I$  and get

$$E_{32}P_{23}E_{31}E_{21}A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } E_{32}P_{23}E_{31}E_{21}b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

We arrived at the desired upper triangular shape. It remains to use back substitution to get  $c = 2$ ,  $b = 1$  and  $a = -1$ .

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  $\mathbf{0}$ ). In other words,  $A$  is invertible for all choices of  $a, b, c, d \in \mathbb{R}$  where  $a \neq 0$ .

3. a) To solve this exercise, we could use the elimination procedure for each system and then read off the solution. But the procedure would never actually eliminate anything (only reorder rows). In particular, the solutions to the system can be read off without using elimination as every row of  $A$  has a unique non-zero element. We first consider the system

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{e}_1.$$

To get the 1 in  $\mathbf{e}_1$ , the third coordinate of  $\mathbf{x}_1$  has to be 1. To get the two zeroes of  $\mathbf{e}_1$ , the other two coordinates of  $\mathbf{x}_1$  have to be zero. Hence, we get  $\mathbf{x}_1 = \mathbf{e}_3$ . Analogously, we obtain  $\mathbf{x}_2 = \mathbf{e}_1$  and  $\mathbf{x}_3 = \mathbf{e}_2$  from the other two systems.

- b) Recall that the inverse  $A^{-1}$  of the  $m \times m$  matrix  $A$ , if it exists, is the unique matrix satisfying  $AA^{-1} = I$ . Now recall our three linear systems from above:

$$A\mathbf{x}_1 = \mathbf{e}_1$$

$$A\mathbf{x}_2 = \mathbf{e}_2$$

$$A\mathbf{x}_3 = \mathbf{e}_3$$

and consider what happens when we arrange the vectors  $(A\mathbf{x}_1), (A\mathbf{x}_2), (A\mathbf{x}_3)$  as columns in a new matrix

$$\begin{bmatrix} | & | & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & A\mathbf{x}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ | & | & | \end{bmatrix} = I.$$

Notice how this corresponds to the column view of matrix multiplication, i.e. we have

$$A \begin{bmatrix} | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & A\mathbf{x}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ | & | & | \end{bmatrix} = I.$$

From this we can conclude that the matrix

$$X = \begin{bmatrix} | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ | & | & | \end{bmatrix}$$

must be the inverse of  $A$ . In particular, we have

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

- c) We proceed as above by first solving the three systems  $D\mathbf{y}_1 = \mathbf{e}_1, D\mathbf{y}_2 = \mathbf{e}_2,$  and  $D\mathbf{y}_3 = \mathbf{e}_3$ . Again, the solutions can be read off directly as  $D$  has a unique non-zero entry in every row. In particular, we get the solutions

$$\mathbf{y}_1 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \end{bmatrix}, \mathbf{y}_2 = \begin{bmatrix} 0 \\ \frac{1}{3} \\ 0 \end{bmatrix}, \mathbf{y}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

As above, we conclude that the matrix

$$Y = \begin{bmatrix} | & | & | \\ \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ | & | & | \end{bmatrix}$$

is the inverse of  $D$  since we have

$$DY = D \begin{bmatrix} | & | & | \\ \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ D\mathbf{y}_1 & D\mathbf{y}_2 & D\mathbf{y}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ | & | & | \end{bmatrix} = I.$$

Concretely, we have  $D^{-1} = Y$ .

- d) Using the same strategy as above, one could now also determine the inverse of  $B$ . But the idea of this exercise is to observe that there is a faster way. In particular, we observe that  $B$  can be obtained from  $A$  and  $D$  as  $DA = B$ . From the lecture we know that  $B^{-1}$  can now be computed as

$$B^{-1} = (DA)^{-1} = A^{-1}D^{-1} = XY = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} & 0 \\ 0 & 0 & 2 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}.$$

4. a) Yes, the inverse of  $A^k$  exists and is given by  $(A^{-1})^k$ . We will argue by induction over  $k$ .

- **Property:** The inverse of  $A^k$  is given by  $(A^{-1})^k$ .
- **Base case:** For  $k = 1$ , the property is true because we are given that  $A^{-1}$  is the inverse of  $A$ .
- **Induction step:** Fix a natural number  $1 \leq k$  and assume that the property is true for this  $k$  (induction hypothesis). We prove that the property is true for  $k + 1$ , i.e. we prove that the inverse of  $A^{k+1}$  is  $(A^{-1})^{k+1}$ . By the induction hypothesis, we have  $A^k(A^{-1})^k = I$ . Hence, from

$$A^{k+1}(A^{-1})^{k+1} = A^k AA^{-1}(A^{-1})^k = A^k(AA^{-1})(A^{-1})^k = A^k I(A^{-1})^k = A^k(A^{-1})^k \stackrel{\text{IH}}{=} I$$

we conclude that the property is indeed true for  $k + 1$ .

b) We prove this by contradiction. Assume for a contradiction that  $A$  has an inverse  $A^{-1}$ . Since  $A$  is nilpotent, we know from Assignment 2 that  $AA^{-1} = I$  is nilpotent too, i.e. there exists some  $k \in \mathbb{N}^+$  such that  $I^k = (AA^{-1})^k = 0$ . But we have  $I^k = I \neq 0$  which is a contradiction.

c) We prove this with the following calculation:

$$A = AI = A(A^3) = A^4 = I.$$

d) Observe that it suffices to find a  $2 \times 2$  matrix  $A \neq I$  with  $A^2 = I$ : indeed, for even  $k = 2\ell$  we then get  $A^k = A^{2\ell} = (A^2)^\ell = I^\ell = I$  as well by using  $A^2 = I$ . Such a matrix is sometimes called self-inverse, since we have  $A^{-1} = A$ . We have seen some self-inverse matrices before in this course. In particular, the  $2 \times 2$  rotation matrix for angle  $\phi = \pi$  is self-inverse. Concretely, this is the matrix

$$A = \begin{bmatrix} \cos(\pi) & -\sin(\pi) \\ \sin(\pi) & \cos(\pi) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

and we can check that indeed  $A^2 = I$  but  $A \neq I$ . Moreover, for odd  $k$  we then have

$$A^k = AA^{k-1} = AI = A \neq I$$

by using that  $k - 1$  must be even.

e) Again, we use rotation matrices. In particular, the rotation matrix with angle  $\pi/2$  should satisfy this. The intuition is that whenever we apply this rotation matrix four times, we rotate one whole turn and hence nothing happens. For a complete solution, we check that this indeed works. For the matrix

$$A = \begin{bmatrix} \cos(\pi/2) & -\sin(\pi/2) \\ \sin(\pi/2) & \cos(\pi/2) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

we get

$$A^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$A^3 = A^2 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$A^4 = A^3 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now consider an arbitrary  $k > 4$ . Then  $k$  can be split as  $k = 4\ell + h$  where  $\ell, h \in \mathbb{N}_0$  and  $0 \leq h < 4$  and we get  $A^k = (A^4)^\ell A^h = A^h$ . We conclude that  $A^k = I$  if and only if  $h = 0$  which is the same as saying that  $A^k = I$  if and only if  $k \equiv_4 0$ .

5. a) Let  $\mathbf{x} \in \mathbb{R}^n$  be arbitrary and assume  $B\mathbf{x} = \mathbf{0}$ . Recall that  $B\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$  where  $x_1, x_2, \dots, x_n$  are the entries of  $\mathbf{x}$  and  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$  are the columns of  $B$ . Hence, in order to prove that the columns of  $B$  are linearly independent, we have to prove  $\mathbf{x} = \mathbf{0}$ .

We achieve this with the calculation

$$\mathbf{x} = I\mathbf{x} = AB\mathbf{x} = A\mathbf{0} = \mathbf{0}.$$

- b) Let  $\mathbf{y} \in \mathbb{R}^n$  be arbitrary and assume  $A\mathbf{y} = \mathbf{0}$ . We want to prove that  $\mathbf{y} = \mathbf{0}$ . By the inverse theorem and using subtask a), we know that there exists  $\mathbf{x} \in \mathbb{R}^n$  with  $B\mathbf{x} = \mathbf{y}$ . Now observe that indeed, we have

$$\mathbf{y} = B\mathbf{x} = BI\mathbf{x} = B(AB\mathbf{x}) = B(A\mathbf{y}) = \mathbf{0}$$

and hence we conclude that the columns of  $A$  are linearly independent.

- c) By applying the matrix  $A$  to  $BA - I$  we get

$$A(BA - I) = ABA - A = IA - A = \mathbf{0}.$$

In particular, we have  $A\mathbf{v} = \mathbf{0}$  for every column  $\mathbf{v}$  of  $BA - I$ . But by the inverse theorem and subtask b), the equation  $A\mathbf{y} = \mathbf{0}$  has a unique solution  $\mathbf{y}$ . And we also know that  $\mathbf{0}$  is a solution since  $A\mathbf{0} = \mathbf{0}$ . By uniqueness, we conclude that we have  $\mathbf{v} = \mathbf{0}$  for every column  $\mathbf{v}$  of  $BA - I$  and hence  $BA - I = \mathbf{0}$ .

6. a) Using the formula for  $2 \times 2$  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  $m \times m$  lower triangular matrix  $L$  with no zeroes on its diagonal. We want to prove that  $L$  is invertible. Let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  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  $i \in [m]$ , there is no way of obtaining  $\mathbf{v}_{n-i}$  from  $\mathbf{v}_{n-i+1}, \dots, \mathbf{v}_m$ : 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  $\mathbf{v}_{n-i+1}, \dots, \mathbf{v}_m$  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  $m \times m$  matrix with a zero on its diagonal. We want to prove that  $L$  is not invertible. Let  $i \in [m]$  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  $i < j \leq m$ . If  $i = m$ , 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  $i < m$ .

Now consider the  $(m-i) \times (m-i)$  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

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

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

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

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

- c) Let  $L$  be an arbitrary  $m \times m$  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  $i, j \in [m]$  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  $\mathbf{u} \in \mathbb{R}^m$  be the  $i$ -th row of  $L^{-1}$  and let  $\mathbf{v} \in \mathbb{R}^m$  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

$$0 = (L^{-1}L)_{ij} = \mathbf{u} \cdot \mathbf{v} = u_1v_1 + \cdots + u_jv_j + \cdots + u_mv_m = u_jv_j \neq 0$$

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  $m \times m$  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  $U^{-1} = ((U^{-1})^\top)^\top = ((U^\top)^{-1})^\top$ .