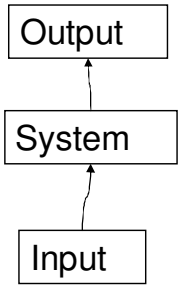
Unsupervised Learning

Unüberwachtes Learnen ist ein im Moment in der Entwicklung begriffenes neues Paradigma. Die Resultate des unüberwachten lernens sind zwar noch nicht Standard und häufig noch recht schwer zu benutzen. Trotzdem ist das unüberwachte Lernen eine grosse Hoffnung für die Zukunft der Lerntheorien und der Neurowissenschaften.

Prinzip:



Wir betrachten ein input output system. Es gibt keinen Lehrer. Die Aufgabe des Systems ist, sich selbst so einzustellen, dass sein output “gut” oder “interessant”. Ist Die definition von “interessant” ist allerdings höchst problematisch und in jeder Studien unterschiedlich

Rekonstruktion des Inputs

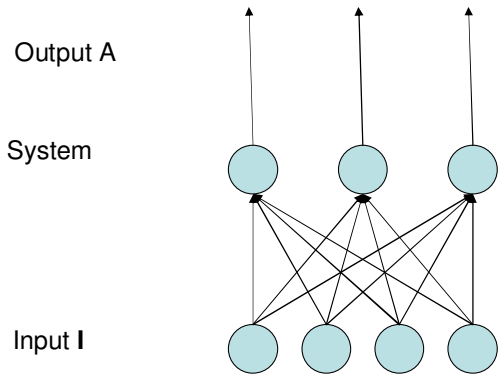
$$\Psi_{error} = \left\langle I_{ist} - I_{reconstructed} \right\rangle_{alle\ daten}$$

$$\Psi_{error} = \left\langle \left(\mathbf{I_{ist}} - \sum_i A_i \mathbf{W}_i \right)^2 \right\rangle_{alle\ daten}$$

Interesting = informative about input

Eine klassische definition von “interessant” oder eben eine klassische wahl einer Zielfunktion ist den Rekonstruktionsfehler des Inputs gegeben den output zu betrachten.

N Linear Neurons



$$A_i = \mathbf{IW}_i$$

Remark: I and W sit in the same space

Wir betrachten jetzt ein lineares System mit N neuronen

1 out of N coding

$$A_i = \mathbf{I}\mathbf{W}_i$$

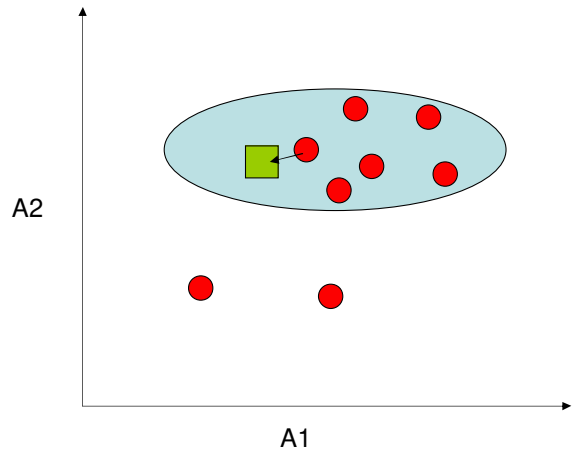
Aber alle A sind auf null gesetzt ausser dem maximalen A

Gradient abstieg auf Ψ_{error} führt zum learning vector quantization

Optimal codebook

Angenommen nur ein Neuron darf codieren

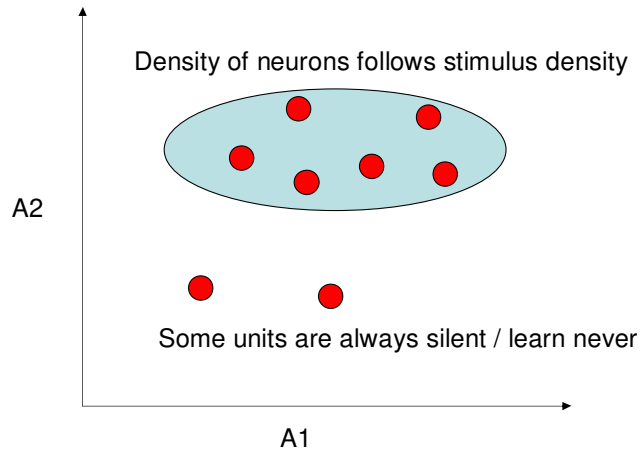
Learning dynamics



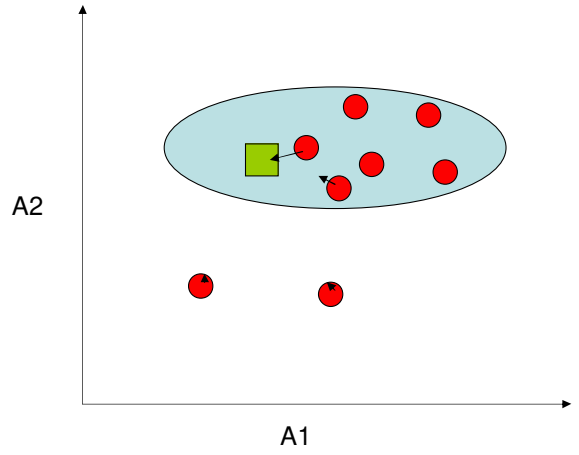
Density of neurons follows stimulus density

Was passiert wenn wir jetzt gradientenabstieg machen

Tote Einheiten

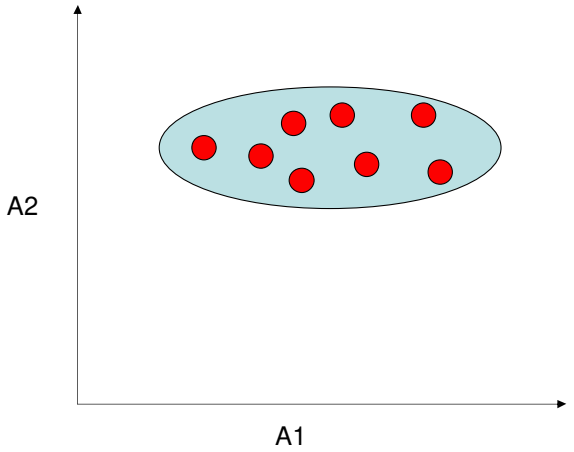


Soft competition



Effectively just smoothing the error surface

Effect



Better density estimation but some bias towards center

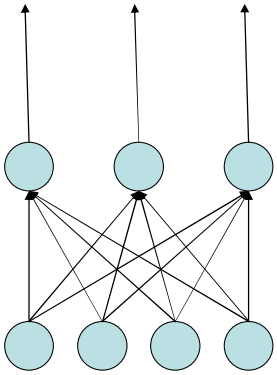
Ψ_{error} Is only approximately optimised

Neurons don't do 1 out of n

Output A

System

Input **I**

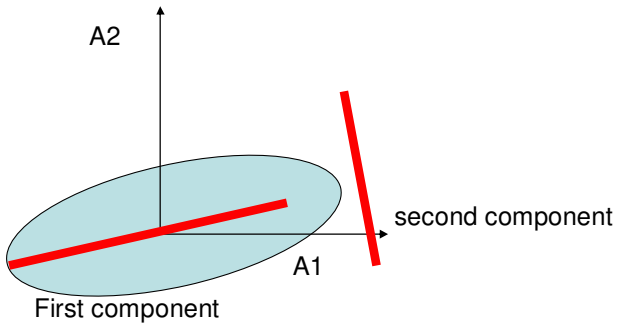


$$A_i = \mathbf{I}\mathbf{W}_i$$

Assume neurons are linear

Principal component analysis

Optimize weights to maximize Ψ_{error}



Alternatives

- Estimate n in an ordered way
- Estimate n together

The same mechanism in a different framework

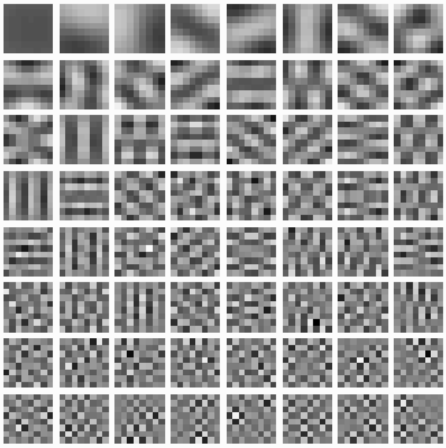
- Constrain all neurons to be uncorrelated
- Constrain all weight vectors to have unitary length
- Maximize the neurons variances

Neurons can do that

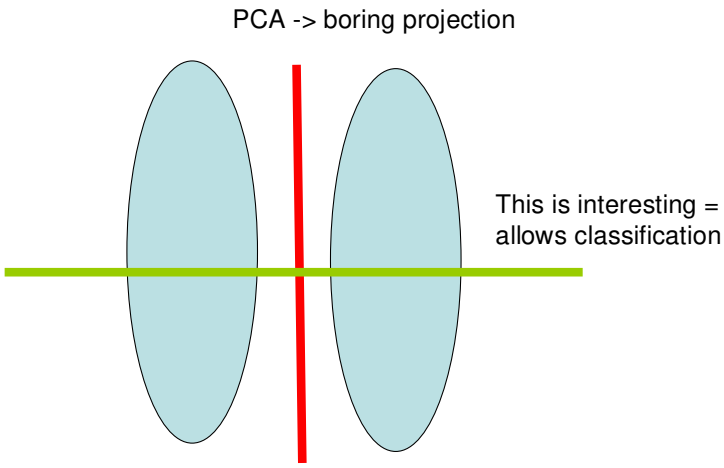
$$\Delta W_{ij} = A_{i(pre)} A_{j(post)} - \alpha W_{ij}$$

= Hebbian Learning !!!

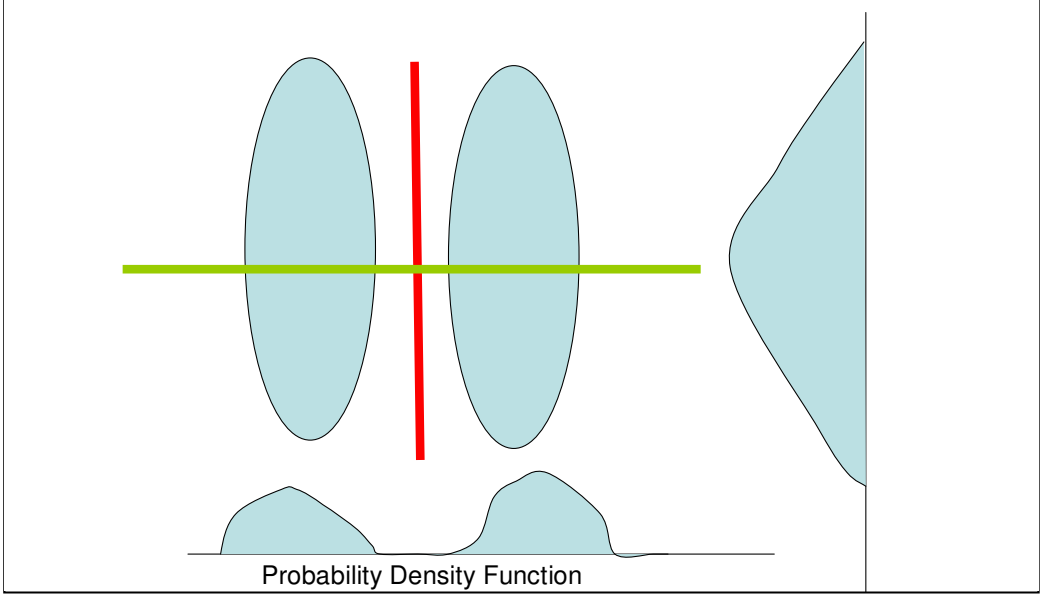
Principal components of natural pictures



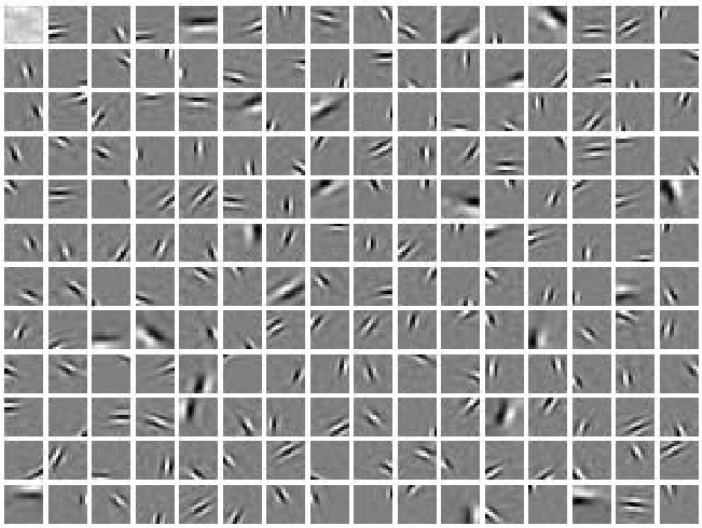
Useless for some problems



Non-gaussianality is interesting (ICA)

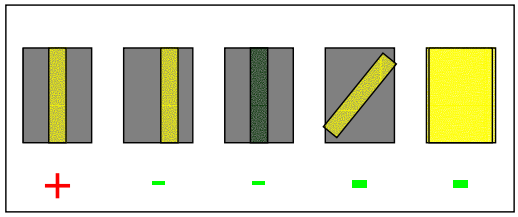


Searching for patterns with supergaussian pdfs



Simple cells

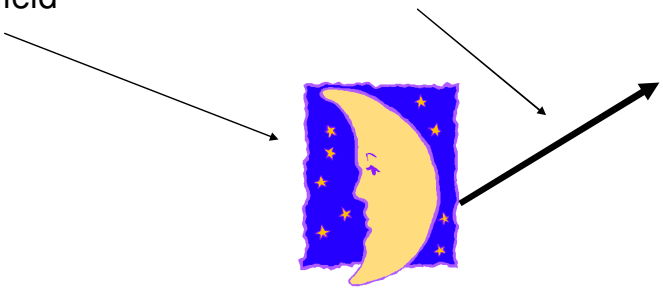
"simple"



Knowing something about the source behavior

Objects rarely enter or leave the visual field

Positions change slowly (if at all)



Leads to stable coding or slow feature analysis

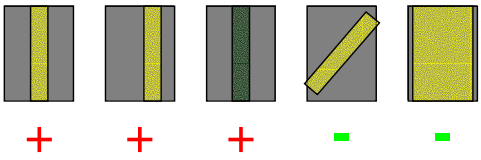
Objective Functions for stable sources

$$\psi_{stable} = -\sum_i \frac{\left\langle \left(s_{i,t} - s_{i,t-\Delta t} \right)^2 \right\rangle}{\mathrm{var}\left(s_{i,t} \right)}$$

$$\psi_{temporal\ coherence} = \sum_i \frac{\left\langle s_{i,t}^{\ 2} s_{i,t-\Delta t}^{\ 2} \right\rangle}{\mathrm{var}^2\left(s_{i,t} \right)}$$

Yields Complex cells when applied to nonlinear cells

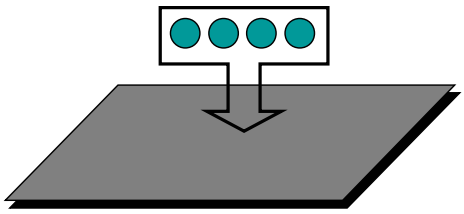
“complex”



Hubel and Wiesel 1962

Complexe Zelle ignorieren die Kontrastpolarität und in gewissem Rahmen die Position eines Stimulus. Sie sind spezifisch für die Richtung eines Balkens und für seine Breite.

Generative Process – alternative interpretation

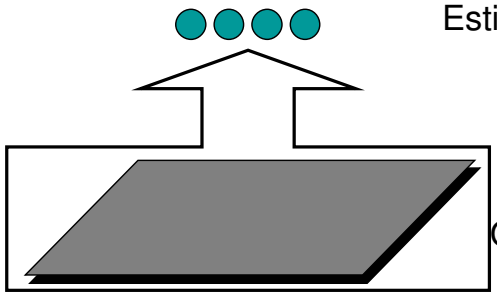


Sources s

Observed picture I

Per definition $I(x, y) = F(s_i, W)$

Regression-
Inverting a generative process



Estimated Sources s

Observed picture I

e.g. Linear Generator

with

$$s = I * V$$

$$V = W^{-1}$$