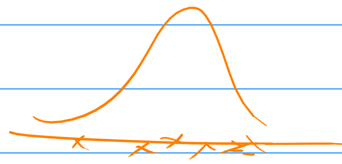
$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$\boxed{\tau = \frac{1}{\sigma^2}} \text{ – precision}$$

$$p(x|\mu,\tau) = \sqrt{\frac{\tau}{2\pi}}\, e^{-\frac{\tau}{2}(x-\mu)^2}$$



$$D = \{x_1, x_2, \dots x_n\} \rightsquigarrow \underline{\mu, \tau}$$

1) τ = const $\underline{\mu}$

$$\underline{p(\mu)} \times p(x|\mu) = \sqrt{\frac{\tau}{2\pi}}\, e^{-\frac{\tau}{2}(x-\mu)^2} \propto p(\mu|x)$$

$\tau_0\mu_0 + \tau\mu$

$p(\mu|\mu_0,\tau_0)$ $-\frac{1}{2}\left(\tau_0\mu^2 - 2\tau_0\mu_0\mu + \tau_0\mu_0^2 + \tau x^2 - 2\mu\tau x + \tau\mu^2\right)$

$$\ln p(\mu|x) = const + \frac{1}{2}\cancel{\ln\tau_0} - \frac{\tau_0}{2}(\mu-\mu_0)^2 + \frac{1}{2}\cancel{\ln\tau} - \frac{\tau}{2}(x-\mu)^2 =$$

$$= const - \frac{\tau_0+\tau}{2}\left(\mu - \frac{\tau_0\mu_0 + \tau x}{\tau_0+\tau}\right)^2$$

$$\frac{1}{\sigma'^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}$$

$$\mathcal{N}(\mu|\mu_0,\tau_0) \xrightarrow{x} \mathcal{N}\left(\mu \,\middle|\, \frac{\tau_0\mu_0 + \tau x}{\tau_0+\tau}, \tau_0+\tau\right)$$

$$\xrightarrow{x_1 \dots x_n} \mathcal{N}\left(\mu \,\middle|\, \frac{\tau_0\mu_0 + \tau\cdot\sum x_i}{\tau_0 + n\cdot\tau}, \tau_0 + n\tau\right)$$

$x^{\alpha-1}(1-x)^{\beta-1}$

2) μ = const, τ ln

$$\underline{\ln p(\tau)} + \frac{1}{2}\ln\tau - \frac{\tau}{2}(x-\mu)^2 + \underline{const}$$

$$p(\tau) = \text{Gamma}(\tau|\alpha_0,\beta_0) = \frac{\beta_0^{\alpha}}{\Gamma(\alpha)}\tau^{\alpha_0-1}e^{-\beta_0\tau}$$

$$\ln p(\tau|x) = const + (\alpha_0-1)\ln\tau - \beta_0\tau + \frac{1}{2}\ln\tau - \frac{\tau}{2}(x-\mu)^2$$

$\tau \sim$ Gamma $\sigma \sim \dots$, $\frac{1}{\sigma^2} \sim$ Gamma

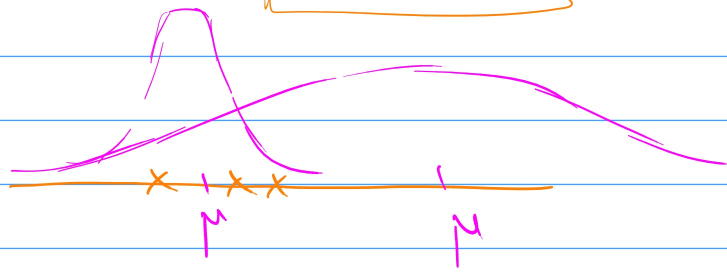
$$\mathrm{Gamma}(\tau\,|\,\alpha_0,\beta_0) \xrightarrow{\;x\;} \mathrm{Gamma}\left(\tau\,\middle|\,\alpha_0+\frac{1}{2},\ \beta_0+\frac{1}{2}(x-\mu)^2\right)$$

$$\xrightarrow{\;x_1 \dots x_n\;} \mathrm{Gamma}\left(\tau\,\middle|\,\alpha_0+\frac{n}{2},\ \beta_0+\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

3) $p(\mu,\tau) = ?$

$$\ln p(x\,|\,\mu,\tau) = const + \frac{1}{2}\ln\tau - \frac{\tau}{2}(x-\mu)^2$$

$$p(\mu,\tau) = \cancel{p(\mu)\,p(\tau)}$$
$$\cancel{\ln p(\mu) + \ln p(\tau)}$$



$$p(\mu,\tau) = p(\mu)\,p(\tau\,|\,\mu) = p(\tau)\,p(\mu\,|\,\tau) =$$

$$\ln p(\mu,\tau\,|\,x) = \qquad = \boxed{\mathrm{Gamma}(\tau\,|\,\alpha_0,\beta_0)\cdot N(\mu\,|\,\mu_0,\lambda_0\tau)}$$

$$= \ln p(\mu,\tau) + \ln p(x\,|\,\mu,\tau) + const = \quad -\frac{\tau}{2}\left(\lambda_0(\mu^2 - 2\mu\mu_0 + \mu_0^2) + x^2 - 2x\mu + \mu^2\right)$$

$$= const + (\alpha_0-1)\ln\tau - \beta_0\tau + \frac{1}{2}\ln(\cancel{\lambda_0}\tau) - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 +$$

$$+ \frac{1}{2}\ln\tau - \frac{\tau}{2}(x-\mu)^2 =$$

$$= const + (\alpha_0-1)\ln\tau - \beta_0\tau + \frac{1}{2}\ln\tau + \frac{1}{2}\ln\tau - \frac{\tau}{2}\left((\lambda_0+1)\mu^2 - 2\mu(x+\mu_0) + \lambda_0\mu_0^2 + x^2\right)$$

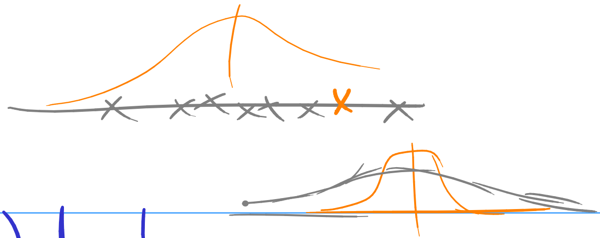
$$N\left(\mu\,\middle|\,\frac{\mu_0+x}{\lambda_0+1},\ \tau(\lambda_0+1)\right)\times \qquad (+\Sigma x,\ \lambda_0+N,\ N)$$

$$= -\frac{\tau(\lambda_0+1)}{2}\left(\mu - \frac{\mu_0+x}{\lambda_0+1}\right)^2 - \frac{\tau}{2}\left(\lambda_0\mu_0^2 + x^2 - \frac{(\mu_0+x)^2}{\lambda_0+1}\right)$$

$$\times\,\mathrm{Gamma}\left(\tau\,\middle|\,\alpha_0+\frac{1}{2},\ \beta_0+\frac{1}{2}\left(\lambda_0\mu_0^2 + x^2 - \frac{(\mu_0+x)^2}{\lambda_0+1}\right)\right) \qquad (N,\ \Sigma,\ \Sigma,\ N)$$

4)

$$\underline{p(\mu, \tau}\,|\, x_1 \ldots x_n)$$



$$p(x \mid x_1 \ldots x_n) = \int p(x, \mu, \tau \mid x_1 \ldots x_n)\, d\mu\, d\tau =$$

$$= \iint \text{Gamma}(\tau) \cdot \mathcal{N}(\mu \mid \tau) \cdot \mathcal{N}(x \mid \mu, \tau)\, d\mu\, d\tau$$

$$= \ldots = \underline{t\text{-Student}}\,(x \mid \underline{\mu}_{MAP}, n))$$

Экспон. семейство — exponential family / natural parameters

$$p(\bar{x} \mid \bar{\eta}) = \underline{h(\bar{x})} \cdot \underline{g(\bar{\eta})} \cdot e^{\bar{\eta}^T \underline{\bar{u}(\bar{x})}}$$

сопр. апр. распр.

$$\ln p(\bar{x} \mid \bar{\eta}) = \ln h(\bar{x}) + \ln g(\bar{\eta}) + \bar{\eta}^T \bar{u}(\bar{x})$$

$$p(x \mid \mu, \tau) = \sqrt{\frac{\tau}{2\pi}}\; e^{-\frac{\tau}{2}(x-\mu)^2} \qquad -\frac{\tau}{2}(x-\mu)^2 = -\frac{\tau}{2}\mu^2 + \tau\mu x - \frac{\tau}{2}x^2$$

$$= \underbrace{\left(\sqrt{\frac{\tau}{2\pi}} \cdot e^{-\frac{\tau}{2}\mu^2}\right)}_{g(\bar{\eta})} e^{\tau\mu x - \frac{\tau}{2}x^2} \qquad \bar{\eta} = \begin{pmatrix} \tau \\ \tau\mu \end{pmatrix} \qquad \bar{u}(x) = \begin{pmatrix} -x^2/2 \\ x \end{pmatrix}$$

Gaussian: $h(\bar{x}) = 1$, $g(\eta) = 4$, $\eta = \begin{pmatrix} \tau \\ \tau\mu \end{pmatrix}$, $\bar{u}(x) = \begin{pmatrix} -x^2/2 \\ x \end{pmatrix}$

Bernoulli: $p(x \mid \theta) = \theta^x (1-\theta)^{1-x} = e^{x \ln\theta + (1-x)\ln(1-\theta)} =$

$$= (1-\theta)\, e^{x\,\underline{(\ln\theta - \ln(1-\theta))}}$$

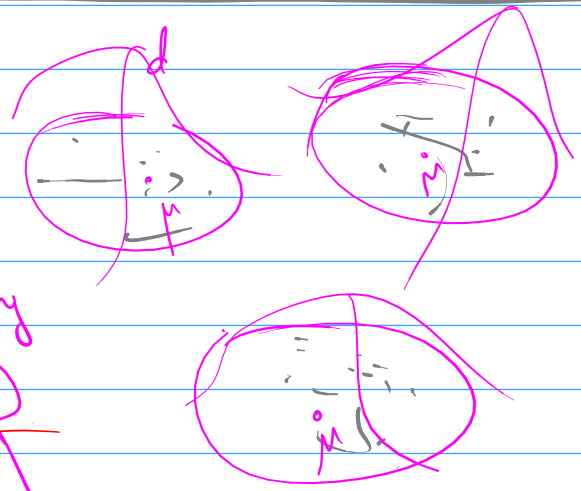
$h(x) = 1$, $g(\eta) = 1-\theta$, $\eta = \ln\frac{\theta}{1-\theta}$, $u(x) = x$

log-odds

$$p(\bar\eta \mid \bar x, D) = f(\bar x, D)\cdot g(\bar\eta)^{D}\, e^{D\,\bar\eta^T \bar x}$$

$$\ln p = \text{const} + D \ln g(\bar\eta) + D\cdot \bar\eta^T \bar x$$

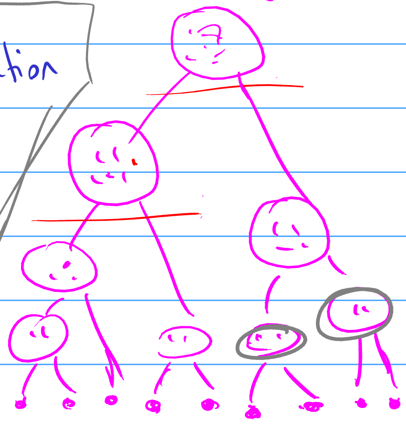
Кластеризация



Unsupervised learning

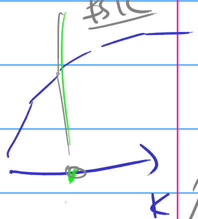
$p(\bar x)$

hierarchical clustering



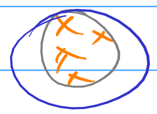
Model selection

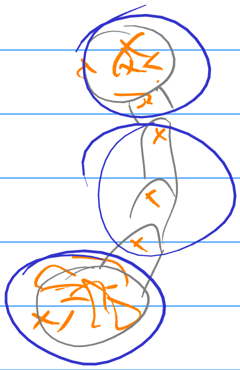
BIC



single-link $d(C_1, C_2) = \min\limits_{x\in C_1,\, y\in C_2} d(x,y)$

complete-link $d(C_1, C_2) = \max d(x,y)$





$D = \{\bar x_1, \bar x_2, \dots, \bar x_N\}$

$\bar\pi = (\pi_1, \pi_2, \dots, \pi_K)$

 $\longrightarrow \bar z = (0 \dots 1 \dots 0) \to$

$z_k = 1$

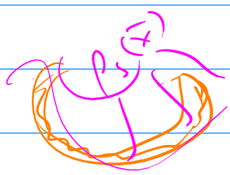
$\to \bar x \sim p_k(\bar x)$

DBSCAN

ε

$p(K)$

Chinese restaurant



$$p(x \mid \bar\pi, \bar\theta_1, \dots, \bar\theta_K) = \pi_1 p_1(\bar x \mid \bar\theta_1) + \dots + \pi_K p_K(\bar x \mid \bar\theta_K)$$

$$p(D \mid \dots) = \prod_{\bar x\in D} p(\bar x \mid \dots) = \prod_{\bar x \in D} \left(\pi_1 p_1(\bar x \mid \bar\theta_1) + \dots + \pi_K p_K(\bar x \mid \bar\theta_K)\right) \xrightarrow[\bar\pi,\, \bar\theta_1, \dots, \bar\theta_K]{} \max$$

unsupervised

D

$p(\bar{x}|\bar{\theta})$

$p(D|\bar{\theta}) \xrightarrow[\bar{\theta}]{} \max$

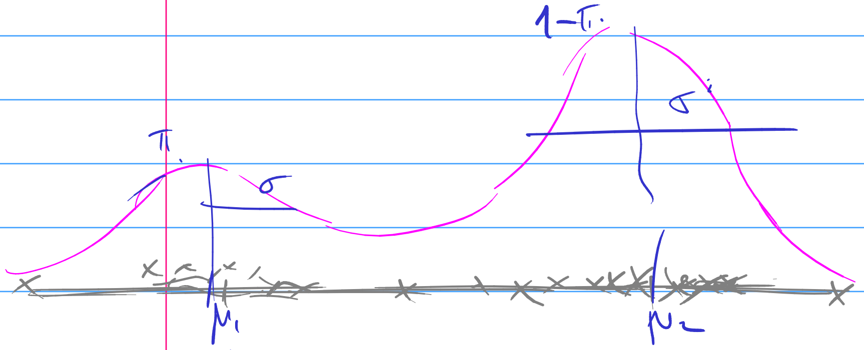
$x_1 \, x_2 \, x_3 \, | \, x_4 - x_5 - x_N$

$p(x_5|x_1,x_2,x_3) =$

$= \int p(x_5, x_4|x_1,x_2,x_3)\, dx_4$

$= \int p(x_5|x_1 - x_4)\, p(x_4|x_1, x_3)\, dx_4$

EM-algorithm



$$p(\bar{x}|\mu_1,\mu_2) = \pi \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu_1)^2} + (1-\pi)\frac{1}{\sqrt{}} e^{-\frac{1}{2\sigma^2}(x-\mu_2)^2}$$

$z_n = 1$ $\qquad$ $z_n = 0$

$$p(D|\pi,\mu_1,\mu_2) = \prod_n \left(\pi N(x_n|\mu_1) + (1-\pi) N(x_n|\mu_2)\right)$$

$(\bar{x}_n, z_n)$

$$p(D,z|\pi,\mu_1,\mu_2) = \prod_n \left(\pi N(x_n|\mu_1)\right)^{z_n} \left((1-\pi) N(x_n|\mu_2)\right)^{1-z_n}$$

$$\ln p(D,z|\ldots) = \qquad \xrightarrow[\pi,\mu_1,\mu_2]{} \max$$

$$= \sum_n \left(z_n \ln\pi + z_n \ln N(x_n|\mu_1) + (1-z_n)\ln(1-\pi) + (1-z_n)\ln N(x_n|\mu_2)\right) =$$

$$= \sum_n \left(z_n \ln\pi + (1-z_n)\ln(1-\pi)\right) + \sum_n z_n \ln N(x_n|\mu_1) + \sum_n (1-z_n)\ln N(x_n|\mu_2)$$

$\rightarrow \mu_1$

$d\pi$

max

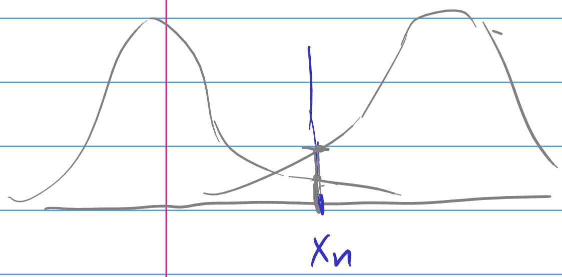
$\rightarrow \mu_2$

$z \longrightarrow \pi, \mu_1, \mu_2$

$$\ln p(D,z|\ldots) = \sum_n \left(\mathbb{E}z_n \ln\pi + (1-\mathbb{E}z_n)\ln(1-\pi)\right) + \sum_n \mathbb{E}z_n \cdot \ln N(x_n|\mu_1) + \sum_n (1-\mathbb{E}z_n)\ln N(x_n|\mu_2)$$

π, μ_1, μ_2 $\quad \boxed{\mathbb{E}[z_n]} = p(z_n = 1 \mid x_n, \pi, \mu_1, \mu_2) =$

$$= \frac{p((z_n=1), x_n \mid \pi, \mu_1, \mu_2)}{p(x_n \mid \pi, \mu_1, \mu_2)} =$$



$$= \frac{\pi N(x_n \mid \mu_1)}{\pi N(x_n \mid \mu_1) + (1-\pi) N(x_n \mid \mu_2)}$$

$\pi, \mu_1, \mu_2 \longrightarrow z$

Expectation - Maximization

EM - algorithm:

$\pi^{(0)}, \mu_1^{(0)}, \mu_2^{(0)} \xrightarrow{\text{E-war}} \mathbb{E} z^{(0)}$

M-war

$\pi^{(1)}, \mu_1^{(1)}, \mu_2^{(1)} \xrightarrow{\text{E-war}} \mathbb{E} z^{(1)}$

M-war

$\pi^{(2)}, \mu_1^{(2)}, \mu_2^{(2)} \xrightarrow{\text{E-war}} \mathbb{E} z^{(2)}$

k-means

k-means

- $\mu_1^{(0)}, \mu_2^{(0)}$
- $z_n^{(0)} := \arg\max \ldots$
- $\mu_1^{(1)}, \mu_2^{(1)}$
-
-

$$p(D \mid \pi, \bar\theta_1 \ldots \bar\theta_k) = \prod_n \sum_k \pi_k \, p_k(\bar x_n \mid \bar\theta_k) \longrightarrow \max$$

$\bar\pi, \bar\theta_1, \ldots, \bar\theta_k$

ACGTGCGCTAG
ACCTGCCCAAG

$$p(s \mid \ldots) = \sum_k \pi_k \, p(s \mid C_k)$$

$\pi_k, \; p_{k,i,a}$

$\{A, C, G, T\}$ $\qquad p_{k,i,a} = p(x_i = a \mid x \in C_k)$

E-war: $\bar x_n$

$$\mathbb{E} z_{nk} = p(\bar x_n \in C_k) = \frac{\pi_k \, p(\bar x_n \mid C_k)}{\sum_\ell \pi_\ell \, p(\bar x_n \mid C_\ell)} = \frac{\pi_k \prod_i p_{k,i,x_i}}{\sum_\ell \pi_\ell \prod_i p_{\ell,i,x_i}}$$

M-шаг: $\pi, p = \arg\max_{\pi, p} p(X, \mathbb{E}z \mid \pi, p)$

$$\prod_{n=1}^{N} \prod_{k=1}^{K} \left(\pi_k \prod_i p_{k,i,x_{ni}} \right)^{\mathbb{E}z_{nk}} \xrightarrow[\pi, p]{} \max$$

$$\sum_n \sum_k (\mathbb{E}z_{nk}) \cdot \ln \pi_k + \sum_n \sum_k (\mathbb{E}z_{nk}) \cdot \sum_i \ln p_{k,i,x_{ni}}$$

$$\sum_{k,i,a} \ln p_{k,i,a} \left(\sum_{n : x_{ni} = a} \mathbb{E}z_{nk} \right)$$

$$\log \mathbb{E}z_{nk} = \underbrace{\log \pi_k + \sum_i \log p_{k,i,x_i}}_{\alpha_k} - \log\left(\sum \ldots\right)$$

np.log(np.sum(...))

$\log(\text{sum}_k(\exp(\alpha_k)))$