Numerical Analysis (089180) prof. Simona Perotto simona.perotto@polimi.it A.Y. 2022-2023 Luca Liverotti luca.liverotti@polimi.it

# Lab 1

September 16, 2022

## 1 Introduction

- how to contact me:
  - email: luca.liverotti@polimi.it
- approx. 10 labs = 20 h
- homework material on webeep.polimi.it

# 2 The lab tool: MATLAB/Octave

MATLAB (MATrix LABoratory) is an interactive software for:

- scientific computing
- statistical analysis
- vector and matrix computations
- graphics
- symbolic computation

Some useful MATLAB/Octave characteristics:

- graphical user interface
- code modularity: functions and scripts
- no needs of compiling
- intuitive graphics commands
- I/O: loading and saving files, variables, and data

Introduction to MATLAB/Octave:

- GUI description
- Basic MATLAB/Octave usage
- Vector and matrices
- Vector operations
- Matrix operations
- Plotting

- Function handles
- Logicals
- Control flow statements
- M-files

## 3 Basic commands

pwd returns the name of the current directory (in this example it is the default directory) ls returns the list of the files in the current directory.

Some real constants are allocated by default:

```
>> pi
ans=3.1416
>> i
ans=0+1i
>> inf
ans=Inf
```

#### Default mathematical functions

Function Meaning
sin, cos, tan sine, cosine, tangent
asin, acos, atan arcsine, arccosine, arctangent
exp exponential
log, log2, log10 natural logarithm, base 2, base 10
sqrt square root
abs absolute value

Matlab provides a large number of standard elementary mathematical functions. To have more information, type the command: help elfun

It is convenient to change the default format used by MATLAB to represent the number. Variables visualization can be changed by using the command format as following:

- format short fixed coma with 5 digits (default format): pi = 3.1416
- format long fixed coma with 15 digits: pi = 3.14159265358979
- format short e mobile coma with 5 digits: pi = 3.1416e + 00
- format long e mobile coma with 15 digits: pi = 3.141592653589793e+00

# 3.1 MATLAB/Octave - Exercises

Exercise 1.1. Define the row vector

$$\underline{v}_k = [1, 9, 25, \dots, (2k+1)^2] \in \mathbb{R}^{k+1}$$

with k=8 using the following strategies:

- 1. A for loop to define one by one each element of the vector.
- 2. The vector syntax to build it in just one shot.

Exercise 1.2. Define a function which, for an input value k, returns the corresponding vector  $\underline{v}_k$  as defined in the previous exercise.

**Exercise 1.3.** Using the function of the previous exercise write another function that returns, for a generic value k, the  $2(k+1) \times 2(k+1)$  matrix

$$M_k = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \sqrt[3]{2} & 0 & 0 & 0 & 0 & \dots & 0 & \mathbf{1} \\ 0 & 0 & \sqrt[3]{2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt[4]{2} & 0 & 0 & \dots & 0 & \mathbf{9} \\ 0 & 0 & 0 & 0 & \sqrt[5]{2} & 0 & \dots & 0 & \mathbf{9} \\ 0 & 0 & 0 & 0 & 0 & \sqrt[6]{2} & \dots & 0 & \mathbf{25} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & \frac{2^{k+1}\sqrt{2}}{2} & 0 \\ \mathbf{1} & \mathbf{1} & \mathbf{9} & \mathbf{9} & \mathbf{25} & \mathbf{25} & \dots & (2k+1)^2 & (2k+1)^2 \end{bmatrix}$$

# 4 Floating-point numbers

### 4.1 Representation of floating-point numbers

Exercise 1.4. Compare the results of the following code segments:

```
Code A

x = 0;

while (x \sim= 1)

x = x + 1/16

end

Code B

x = 0;

while (x \sim= 1)

x = x + 0.1

end
```

Solution Exercise 1.4. Code A works as expected: the **while** loop is repeated 16 times, and the final value of x is 1.

Instead, code B does **not** work as expected, and results in an infinite loop.

#### 4.2 Floating-point arithmetic

**Definition 1.1 (Machine epsilon).** The smallest positive machine number that, if added to 1, gives a result greater than 1.  $eps = 2^{-t}$  Note that eps is different from realmin. In MATLAB/Octave you can get the machine number by using the eps command.

Exercise 1.5 (Machine epsilon). Find the machine epsilon by implementing an ad hoc procedure. Comment and justify the obtained results.

Exercise 1.6 (Overflow and associative property).

```
realmax

a = 1.0e+308;

b = -a;

c = 1.1e+308;

(a + b) + c

(a + c) + b
```

Exercise 1.7 (Numerical cancellation). Consider the following function

$$f(x) = \frac{e^x - 1}{x}$$

- a. Evaluate f(x) for values of x around zero (try with  $x_k = 10^{-k}$ ,  $k \in [1, 20]$ ). What do you obtain? Explain the results
- b. Propose an approach for fixing the problem. (Hint: Use Taylor expansions to get an approximation of f(x) around x = 0.)

c. How many terms in the Taylor expansion are needed to get double precision accuracy (16 decimal digits)  $\forall x \in [0, 1/2]$ ?

Exercise 1.8 (Error propagation). The sequence

$$1, \frac{1}{3}, \frac{1}{9}, \ldots, \frac{1}{3^n}, \ldots$$

can be generated with the following recursive relations

$$\begin{cases} p_n = \frac{10}{3} p_{n-1} - p_{n-2} \\ p_1 = \frac{1}{3}, p_0 = 1 \end{cases} \qquad \begin{cases} q_n = \frac{1}{3} q_{n-1} \\ q_0 = 1. \end{cases}$$

- a. Implement the two relations in order to generate the first 100 terms of the sequence.
- b. Study the stability of the two algorithms and justify the obtained results.

#### A Homework

**Homework 1.1.** Find the minimum positive number representable in MATLAB/Octave by implementing an ad hoc procedure. Compare with realmin.

**Homework 1.2.** a. Use Taylor polynomial approximation to avoid the loss of significance errors in the following function when x approaches 0

$$f(x) = \frac{1 - \cos(x)}{x^2}$$

b. Reformulate the following function g(x) to avoid the loss of significance error in its evaluation for increasing values of x towards  $+\infty$ 

$$g(x) = x\left(\sqrt{x+1} - \sqrt{x}\right).$$

**Homework 1.3.** We can compute  $e^{-x}$  around x = 0 using Taylor polynomials in two ways, either using

$$e^{-x} \approx 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \cdots$$

or using

$$e^{-x} = \frac{1}{e^x} \approx \frac{1}{1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots}.$$

Which approach is the most accurate?

Homework 1.4. Consider the following integral

$$I_n(\alpha) = \int_0^1 \frac{x^n}{x + \alpha} dx, \quad \forall n \in \mathbb{N}, \alpha > 0.$$

- a. Give an upper bound for  $I_n(\alpha)$ ,  $\forall n \in \mathbb{N}, \alpha > 0$ .
- b. Prove the following recursive relation between  $I_n(\alpha)$  and  $I_{n-1}(\alpha)$ :

$$\begin{cases} I_n(\alpha) = -\alpha I_{n-1}(\alpha) + \frac{1}{n} \\ I_0(\alpha) = \ln\left(\frac{\alpha+1}{\alpha}\right) \end{cases}$$

- c. Employing the previous relation, compute  $I_{40}(\alpha=8)$  and comment the obtained results.
- d. Write a numerically stable recursive relation for  $I_{40}(\alpha = 8)$ .

Homework 1.5. Given the following sequence:

$$\begin{cases} x_{n+1} = 2^{n+1} \left[ \sqrt{1 + \frac{x_n}{2^n}} - 1 \right] \\ x_0 > -1 \end{cases}$$

for which  $\lim_{n\to+\infty} x_n = \ln(1+x_0)$ .

- a. Set  $x_0 = 1$ , compute  $x_1, x_2, \ldots, x_{71}$  and explain the obtained results.
- b. Transform the sequence in an equivalent one that converges to the theoretical limit.