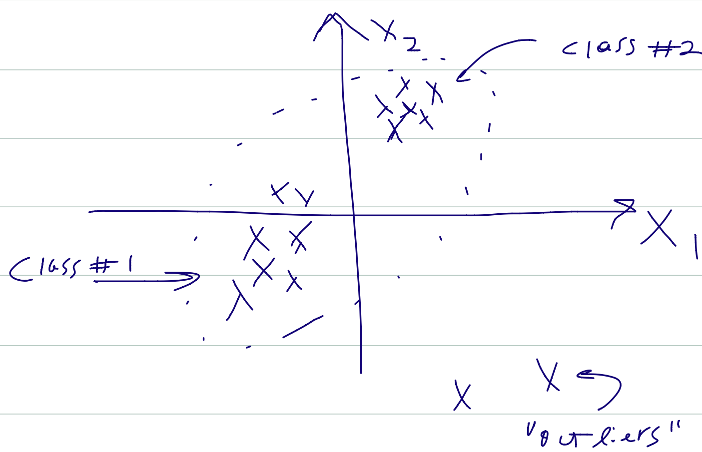
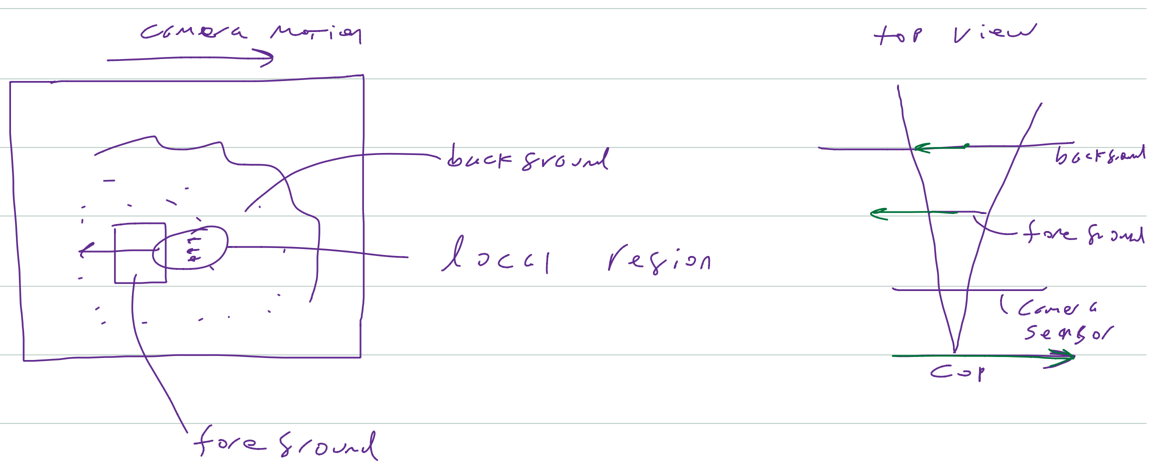
- A3, out Thursday
- Project status report, 10 days
- Office hours 3-5pm Friday
- Tutorial Wednesday next week.

Mixture Model

Represent data coming from multiple sources, but each has simple (e.g. Gaussian) dist'n

e.g. multiple objects (lines, surfaces, polynomial fits, motion)





K-Means

0: Choose K exemplars

1: assign points to closest exemplar "assign ownership"

2: re-estimate mean values "minimize squared error" based on current assignment

3: iterate and pray

Claim: SSE always non-increasing but may get stuck in local minima.

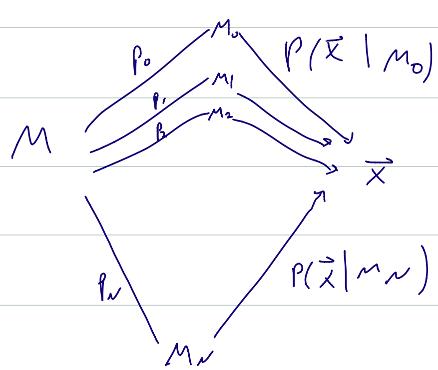
Mixture model

$$P(\vec{x}|M) = \sum_{n=0}^{N} \pi_n P(\vec{x}|M_n) \qquad \sum_{n=0}^{N} \pi_n = 1$$

prob of point $\vec{x}$ given all models

$P(x|M_0)$ = "outlier process"

Generative Process

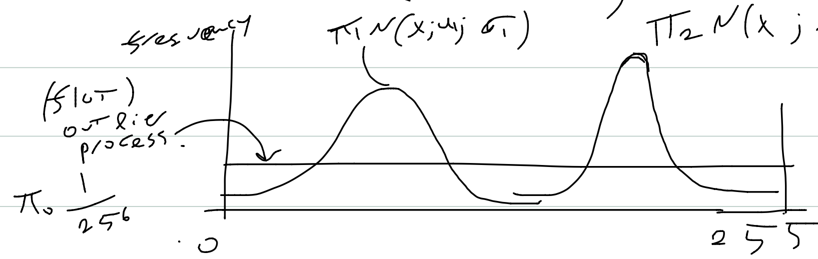


Problem: Given data $\vec{x}_1, \dots \vec{x}_K$ ← no. of data points

find model parameters & mixing properties

$[\mu_0, \mu_1, \dots \mu_N]$ $[\pi_0, \pi_1, \dots \pi_N]$

Example: Pixel (greyscale) intensities



$1 = \pi_0 + \pi_1 + \pi_2$

$$P(x \mid \mu_n) = N(x; \mu_n, \sigma_n) = \frac{1}{\sqrt{2\pi}\,\sigma_n} \exp\left(\frac{-1}{2\sigma_n^2}(x-\mu_n)^2\right)$$

$$P(x) = \pi_1 N(x; \mu_1, \sigma_1) + \pi_2 N(x; \mu_2; \sigma_2) + \pi_0 / 256$$

Optimize by EM algorithm — "Expectation Maximization"

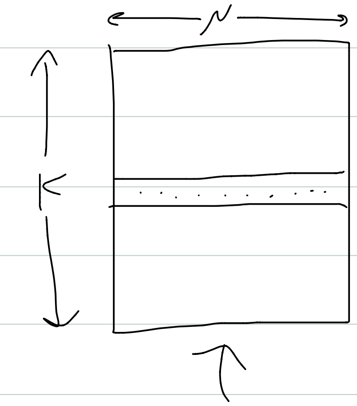
0. Guess some initial values π_n, μ_n, σ_n

1. E-step suppose model is correct

   Estimate "ownership" of each data sample to each model.

$$\tau_n(x_k) = \tau_{nk} \triangleq \frac{\pi_n P_n(x_k; \mu_n, \sigma_n)}{\sum_{n'} \pi_{n'} P_{n'}(x_k, \mu_{n'}, \sigma_{n'})}, \quad \forall k \; \sum_n \tau_{nk} = 1$$

$$0 \le \tau_{nk} \le 1$$



2. <u>M-step</u> maximize likelyhood of data as if ownership were correct.

$$\mu_n' \leftarrow \frac{\sum_{k=1}^{K} \tau_{nk} x_k}{\sum_{k=1}^{K} \tau_{nk}}$$

Recall $\hat{\mu} \leftarrow \frac{1}{K} \sum_{k=1}^{K} x_k$

standard estimate for mean

$$\sigma^2 \leftarrow \sum_{k=1}^{K} (x_k - \mu_k)^2 / (K-1)$$

$$(\sigma_n')^2 \leftarrow \sum_{k=1}^{K} \tau_{nk} (x_k - \mu_n')^2 \Big/ \sum_{k=1}^{K} \tau_{nk}$$

$$\Sigma = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix}$$

$$\pi_n' \leftarrow \sum_{k=1}^{K} \tau_{nk} / K$$

<u>Iterate and pray</u>

<u>Typical heuristics</u>

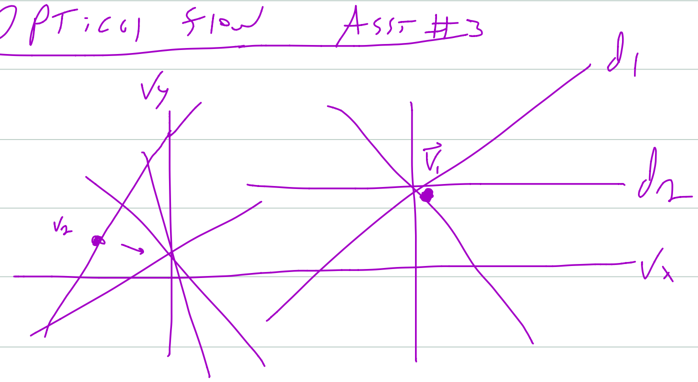
- Ransac to choose initial owner
- fix π's and $\Sigma = \sigma I$ diagonal + fixed variance σ

  <u>Tune μ</u>
- tune π
- tune Σ in general - non isotropic

- split and merge
- multiple runs w/ different start points
- stochastic search

advantaged data

Optical flow Asst #3



Proof of Convergence

$$\mathcal{L} = \prod_{k=1}^{K} p(\vec{x}_i \mid M)$$

likelihood of data

whole model

$$P(\vec{x}_k \mid M) = \sum_{n=0}^{N} \pi_n P(x \mid M_n)$$

Radford Neal + Geoff Hilton – EM algorithm

Can show EM is a local maximum of $\mathcal{L}$

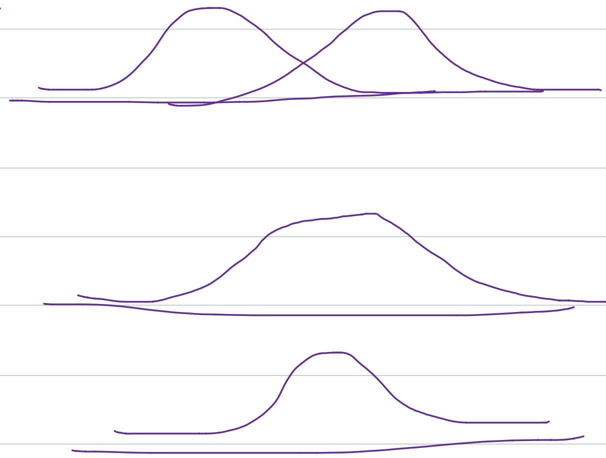
$(\mu_1, \sigma_1, \mu_2 \sigma_2, \ldots, \mu_N \sigma_N)$

$$\mathcal{L}(\vec{x}, \pi, M) = \prod_{k=1}^{K} \left[\sum_{n=0}^{N} \pi_n P(x_k \mid M_n) \right]$$

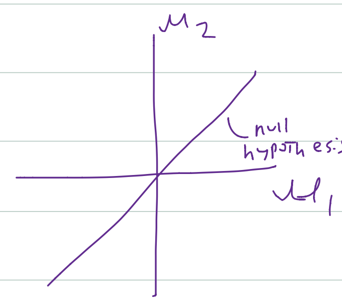
$\nabla \mathcal{L}$

How many Models?

1d problem



at which point can you say these are two or one peak in the data.



MacKay
"Bayesian interpolation"