D-INFK | Linear Algebra | HS 2024
Bernd Gärtner
Robert Weismantel

Assignment 9

Submission Deadline: **26 November, 2024** at 23:59
Course Website: https://ti.inf.ethz.ch/ew/courses/LA24/index.html

Exercises

You can get feedback from your TA and bonus points for Exercise 2 by handing in your solution as pdf via Moodle before the deadline.

1. Gram-Schmidt (in-class) (★☆☆)

This task includes Challenge 13 from the lecture notes.

Consider the invertible matrices

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

and

$$B = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 4 & 5 & 6 \\ 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 9 \end{bmatrix}.$$

a) Apply the Gram-Schmidt process to the columns of A.

b) Write down a QR-decomposition of A.

c) Apply the Gram-Schmidt process to the columns of B.

d) Is it always true that the Gram-Schmidt process on the columns of an upper triangular $n \times n$ matrix with non-zero diagonal entries yields the canonical basis $\mathbf{e}_1, \dots, \mathbf{e}_n$? Provide a proof or counterexample.

2. Preserving inner products (bonus, hand-in) (★☆☆)

Let $Q \in \mathbb{R}^{m \times m}$ be an arbitrary matrix satisfying $(Q\mathbf{v})^\top (Q\mathbf{w}) = \mathbf{v}^\top \mathbf{w}$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$. Prove that Q is orthogonal.

3. Orthogonal 2×2 matrices (★★☆)

Recall 2×2 rotation matrices from Example 5.4.4. It is shown in this example that 2×2 rotation matrices are orthogonal.

a) Find an orthogonal 2×2 matrix that is not a rotation matrix.

b) Consider an arbitrary 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Prove that if A is orthogonal, then we have $|ad - bc| = 1$.

c) Prove that the converse is not true, i.e. find values for a, b, c, d such that A is not orthogonal but we still have $|ad - bc| = 1$.

4. Fitting a parabola (★☆☆)

This task includes Challenge 8 from the lecture notes of the second part of the course.

Assume we are given $m \geq 3$ distinct datapoints $(t_1, b_1), \ldots, (t_m, b_m)$ where $t_k, b_k \in \mathbb{R}$ for all $k \in [m]$ (distinct means that we have $t_i \neq t_j$ for all $i \neq j$ with $i, j \in [m]$). Using the least squares method, we want to find a parabola described by three parameters $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}$ such that we have

$$b_k \approx \alpha_0 + \alpha_1 t_k + \alpha_2 t_k^2$$

for all $k \in [m]$. More concretely, we want to solve the optimization problem

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^3} ||A\boldsymbol{\alpha} - \mathbf{b}||^2 = \min_{\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}} \sum_{k=1}^{m} (b_k - (\alpha_0 + \alpha_1 t_k + \alpha_2 t_k^2))^2$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad A = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & t_m^2 \end{bmatrix}.$$

a) Compute the matrix $A^\top A$.

b) Prove that for $A^\top A$ to be diagonal, we must have $t_k = 0$ for all $k \in [m]$. Note that this is an uninteresting case which is actually excluded by the assumption $m \geq 3$ and the assumption that our datapoints are distinct.

5. Fitting a circle (★★☆)

a) Consider the following points

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} -\frac{2}{3} \\ \frac{4}{3} \end{bmatrix}, \mathbf{p}_4 = \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{2} \end{bmatrix} \in \mathbb{R}^2$$

in the plane. We want to find a circle C_r with origin $\mathbf{0}$ and radius $r \in \mathbb{R}^+$ such that the sum of the quadratic distances of the points to the circle is minimized. Note that the quadratic distance of a point $\mathbf{p} \in \mathbb{R}^2$ to the circle C_r is $(r - ||\mathbf{p}||)^2$. Find the optimal value of r for the four points above.

Note that the interesting thing here is to find a formula for r in terms of $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$. The actual numerical answer is of secondary interest, i.e. you are not expected to simplify the value you get for r as much as possible.

b) Generalize the result from a) to a formula for r in terms of n datapoints $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^2$.

6. Permutation matrices (★★☆)

This exercise includes Challenge 12 from the lecture notes.

Let $P \in \mathbb{R}^{n \times n}$ be a permutation matrix for some $n \geq 1$. In particular, P has the form

$$P = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{e}_{p(1)} & \mathbf{e}_{p(2)} & \cdots & \mathbf{e}_{p(n)} \\ | & | & \cdots & | \end{bmatrix}$$

where $p : [n] \to [n]$ is a bijective function (the permutations of $[n]$ are exactly the bijective funtions $p : [n] \to [n]$). Prove that there exists $k \in \mathbb{N}$ with $P^k = I$.