Connectionism 6

Dimension Reduction
Kohonen Maps
Emergence

Dealing with High Dimensional Data

An input pattern may be an image, a sound, or an arbitrarily large vector

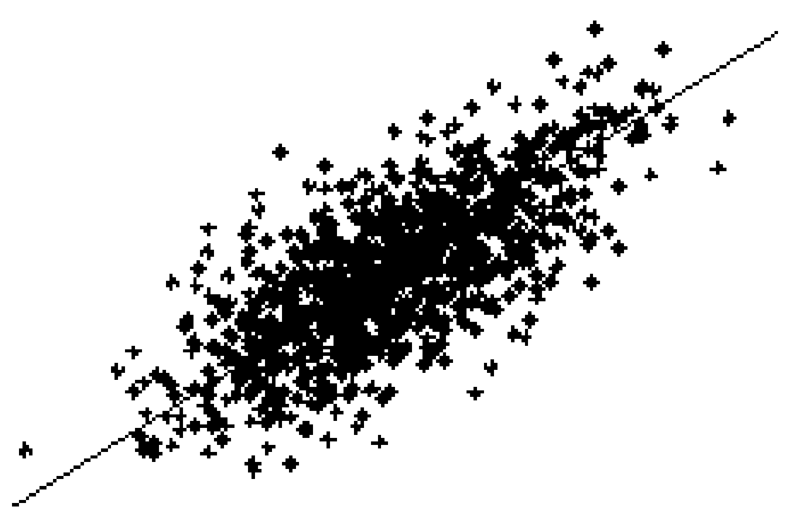
If we use a vector of length n to describe anything (input, hidden representation, output), we have a point in n-dimensional space.

Visualizing such spaces is hard. Thinking about them can warp your brain (at first).

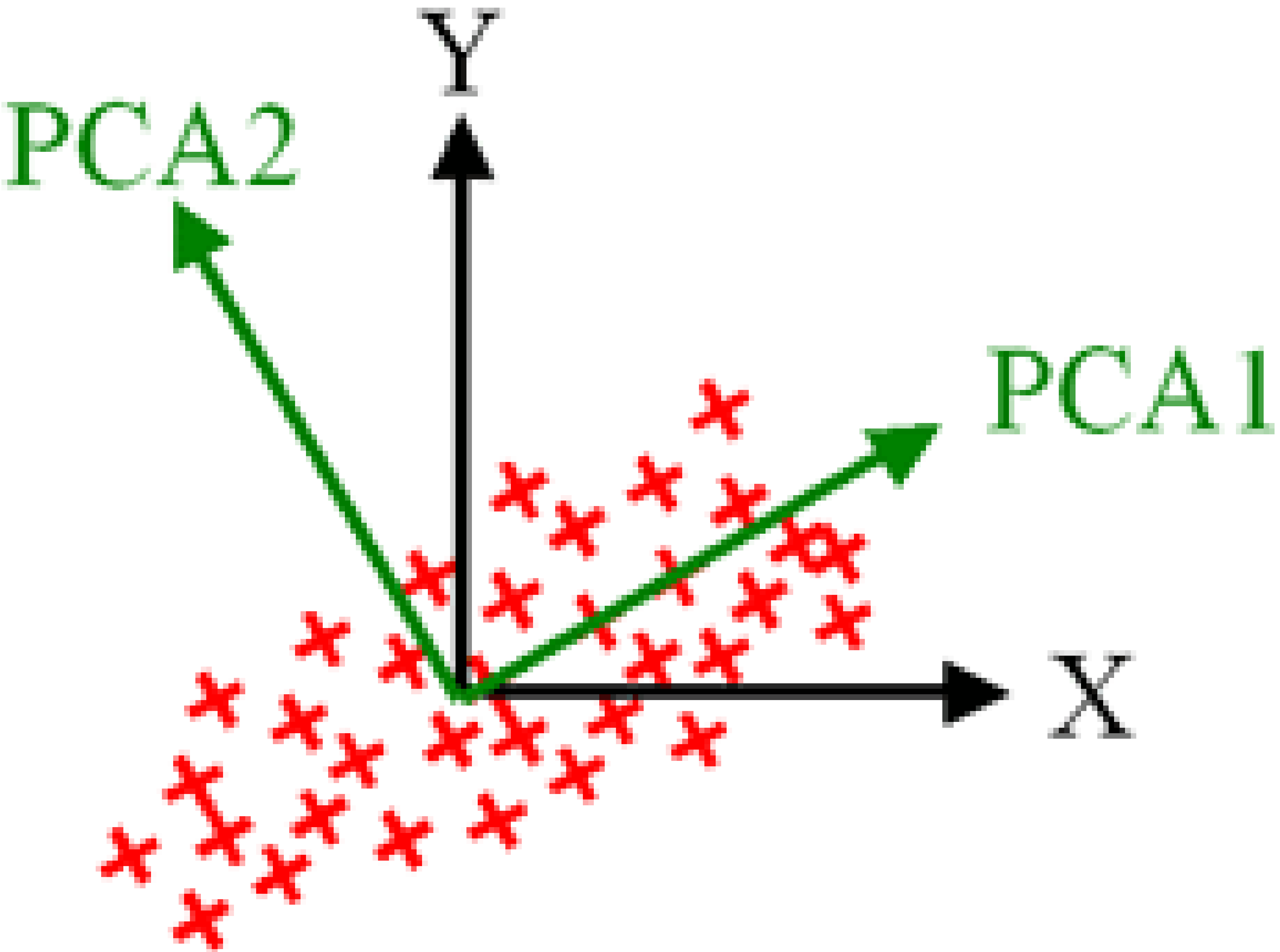
Mathematicians find this very easy (infinite dimensional spaces anyone?)

Our autoassociator showed us one way of reducing the dimensionality of a set of patterns.

A commonly used statistical/mathematical technique is to do Principal Components Analysis (PCA).



2-d set, showing principal axis of variation



A principal component decomposition

Need slide showing projection

Informal account of PCA

[1] Center the data, and preferably scale it
(= convert to standard scores)

[2] Rotate the coordinate axes so that axis 1 captures as much variation as possible

[3] Rotate remaining axes until each successive axis captures as much variation as possible

[4] Repeat

Excellent tutorial on PCA:
http://www.cs.mcgill.ca/~sqrt/dimr/dimreduction.html (may have java issues)

Good overview in text here:
http://georgemdallas.wordpress.com/2013/10/30/principal-component-analysis-4-dummies-eigenvectors-eigenvalues-and-dimension-reduction/

PCA and our AUTOASSICIATOR

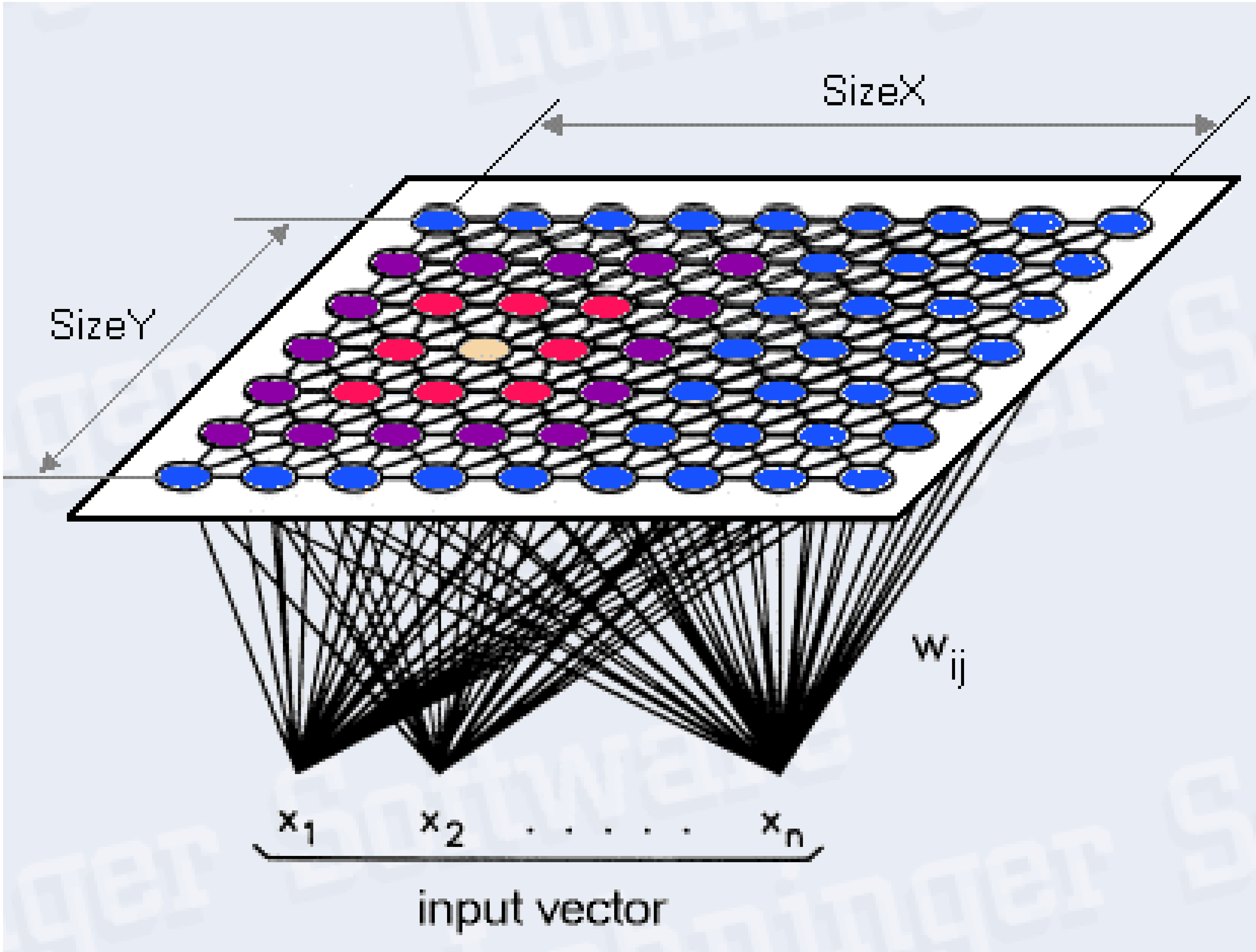
Q: Isn't PCA what our autoassociator is doing?

A: Close. It is doing a non-linear transformation. One informal way to think of this is to allow 'bendy' axes

PCA is quick, formally clean, and widely used. We will use it later to analyze hidden unit representations

Kohonen's Self Organizing Maps

Another dimension-reducing technique
Also non-linear
Powerful, simple, and biologically not implausible


SizeX
SizeY
w_{ij}
x_1 x_2 . . . . . x_n
input vector

The goal of a Kohonen network is to map input vectors (patterns) of some arbitrary dimension *N* onto a discrete map with 1 or 2 dimensions.

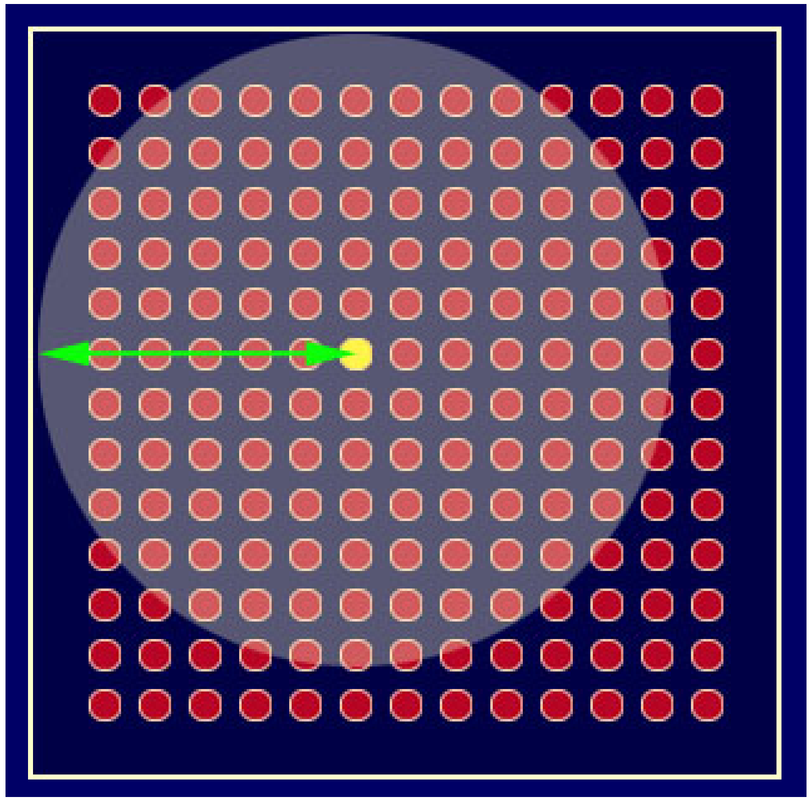
Patterns close to one another in the input space should be close to one another in the map. i.e. they should be topologically ordered.

A Kohonen network is composed of a grid of output units and *N* input units. The input pattern is fed to each output unit.

The input lines to each output unit are weighted. These weights are initialized to small random numbers.

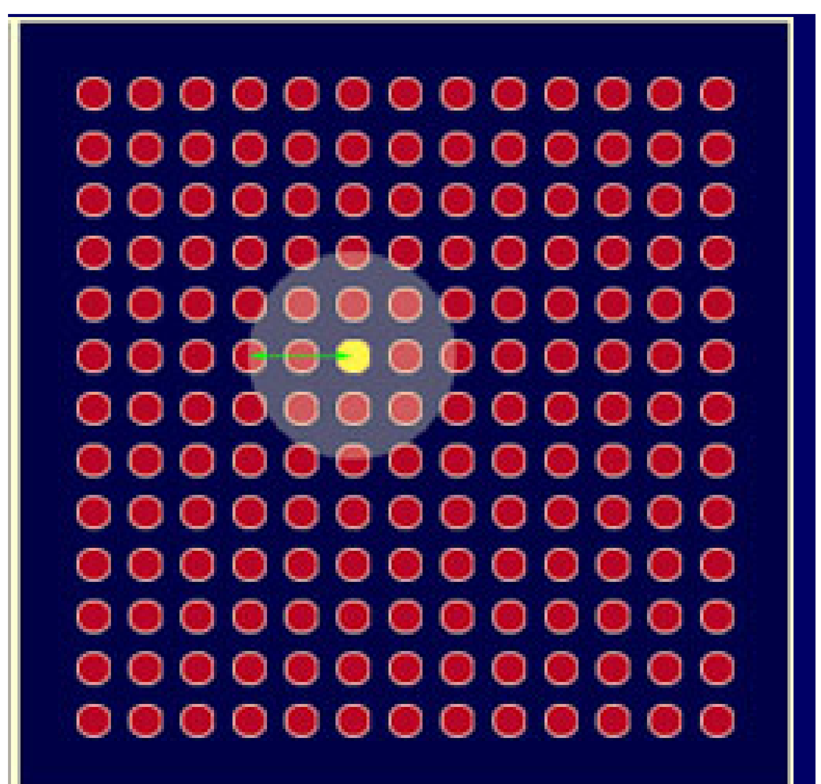
Training Algorithm

- Initialize all weights randomly
- Loop until Δw is negligible:
  - for each input pattern
    - present pattern
    - find ‘winning unit’ + its neighbours
    - update wts to these units
  - Reduce neighborhood size



Neighborhood of a best matching unit

...decreases over time



Demo

http://www.cis.hut.fi/research/javasomdemo/index.html

http://www.eee.metu.edu.tr/~alatan/Courses/Demo/Kohonen.htm

http://fbim.fh-regensburg.de/~saj39122/jfroehl/diplom/e-sample-applet.html

Before the next class:

Read Deb Roy et al, 2009

Read Schafer and Mareschal (2001)

Finish textbook, Chpt 3!

Topics

- Innateness and modularity
- Emergent properties
- General learning principles and the emergence of domain specific representations
- Face recognition in infants
- Vocabulary development
- Linguistic environment, Deb Roy
- Learning the past tense

The Modularity of Mind

- Jerry Fodor (1983) "The Modularity of Mind: An Essay on Faculty Psychology" MIT Press
- The incredible abilities of the neonate:
  - tell speech from other sounds (day 0)
  - tell mother tongue from unrelated languages (day 4)
  - surprise if 2 objects occupy the same location (3 months)
  - surprise if one solid passes through another (4 months)
  - show naive understanding of gravity and mass (7 months)
  - learn languages really really quickly, despite impoverished input

Development and Learning

- What is given, and what acquired?
- Elman's claim: "innate predispositions...channel...attention to certain aspects of the environment over others"
- Little role for inbuilt representations
- representations emerge from brain—environment and brain—brain interactions
- Why do humans (as opposed to ducks, donkeys, etc) have such a long period of learning and maturation?
- trade-off between plasticity and (minimal) predispositions to allow optimal adaptation

Elman's hypothesis

...we show how domain-specific representations can emerge from domain-general architectures and learning algorithms

...and how these can ultimately result in a process of modularization as the **end product** of development, rather than its starting point

Where does structure come from?

- It's in the genes??????
- It's imposed, or found??????
- The developing system, and the environment in which it develops together bring forth the structure
- .....emergence....

How is it possible to talk of knowledge and structure, as if they were interchangeable?

Emergence

Properties that emerge at a system level, that are not derivable from the individual components.

Elman is *very* loose in his use of the term. Beware.

Examples of Emergent Phenomena

- Soap bubbles sphericity
  - minimization of surface tension
  - general principle of energetic minimization
- Hexagonal bee cell design
  - maximizes spatial exploitation
  - general principle of economy
- Simple organizational principles, under appropriate environmental conditions, generate structure and order

The following three descriptions come from the excellent book/website found at http://www.patternsinnature.org/index.html

Each water molecule in a body of water interrelates with neighboring molecules by way of by various intermolecular forces that results in the molecules being attracted to each other. In the body of a liquid, each molecule is pulled equally in all directions by neighboring molecules, resulting in a net force of zero. But, at the surface of the liquid, the molecules are pulled inwards by other molecules deeper inside the liquid because the surface molecules are not attracted as intensely by the molecules of air (or another liquid) that surround the water drop. This net difference in attractive forces is called surface tension.

This surface tension at a water drop's surface results in the formation of the smallest area for the given volume of whatever is contained. The resulting shape is a sphere.

A bubble is sustained as a result of the balance between the internal gas pressure of the bubble and the forces associated with surface tension. The internal pressure that tends to push the sphere apart is counteracted by the surface tension acting around the surface of the sphere to provide a net force that pushes inward.

Using their bodies as a template, bees make each close packed wax cell cylindrical, like a tube. Like glass, beeswax becomes increasingly fluid as it is heated. The mobility of one wax particle with respect to another changes significantly at specific transition temperatures. The bees raise the temperature of the wax to 37-40 degrees Centigrade permitting the wax to have more malleability . At this temperature the cells take on their hexagonal shape due to their compression by the six closest neighbors in the packed arrangement of cells.

The entire process of creating a honeycomb pattern in a hive, therefore, comes from a combination of physical laws and bee behavior that takes advantage of these laws. The pattern of bee behavior is just as important as the pattern of physical laws that is employed to create the comb structure

Often, some physical quantity is being minimized: surface area, energy expenditure, oxygen consumption, etc....

The quantity being minimized is a function of the system as a whole.

- Rayleigh-Bernard convection: discontinuity and phase transition

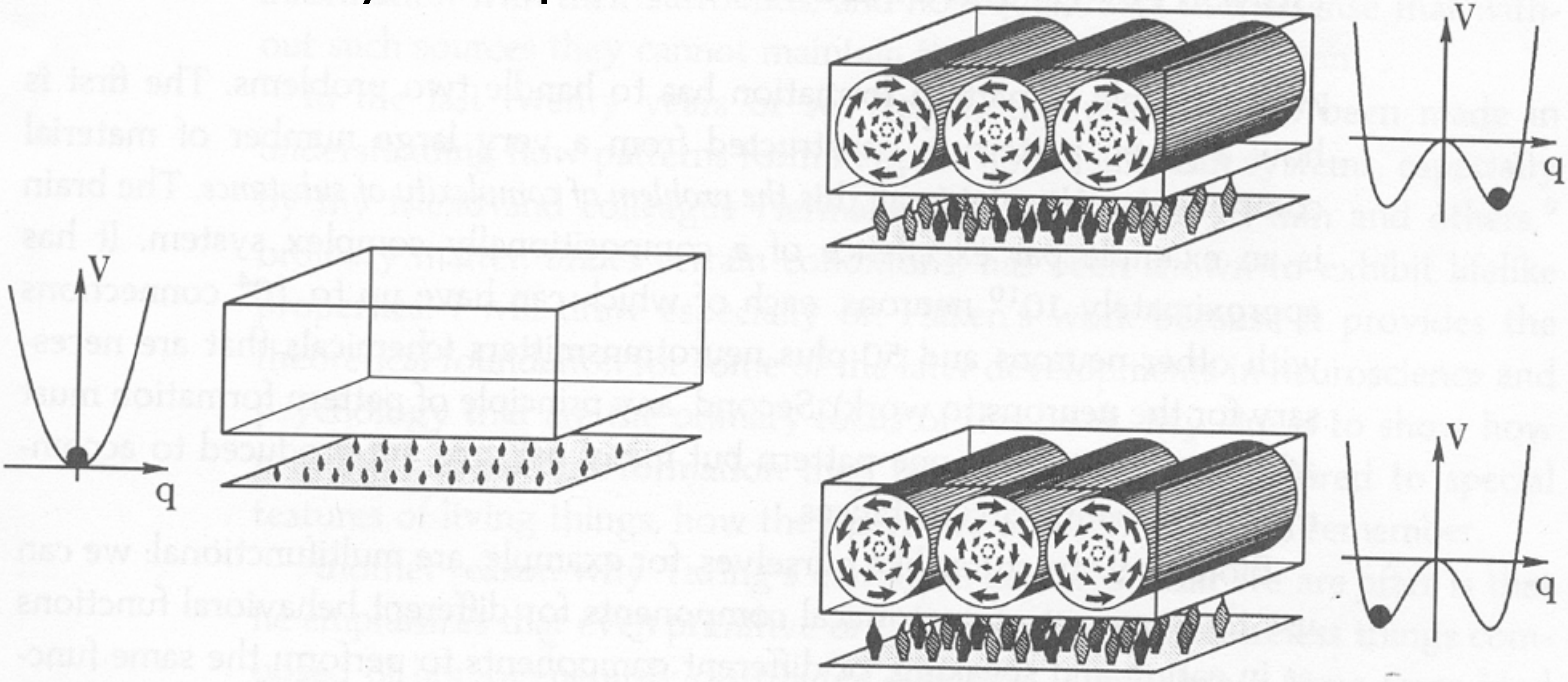


Figure 1.1 (*Left*). A layer of liquid heated weakly from below displays no macromotion. The liquid is in a rest state, as indicated by the ball resting in the minimum of the potential well. (*Right*). At a critical value of the temperature gradient the liquid displays macrosopic rolling motions. The ball can assume one of two possibilities, rolls rotating in one direction or the other.

Minimization of metabolic cost, indexed by oxygen consumption

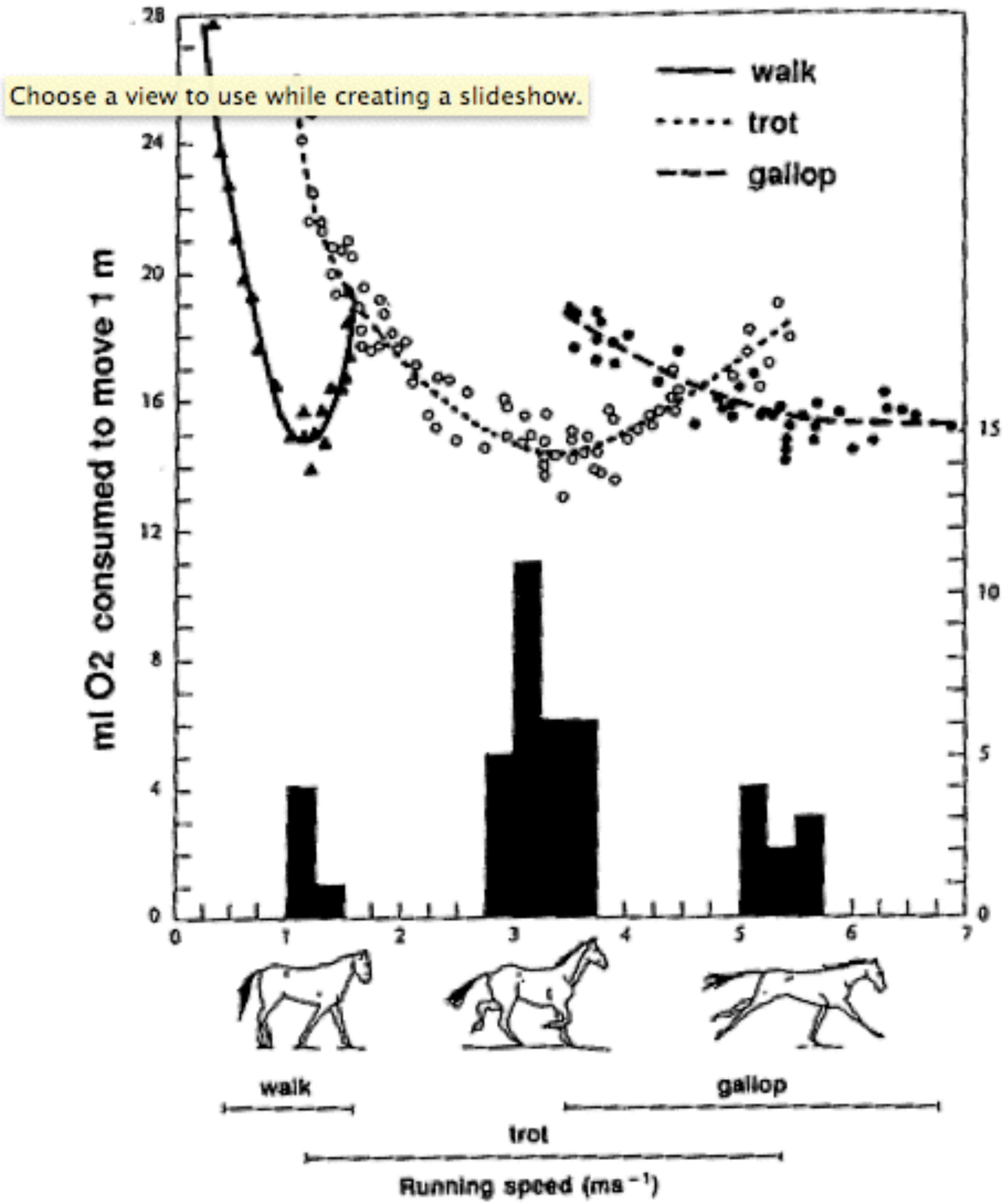


Figure 3.1 Oxygen consumption per meter moved and preferred speed (histograms) of walk, trot, and gallop of ponies. (Adapted from reference 4. Reprinted with permission)

Hoyt and Taylor, Nature, 1981