

Modeling with Matlab

AIV M1 bootcamp

Antoine FRÉNOY, Dusan MISEVIC
frenoy@gmx.com, dule@alife.org

September 2012

The course material is available on <http://perso.crans.org/frenoy/matlab2012.html>.

Basics of Matlab programming, basic algorithmic

A. The very beginning

1. Start matlab.
2. Calculate $32.4/23.145 * 8.3$, store the result in the variable `x`.
3. Type `doc sin` in the command window.
4. Calculate $\sin(\frac{4 \times \pi}{5})$, store the result in the variable `y`.
5. Write matlab code to swap the content of the variables `x` and `y`.

B. Functions and very simple algorithms

1. Write a matlab function `m2sum` that takes two numbers and returns their sum. *Look in the help: **function**.*
2. Write a matlab function `sequence` that takes an integer `n` and returns the n^{th} element of the sequence (u_n) defined by:

- $u_0 = 1$
- $u_{n+1} = u_n + 3 * n - 4$

*Look in the help: **for**.*

C. Prime numbers

1. Write a matlab function `mgcd` that takes two integers and returns their greater common divider. *Look in the help: **while**.*
2. Write a matlab function `misprime` that takes a positive integer and tells whether it is a prime number. *Look in the help: **if**, **sqrt**, **floor**, **break**.*

D. Algorithmic on vectors

1. Write a matlab function `mini` that takes a vector and returns its lowest element. *Look in the help: **Inf**.*
2. Write a matlab function `mini2` that takes a vector and returns its lowest element and the index of this element.
3. Write a matlab function `reverse` that takes a vector and reverses it (the first element becomes the last one and so on). *Look in the help: **size, length, zeros**.*
4. Write a matlab function `incsort` that takes a vector and sorts it in ascending order.

E. Text printing

1. Write a matlab function `nicetext` that takes a number and prints a sentence about this number, for example: “The number is 15, thank you !”. *Look in the help: **fprintf**.*

F. Distances

1. The Hamming distance between two strings of equal length is the number of positions at which the corresponding symbols are different. Write a matlab function `hamming` that takes two vectors and returns their hamming distance. *Look in the help: **assert**.*
2. The Levenshtein distance between two strings is defined as the minimum number of edits needed to transform one string into the other, with the allowable edit operations being insertion, deletion, or substitution of a single character. Write a matlab function `levenshtein` that takes two vectors and returns their Levenshtein distance. *Look in the help: **return**.*

G. Loading and writing data

1. Load the variable `m` contained in the file `load1.mat`. Sum the elements of `m` and put the result in a variable `s` that you will store in the file `save1.mat`. *Look in the help: **load, sum, save**.*
2. Load the file `data1.xls` and sum the third line of numerical values. *Look in the help: **xlsread**.*

H. Working with matrices

1. Load the matrix `A` from the file `a.mat`.
2. Multiply element by element the third and fifth column of `A`, ie create a column vector `z` such that $\forall i, z(i) = A(i, 3) * A(i, 5)$. *Look in the help: **arith**.*
3. Create a column vector that contains the sum of each line of `A`.
4. Sum all the elements of `A` that belong to an even column.
5. Count the number of elements of `A` that are greater than 2. *Look in the help: **find, numel**.*

I. Randomness

1. Create a 10 by 15 matrix filled with uniformly distributed random numbers between 1 and 4. *Look in the help:* **rand**.
2. Create a 100 by 2 matrix filled with normally distributed random numbers with mean -2 and standard deviation 1. *Look in the help:* **randn**.

Plotting

J. 2D curves

1. Plot the graph of the function $x \mapsto 3 * x^2 - 6 * x + 1$ over $[0,10]$. *Look in the help:* **plot**.
2. Plot the graphs of the functions \sin and \cos^2 over $[0,5]$ with different colors. Add a legend to identify each function. *Look in the help:* **legend**.

K. Histograms

1. Create a vector **A** filled with 10000 uniformly distributed random numbers between 0 and 10. Plot the distribution of **A**. *Look in the help:* **hist**.
2. Same question with normally distributed random numbers with mean 8 and standard deviation 2.

L. Points

1. Load the file **points.mat**. It contains a list **l** of 2D points. Plot these points using “x” as marker on a 2D graph.
2. Zoom to plot only $x \in [0, 5]$ and $y \in [-1, 1]$. *Look in the help:* **xlim**, **ylim**.
3. Name the x-axis “time” and the y-axis “sales”. *Look in the help:* **xlabel**, **ylabel**.
4. The file **points.mat** also contains standard error for each point in variable **err**. Plot the same graph as in q. 1 with error bars. *Look in the help:* **errorbar**.

M. Advanced plotting

1. Let's consider the function $f(x, y) = \sin(6 * x) + y$. Plot f over $[0, 1] \times [0, 1]$ as a 3D graph, ie $z = f(x, y)$. *Look in the help:* **surf**.
2. Plot the same function as a color 2D graph, ie (x, y) is colored according to the value of $f(x, y)$. Add a color scale. *Look in the help:* **pcolor**, **colorbar**.
3. Put the last two graphes on the same figure. *Look in the help:* **subplot**.
4. Export this figure as a pdf file named **f1.pdf**. *Look in the help:* **print**.

A few mathematical problems

N. Collatz problem

A Collatz sequence is defined by:

- $u_0 = a$ with $a \in \mathbb{N}^*$
- if u_n is even then $u_{n+1} = u_n/2$
- else $u_{n+1} = 3 \times u_n + 1$

1. Write a matlab function `collatzseq` that takes a and n and returns u_n for $u_0 = a$.
2. Write a matlab function `collatz` that takes a and returns the smallest n such that $u_n = 1$ (with $u_0 = a$) if such a n exists.
3. Do you think that such a n exists for any a ?

O. Intermediate value theorem

Let's define $f(x) = x^3 + 2 * x^2 - x - 1$.

1. Write a matlab function `f` that takes x and returns $f(x)$.
2. Calculate $f(0)$.
3. Calculate $f(1)$.
4. Plot the graph of f over $[0, 1]$.
5. Write a matlab function `sol` that takes e and returns an approximate solution with an error smaller than e of $f(x) = 0$ over $[0, 1]$.
6. Explain with mathematical arguments why we know that such a solution exists.

Probability and statistics

P. Monty Hall

From Wikipedia:

The Monty Hall problem is a famous problem in probability. The problem is based on a television game show from the United States, Let's Make a Deal. It is named for the host of this show, Monty Hall. In the problem, there are three doors. A car (prize of high value) is behind one door and goats (booby prizes of low value) behind the other two doors. First, the player chooses a door but does not open it. Then the host, who has knowledge of what is behind every door, opens a different door which he is certain has a goat behind it (opening either door with equal chances if the car is behind the player's door). Last, the host lets the player choose whether to keep what is behind the first door or to change choices to the third door (the one the host did not open). The rules of the problem are that the host has to open a door with a goat behind and has to let the player switch. The question is whether changing choices increases the chances of getting the car.

1. Do you think that changing choices increases the chances of getting the car ?
2. Write matlab code to simulate the problem. Was your guess right ?
3. Can you explain the result you get with a mathematical proof ?

Q. Basic statistics

1. We randomly picked 20 men and 15 women among the world population, and recorded their size in the file `humansizes.mat`. Perform a statistical test that will estimate whether men are statistically taller than women. *Look in the help: `ttest2`.*
2. We did the same for Northern Gannet, with 30 females and 28 males, and saved the results in the file `birdsizes.mat`. Perform a statistical test that will estimate whether there is a size difference between males and females.

R. Central limit theorem

1. Read http://en.wikipedia.org/wiki/Central_limit_theorem. *Look in the help: `web`.*
2. Choose a distribution D among the ones that you can easily generate with the Statistics Toolbox (*Look in the help: `random`*). Produce a nice plot to illustrate this distribution.
3. We want you to write matlab code that will numerically test this theorem for the distribution you choosed. We suggest that you write a function that:
 - Takes `n`, the number of independent random variables, as an argument.
 - Picks `n` independent random variables following the distribution D that you choosed. For us, an independent variable will be represented by a finite but large vector of values. You will have to choose this number of values or to put it as an argument of the function.
 - Sums these variables.
 - Produces a plot that shows the distribution of this sum.
 - Uses a statistical test to estimate whether the sum is normally distributed with mean and variance being the ones given by the central limit theorem. You will have to browse the help to find the appropriate test.

Differential equations

S. First order

Use matlab to get a numerical solution (*ie* a 2D curve) of the following differential equations, and use your math skills to get a formal solution when you can. We will always solve over $[0,1]$ and take $y(0)=1$. *Look in the help: `ode45`.*

1. $y' - 3y = 2$
2. $y' + 2y = e^{2x}$
3. $y' - 5y = e^{5x}$

4. $y' - y = \sin(x)$
5. $(2 + x)y' = 2 - y$
6. $(1 + x)y' + y = \cos(x)$
7. $(1 + x)y' + y = (1 + x)\sin(x)$
8. $y' + y = \sin(x) + 3\sin(2x)$

T. Second order

Use matlab to get a numerical solution (*ie* a 2D curve) of the following differential equations, and use your math skills to get a formal solution when you can. We will always solve over $[0,1]$ and take $y(0)=1$ and $y'(0)=-1$.

1. $y'' - 2y' + 2y = xe^x$
2. $y'' - 4y' + 4y = 2(x - 2)e^x$
3. $y'' - 4y' + 13y = 10\cos(2x) + 25\sin(2x)$
4. $y'' + 3y' + 2y = \frac{x-1}{x^2+1}e^{-x}$
5. $y'' + y = 3x^2 - 5x - 2$
6. $y'' - y' - e^{2x}y = e^{3x}$ (suggestion: $u = e^x$)

U. System of equations

1. Use matlab to solve the following sytem over $[0,100]$ with parameters $r = 0.3$ and $m = 0.1$.
 - $a' = r*\sin(t)*a + m*a*b - m*a*c$
 - $b' = r*\cos(t)*b + m*b*c - m*b*a$
 - $c' = r*\sin(t/2)*c + m*c*a - m*c*b$

Analyzing a large set of data

Description of the raw data

The file `influenza.xls` contains data about mortality by influenza in United States for European American population between 1900 and 2006. “EAM” refers to European American Males, and “EAF” refers to European American Females. “Raw Data” is the number of deaths from influenza at each time point and in each age category, “Population” is the number of people alive at each time point and in each age category, “1 minus TOT” is 1 minus the mortality by any cause at each time point and in each age category.

We want to analyze these data.

Loading the data

1. Read the whole exercise before choosing which data you will store and how you will store them.
2. Be careful about possibly missing data.
3. You will probably have to make approximations, interpolations, ...

Output

1. A first kind of graph we want to extract from these data would show us, for a given time point (parameter), what is the mortality by influenza (y-axis) for each age category (x-axis).
2. A second kind of graph we want to extract from these data would show us, for a given age category (parameter), what is the mortality by influenza (y-axis) at each time point (x-axis).
3. A third kind of graph we want to extract from these data would show us, for people born in a given year (parameter), what is the mortality by influenza (y-axis) at each age (x-axis).

Plot one graph of each category. Do not forget to label the axes of your plot, add a legend, ...

Game

Choose one of the following games, and implement it !

Mastermind

You should first read [http://en.wikipedia.org/wiki/Mastermind_\(board_game\)](http://en.wikipedia.org/wiki/Mastermind_(board_game)).

1. The number of colors, the number of holes in each row and the total number of rows should be parameters.
2. You should first implement the codemaker part (the central function would be the one taking the correct pattern and the guess made by the codebreaker, and returning the feedback to give to the codebreaker).
3. Then implement the game with the user (the codebreaker) entering his guess as text.
4. Make a nice graphical interface.
5. Bonus: write an artificial intelligence that plays wisely !

Minesweeper

You should first read [http://en.wikipedia.org/wiki/Minesweeper_\(video_game\)](http://en.wikipedia.org/wiki/Minesweeper_(video_game)).

1. The size of the grid and the distribution of mines should be parameters.
2. We suggest that you first implement everything without graphics.
3. Then implement the graphical interface.
4. Bonus: write an artificial intelligence that plays wisely !

Sea Battle

You should first read [http://en.wikipedia.org/wiki/Battleship_\(game\)](http://en.wikipedia.org/wiki/Battleship_(game)).

1. The size of the grid, the number of boats, ... should be parameters.
2. We suggest that you first implement everything without graphics.
3. Do not forget that the players should be able to choose their position of his boats.
4. Implement the graphical interface for a 2 players mode.
5. Bonus: write an artificial intelligence that plays against a human player.

Snake

You should first read [http://en.wikipedia.org/wiki/Snake_\(video_game\)](http://en.wikipedia.org/wiki/Snake_(video_game)).

1. The size of the grid, ... should be parameters.
2. A difficulty is that the game is real time, so your code has to be fast enough.
3. The game is not really complicated, so we expect a lot of eye candy and extra features (several level with increasing difficulty, a level editor, ...) !

Tetris

You should first read <http://en.wikipedia.org/wiki/Tetris>.

1. The size of the grid, ... should be parameters.
2. A difficulty is that the game is real time, so your code has to be fast enough.
3. The game is not really complicated, so we expect a lot of eye candy and extra features (several level with increasing difficulty, a level editor, ...) !

Pong

You should first read <http://en.wikipedia.org/wiki/Pong>.

1. The size of the grid, ... should be parameters.
2. A difficulty is that the game is real time, so your code has to be fast enough.
3. Write a 2 players mode.
4. Write a mode with an artificial intelligence that plays against a human player.

Tic Tac Toe

You should first read <http://en.wikipedia.org/wiki/Tic-tac-toe>.

Other ?

If you have an other idea of a game you would like to implement, talk to us !

Image analysis

Choose and implement one of the following projects.

Seam Carving

The following links explain an automated image resizing algorithm called seam carving. Watch the Youtube video, read the Wikipedia page. The third link is the paper by the inventors of this technique which provides more details if you need.

- <http://www.youtube.com/watch?v=6NcIJXTlugc>
 - http://en.wikipedia.org/wiki/Seam_carving
 - <http://perso.crans.org/frenoy/matlab2012/seamcarving.pdf>
1. Try to summarize with your words what is seam carving and what are the interests and weaknesses compared to other resizing techniques.
 2. We want you to implement seam carving with matlab ! Here are the steps we suggest:
 - Identify the arguments of the matlab function.
 - Load the image (*Look in the help: **imread**.*).
 - Compute the energy matrix.
 - Choose and implement an algorithm to find the best seam to remove.
 - Perform the removal and update the energy matrix.
 - Loop the last two steps.
 - Save the image (*Look in the help: **imwrite**.*).
 3. Test on several images (<http://images.google.com/>).
 4. Every improvement or innovation you can do is welcome ! For example you can try to use other energy functions, to improve the performance of your program, to add a user interface ...

Detecting bacteria

You can find microscope images coming from our lab on <http://perso.crans.org/frenoy/matlab2012/microscope/>. We want you to write matlab code to detect the bacteria present on these images. Many functions from the Image Analysis toolbox could be useful, read the documentation !

Digit recognition

Write matlab code that recognizes handwritten digits from an image. This is a very open problem, you will have to look by yourself for digit recognition algorithms. Use the web !