Prof. Dr. Jens Honore Walther
Dr. Georgios Arampatzis
ETH Zentrum, CLT
CH-8092 Zürich

Exam

Issued: Monday, 03.08.2020

Exam directives. In order to pass the exam, the following requirements have to be met:

- Read carefully the first two pages of the exam. Write your name and Legi-ID where requested. Before handing in the exam, **PUT YOUR SIGNATURE ON PAGE 2**.
- Clear your desk (no cell phones, cameras, etc.): on your desk you should have your Legi, your pen, paper and your notes.
- Calculators are not permitted.
- If necessary the teaching assistants will give you additional paper sheets. On the top-right corner of every page write your complete name and Legi-ID.
- The personal summary consists of no more than 4 pages (2 sheets). The personal summary can be handwritten or machine-typed. In case it is machine-typed, the text has to be single-spaced and the font size has to be at least 8 pts.
- You can answer in English or in German; the answers should be handwritten and clearly readable, written in blue or black - do NOT write anything in red or green. Only one answer per question is accepted. Invalid answers should be clearly crossed out.
- If something is disturbing you during the exam, or it is preventing you from peacefully solving the exam, please report it immediately to an assistant. Later notifications will not be accepted.
- You must hand in: the exam cover, the sheets with the exam questions and your solutions. The exam cannot be accepted if the cover sheet or the question sheets are not handed back.
- **You need 100 points out of a 120 for a 6.**

Family Name:	
Name:	
Legi-ID:	

Question	Maximum score	Score	TA 1	TA 2
1				
2				
3				
1-3	30			
4	15			
5	25			
6	15			
7	15			
8	20			
Total	120			

With your signature you confirm that you have read the exam directives; you solved the exam without any unauthorized help and you wrote your answers following the outlined directives.

Signature: ______________________

GOOD LUCK!

Multiple Choice Questions [30 Points]

There is a total of 20 correct answers over all multiple choice questions. Each correct answer is awarded 1.5 point. Any missing answer is awarded 0 point. **For each wrong answer,** 1.5 **points are deducted**. The minimal number of points for your answers to all multiple choice questions is 0.

Question 1: Nonlinear solvers

a) Which statement(s) about the convergence of Newton's method is/are correct? (Assume a good initial condition has been chosen)

- ☐ For $f(x) = x^2 - 3x + 2$ the method converges quadratically.
- ☐ For $f(x) = x^{1/3}$ the method converges linearly.
- ☐ For $f(x) = x^2 - 4x + 4$ the method converges quadratically.
- ☐ For $f(x) = x$ the method converges in one step.

b) Which of the following statement(s) about nonlinear solvers is/are correct?

- ☐ The Bisection method is guaranteed to converge.
- ☐ The order of convergence of nonlinear solvers cannot be estimated numerically.
- ☐ The Bisection method converges faster the Newton's method.
- ☐ The Secant method can only be used when the closed form of $f'(x)$ is known.

c) You goal is to find the minimum of the function $g(x, y) = x^2 + 2xy + y^3$. You choose to use the Newton-Raphson method: $\mathbf{x_{k+1}} = \mathbf{x_k} - J^{-1}(\mathbf{x_k})F(\mathbf{x_k})$. Select the correct expressions for $F(\mathbf{x})$ and $J(\mathbf{x})$: :

- ☐ $F(\mathbf{x}) = \begin{pmatrix} x^2 + 2xy + y^3 \end{pmatrix}$, $J(\mathbf{x}) = \begin{pmatrix} 2x + 2y \\ 2x + 3y^2 \end{pmatrix}$
- ☐ $F(\mathbf{x}) = \begin{pmatrix} 2x + 2y \\ 2x + 3y^2 \end{pmatrix}$, $J(\mathbf{x}) = \begin{pmatrix} 2 & 6y \\ 2 & 2 \end{pmatrix}$
- ☐ $F(\mathbf{x}) = \begin{pmatrix} 2x + 3y^2 \\ 2x + 2y \end{pmatrix}$, $J(\mathbf{x}) = \begin{pmatrix} 2 & 6y \\ 2 & 2 \end{pmatrix}$
- ☐ $F(\mathbf{x}) = \begin{pmatrix} 2x + 2y \\ 2x + 3y^2 \end{pmatrix}$, $J(\mathbf{x}) = \begin{pmatrix} 2 & 2 \\ 2 & 6y \end{pmatrix}$

d) The update rule for finding the root of the equation $x^2 - R = 0$ using Newton's method is:

- ☐ $x_{i+1} = \frac{x_i}{2}$
- ☐ $x_{i+1} = \frac{1}{2}(x_i + \frac{R}{x_i})$
- ☐ $x_{i+1} = \frac{1}{2}(x_i + \frac{R}{2})$
- ☐ $x_{i+1} = \frac{3x_i}{2}$

Question 2: Interpolation with Lagrange Polynomials and Cubic Splines

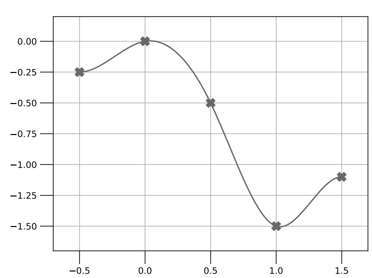
a) Which of the following statement(s) about Lagrange interpolation is/are correct?

- ☐ The resulting function of a Lagrange interpolation through 3 sample points of the function $f(x) = \frac{1}{3}x^3 + 5x + 2$ is an exact fit to the original function $f(x)$.
- ☐ The first Lagrange basis function through the points $\{x, y\}_i = \{2, 0.5\}, \{3, 1\}$ is $l_1(x) = \frac{x-3}{-0.5}$.
- ☐ When fitting noisy data, in general, the resulting function $f(x)$ from the Lagrange interpolation diverges from the Least-Squares solution.
- ☐ If we gather experimental data from a process following a 3rd order polynomial where noise can be present and fit the data using Lagrange interpolation, we will always acquire the same interpolating function for a number of samples points $N \geq 4$.

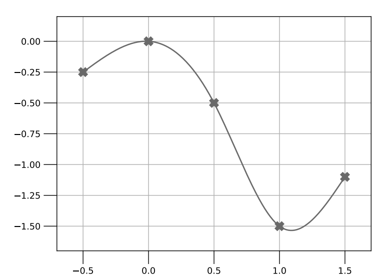
b) The figures below show three different cubic spline fits for the same data points. The only difference between them is the boundary condition, applied to both ends for every case.

Select the correct combination that matches the letter of the picture to the boundary conditions it was generated with:

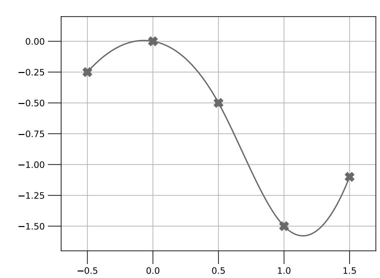
- ☐ (a): parabolic runout, (b): natural, (c): clamped
- ☐ (a): clamped, (b): parabolic runout, (c): natural
- ☐ (a): natural, (b): clamped, (c): parabolic runout
- ☐ (a): clamped, (b): natural, (c): parabolic runout



(a)



(b)



(c)

c) Which of the following statement(s) is/are correct for Cubic Spline Interpolation through 6 points $\{x_i, y_i\}_{i=1,\ldots,6}$, with clamped boundary conditions at both ends?

- ☐ The first derivatives of the fitted curves are continuous at the interior data points.
- ☐ The matrix system we need to solve can be reduced to a 4x4 tridiagonal system.
- ☐ If we work with equidistant segments, the coefficients A_i, B_i, C_i, D_i of the matrix system to solve are the same for all interior points.
- ☐ The second derivatives at the end points are continuous.
- ☐ The second derivatives are polynomials of order 2.
- ☐ The interpolation function will be a piecewise cubic polynomial, defined over 6 segments.

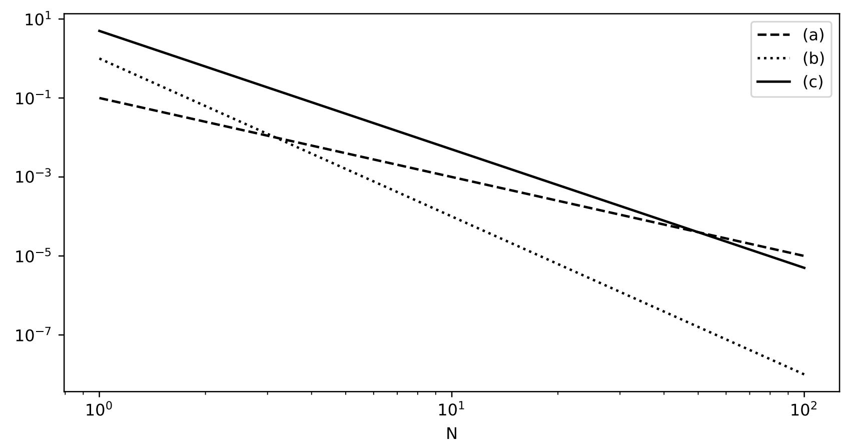
Question 3: Numerical Integration

a) Which of the following statement(s) about numerical quadrature is/are correct?

- ☐ Simpson's rule is able to perfectly integrate a polynomial of order 4 or less.
- ☐ The order of accuracy of a numerical quadrature scheme over an interval $[a, b]$, split into N sub-intervals, is one less than the order of accuracy on one of the sub-interval $[x_i, x_{i+1}]$.
- ☐ The mid point rule has a higher order of accuracy than the trapezoidal rule.
- ☐ When using Simpson's rule over an interval $[a, b]$, to reduce the error on I by a factor 1000, 10 times more data points must be used.

b) The figure below shows the evolution of the error with respect to the number of sub-intervals N for three quadrature rules over an interval $[a, b]$. Select the correct combination of quadrature rule and plot label:

- ☐ (a): Trapezoidal, (b): Simpson, (c): Rectangle
- ☐ (a): Simpson, (b): Rectangle, (c): Trapezoidal
- ☐ (a): Rectangle, (b): Simpson, (c): Midpoint
- ☐ (a): Midpoint , (b): Trapezoidal, (c): Rectangle



c) Which of the following statement(s) about Romberg integration is/are correct?

- ☐ I_1^n is 3rd order of accuracy.
- ☐ I_1^n is 4th order of accuracy.
- ☐ If you have already computed I_0^{2n}, you need n additional function evaluations to compute I_0^n.
- ☐ The number of function evaluations for I_1^2 is identical as that of I_1^1.
- ☐ The number of function evaluations for I_2^2 is twice as that of I_2^1.
- ☐ To compute I_3^1, four intervals are sufficient, that is, we only need to compute I_0^1, I_0^2, I_0^4.

d) Which of the following statement(s) about Gauss quadrature is/are true?

- ☐ The two-point Gauss quadrature is identical to the trapezoidal rule.
- ☐ The two-point Gauss quadrature is identical to Simpson's rule.
- ☐ Gauss quadrature is a method to find the optimal number of quadrature points on an interval $[a, b]$.
- ☐ The two-point Gaussian quadrature recovers the integration of third order polynomial exactly.

e) Which of the following statement(s) about adaptive integration is/are correct?

- ☐ The error $\epsilon(h/2) \approx G(h/2) - G(h)$, coming from Richardson extrapolation, can be used as a stopping criterion for the refinement process.
- ☐ F must have an analytical expression to estimate the local numerical error.
- ☐ The adaptive refining procedure should stop automatically without an external criterion.
- ☐ Refining the grid locally does not change the order of accuracy of the underlying integration scheme.

f) For Monte Carlo Integration $I \approx I_M = \frac{1}{M}\sum_{i=1}^{M} f(x^{(i)})$ which of the following statement(s) is/are true?

- ☐ Monte Carlo Integration suffers the curse of dimensionality and thus can only be applied for low dimensional problems.
- ☐ The error is estimated by the variance of the Monte Carlo Integral $\varepsilon_M = \sqrt{\mathrm{Var}\left[I_M\right]}$.
- ☐ For Monte Carlo integration the error estimate behaves as $\varepsilon_M = \mathcal{O}(M^{-1/2})$.
- ☐ The error for Monte Carlo integration is computed as the distance between the true integral I and the Monte Carlo estimate $\varepsilon_M = |I - I_M|$.

Numerical Problems [60 Points]

The following exercises involve calculations similar to the ones performed in the lecture and the exercises. **Please use one separate sheet of paper for each of the questions.**

Question 4: Linear Least Squares [15 Points]

You have gathered the following data:

t	1	2	3
x	2	5	10

Based on you knowledge of kinematics you expect the relation between x and t to follow:

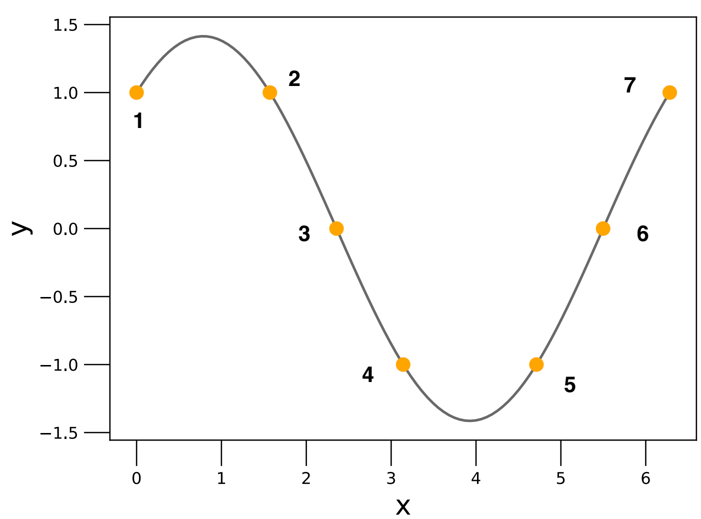
$$x(t) = \frac{1}{2}gt^2 + x_0$$

a) Using a linear least square fit, estimate the value of the intercept x_0 and the gravitational acceleration g.

b) Assume that the gravitational field varies as $g(t) = sin(w_g t)$. Keeping the same kinematic formulation with the new expression for g, can you estimate g, x_0 and w_g using linear least squares? Justify you answer.

Question 5: Minimum of a graph function [25 Points]

You are given the following graph of a function $f(x)$ that is a linear combination of the functions $\cos(x)$ and $\sin(x)$.



Assume you are given a set of points $\{x_i, y_i\}$, given in the Table below, where $y_i = f(x_i)$. Further assume, the function is 2π-periodic.

i	1	2	3	4	5	6	7
x_i	0	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π	$\frac{3\pi}{2}$	$\frac{7\pi}{4}$	2π
y_i	1	1	0	-1	-1	0	1

We want to fit the data and to find the minimum of the function $f(x)$. Two interpolation methods are available to do so: Lagrange interpolation and Cubic Spline interpolation.

a) Using the Lagrange interpolation method and all the available data, what is the degree of the resulting polynomial? Are there any dangers in fitting this kind of data with Lagrange interpolation?

b) Using Cubic Spline interpolation, specify the boundary conditions that you would use to fit the data from the function $f(x)$. Write down the equations you would need to solve for the second derivatives on the boundaries.

c) Using a smart selection of data points, fit the data using Lagrange interpolation and calculate the minimum of $f(x)$.

   A second order curve is already a good approximation of the local behaviour of $f(x)$ around the minimum.

Question 6: Inverse Transform Sampling [15 points]

Assume you are given the probability density function for the Cauchy distribution

$$p(x) = \frac{1}{\pi}\frac{1}{(1+x^2)}, \quad \forall x \in \mathbb{R} \tag{1}$$

In class you learned that you can obtain samples from a distribution by using the inverse transform method.

a) Compute the cumulative distribution function $F(x)$ for the Cauchy-distribution

$$F(x) = \int_{-\infty}^{x} p(x')\mathrm{d}x' \tag{2}$$

   You might want to use $x = \tan(\theta)$ as substitution.

b) Compute the inverse $F^{-1}(u)$ of the cumulative distribution function $F(x)$ found in the previous subquestion.

Question 7: Richardson Extrapolation [15 points]

Finite differences are used to approximate derivatives in computer simulations. The central difference approximation of $f'(x)$ is expressed as:

$$f'(x) \approx D(h) := \frac{f(x+h) - f(x-h)}{2h} \tag{3}$$

a) Using the Taylor expansion $f(x+h) = f(x) + hf'(x) + h^2\frac{f''(x)}{2} + h^3\frac{f'''(x)}{6} + \mathcal{O}(h^4)$, find the order of accuracy of the central difference approximation.

b) Use Richardson extrapolation to derive a new method that is fourth order accurate.

   You can express the new approximation as a function of $D(h)$.

Pseudocode [20 Points]

In the last part of the exam you are asked to write a pseudo-code showing implementation details for some of the algorithms learned in class. For each of your pseudocodes, make sure to specify the **input**, **output** and **steps**. As a template please us the following example computing the Fibonacci series.

Algorithm 1 Fibonacci Series

Input:

N, {number of elements to compute}
$n_{\max}$, {threshold to stop computation}

Output:

$\vec{F}$, {vector containing Fibonacci numbers}

Steps:

```
F[0] ← 0
F[1] ← 1
n ← 2
while n < N + 3 do
    F[n] ← F[n − 1] + F[n − 2]
    if F[n] > n_max then
        break
    end if
    n ← n + 1
end while
```

Question 8: Artificial Neural Network Training Loop [20 Points]

You want to predict the new daily cases of COVID-19 using a neural network. Assume your colleagues have already implemented a functioning neural network model with the following functions:

- `forward_pass(x)`: Given an input `x`, returns the output `y` through the neural network.
- `compute_loss(x,y)`: Computes and returns the loss (L2 norm in this case) between 2 entries `x` and `y`.
- `compute_gradients(x,y)`: Computes the gradient of `x` with respect to `y`, i.e. $\frac{\partial x}{\partial y}$.
- `update_weights(x)`: Updates the weights of the model by `x` $(\mathbf{w}^{k+1} = \mathbf{w}^k + \mathbf{x})$.

Additionally, the following helper functions have been implemented:

- `split_dataset(X,r)`: Given a dataset `X` of shape `[N,n_dim]`, where `N` is the number of samples, and `n_dim` the dimension of a sample, splits the dataset into 2 subsets of relative size r and $1 - r$.
- `shuffle(X,Y)`: Shuffles the content of `X` and `Y`.
- `get_batch(X,i,n_batch)`: Returns the i^{th} batch of size `n_batch` from the dataset `X`.

You are also given the input dataset $X = [x_1, x_2, ..., x_N]$ where $x_i \in \mathbb{R}^d$ and the target dataset $Y = [y_1, y_2, ..., y_N]$ where $y_i \in \mathbb{R}$. The goal of the ANN is to learn the mapping $y = f_{ANN}(x)$ between the input and the target.

a) Using the functions defined above and the given datasets, write the full training loop of the ANN as a pseudocode. Make sure you that:

- The number of training epochs, the batch size and the learning rate are defined as input.
- The network is trained using batch SGD.
- You have a mechanism to prevent overfitting.

b) Briefly describe overfitting and its dangers.