MATLAB Fundamentals

Heidelberg-EMBL
12-15/09/2016

Tutor: Karin Sasaki
(EMBL Centre for Biological Modelling)

* All icons are from The Noun Project

Bio-IT

Bio-IT

The EMBL Computational Centres

Modelling

CBM

Networks

CBNA

Statistics

CSDA

EMBL-HD’s community for scientific computer users

- Courses and workshops
- Resources
- Networking and social events
- Web portal: https://bio-it.embl.de

Interdisciplinary research and support for **ALL** of EMBL.

- Consulting: advice or hands-on help
- Collaboration
- Training
- Online resources

Introductions

Introduce your neighbour!

Find out:

1. Their name.

2. Something they think they are good at/enjoy doing.

3. If they were a component of a cell, what would they be and why.

4. How they think MATLAB will/has help their work.

Aim of the course

1. MATLAB language
2. Basics of programming
3. Import and visually analyse data
4. More practice:
    - Build, simulate and explore mathematical models of biological systems
    - Image processing.

Logistics

- Morning open session: 10 am -12 noon
- Afternoon practical sessions: 1 m - 5/6 pm
    - is ok if you need to leave!
- Informal course, so ask questions throughout.
    - Color coded help: Green for everything is ok, Pink for need help
- Break at 3:30
- Any questions?

Introductions and expectation

What is your name?

What do you do?

Why do you want to learn MATLAB?

Course content

- Get familiarised:
    - MATLAB syntax and functions
    - Basics of programming
- Learn to:
    - Import, visualise, explore and export data
    - Build, simulate and explore mathematical models of biological systems with MATLAB

MATLAB

- MATLAB is a language for technical computing
- MATLAB (Matrix Laboratory) is based on matrix (**array**) operations
- It integrates computation, visualisation, and programming
- Interactive environment
- MATLAB is a numerical software

MATLAB Windows, variables and output format

MATLAB Windows, variables and output format

exercise 2.1 in pdf

Arrays

MATLAB was originally written to ease dealing with tools of linear algebra – vectors and matrices.

Array - is a multi dimensional grid of data.

Single number: is a 1 x 1 array.

Column vector: a m x 1 array.

$$\boldsymbol{x} = \underbrace{\begin{pmatrix} x1 \\ x2 \\ x3 \\ x4 \\ x5 \end{pmatrix}}_{1} \updownarrow m$$

Row vector: a 1 x n array.

$$\mathbf{y} = \underbrace{(y_1 \;\; y_2 \;\; y_3 \;\; y_4 \;\; y_5)}_{n} \updownarrow 1$$

Matrix: a m x n array.

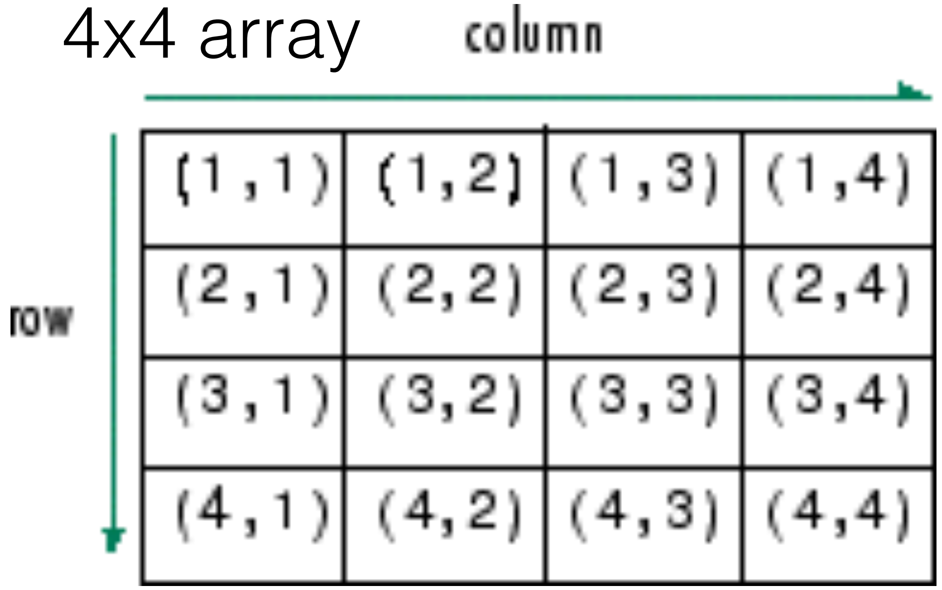
$$\boldsymbol{A} = \underbrace{\begin{pmatrix} -3 & 0 & 2 & 3 & -3 \\ -5 & -1 & 3 & 0 & -3 \\ 2 & 3 & -5 & 2 & 2 \\ -1 & 0 & 2 & -1 & -2 \\ 4 & -3 & -1 & -2 & 0 \end{pmatrix}}_{n} \updownarrow m$$

MATLAB stores data all in arrays

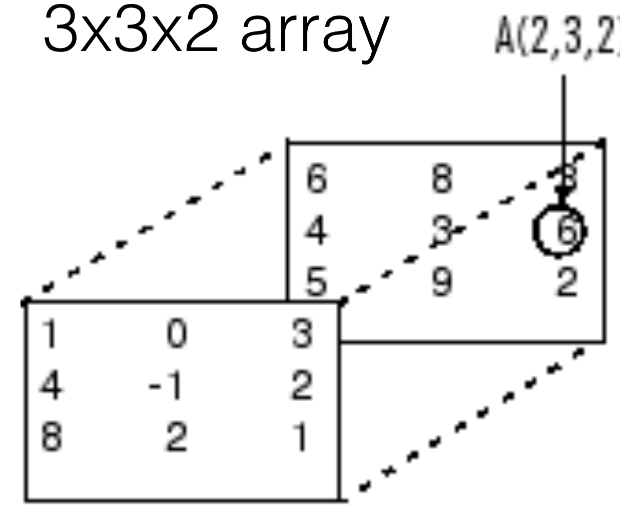
So working with arrays is fundamental to working with MATLAB

Arrays

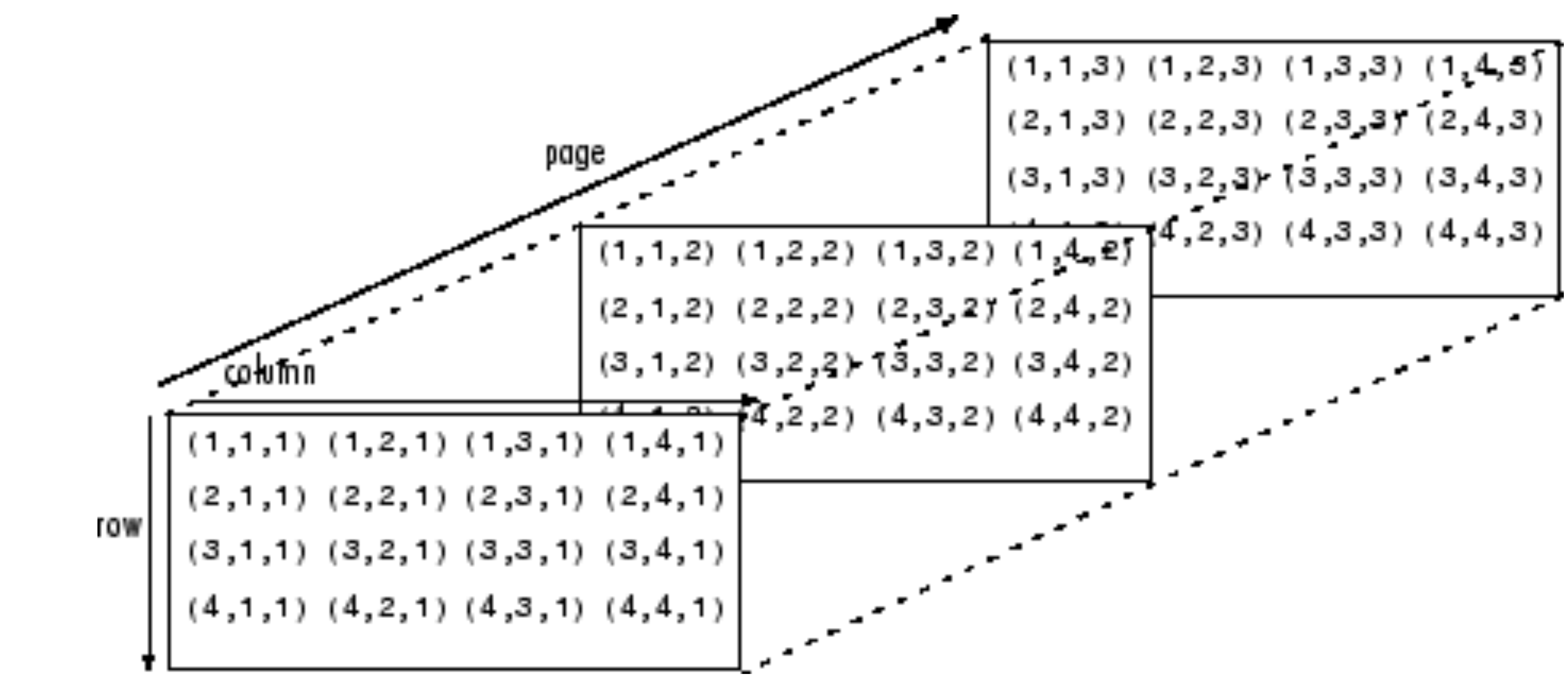
4x4 array



3x3x2 array

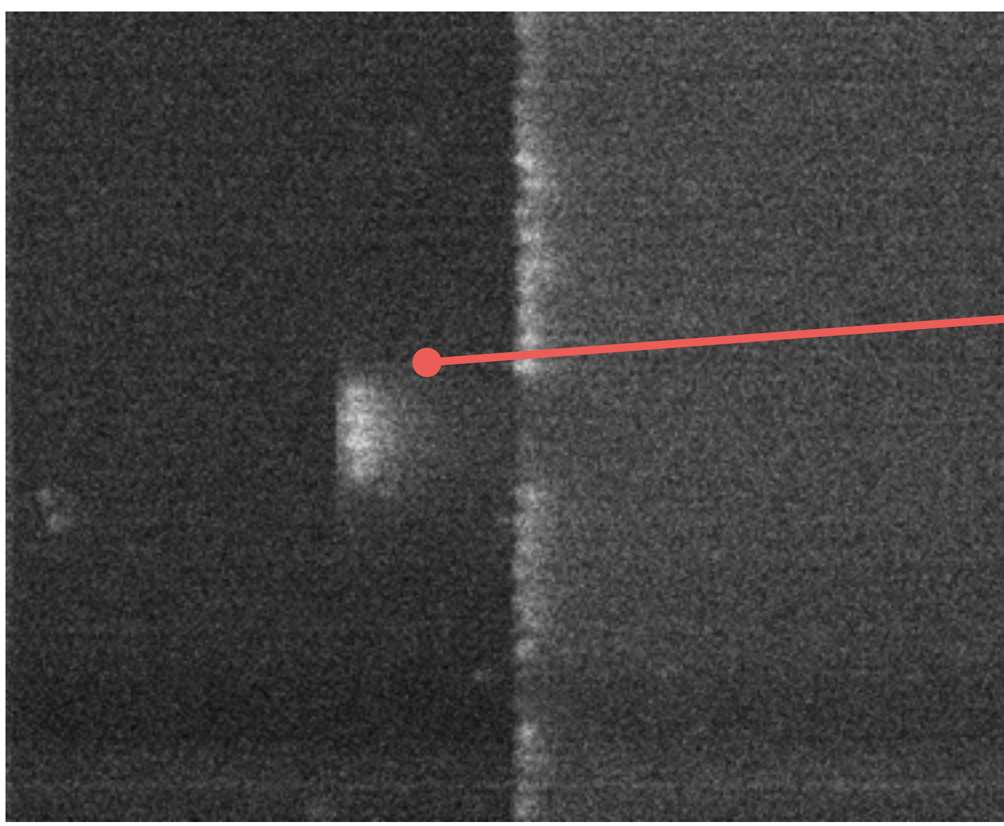



```
A(:,:,1) =

     1     0     3
     4    -1     2
     8     2     1

A(:,:,2) =

     6     8     3
     4     3     6
     5     9     2
```



Arrays - Example 1
Image storage

Image: set of data that is real-valued, ordered, represents color and intensity



73	61	55	27	137	112	121	197	239
9	25	131	55	124	147	173	133	135
32	42	86	76	144	68	143	94	178
26	76	57	78	91	87	51	176	148
42	60	95	90	95	36	150	158	122
38	26	84	65	51	49	106	66	119
32	48	78	28	24	19	94	127	62
61	128	88	92	55	99	110	126	127
55	70	57	63	59	101	118	90	88
54	26	38	31	67	78	31	127	107
67	57	81	70	83	142	143	99	88
103	48	78	119	61	55	120	139	201

Arrays - Example 2
Stoichiometry matrix

E + S → ES
k_1*E*S
E--; S--; ES++

ES → E + S
k_2*ES
ES--; E++;S++

ES → E + P
k_3*ES
ES--; E++; P++

	E	S	ES	P
E + S → ES	-1	-1	1	0
ES → E + S	1	1	-1	0
ES → E + P	1	0	-1	1

Arrays - Other examples

- Can you think of other examples?

Operations on Arrays

Help

Arrays

exercise 2.2 in pdf

Data types/classes

Data types

exercise 2.3 in pdf

*

In-built Functions

In-built Functions

exercise 2.4 in pdf

Scripts, logical operators and control of flow

Relational/Logical Operator

<table>
<tr><td><</td><td>Less than</td></tr>
<tr><td><=</td><td>Less than or equal to</td></tr>
<tr><td>></td><td>Greater than</td></tr>
<tr><td>>=</td><td>Greater than or equal to</td></tr>
<tr><td>==</td><td>Equal to</td></tr>
<tr><td>~=</td><td>Not equal to</td></tr>
<tr><td>&</td><td>Logical AND</td></tr>
<tr><td>|</td><td>Logical OR</td></tr>
</table>

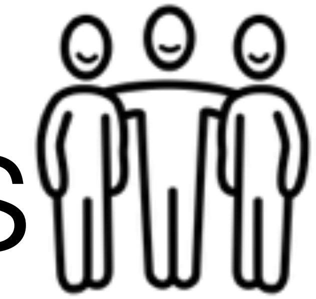
Control flow

	Description	Language Synthanx	Use for:
if, elseif, else	Execute statements if condition is true	if logical operator (true) statements elseif logical operator (true) statements else statements end	• Conditional assignment • Compare arrays • Test for equality • Evaluate various conditions
for	Execute statements a specified number of times	for condition statements end	• Assign matrix values • Decrement values • Execute statements for specificed values
while	Repeat execution of statements while condition is true	while logical operator (true) statements end	• Repeat statements until expression is false

- Loops have always been slow in MATLAB, so avoid loops if possible.; Vectorisation still speeds things up.

* Examples adapted from the Mathworks website

Scripts, logical operators and control of flow



exercises 2.5 and 4.3 in pdf

Array operations
vs
loops

Functions

Functions

exercise 4.1 in pdf

Scripts vs Functions

Scripts vs Functions

exercise 4.2 in pdf

Debugging tools

Debugging

exercise 4.4 in pdf

Contractile actin network

80μm in diameter

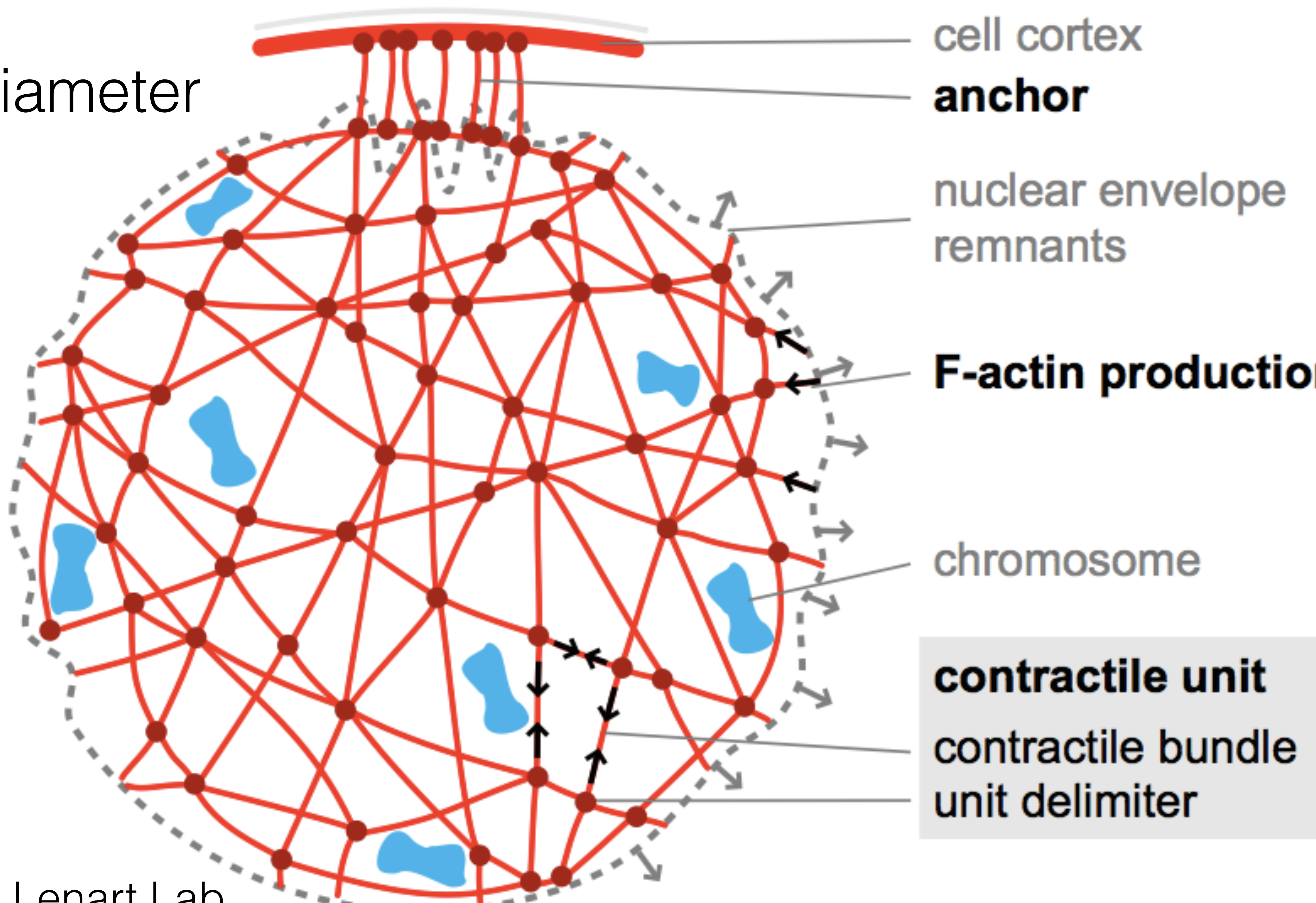
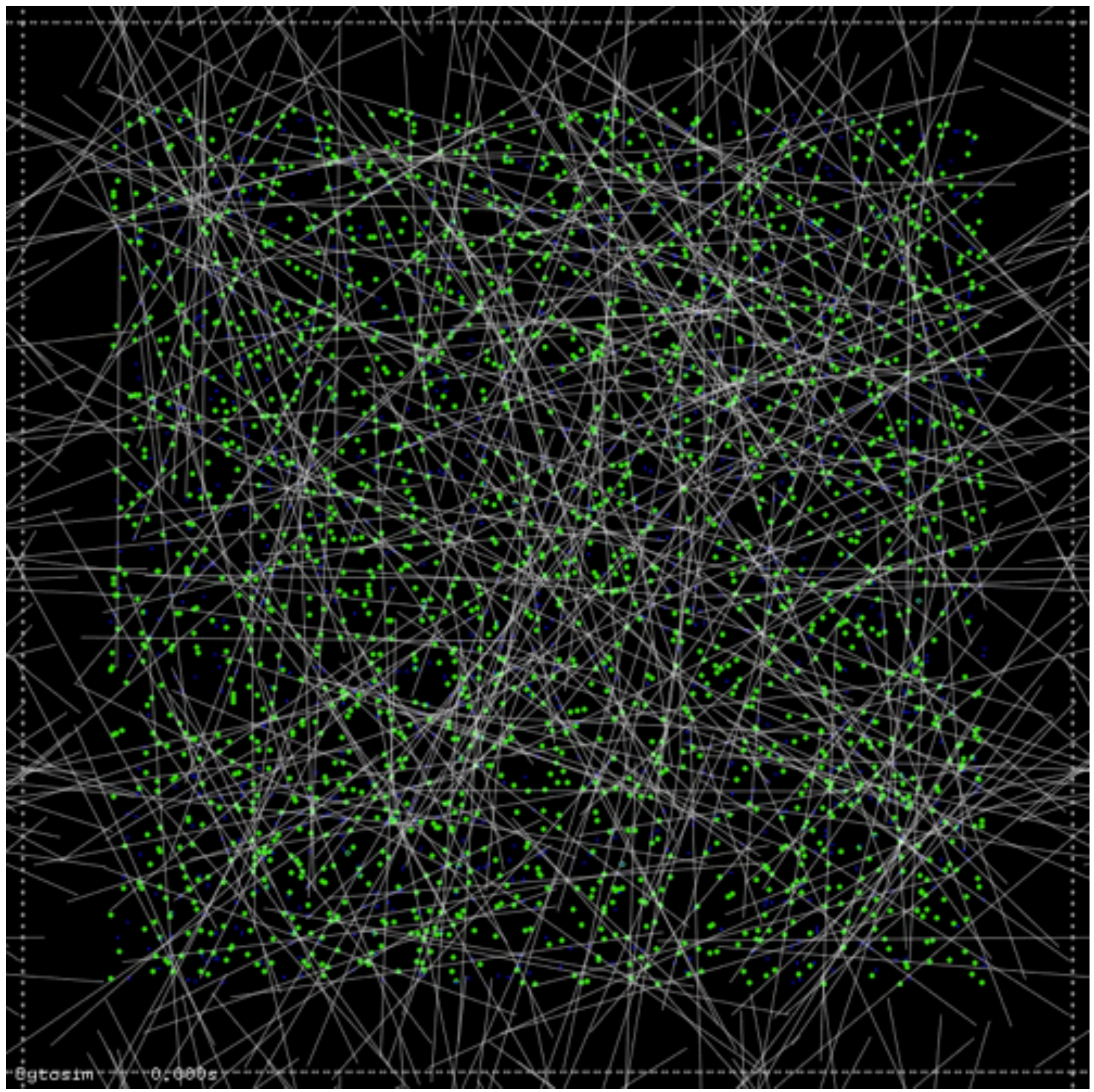


Image from the Lenart Lab

Contractile actin network

The nucleus is (0,0)



Cytosim

Q3. Is the network contracting?

Q4. Can you quantify the behaviour?

Lipid-mediated recruitment of proteins to specific membranes

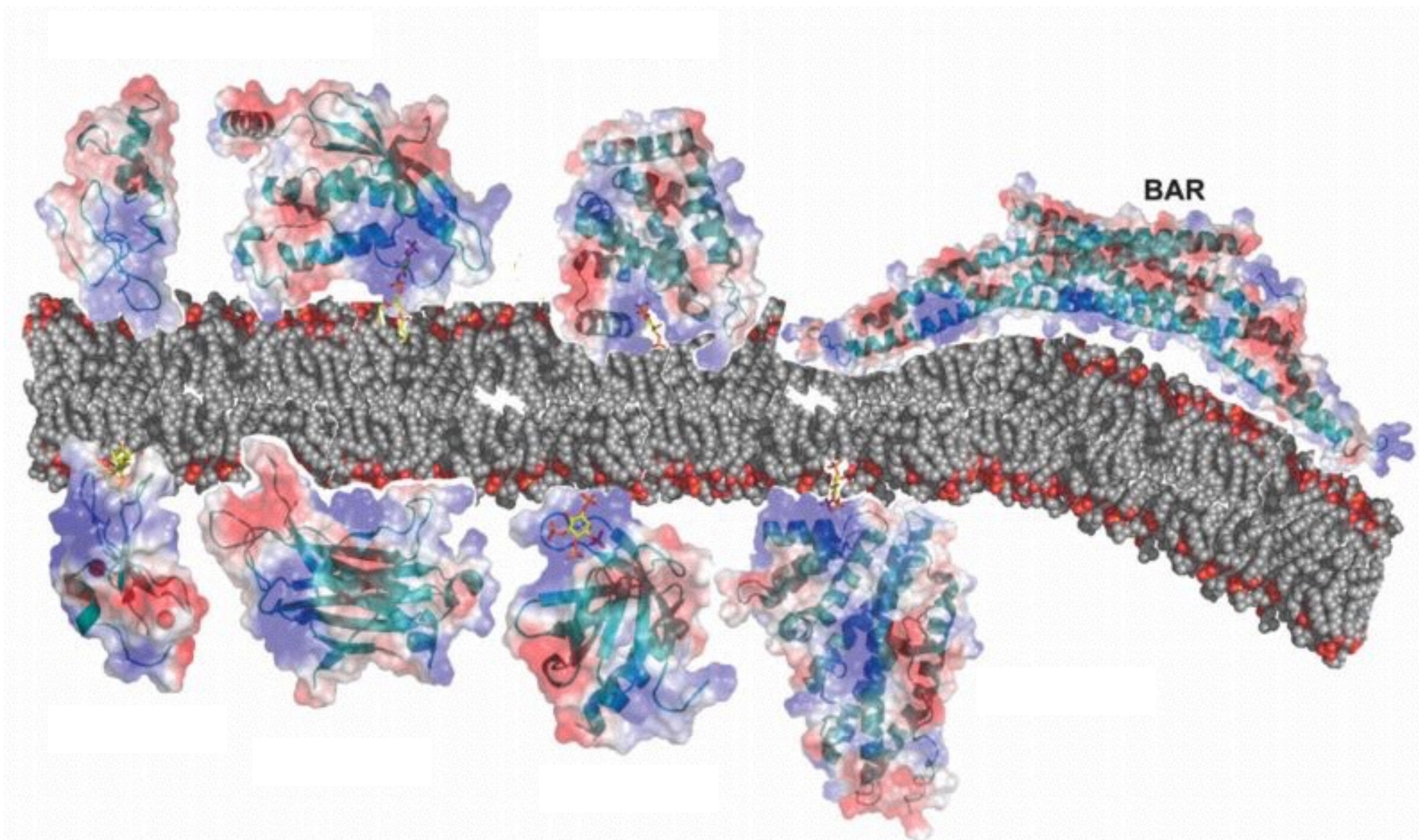


Image from Prof. Wonhwa Cho website.

Lipid-mediated recruitment of proteins to specific membranes

Measured membrane-binding properties of some of the most common phosphoinositide-binding targets

We want to answer two questions

Q1. Do we observe cooperativity between lipids?

Q2. Is cooperativity inhibitory or enhancing for protein recruitment?

Importing data interactively/programatically

Importing data

exercise 6.1 in pdf

Plotting interactively/ programatically

Plotting

plot(x, y)	Plots graph with x and y being vecotors with the coordinates in the x and y axes
xlabel(‘X’) xlabel(‘Y’)	names the axes
title(‘Plot’)	Names the figure
xlabel(‘range’) ylabel(‘variable’)	Names the axes
legend(‘name1’…)	Graph legend for lines and patches
hold on hold off	Retains current graph in figure Transcribes current graph in figure
figure (1)	Creates figure 1
saveas(h, filename)	Saves figure with handle h to file filename

Color specifier	**Colors**	**Line specifier**		**Marker specifier**	
r	Red	‘-'	Solid line (default)	'+'	Plus sign
g	Green	‘--'	Dashed line	'o'	Circle
b	Blue	‘:'	Dotted line	'*'	Asterisk
c	Cyan	‘-.'	Dash-dot line	'.'	Point
m	Magenta			'x'	Cross
y	Yellow			square' or 's'	Square
k	Black			'diamond' or 'd'	Diamond
w	White				

Plotting (interactively/programatically)

exercise 6.2 in pdf

More practice!

Mathematical descriptions of
the time behaviour of a spatially
homogenous chemical system

Image processing

Mathematical descriptions of the time behaviour of a spatially homogenous chemical system

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a system using mathematical concepts and language.”

variables

[x]

Vmax

Kd

EC_{50}

length

$t_{1/2}$

relationships

$$K_d = \frac{[A] \cdot [B]}{[AB]}$$

$$d[X]/dt = k \cdot [Y]^2$$

$$\sum_i [X]_i - F(t) = 0$$

$$k(t) \sim N(k, \sigma^2)$$

If $\text{mass}_t > \text{threshold}$
then $\text{mass}_{t+\Delta t} = 0.5 \cdot \text{mass}$

constraints

[x]>0

Energy conservation

Boundary conditions (v < upper limit)

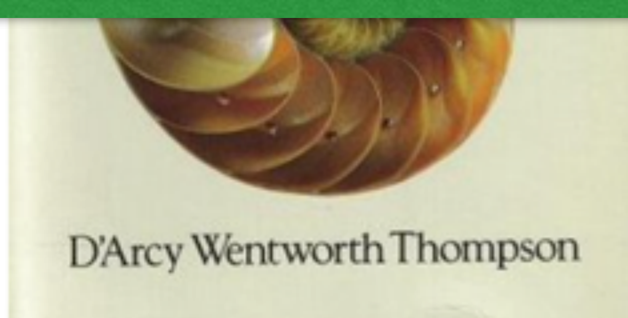
Objective functions (maximise ATP)

Initial conditions

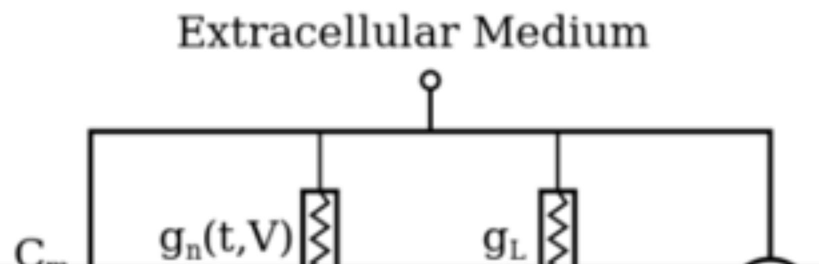
Different types: Dynamical models, logical models, rule-based models, multi-agent models, statistical models, etc.

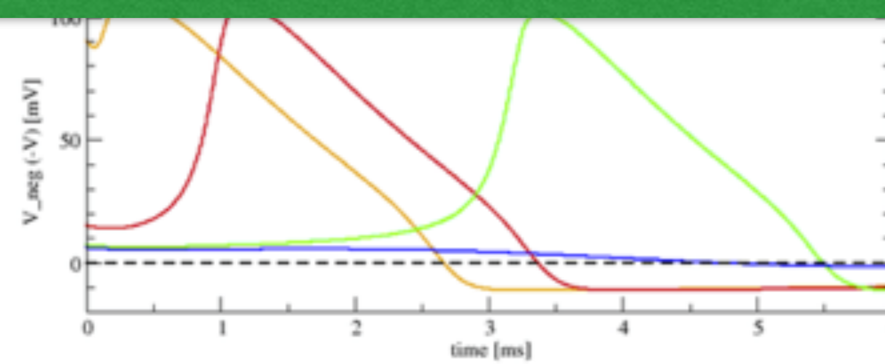
Why using mathematical models?

Describe	Explain	Predict

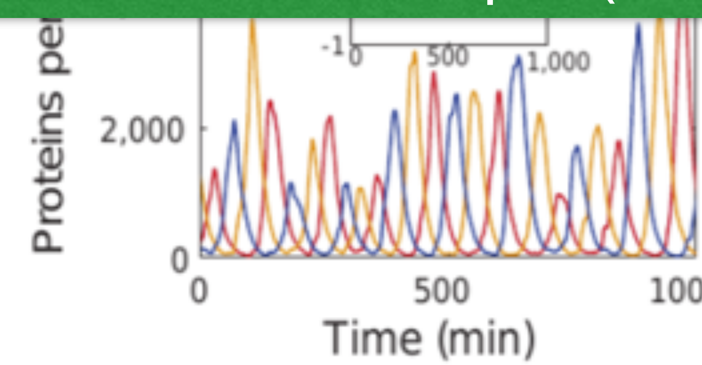


1917





1952



2000

"Not unique! Several system structures may generate the same properties."

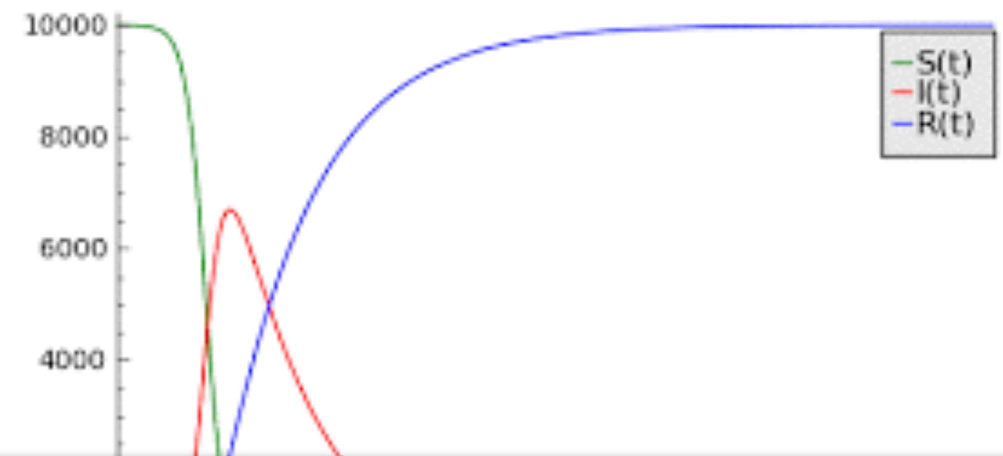
Michael Stumpf (ICL)

Modelling the spread of HIV in a society

Necessary constituents?

- ✓ Susceptible - S
- ✓ Infected/infective - I
- ✓ Recovered/no longer susceptible - R

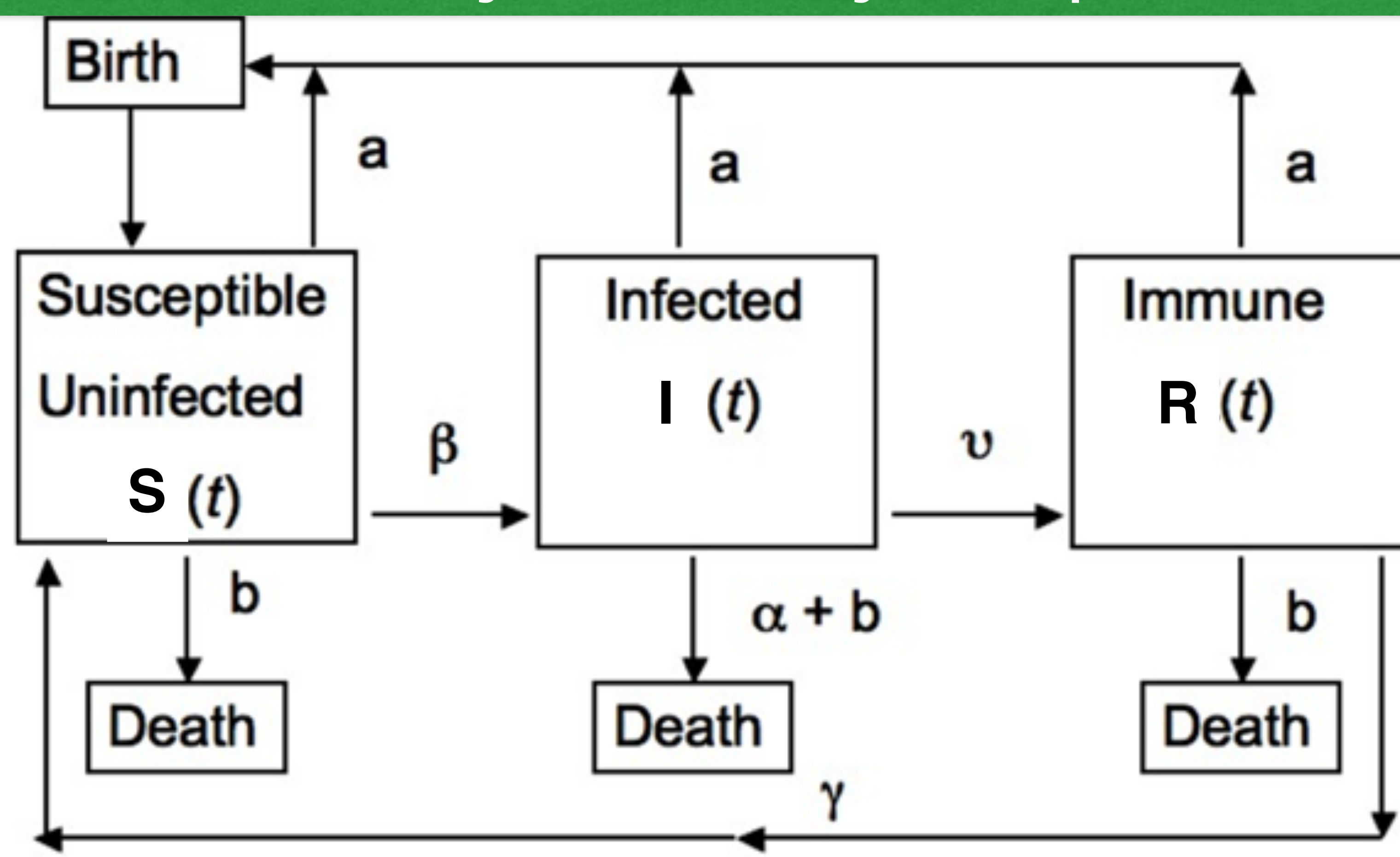
Emerging properties



Naming variables carefully is very important!!

Relationships

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta\frac{SI}{N} \\
\frac{dI}{dt} &= \beta\frac{SI}{N} - \gamma I \\
\frac{dR}{dt} &= \gamma I \\
N &= S + I + R
\end{aligned}
$$



Why using mathematical models?

Mathematical modelling can be very helpful for discovering and understanding biological processes and organisation principles, because:

- it forces the investigator to formulate hypotheses and insights in a clear-cut and formal way
- it may allow for the representation and evaluation of system compounds that are experimentally not accessible
- it allows to explore many scenarios or parameter values in less time and cheaper than in experiments
- it may help to extract structural dependencies or mathematical and physical relation that are hard to find by biological intuition, e.g. + and - loops

Deterministic mathematical description of the time behaviour of a spatially homogenous chemical system

Assumptions:

1. chemical reactions are continuous rate processes

2. chemical reactions are a predictable process, based only on the initial conditions

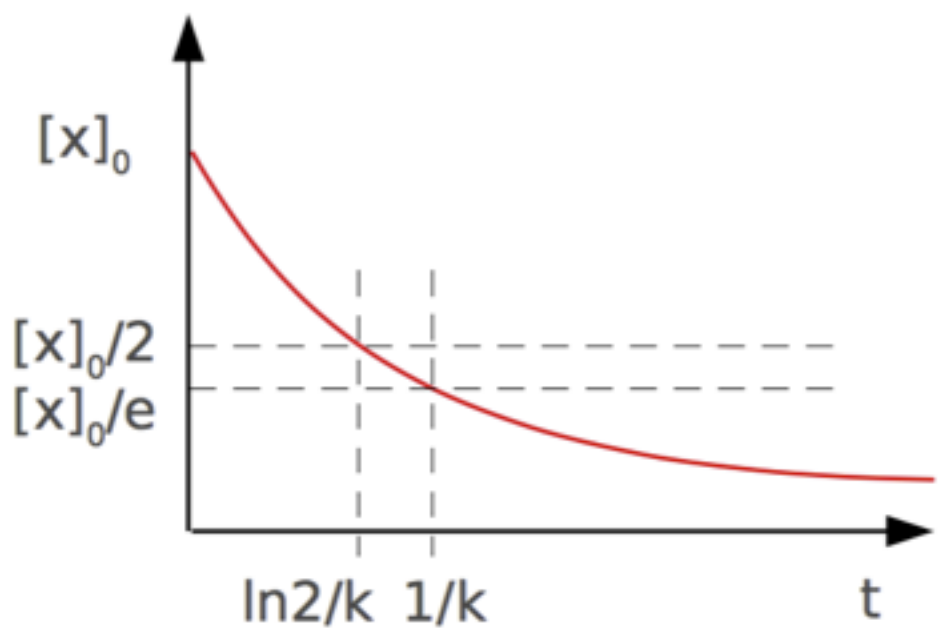
➡ governed by a set of coupled, ordinary differential equations (ODEs)

Differential Equations

A differential equation is a mathematical equation that relates some function with its derivatives.

(In biology) the functions represent physical quantities, the derivatives represent their rates of change, and the equation defines a relationship between the two.

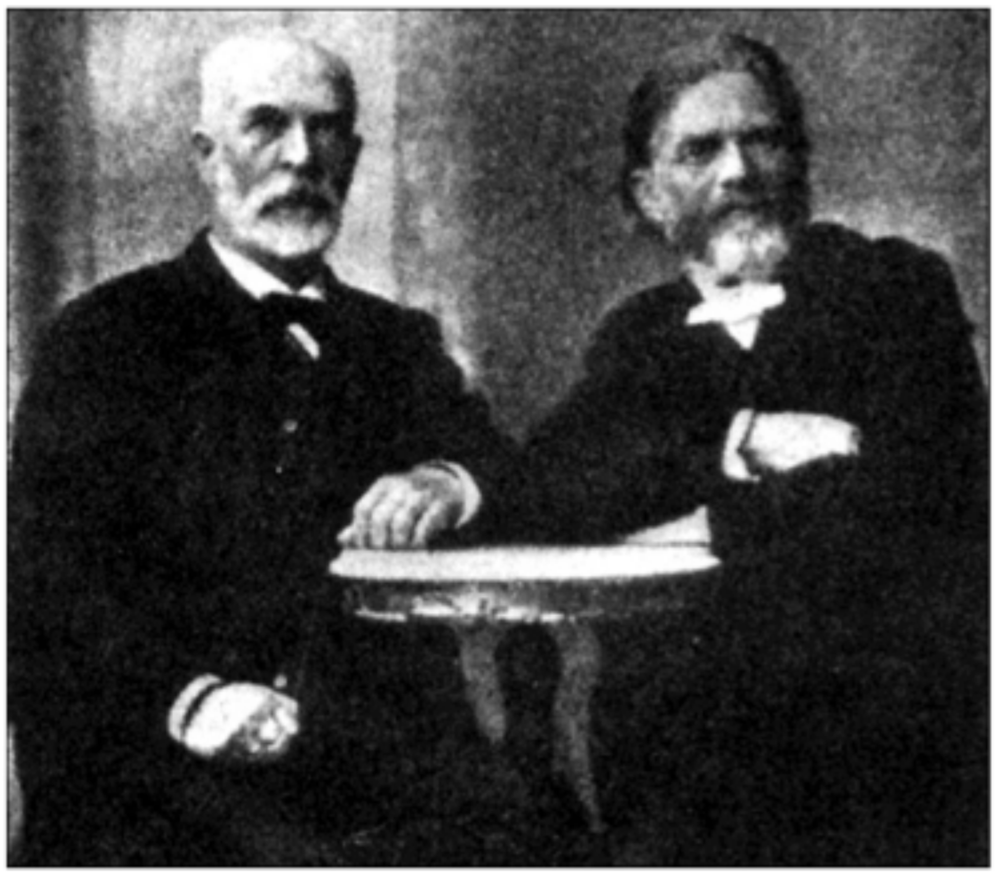
$$x \xrightarrow{k} y$$



$$\boxed{\boxed{\frac{d[x]^{(t)}}{dt}} = -1 \cdot k \cdot \boxed{[x]^{(t)}}}$$

Law of Mass Action

Waage and Guldberg (1864)



activity

rate-constant

stoichiometry

$$v = k \cdot \prod_i [X_i]^{n_i}$$

velocity

When the numbers of catalyst and substrate molecules are usually in the same order of magnitude, using mass action kinetics would make sense, since the reaction depends on those concentrations.

Reversible reaction

$$2x \underset{k2}{\overset{k1}{\rightleftharpoons}} y$$

is equivalent to

$$2x \rightarrow y; v1 = k1 \cdot [x]^2$$

$$y \rightarrow 2x; v2 = k2 \cdot [y]$$

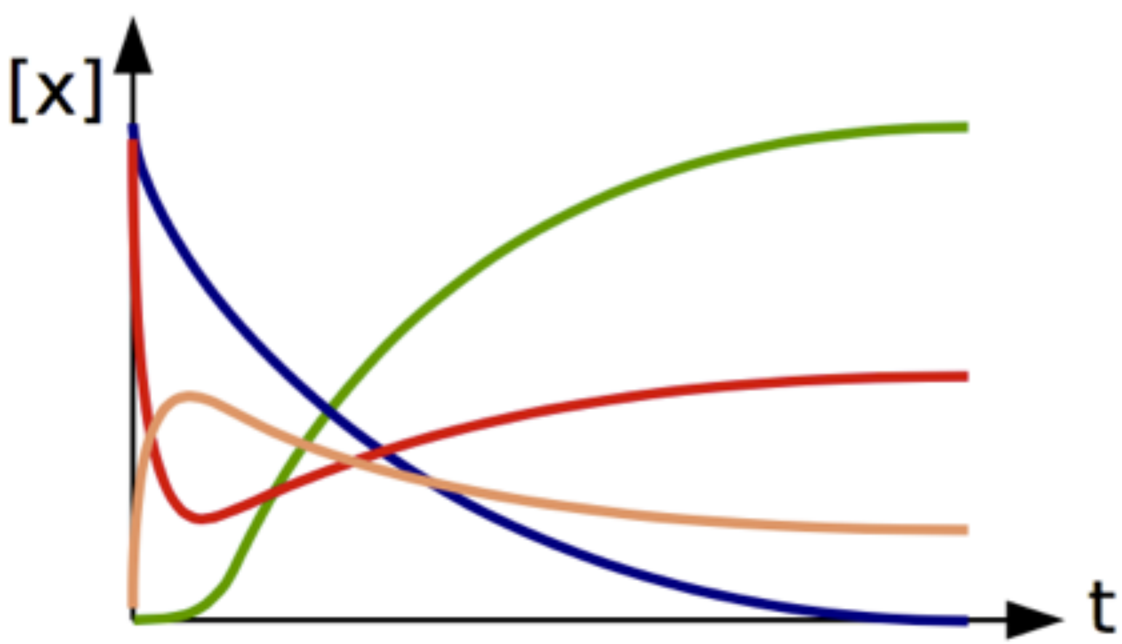
$$\frac{d[x]}{dt} =$$

$$\frac{d[y]}{dt} =$$

Example of an enzymatic reaction

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

$$\begin{aligned}
d[E]/dt &= -k_1[E][S] && +k_2[ES] && +k_3[ES] \\
d[S]/dt &= -k_1[E][S] && +k_2[ES] && \\
d[ES]/dt &= +k_1[E][S] && -k_2[ES] && -k_3[ES] \\
d[P]/dt &= && && +k_3[ES]
\end{aligned}$$



Not feasible in general

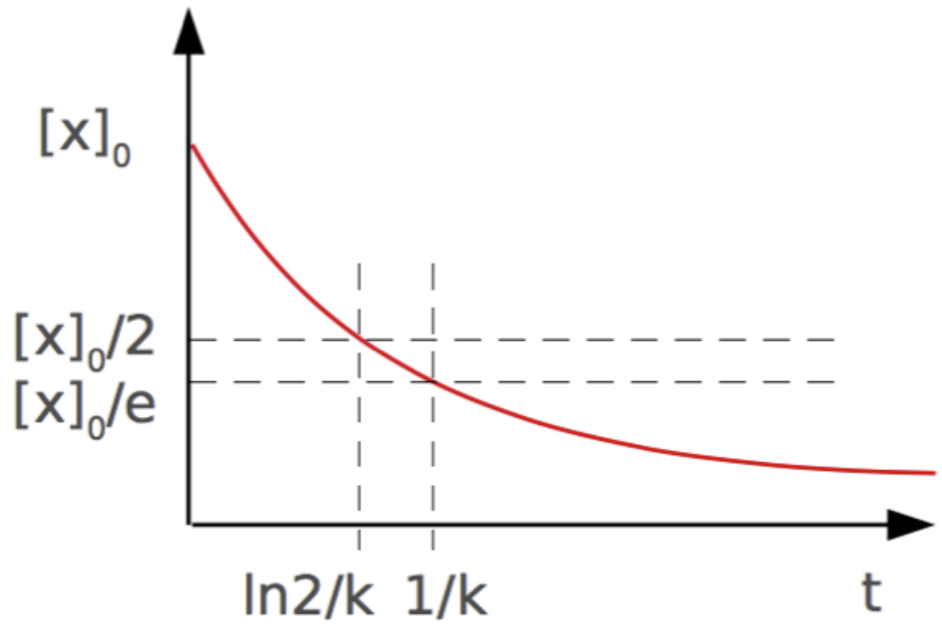
→ Numerical integration

(In biology) the functions represent physical quantities, the derivatives represent their rates of change, and the equation defines a relationship between the two.

Describe

Explain

$$x \xrightarrow{k} y$$



$$\boxed{\boxed{\frac{d[x]\text{(t)}}{dt}} = -1 \cdot k \cdot \boxed{[x]\text{(t)}}}$$

Predict

$$x(t) = \quad ?$$

Analytical (exact) solution

$$x(t) = [x]_0 \cdot e^{-kt}$$

Numerical approximation to solution

Others

Exercises

Exercises:

1. Write an ODE model of each of these systems
   - use mass action
2. Simulate the model in python

Decay: $A \xrightarrow[k1]{} \varnothing$

Equilibrium: $A \xrightarrow[k2f]{} B$ $\quad$ $B \xrightarrow[k2r]{} A$

Enzymatic reaction: $S + E \xrightarrow[k3SE]{} SE$ $\quad$ $SE \xrightarrow[-k3SE]{} S + E$ $\quad$ $SE \xrightarrow[-k3P]{} E + P$

Gene expression (central dogma):

Reactions: $G \xrightarrow[k4G]{} G + M$ $\quad$ $M \xrightarrow[k4M]{} M + P$ $\quad$ $M \xrightarrow[k4Md]{} \varnothing$ $\quad$ $P \xrightarrow[k4Pd]{} \varnothing$

Gene regulation:

Add a feedback in which the Protein inhibits transcription by binding.

Additional reactions: $G + P \xrightarrow[k4if]{} GP$ $\quad$ $GP \xrightarrow[k4ir]{} G + P$ $\quad$ **(GP is inactive)**

Challenges:

Conceive an oscillatory system involving gene expression.

Conceive an oscillatory system involving the phosphorylation of proteins.

Develop a program which takes a list of reactions, and parameters as arguments.

Image processing

exercise 8.2 in pdf

Next steps

- Difference between decimal numbers and floating point numbers.
- **Algorithmic analysis:** How to define efficiency of an algorithm and why does this matter.
- **Performance tuning:** Useful Matlab tools for understanding your code: `mlint, tic/toc, profile on/off/report.`
- Learn about **numerical approaches for searching** (e.g. exhaustive enumeration, bisection search) and **numerical methods for approximations** to equations (e.g. Euler method, Newton method).
- Learn how to write **recursive programs** and why they might be useful. Use examples such as solving the problem of the Tower of Hanoi and calculating the Fibonacci numbers.

Final comments

- For MATLAB, think *ARRAYS*
- Carefully name variables and files; choose names that already explain the program/variable
- Comment! (spend almost as much time as programming).
- Save, save save!
- More tips: http://www.matlabtips.com

Final comments

- **Other apps:** http://uk.mathworks.com/discovery/matlab-apps.html
- **If you don't know how to do something:**
    - Find previous examples (Google) and edit them
    - See the MATLAB documentation and videos
    - See MATLAB examples
      http://www.mathworks.com/examples/matlab
      MATLAB Central
    - For more mathematical questions see/ask in mathsoverflow.net
- **Alternatives:**
    - SciLab
- **Feedback - Please fill in the online form!**