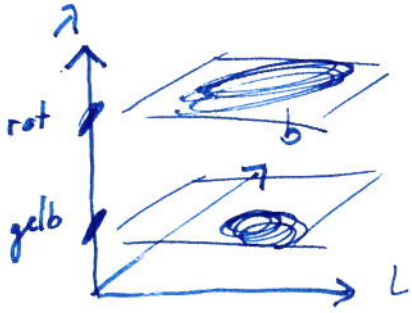
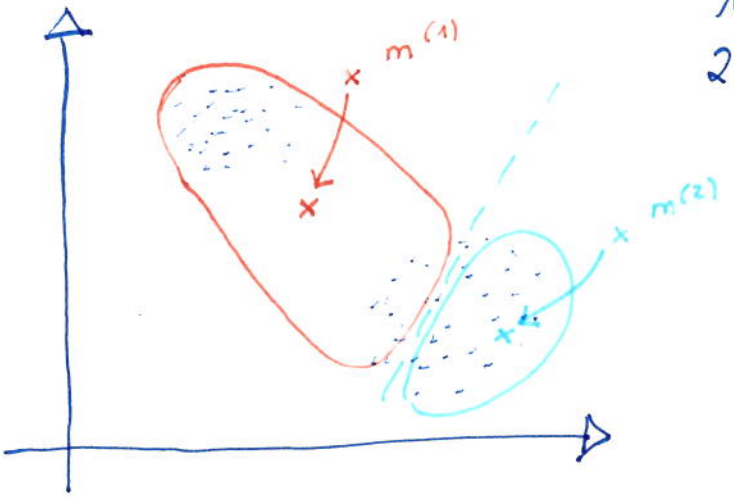
Maschinenlernen (mit MATLAB ...) 7.10.2014

Bananen: „gelb", „krumm/lang"

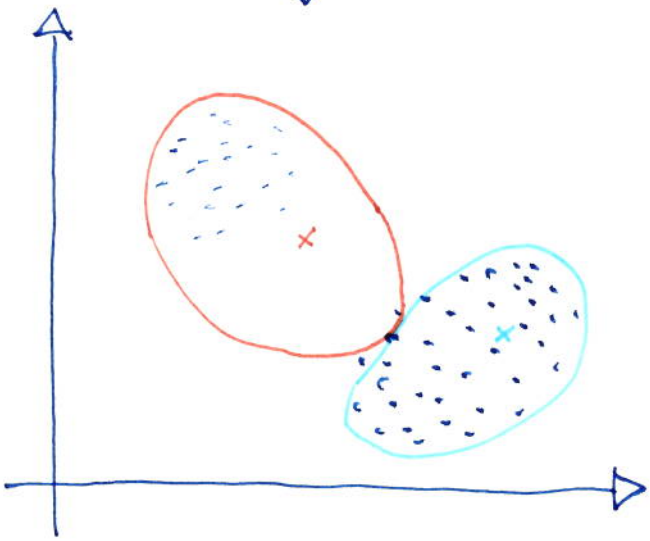
Äpfel: „rot", „rund"



Beispiel



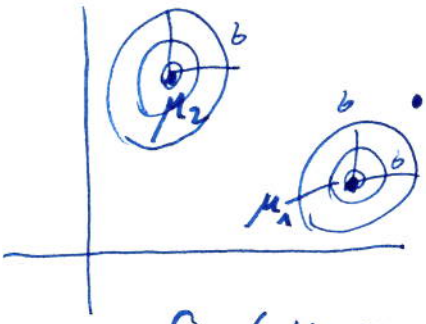
1. Wähle # Cluster K
2. Wähle zufällige Cluster-Zentren
3. repeat
   → „assignment": „Zuständigkeit"
   → „update"



(Cluster verschieben sich in jedem Schritt)

Programm versagt bei best. Beispielen zB  usw.

(es sucht weis runde Cluster)



$$f_{x|\theta} = \sum_{k=1}^{2} \pi_k \frac{1}{\sqrt{2\pi}\, b_k} \exp\left(-\frac{(x-\mu_k)^2}{2b^2}\right)$$

(2)

$\{x_n\}_{n=1}^{N}$ Daten $\qquad \sum_{k=1}^{2} \pi_k = 1$

$\theta = (\mu_1, \mu_2, b)$ $\qquad \{k_n\}_{n=1}^{N}$ Label „Verteilung 1/2"

$$\mathbb{P}(k_n = 1 \mid x_n, \theta)$$

$$= \frac{f_{x|\theta}(x_n \mid k_n = 1)\, \pi_1}{f_{x|\theta}(x_n \mid k_n = 1)\, \pi_1 + f_{x|\theta}(x_n \mid k_n = 2) \cdot \pi_2 = f_{x|\theta}(x_n)}$$

$$f_{x|\theta}(x_n \mid k_n = 1) = \frac{1}{\sqrt{2\pi}\, b} \exp\left(-\frac{(x_n - \mu_1)^2}{2b^2}\right)$$

$f_{\mu_1, \mu_2 | \{x_n\}, b}$ ← Davon ein Max. finden

likelihood

$$= \boxed{f_{\{x_n\} | \mu_1, \mu_2, b} \cdot \text{„prior } \mu_1, \mu_2\text{"}}$$

$f_{\{x_n\}|b}$ ← konstant

nachdiff.

$$L := \log f_{\{x_n\} | \mu_1, \mu_2}$$

$$\frac{\partial}{\partial \mu_u} L = \left(\frac{\pi_u \dots \exp(\dots)}{\sum}\right) \leftarrow \text{1 Term} \leftarrow \text{2 Terme} \cdot \left(\frac{x_n - \mu_u}{b^2}\right)$$

$$\frac{\partial}{\partial \mu_u} L = \sum_n \frac{(x_n - \mu_u)}{b^2} f_{\mu_u | x_n}$$

(Finde Nullstellen der Ableitung mit (Quasi-) Newton)

↑ Daten

$$\frac{\partial^2}{\partial \mu_u^2} L \approx -\sum_n f_{\mu_k | x_n} \frac{1}{b^2}$$

$$\mu_{k,neu} = \frac{\sum f_{\mu_k | x_n} \cdot x_n}{\sum f_{\mu_k | x_n}}$$

so berechnet man den neuen Schwerpunkt

Schätzer für den bisherigen Cluster