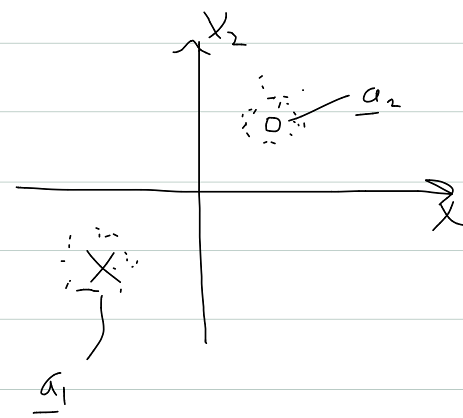
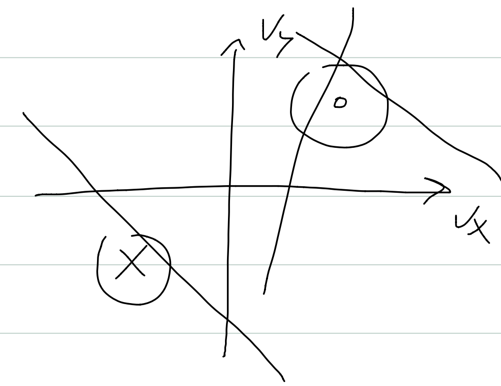
Assign #3

Th. 5-6 PM   A.I. lab   Tutorial

(8 input for starting mixtures)





Constraint lines

Let $e_k = C_k^T(\underline{a}, 1)$

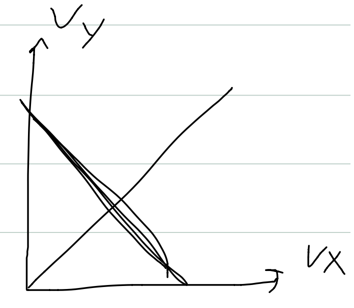
LSQ $\min \sum_{k=1}^{K} e_k^2$ Min sum squared distance to line C_k

" squared deviation from BCC of optical flow.

$e_k = (C_{kx}, C_{ky}, C_{xT}) \cdot (a_1, a_2, 1)$ optical flow variables.

$(I_{kx}, I_{ky}, I_{kT}) \cdot (V_x, V_y, 1) = 0$

$\boxed{I_{kx} V_x + I_{ky} V_y + I_t = 0}$ BCC



LSQ derivation

$$\mathcal{E}(\underline{a}) = \frac{1}{2}\sum_k e_k^2 \qquad e_k = (c_{kx}a_1 + c_{ky}a_2 + c_{kT})$$

$$\frac{\partial}{\partial \underline{a}}\mathcal{E}(\underline{a}) = \boxed{\sum_k e_k \frac{\partial e_k}{\partial \underline{a}} = 0} \qquad \underline{a} = (a_1, a_2)^T$$

$$\boxed{\begin{aligned} \frac{\partial}{\partial a_1} &\rightarrow \sum_k e_k c_{kx} = 0 \\ \frac{\partial}{\partial a_2} &\rightarrow \sum_k e_k c_{kx} = 0 \end{aligned}}$$

2 eqns

2 unkn (a_1, a_2)

All you need to know

$$\underbrace{\begin{bmatrix} \sum_k w_k c_{kx}^2 & \sum_k w_k c_{kx} c_{ky} \\ \sum_k w_k c_{kx} c_{ky} & \sum_k w_k c_{ky}^2 \end{bmatrix}}_{\underline{\underline{C}}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}}_{\underline{a}} = \underbrace{\begin{bmatrix} -\sum w_k c_{kx} c_{kT} \\ -\sum w_k c_{ky} c_{kT} \end{bmatrix}}_{\underline{b}}$$

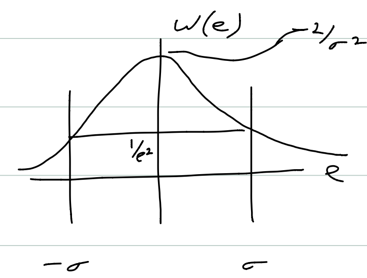
$$\hat{\underline{a}} = \underline{\underline{C}}^{-1} b \qquad \text{for LSQ } w_k = 1$$

$$\left.\underbrace{\begin{bmatrix} c_{1x} & c_{1y} & c_{1T} \\ c_{2x} & c_{2y} & c_{2T} \\ & \vdots & \\ c_{kx} & c_{ky} & c_{kT} \end{bmatrix}}_{w_k}\right\} \rightarrow \hat{a}$$

Robust

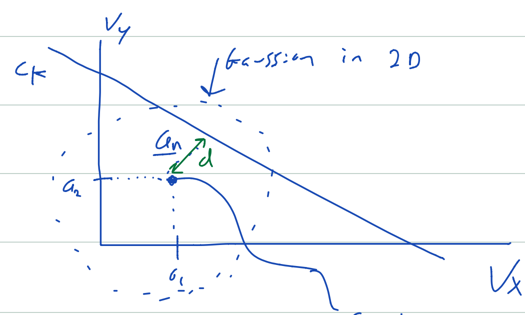
$$\min \sum_k \rho(e_k) \qquad \rho(e) = \frac{e^2}{\sigma^2+e^2} \qquad \frac{d}{de}\rho(e) = \frac{e\,2\sigma^2}{(\sigma^2+e^2)^2}$$

$$W(e_k) = \frac{\rho(e)}{e} = \frac{2\sigma^2}{(\sigma^2+e^2)^2}$$



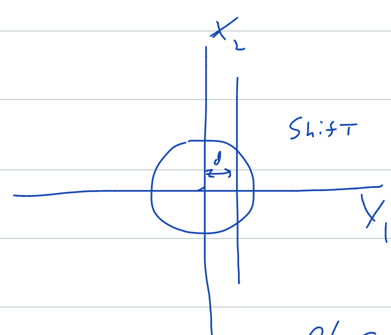
Mixture Models

$$\min \sum_k \tau_k e_k^2$$



What is $P(\underline{c}_k \mid a_n)$?

$\underline{a}$ is isotropic gaussian with variance σ^2.



Shift to origin and rotate

$$P(c_k \mid a_n) = \iint N\left((x_1,x_2); 0; \begin{bmatrix}\sigma & 0\\ 0 & \sigma\end{bmatrix}\right) dx_1\,dx_2$$

$$= N(d, 0, \sigma) \underbrace{\int N(x_2; 0; \sigma)\, dx_2}_{\to\ 1}$$

$$= N(d; 0; \sigma)$$ Gaussian of normal distance to line

"Gaussian Processes"
"Gaussian identities"
"Matrix identities"

Mixture Model

Note: we have not shown that update step(s) (weighted LSQ) are Maximum Likelihood of probability.

Consider data $X = [x_1, x_2, \ldots, x_K]$ $\quad x_k \in \mathbb{R}^d$ dim d

Likelihood of data given model θ

$$\mathcal{L}(X|\theta) = P(X|\theta) = \prod_{k=1}^{K} P(x_k|\theta)$$

iid points x

$$\text{Let } \mathcal{L}(\theta) = \sum_{k=1}^{K} \log P(x_k|\theta)$$

Maximum Likelyhood soln $\boxed{\frac{\partial}{\partial\theta}\log\mathcal{L}(\theta) = 0}$

$$\log\mathcal{L}(\theta) = \sum_{k=1}^{K} \log P(x_k|\theta) \qquad P(x_k|\theta) \sim \mathcal{N}(x_k, \mu, \Sigma)$$

$$\theta = [\mu, \Sigma]$$

$$\alpha = \sum_{k=1}^{K} (x_k - \mu)\Sigma^{-1}(x_k - \mu)$$

$$\mathcal{N}(x; \mu, \Sigma) = (2\pi)^{1/2}|\Sigma|^{1/2} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

$\frac{\partial}{\partial\mu}\log\mathcal{L} = 0$ when $\hat{\mu} = E(x_k)$

$\frac{\partial}{\partial\Sigma}\log\mathcal{L} = 0$ when $\hat{\Sigma} = E(x_k x_k^T)$

$d\times1$, $1\times d$, $d\times d$ matrix

$$\boxed{\begin{aligned}\hat{\mu} &= \frac{1}{K}\sum_k x_k \\ \hat{\Sigma} &= \frac{1}{K}\sum_k x_k x_k^T\end{aligned}}$$

Define "Gaussian Model"

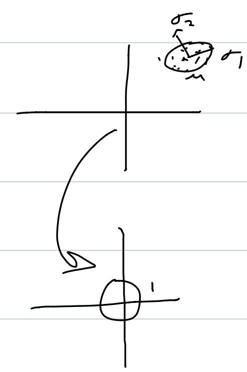
True value is μ

Measured values are $X_k \sim \mathcal{N}(\mu, \Sigma)$ Gaussian
with covariance matrix Σ

Key fact

MLE of Gaussian Model $\longleftrightarrow$ (Weighted)* LSQ fit to data

computationally equiv



$\tilde{X}_k \leftarrow \Sigma^{-1/2}(X_k - \mu)$

Whitening the data

Claim w/o proof

EM steps will (locally) maximize

$$\mathcal{L} = \prod_{k=1}^{K} \left(\sum_{n=1}^{N} \pi_n P(G_k | G_n) \right)$$

$\uparrow$ iid samples.