May2023: Information Page

How?
The main option is to write your answers in the boxes in Inspera. As a backup option, if you need to, you can submit hand-written solutions on paper "the normal way". In part A you are hopefully able to answer in Inspera only. If you submit paper-answers, make a note about it in Inspera in the corresponding question. For part B, it might be helpful to write some solutions on paper (if you do, clearly identify the sheets with your number).

Grades
The exam is divided into two parts, part A and part B. Part A is related to grade 3 and part B to grade 4 and 5.

- On Part A there are 7 questions: 2 for concepts, 2 for algorithms, and 3 for analysis, each worth roughly 1 point per question.
- In part B there are 3 problems, and the points vary between 4 and 8 points per problem.

The grades are (subject to modifications):

- Grade 3: You must get at least 3,5 out of the 7 possible points. Further, you must get at least 1 point in each of the 3 categories.
- Grade 4: You must fulfill all requirements for grade 3 + reach at least 12 points for partA and partB combined.
- Grade 5: You must fulfill all requirements for grade 3 + reach at least 18 points for partA and partB combined.
  (In total, partA and partB give 25 points.)

Tools available

The tools available in Inspera are

- **Online Python:** The idea is that you are able to use Python as a pocket calculator, very much in the same way as on the problem solving sessions. It will be enough to be able to use numpy.
- **Numpy Cheat Sheet** is available as a link
- **Numpy reference manual** is available as a link
- **Formula sheet**

You should be able to find these resources at the bottom part of Inspera.

Tools to take with you
Pocket calculator is allowed, so you can bring it to the exam if you want to. It's not necessary though, you can use Python instead.

Good Luck!

1 May2023: PartA: Concept-models

Classify the following items:

Please match the values:

	Deterministic model	Stochastic model	Deterministic method	Stochastic method
R,F integer-valued $R \xrightarrow{\alpha} 2R$ $R + F \xrightarrow{\beta} 2F$ $F \xrightarrow{\gamma} \emptyset$	○	○	○	○
Monte Carlo simulation	○	○	○	○
SSA algorithm	○	○	○	○
$\frac{dF}{dt} = \beta FR - \gamma F$ $\frac{dR}{dt} = \alpha R - \beta FR$	○	○	○	○
QR iteration	○	○	○	○

Maximum marks: 1

2 May2023: PartA: Concept-gaps

Make the following statements as correct as possible. The same term may (but does not have to) be used multiple times.

Help

1. The [ ] computes all eigenvalues of a matrix A .
2. Brownian motion $B(t)$ is normally distributed with [ ] t.
3. If an $n \times n$ matrix is [ ] , then it is non-singular.
4. The number of singular values of a matrix depends on its [ ]
5. Using Gram-Schmidt, one can compute the [ ] of a matrix.

orthogonal	symmetric	singular value decomposition
QR decomposition	variance	SSA algorithm
rank	inverse power method	condition number
power method	standard deviation	dimensions
QR iteration	diagonal	

Maximum marks: 1

3 May2023: PartA: Algo-MC

Felicity buys several things to eat each day. In the morning she buys a croissant at the local supermarket for 10 sek. For lunch she goes to a buffet place where the price depends on the weight. We assume that the price she pays each day follows the uniform distribution $U(90, 110)$ (in sek). In the afternoon she gets some kind of cake. Her choice varies. We assume that the price can be modeled by a given probability density function (pdf) $f(x)$. Assume there exists a function sample_f(), which returns one random number (in sek) that is distributed according to the pdf $f(x)$ every time that it is called.

Using Monte-Carlo simulation, estimate how much money Felicity spends on food outside her home on average each day.

Write pseudo code for this problem here.

Fill in your answer here

Maximum marks: 1

4 May2023: PartA: Algo-LS

You are given the data set

x	-5	-1	0	2	4
y	1	-5	6	-4	2
z	-5	2	12.6	-1	0.5

Use the ansatz $p(x, y) = c_0 + c_1 x + c_2 y$ to fit the data in z, formulate the normal equations, and solve them. Write down the polynomial p that you get.

Note: Your answer should include the matrix A, and the matrix involved in the normal equations, as well as the right hand side. One easy way to write a matrix in the text field is using new lines and spaces such as

3x3 identity matrix =

1 0 0

0 1 0

0 0 1

Alternatively, it is fine to do some computations in python (eg matrix multiplications). You can also copy and paste the python syntax for formulating matrices here.

Fill in your answer here

Maximum marks: 1

5 May2023: PartA: Ana-MC

We use Monte-Carlo simulation to estimate $G = \int_0^1 g(x)\,dx$ with $g(x) = \frac{6}{1+x^2}$. We sample N samples $X_i, i = 1, \ldots, N$, from the uniform distribution $U(0,1)$ and estimate the G as $\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} g(X_i)$.
We also compute
$\frac{1}{99}\sum_{i=1}^{100}\left(g(X_i) - \hat{\mu}_{100}\right)^2 = 1.21$ and $\hat{\mu}_{100} = 4.75$
Use that information to determine the interval (with 2 decimals), within which the true integral value G lies with 95% probability using N=100. Which theorem is behind the approach you are using?

Fill in your answer here

Maximum marks: 1

6 May2023: PartA: Ana-SVD

The SVD of a matrix A is given by

$$U = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 & 0 & -\frac{1}{\sqrt{3}} \\ 0 & 0 & 1 & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \Sigma = \begin{pmatrix} 2\sqrt{3} & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} V = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Based on this information, answer the following questions:

a) What is rank(A)?

- 2
- 3
- 4
- 5
- 1
- 0
- Not possible to decide based on the given information

b) What is $cond_2(A)$? (The condition number measured in the 2-norm)

- ∞
- $4\sqrt{3}$
- 2
- $\sqrt{3}$
- Not possible to decide based on the given information
- $2\sqrt{3}$

c) The rank-1-matrix that best approximates $\boldsymbol{A}$ in the 2-norm is given by

- $\begin{pmatrix} 2 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 \\ -\sqrt{2} & 0 & 0 & \sqrt{2} & 0 \end{pmatrix}$
- $\begin{pmatrix} 2 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\sqrt{2} & \sqrt{2} & 0 & 0 & 0 \end{pmatrix}$
- $\begin{pmatrix} 2 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$
- $\begin{pmatrix} 2 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix}$

Maximum marks: 1.02

7 May2023: PartA: Ana-LS

You are given the following measurement data

x	500	1000	1500	2000	2500
y	2	8	10	13	20

Your goal is to find a quadratic polynomial that fits the given data best in a least-squares sense (representing y as a function of x). We want to set up the resulting overdetermined system $Ac = b$ in such a way that the condition number of A becomes as small as we can make it. Describe how you achieve that and write down the resulting matrix A with all its entries.

Note: For any computations you might do, use 1 decimal (1 number after the dot). You can use python as a pocket calculator.

Note: an easy way to write a matrix, e.g. the 3x3 identity matrix is using new lines and spaces
1 0 0
0 1 0
0 0 1
You can also use python syntax.

Fill in your answer here

Maximum marks: 1

8 May2023: PartB: Higher grades-methods

(4 points) Tom has learned about different methods for solving least-squares problems. He wants to test them on a specific matrix $A \in \mathbb{R}^{100\times 15}$ together with a vector $b \in \mathbb{R}^{100}$ and wants to solve the overdetermined system $Ax = b$. He has designed his problem such that for the exact solution $x_{15} = 1.000000000000000$. He tests 4 different methods and produces the following results:

Method 1	$x_{15} = 1.000000131408709$
Method 2	$x_{15} = -0.422476064944815$
Method 3	$x_{15} = 0.981684027615913$
Method 4	$x_{15} = 1.000000131409875$

Additionally, Tom knows that $cond_2(A) = 2.3 * 10^{10}$. Tom does not know how to properly compute running times and therefore cannot test it himself. Based on what he has read, he expects that for large problems

runtime(Method 2) < runtime(Method 1), runtime(Method 3) < runtime(Method 4).

The methods that Tom has implemented are (in random order)

- (A) approach using QR decomposition based on Gram-Schmidt
- (B) approach using QR decomposition based on Householder
- (C) approach using A^+ based on a SVD
- (D) setting up and solving normal equations.

Identify which method corresponds to which approach, i.e., map methods 1,2,3,4 to the letters (A),(B),(C),(D) and **explain** why you think that this is the correct mapping.

Fill in your answer here

Maximum marks: 4

9 May2023: PartB: Higher grades-matrix

A) (4 points) We store our data in a matrix $A \in \mathbb{R}^{m \times n}$ with $m = 10^3$ and $n = 250$. All entries use double precision. We know that the rank of the matrix is $r = 10$. We want to exploit this to save memory, i.e., instead of storing the original matrix A we want to find a way to store exactly the same information but with less memory demand. (1) Describe how you would do that and (2) compute the number of bytes we need to store this new information (remember that we want to store as little information as possible).
Note: To store one double precision number we need 8 bytes.

Fill in your answer here

B) (4 points) For which value(s) of $b \in \mathbb{R}, -3 < b < 3,$ is for the following matrix

$$\begin{pmatrix} 1 & b \\ b & 9 \end{pmatrix}$$

the singular value $\sigma_2 \geq 1$? **Justify** your response. No python allowed for this problem.

(If you write your solution on a piece of paper, make a note here about that.)

Fill in your answer here

Maximum marks: 8

10 May2023: PartB: Higher grades-Pareto

(6 points) The Pareto distribution is widely used to describe observations in social and economical sciences, quality control and many others. Its CDF (cumulative distribution function) $F(x)$ is defined by a parameter $\alpha > 0$ and is given by

$$F(x) = \begin{cases} 1 - x^{-\alpha} & \text{for } x \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose someone tried to model this distribution by, first, sampling numbers u from the Uniform distribution $U(0,1)$ and, then, applying the Inverse Transform method to obtain Pareto samples x. The results were approximated to two decimals and look as follows:

u	x
0.13	1.10
0.82	3.18
0.36	1.36
0.67	2.11
0.25	1.21

Find the parameter value α that best fits these samples, with an accuracy of two decimals (2 numbers after the dot). You can use python for the computations that you need to do. Explain all your steps and provide the code if you use any. (If you write on paper, mention it here.)

Fill in your answer here

Maximum marks: 6