233 Computational Techniques

Problem Sheet for Tutorial 5

Problem 1
Solve the following system of equations using Gauss-Jordan elimination. Identify basic variables. Express all solutions in terms of non-basic variables. Determine the space of solutions and verify the result.

$$\begin{array}{rcrcrcrcrcr} 2x_1 & + & x_2 & - & x_3 & + & 2x_4 & - & x_5 & = & -2 \\ 4x_1 & + & 2x_2 & & & + & 3x_4 & - & 2x_5 & = & 2 \\ x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 3 \end{array}$$

Problem 2
(a) Find the *Cholesky factorization* of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & -5 \\ -1 & -5 & 6 \end{bmatrix}.$$

(b) Then solve $\boldsymbol{Ax} = \boldsymbol{b}$ for $\boldsymbol{b} = [1, -3, 6]^T$ by forward and backward substitution, using the triangular shape of the factorization matrices. (See the end of section 3.5.1 in the lecture notes.)

Problem 3
Let $\boldsymbol{a}_1 = [1, 2, 2]^T$, $\boldsymbol{a}_2 = [1, 0, 1]^T$, and

$$P = \{x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 \ : x_1, x_2 \in \mathbb{R}\}$$

be the plane in $\mathbb{R}^3$ spanned by the vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.
(a) Find the matrix $\boldsymbol{A} \in \mathbb{R}^{3\times 2}$ of which P is the range space. Then determine the vector $\boldsymbol{p}$ in P with the shortest ℓ_2 distance from the vector $\boldsymbol{b} = [3, -1, 1]^T$, using the normal equations with Cholesky factorization.
(b) What are the range and nullspace components $\boldsymbol{b}_R$ and $\boldsymbol{b}_N$ of the vector $\boldsymbol{b}$? Check that they are orthogonal.
(c) What is the shortest distance of $\boldsymbol{b}$ from P?
(d) Alternative geometric approach: Determine $\boldsymbol{b}_N$ and $\boldsymbol{b}_R$ using orthogonal projection onto the nullspace of the relevant matrix $\boldsymbol{A}^T$. *Hint:* The cross product is a good tool to find this nullspace.

Solution

Problem 1
In tableau notation, a possible sequence of steps is the following (with the third equation as the first row; pivot elements underlined):

$$\left[\begin{array}{ccccc|c} \underline{1} & 1 & 1 & 1 & 1 & 3 \\ 2 & 1 & -1 & 2 & -1 & -2 \\ 4 & 2 & 0 & 3 & -2 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & \underline{-1} & -3 & 0 & -3 & -8 \\ 0 & -2 & -4 & -1 & -6 & -10 \end{array}\right] \longrightarrow$$
$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & 1 & -2 & 5 \\ 0 & 1 & 3 & 0 & 3 & 8 \\ 0 & 0 & \underline{2} & -1 & 0 & 6 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 3/2 & 3 & -1 \\ 0 & 0 & 1 & -1/2 & 0 & 3 \end{array}\right] .$$

The first three columns of the final tableau contain the three unit vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$. (A different choice of pivot elements would have produced them in other columns.) So x_1, x_2 and x_3 are the basic variables, and we can read off

$$x_1 = 1 + 2x_5 , \qquad x_2 = -1 - \frac{3}{2}x_4 - 3x_5 , \qquad x_3 = 3 + \frac{1}{2}x_4 ,$$

where the non-basic variables x_4 and x_5 can take arbitrary values. Hence the set of solutions is

$$\left\{\boldsymbol{x} = [1, -1, 3, 0, 0]^T + x_4[0, -3/2, 1/2, 1, 0]^T + x_5[2, -3, 0, 0, 1]^T \ : x_4, x_5 \in \mathbb{R}\right\} .$$

(Note that this is *not* a vector space! For instance, the zero vector $\boldsymbol{x} = \boldsymbol{0}$ is not a solution. In general, for a given matrix $\boldsymbol{A}$ and a given vector $\boldsymbol{b}$ (with dimension equal to the row dimension of $\boldsymbol{A}$) the set $\{\boldsymbol{x} : \boldsymbol{Ax} = \boldsymbol{b}\}$ is a vector space if and only if $\boldsymbol{b} = \boldsymbol{0}$.)

Problem 2
(a) The aim is to split $\boldsymbol{A}$ in the following way:

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & -5 \\ -1 & -5 & 6 \end{bmatrix} = \boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^T = \begin{bmatrix} \ell_{11} & & \\ \ell_{21} & \ell_{22} & \\ \ell_{31} & \ell_{32} & \ell_{33} \end{bmatrix} \begin{bmatrix} \ell_{11} & \ell_{21} & \ell_{31} \\ & \ell_{22} & \ell_{32} \\ & & \ell_{33} \end{bmatrix} .$$

With the usual sign convention $\ell_{ii} > 0$ for the diagonal elements of $\boldsymbol{L}$, one obtains
• from the first column of $\boldsymbol{A}$:

$$1 = \ell_{11}^2 \Rightarrow \ell_{11} = 1 , \qquad 1 = \ell_{21}\ell_{11} = \ell_{21} , \qquad -1 = \ell_{31}\ell_{11} = \ell_{31} ,$$

• from the second column of $\boldsymbol{A}$, starting with a_{22} (as $\boldsymbol{A}$ is symmetric, the equation for a_{21} is the same as the one for a_{12}, which we have just solved)

$$5 = \ell_{21}^2 + \ell_{22}^2 = 1 + \ell_{22}^2 \Rightarrow \ell_{22} = 2 , \qquad -5 = \ell_{31}\ell_{21} + \ell_{32}\ell_{22} = -1 + 2\ell_{32} \Rightarrow \ell_{32} = -2 ,$$

- and from the third column (only the last element can give anything new),

$$6 = \ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2 = 1 + 4 + \ell_{33}^2 \Rightarrow \ell_{33} = 1 \ .$$

So
- the factorization was successful (which implies that $\boldsymbol{A}$ is positive definite), and
- the Cholesky factor of $\boldsymbol{A}$ is

$$\boldsymbol{L} = \begin{bmatrix} 1 & & \\ 1 & 2 & \\ -1 & -2 & 1 \end{bmatrix} . \tag{1}$$

(b) The idea is to solve $\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ in two steps by defining $\boldsymbol{y} := \boldsymbol{L}^T\boldsymbol{x}$ and solving (1) $\boldsymbol{L}\boldsymbol{y} = \boldsymbol{b}$ by forward substitution and (2) $\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{y}$ by backward substitution. So:

$$\begin{bmatrix} 1 \\ -3 \\ 6 \end{bmatrix} = \boldsymbol{b} = \boldsymbol{L}\boldsymbol{y} = \begin{bmatrix} 1 & & \\ 1 & 2 & \\ -1 & -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_1 + 2y_2 \\ -y_1 - 2y_2 + y_3 \end{bmatrix} \Rightarrow \boldsymbol{y} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \boldsymbol{y} = \boldsymbol{L}^T\boldsymbol{x} = \begin{bmatrix} 1 & 1 & -1 \\ & 2 & -2 \\ & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_2 - 2x_3 \\ x_3 \end{bmatrix} \Rightarrow \boldsymbol{x} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} .$$

Problem 3
(a) The matrix is

$$\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2] = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 2 & 1 \end{bmatrix} .$$

- We seek $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$ minimizing $\|\boldsymbol{b} - \boldsymbol{p}\|_2^2 = \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2$, where $\boldsymbol{x} = [x_1, x_2]^T$.
- Any such $\boldsymbol{x}$ will be a solution of the normal equation, $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$. So we compute $\boldsymbol{A}^T\boldsymbol{A}$ and apply Cholesky factorization:

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} 9 & 3 \\ 3 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 3 & \\ 1 & 1 \end{bmatrix}}_{\boldsymbol{L}} \underbrace{\begin{bmatrix} 3 & 1 \\ & 1 \end{bmatrix}}_{\boldsymbol{L}^T} .$$

(As the factorization was successful, we know that the solution $\boldsymbol{x}$ is unique.)
- Using the factorization of $\boldsymbol{A}$, we can write this equation as $\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ and hence solve it in two steps, (1) $\boldsymbol{L}\boldsymbol{y} = \boldsymbol{A}^T\boldsymbol{b}$, and (2) $\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{y}$:

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \boldsymbol{A}^T\boldsymbol{b} = \boldsymbol{L}\boldsymbol{y} = \begin{bmatrix} 3 & \\ 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3y_1 \\ y_1 + y_2 \end{bmatrix} \Longrightarrow \boldsymbol{y} = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \boldsymbol{y} = \boldsymbol{L}^T\boldsymbol{x} = \begin{bmatrix} 3 & 1 \\ & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ x_2 \end{bmatrix} \Longrightarrow \boldsymbol{x} = \begin{bmatrix} -2/3 \\ 3 \end{bmatrix} .$$

(Alternatively, $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ can also be solved by Gauss elimination.)

• For the vector $\boldsymbol{p}$ we obtain

$$\boldsymbol{p} = -(2/3) \cdot [1, 2, 2]^T + 3 \cdot [1, 0, 1]^T = [7/3, -4/3, 5/3]^T \ .$$

(b)

$\boldsymbol{b}_R = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{p}$, since this is how the minimizing solution $\boldsymbol{x}$ of the least squares problem is constructed. The component of $\boldsymbol{b}$ in the nullspace of $\boldsymbol{A}^T$ is $\boldsymbol{b}_N = \boldsymbol{b} - \boldsymbol{b}_R = [2/3, 1/3, -2/3]^T$. The dot product of $\boldsymbol{b}_R$ and $\boldsymbol{b}_N$ is $\boldsymbol{b}_R^T\boldsymbol{b}_N = 14/9 - 4/9 - 10/9 = 0$, so they are indeed orthogonal.

(c) The minimal distance is $\|\boldsymbol{b}_N\|_2 = \sqrt{4/9 + 1/9 + 4/9} = 1$.

(d) By definition, $\boldsymbol{b}_R \in \text{range}(\boldsymbol{A}) = P$ and $\boldsymbol{b}_N \in \text{null}(\boldsymbol{A}^T)$. As $\text{range}(\boldsymbol{A}) \perp \text{null}(\boldsymbol{A}^T)$ and $\text{null}(\boldsymbol{A}^T)$ must be one-dimensional, it must be spanned by the cross-product of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$,

$$\text{null}(\boldsymbol{A}^T) = \mathbb{R}\, \boldsymbol{a}_1 \times \boldsymbol{a}_2 = \mathbb{R} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \mathbb{R} \underbrace{\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}}_{\boldsymbol{s}} ,$$

and $\boldsymbol{b}_N = t\boldsymbol{s}$ for some $t \in \mathbb{R}$. From $\boldsymbol{b} = \boldsymbol{b}_R + \boldsymbol{b}_N = \boldsymbol{b}_R + t\boldsymbol{s}$, we find the value of t by taking the dot product with $\boldsymbol{s}$, since as an element of $\text{null}(\boldsymbol{A}^T)$, $\boldsymbol{s}$ is orthogonal to $\boldsymbol{b}_R$:

$$\boldsymbol{s}^T\boldsymbol{b} = t\|\boldsymbol{s}\|_2^2 \quad \Longrightarrow \quad t = \frac{\boldsymbol{s}^T\boldsymbol{b}}{\|\boldsymbol{s}\|_2^2} = \frac{1}{3} \ .$$

So $\boldsymbol{b}_N = [2/3, 1/3, -2/3]$, and $\boldsymbol{b}_R = \boldsymbol{b} - \boldsymbol{b}_N = [7/3, -4/3, 5/3]$, as above.