MATHS

Msc Bioinformatics for Health Sciences

Eduardo Eyras
Pompeu Fabra University

2016-2017

http://comprna.upf.edu/courses/Master_MAT/

Program

- **Linear Algebra**

  Vector space
  Linear maps, matrix representaitons
  Projections
  Eigenvalues, eigenvectors and Matrix diagonalization
  Singular Value Decomposition and PCA

- **Optimization**

  Analytical optimization, Partial derivatives, Hessian matrix
  Optimization with constraints, Lagrange multipliers
  Approximate methods
  Support Vector Machines: optimization of hyperplanes

http://comprna.upf.edu/courses/Master_MAT/

Program

- **Linear Algebra**

A bit of mathematic formalism, how to perform simple demonstrations, matrix diagonalization (procedure, what it means), applications

- **Optimization**

Functions of many variables, partial derivatives, how to find maxima/minima of a function of many variables in general and using constraints, heuristics, applications

http://comprna.upf.edu/courses/Master_MAT/

Example

Suppose we divide a population of individuals (plants, animals,) into only two classes: children and adults.

Let c_n denote the number of children at time step n and a_n the number of adults. Let's assume that the population evolves according to the following rules:

$$c_{n+1} = \frac{1}{8} c_n + 6 a_n$$

$$a_{n+1} = \frac{1}{5} c_n$$

We can denote the whole population by a vector:

$$\mathbf{p}_n = \begin{bmatrix} c_n \\ a_n \end{bmatrix}$$

Example

So you can re-write the population model in matrix annotation:

$$c_{n+1} = \frac{1}{8}c_n + 6a_n$$

$$a_{n+1} = \frac{1}{5}c_n$$

$$\mathbf{p}_n = \begin{bmatrix} c_n \\ a_n \end{bmatrix}$$

$$\mathbf{p}_{n+1} = \begin{bmatrix} \frac{1}{8} & 6 \\ \frac{1}{5} & 0 \end{bmatrix} \mathbf{p}_n$$

Example

$$\mathbf{p}_{n+1} = \begin{bmatrix} \frac{1}{8} & 6 \\ \frac{1}{5} & 0 \end{bmatrix} \mathbf{p}_n$$

Questions one can answer:

Given the distribution at time step n=3, find the stage distribution at a later time step n=4.

Given the stage distribution at n =4, find the distribution at n=3.

Find a general formula to express population change during Δs steps

Is the population stable over time? I.e. after many time steps, is the distribution reaching a stationary point?

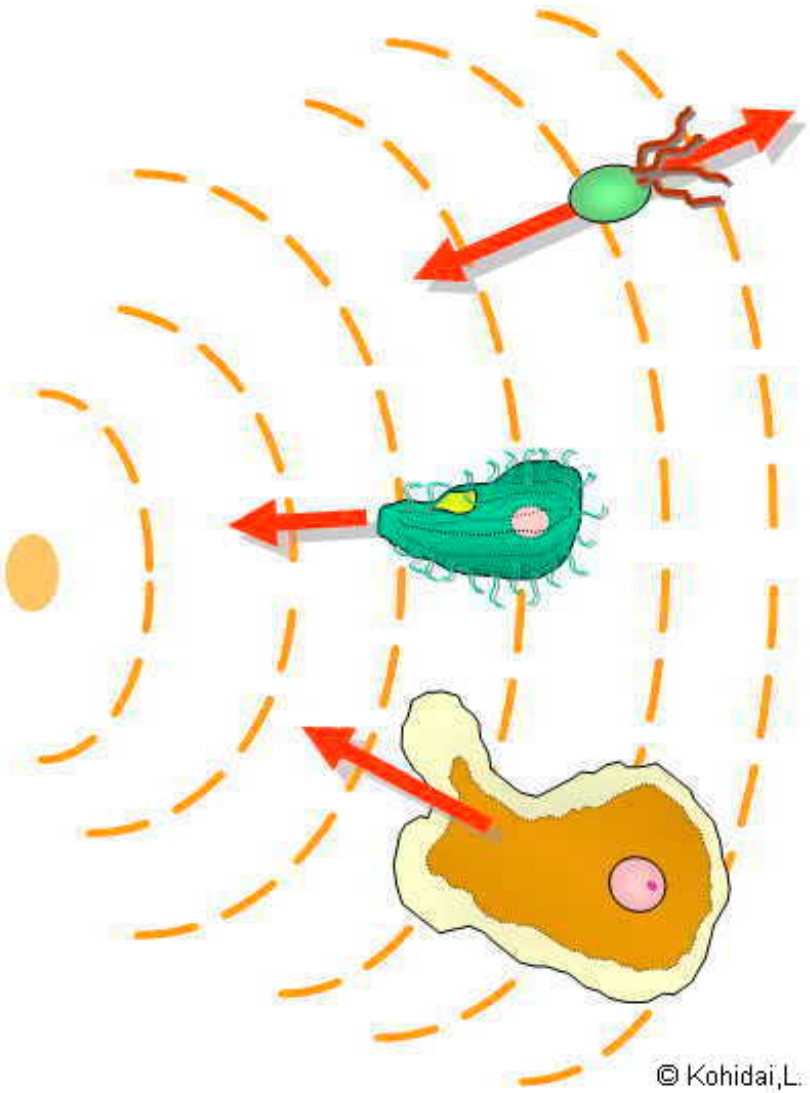
Example

Imagine an organism that moves on a surface (2D).

It might use local cues to orient itself and move towards sites that have higher nutrient levels.

This type of motion is called chemotaxis.

One could represent the distribution of nutrients in the plane with a function of two variables *f(x,y)*



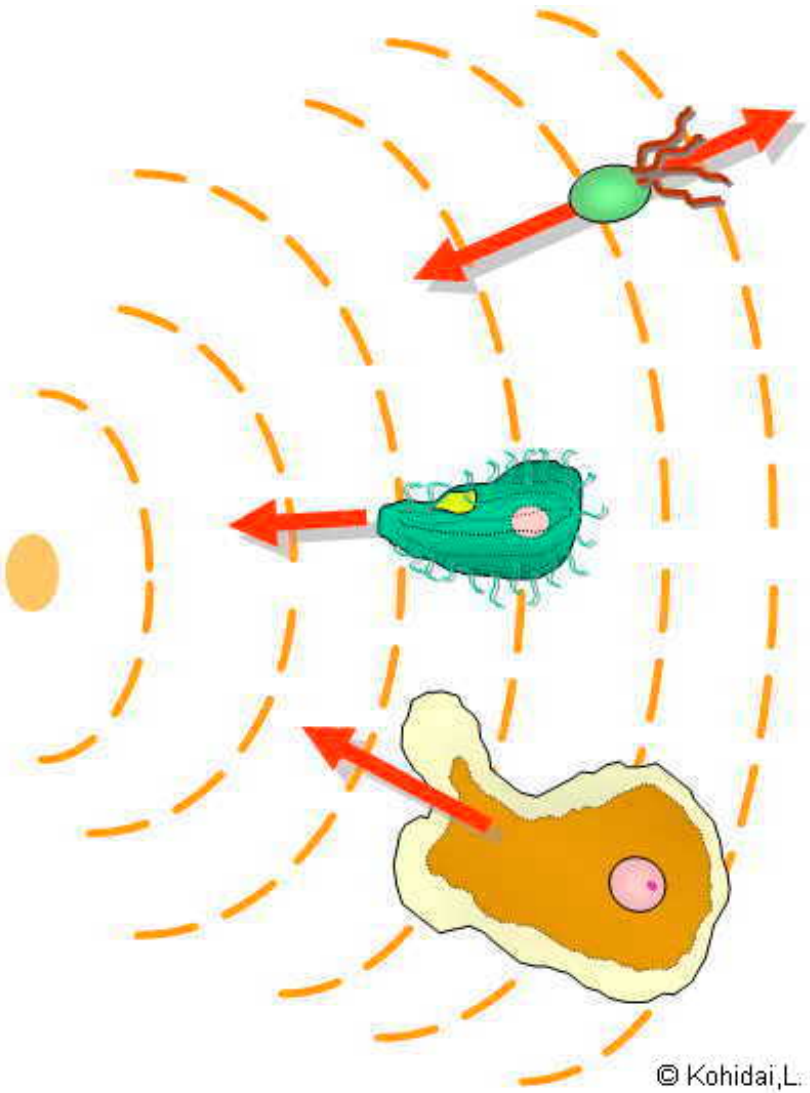
Example

Questions on can answer:

Can we find points/lanes/regions of maximum/minimum nutrient content?

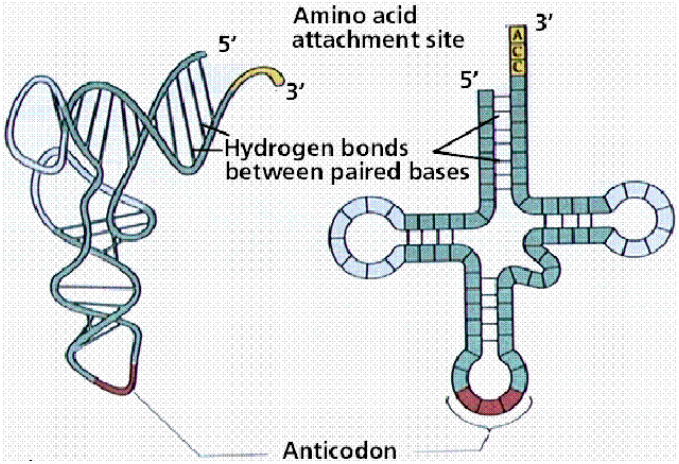
Are these local? Or global?

A highly efficient chemotactic organism would be orthogonal to the level lines of the nutrient distribution. That is: lines of equal concentration. Can we find the efficient trajectories?



Example

Structural conformations of molecules

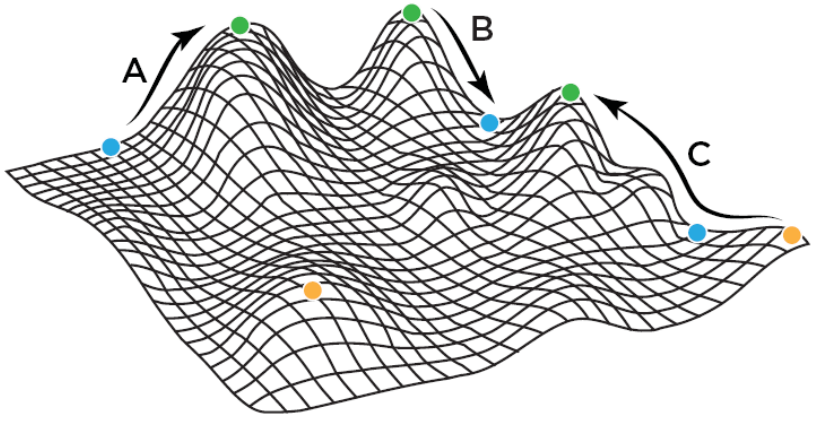


Studied in terms of stability of folding energies. E.g. energy functions with constraints

What is the landscape of stable structures?

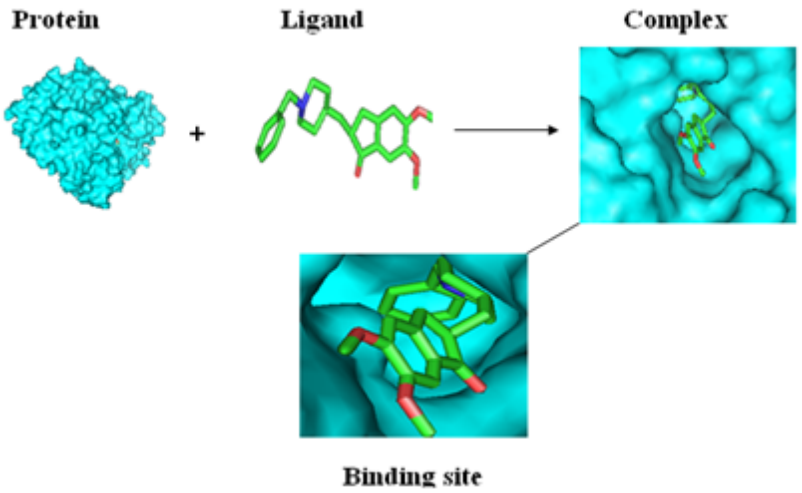
Are there multiple possible conformations?

How does the stability varies with e.g. point mutations?

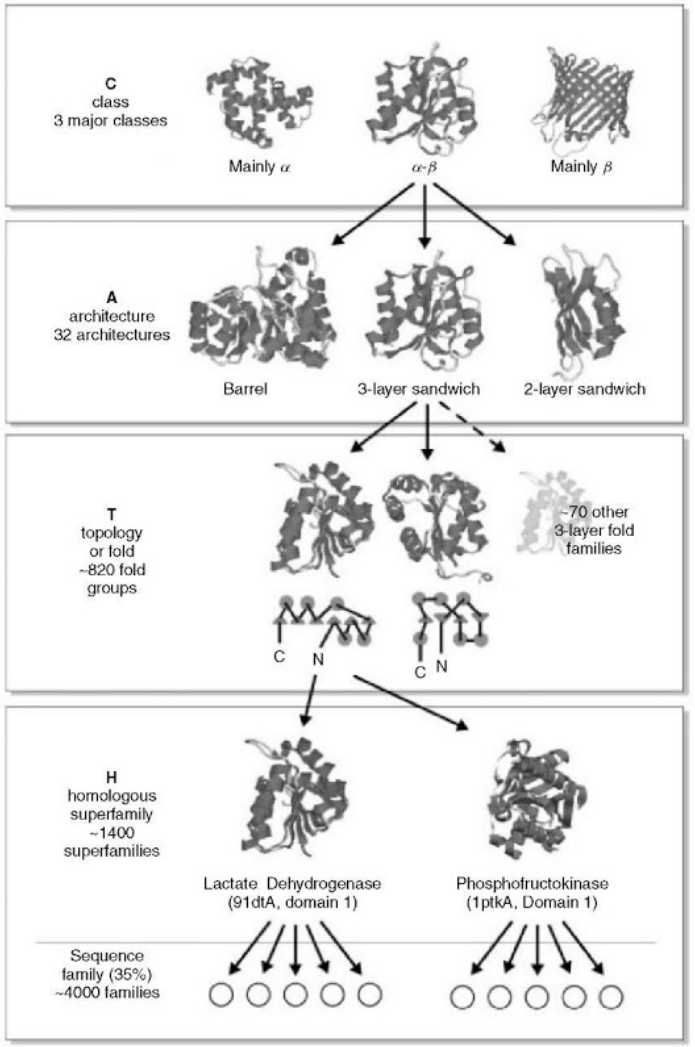


Example

Molecular docking, comparison of structures, rotation and translation of coordinates in space



How do I compare two molecular structures using coordinates in space?

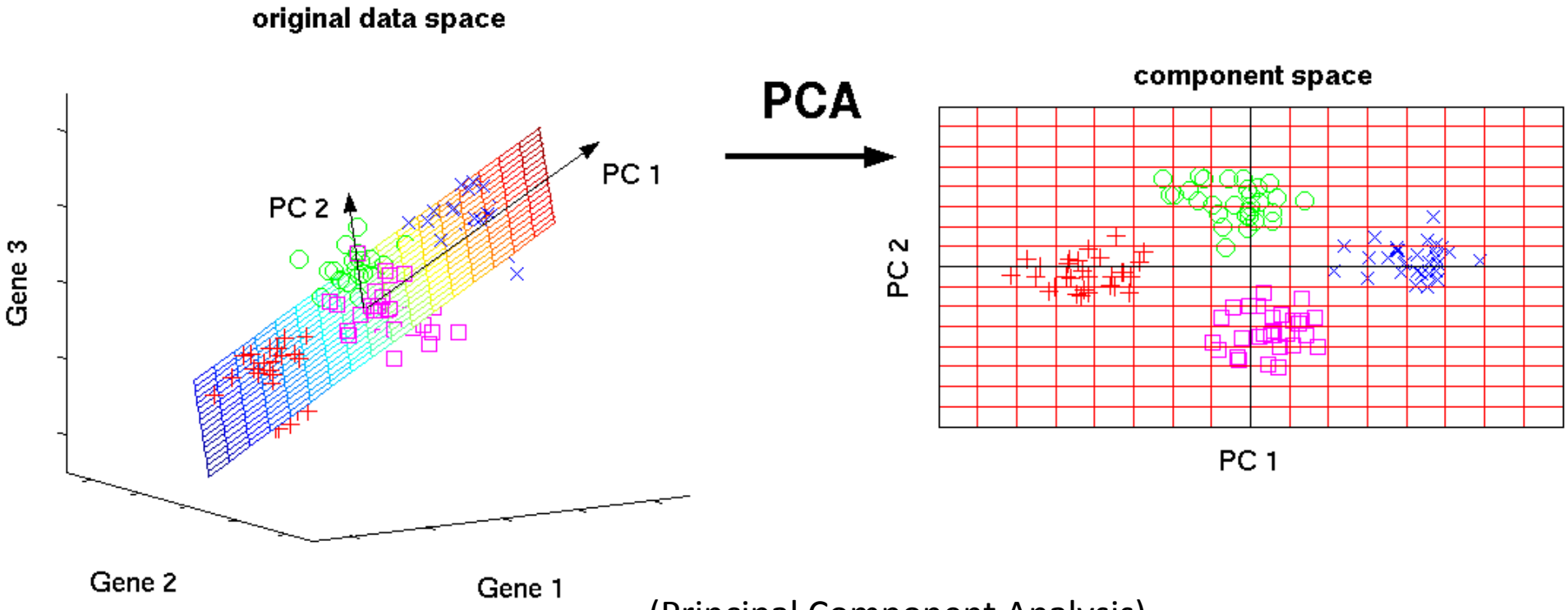


Example

Data analysis.

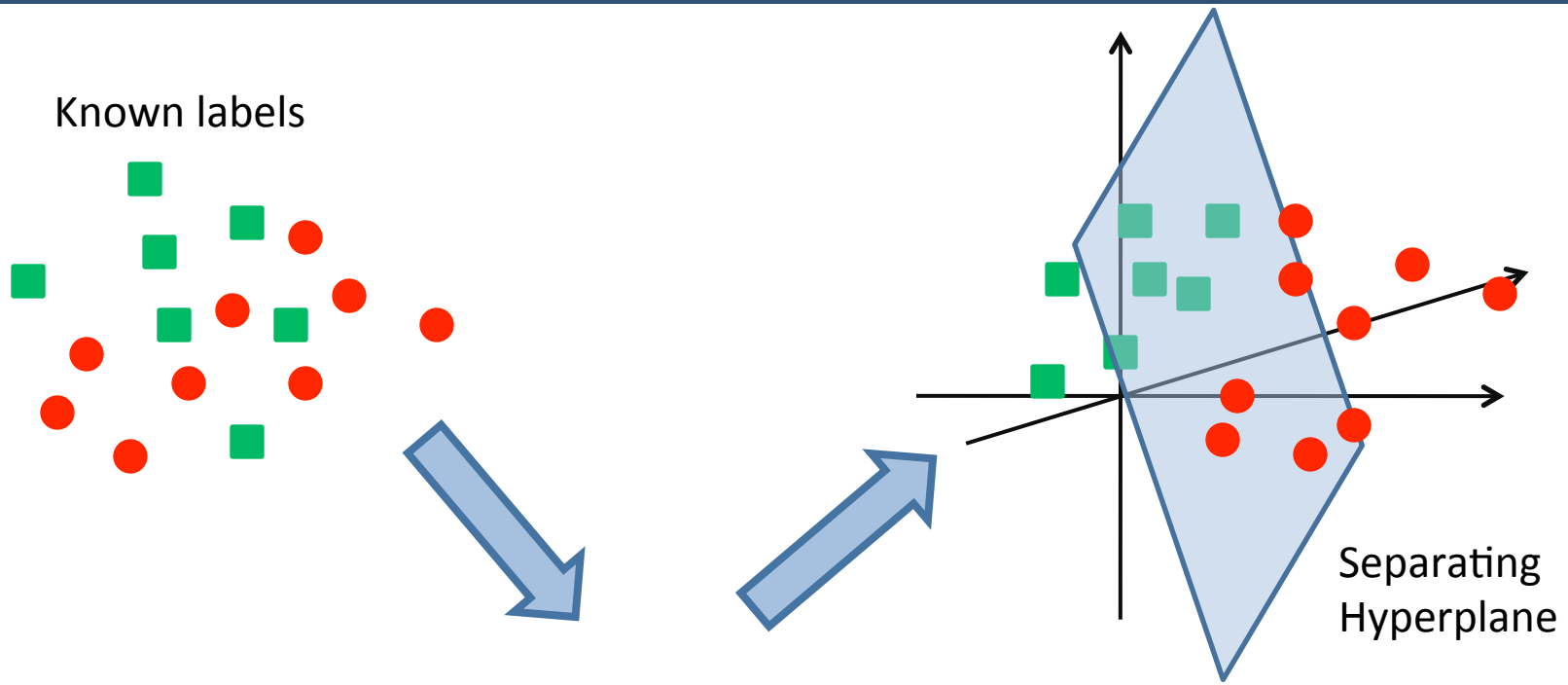
For instance: samples = vectors in a parameter space (genes, experiments, properties)

Can I find a different parameter space that highlights better differences and commonalities between samples?



(Principal Component Analysis)

What are Support Vector Machines (SVM)?

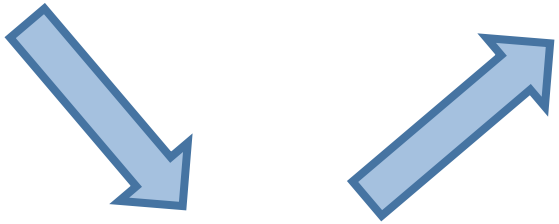


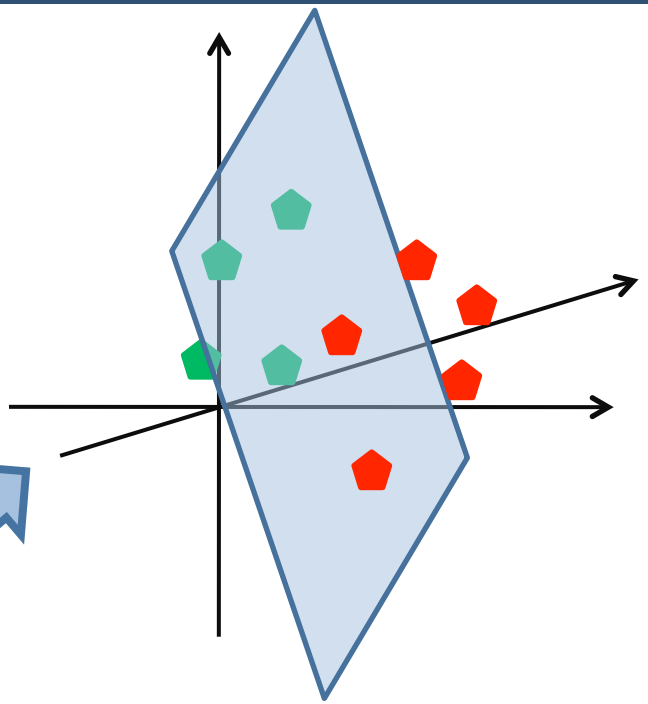
	feature_1	feature_2	feature_3	…	feature_n	
example1	0.4552	6.0000	-0.3657		4.3545	●
example2	-0.2414	3.2027	0.0456		-1.3450	■
example3	0.1256	-1.4539	0.0012		3.2456	■
…						

Combine the features and project the data in a n-dimensional space

What are Support Vector Machines (SVM)?

New examples





	feature_1	feature_2	feature_3	…	feature_n
New example1	0.3451	-3.5431	-0.4566		1.3456
New example2	0.0487	2.5644	0.2100		-2.0567
New example3	-1.1209	-1.4492	0.0024		1.5687
…					

Combine the features and project the data in a hyperplane to classify the data

Classes:

Theory and problems

Evaluation:

Homework from each part
Final exam

http://comprna.upf.edu/courses/Master_MAT/