RVM

$$p(y|\bar{w},\bar{x}) = \mathcal{N}(y|\bar{w}^T\bar{x}, \beta^{-1})$$

$$p(\bar{w}|\bar{\alpha}) = \prod_i \mathcal{N}(w_i|0, \alpha_i^{-1}) \qquad A = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \ddots \; \alpha_d \end{pmatrix}$$

$$p(\bar{w}|D, \bar{\alpha}, \beta) = \mathcal{N}(\bar{w}|\bar{\mu}, \Sigma)$$

$$\Sigma^{-1} = \underbrace{A} + \beta X^T X, \qquad \mu = \beta \Sigma X^T \bar{y}$$

$$p(\bar{y}|X, \bar{\alpha}, \beta) = \int p(\bar{y}|\bar{w}, X, \beta)\, \underline{p(\bar{w}|\bar{\alpha})}\, d\bar{w} =$$

$$= \left(\frac{\beta}{2\pi}\right)^{N/2} \cdot \frac{1}{(2\pi)^{d/2}} \cdot \prod_i \sqrt{\alpha_i} \cdot \int e^{-E(\bar{w})} d\bar{w}$$

$$E(\bar{w}) = \underline{\frac{1}{2}\bar{w}^T A \bar{w} + \frac{1}{2}(\bar{y} - X\bar{w})^T(\bar{y} - X\bar{w})}$$

1) $$E(\bar{w}) = \underline{\frac{1}{2}(\bar{w} - \bar{\mu})^T \Sigma^{-1}(\bar{w} - \bar{\mu})} + E(\bar{y}) = \frac{1}{2}\left(\beta \bar{y}^T \bar{y} - \bar{\mu}^T \Sigma^{-1} \bar{\mu}\right),$$

$$E(\bar{y}) = \frac{1}{2}\bar{y}^T C^{-1} \bar{y}, \quad C = \beta^{-1} I + X A^{-1} X^T$$

$$\underline{\log p(y|X, \bar{\alpha}, \beta)} = \frac{N}{2}\log\beta - \frac{d}{2}\log 2\pi + \underline{\frac{1}{2}\sum \log \alpha_i} - E(\bar{y}) -$$

$$- \frac{1}{2}\underline{\log\det\Sigma} - \frac{N}{2}\log 2\pi \xrightarrow[\bar{\alpha}, \beta]{} \max$$

2) $$\frac{\partial \log\det S}{\partial x} = \mathrm{Tr}\left(S^{-1} \frac{\partial S}{\partial x}\right), \qquad \frac{\partial S^{-1}}{\partial x} = -S^{-1}\frac{\partial S}{\partial x} S^{-1}$$

$$= -\mathrm{Tr}\left(\frac{\partial S^{-1}}{\partial x} \cdot S\right)$$

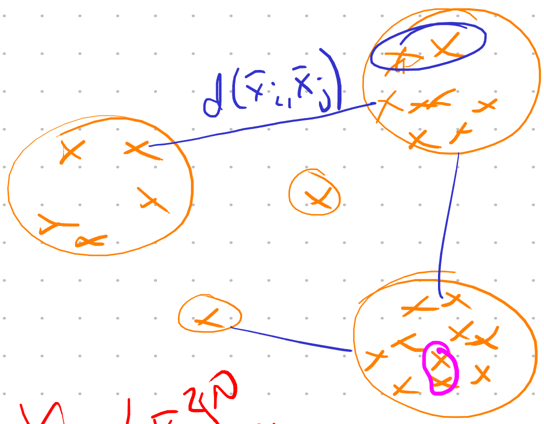
$$\frac{\partial \log\det\Sigma}{\partial \alpha_i} = -\underline{\Sigma_{ii}} \qquad \frac{\partial \log\det\Sigma}{\partial \beta} = \underline{-\mathrm{Tr}(X^T X \cdot \Sigma)}$$

$$\frac{\partial \log p(y|X, \bar{\alpha}, \beta)}{\partial \alpha_i} = \frac{1}{2\alpha_i} - \frac{\mu_i^2}{2} - \frac{1}{2}\Sigma_{ii} = 0$$

$$\alpha_i = \frac{1 - \alpha_i \Sigma_{ii}}{\mu_i^2} \qquad \beta^{-1} = \frac{(\bar{y} - X\bar{\mu})^T(\bar{y} - X\bar{\mu})}{\sum_i \alpha_i \Sigma_{ii}}$$

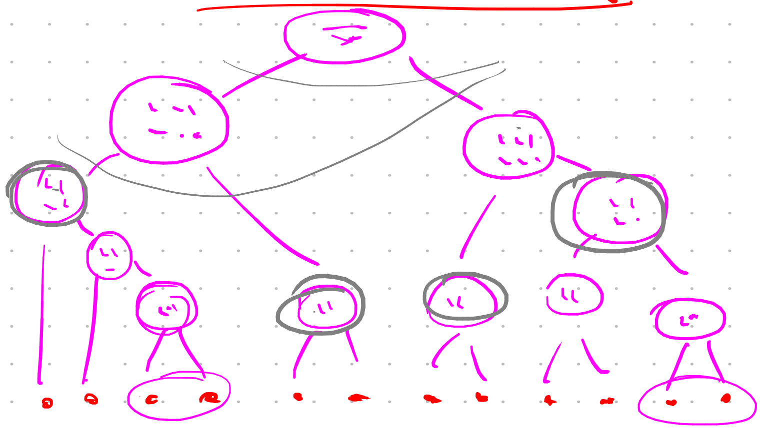
① **Clustering** — unsupervised learning

$\bar{x} \in \mathbb{R}^d$



$X = \{\bar{x}_n\}_{n=1}^{N}$

Agglomerative clustering



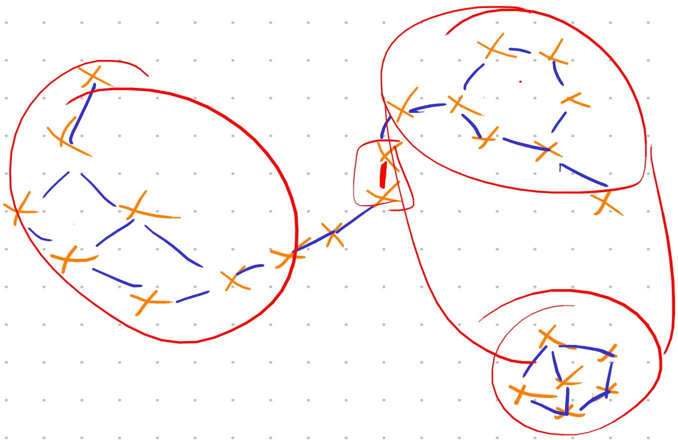
Single-link clustering

$$d(C_1, C_2) = \min_{\substack{\bar{x}_1 \in C_1 \\ \bar{x}_2 \in C_2}} d(\bar{x}_1, \bar{x}_2)$$

Complete link

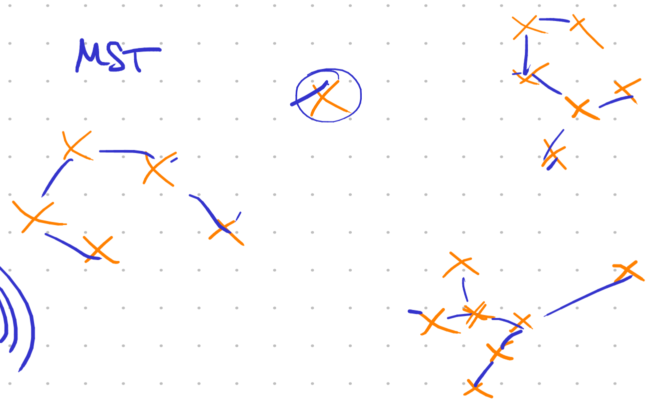
— " — $\max d(\bar{x}_1, \bar{x}_2)$

avg



Graph theory

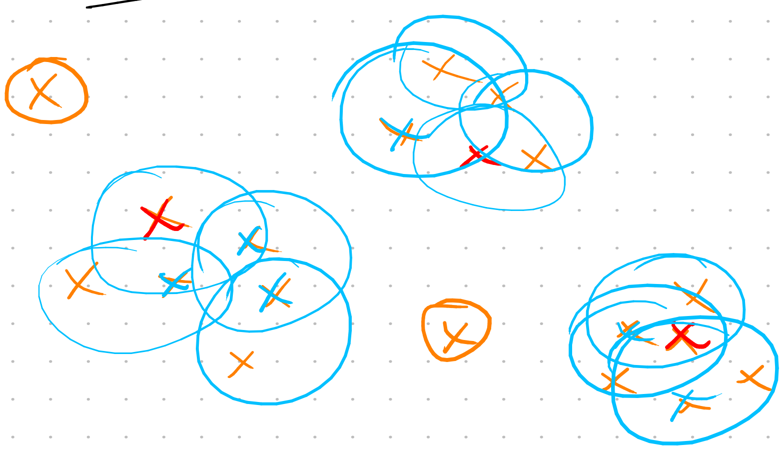
MST



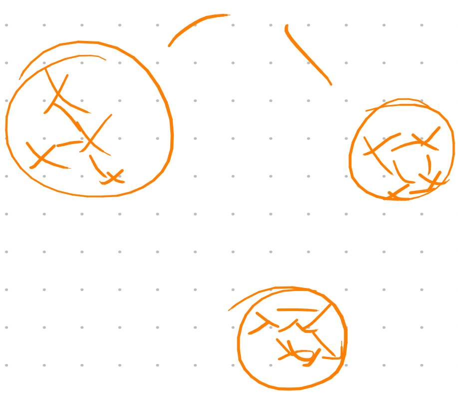
$G_{complete} = (X, E = (d(\bar{x}_i, \bar{x}_j)))$

DBSCAN

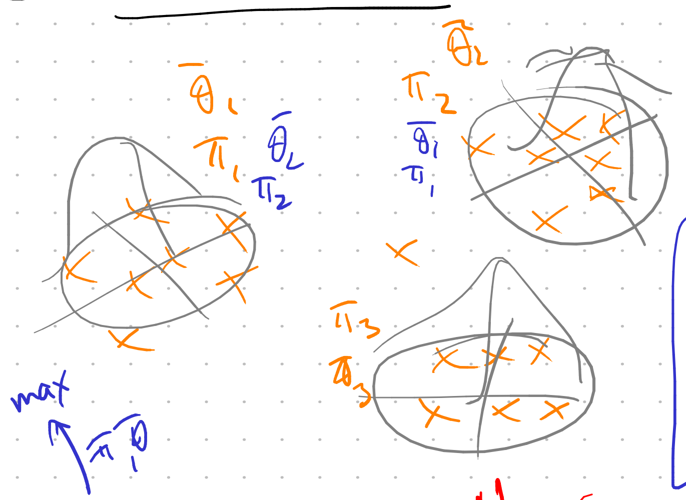
ε



BIRCH



(2) Mixture models



$$p(\bar x) = \pi_1 \underbrace{p_1(\bar x)}_{p_k(\bar x|\bar\theta_k)} + \dots + \pi_K p_K(\bar x)$$

$$\sum \pi_k = 1$$

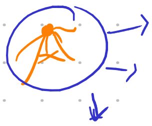
LDA, QDA:

$$p(C_k|\bar x) \propto p(C_k)\, p(\bar x|C_k)$$

$$p(D|\dots) = \prod_n \prod_k \left(\pi_k p_k(\bar x_n|\bar\theta_k)\right)^{z_{nk}}$$

$$p(D|\bar\pi, \bar\theta_1, \dots, \bar\theta_K) = \prod_{n=1}^{N} \left(\pi_1 p_1(\bar x_n|\bar\theta_1) + \dots + \pi_K p_K(\bar x_n|\bar\theta_K)\right)$$

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



$$Z = \{\bar z_n\}_{n=1}^{N}$$ — latent variables

$$\bar z_n = (0 \dots 1 \dots 0)$$

$$z_{nk} = [\bar x_n \in C_k]$$

 $\bar\pi$

$\Downarrow$

k — # кластера

$\Downarrow$

$\bar x \sim p_k(\bar x|\bar\theta_k)$

$$p(X, Z|\bar\pi, \bar\theta_1, \dots, \bar\theta_K) = \prod_n p(\bar x_n, \bar z_n|\bar\pi, \bar\theta_1, \dots, \bar\theta_K) =$$

$$= \prod_{n=1}^{N} p(\bar z_n|\bar\pi)\, p(\bar x_n|\bar z_n, \bar\theta_1, \dots, \bar\theta_K) =$$

$$= \prod_{n=1}^{N} \prod_{k=1}^{K} \left(p(\bar z_n = k|\bar\pi)\, p(\bar x_n|\bar\theta_k)\right)^{z_{nk}} =$$

$$= \prod_n \prod_k \left(\pi_k p(\bar x_n|\bar\theta_k)\right)^{z_{nk}}$$

$$\log p(X, Z|\bar\pi, \bar\theta_1, \dots, \bar\theta_K) = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(z_{nk} \log \pi_k + z_{nk} \log p(\bar x_n|\bar\theta_k)\right)$$

$$= \sum_{k=1}^{K} \left(\sum_n z_{nk}\right) \log \pi_k + \sum_{k=1}^{K} \left(\sum_n z_{nk} \log p(\bar x_n|\bar\theta_k)\right)$$

$$\sum \pi_k = 1$$

$$\pi_k^* = \frac{\sum_n z_{nk}}{\sum_k \sum_n z_{nk}} = N$$

$$p(\bar{x}_n | \bar{\theta}_k) = \mathcal{N}(\bar{x}_n | \bar{\mu}_k, \Sigma_k)$$

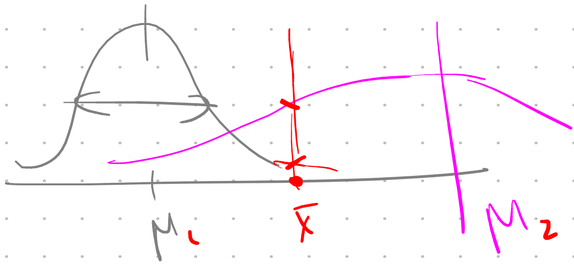
$$\bar{\mu}_k^* = \mathrm{Avg}(\bar{x}_n : z_{nk} = 1), \quad \Sigma_k^* = \_\_$$

③ Expectation – Maximization algorithm

$X, Z, \bar{\theta}$ $\qquad p(X|\bar{\theta}) \xrightarrow{\bar{\theta}} \max$ — трудно

$\qquad\qquad\qquad p(X, Z|\bar{\theta}) \xrightarrow{\bar{\theta}} \max$ — легко

(E-шаг): fix $\bar{\theta}$, find $\mathbb{E}[Z]$:



$$\mathbb{E}[z_{nk}] = p(z_{nk}=1) = p(C_k | \bar{x}_n, \bar{\theta}) =$$

$$= \frac{p(C_k|\bar{\theta})\, p(\bar{x}_n | C_k, \bar{\theta})}{\sum_{k=1}^{K} p(C_k|\bar{\theta})\, p(\bar{x}_n|C_k, \bar{\theta})} = \frac{\pi_k \mathcal{N}(\bar{x}_n | \bar{\mu}_k, \Sigma_k)}{\sum_{l=1}^{K} \pi_l \mathcal{N}(\bar{x}_n | \bar{\mu}_l, \Sigma_l)}$$

(М-шаг): fix $\mathbb{E}[Z]$, $\quad \underline{\mathbb{E}_Z[\log p(X, Z|\bar{\theta})]} \xrightarrow{\bar{\theta}} \max$

$$\mathbb{E}_Z[\log p(X, Z|\bar{\theta})] = \mathbb{E}_Z\Big[\sum_n \sum_k z_{nk}\big(\log \pi_k + \log p(\bar{x}_n|\bar{\theta}_k)\big)\Big] =$$

$$= \sum_n \sum_k \mathbb{E}[z_{nk}] \cdot \big(\log \pi_k + \log p(\bar{x}_n|\bar{\theta}_k)\big) =$$

$$= \underbrace{\Big[\sum_k \Big(\sum_n \mathbb{E}[z_{nk}]\Big) \cdot \log \pi_k\Big]}_{\pi} + \sum_k \underbrace{\Big[\sum_n \mathbb{E}[z_{nk}] \log p(\bar{x}_n|\bar{\theta}_k)\Big]}_{\bar{\theta}_k}$$

4) EM в общем виде

$X, Z, \bar\theta \qquad p(X|\bar\theta) \to \max_{\bar\theta}$

$$p(X|\bar\theta) = \int p(X,Z|\bar\theta)\,dZ$$

$$\boxed{Q(\bar\theta, \bar\theta^{(m)}) = E_{p(Z|X,\theta^{(m)})}\left[\log p(X,Z|\bar\theta)\right]}$$

E-шаг

$$\boxed{\bar\theta^{(m+1)} = \operatorname*{argmax}_{\bar\theta} Q(\bar\theta, \bar\theta^{(m)})}$$

M-шаг

$$Q(\bar\theta, \bar\theta^{(m)}) = \int p(Z|X,\theta^{(m)}) \log p(X,Z|\bar\theta)\,dZ$$

Цель: $p(X|\bar\theta^{(m+1)}) \geq p(X|\bar\theta^{(m)})$

$$p(X|\bar\theta) \;?\; p(X|\bar\theta^{(m)})$$

$$\log p(X|\theta) - \log p(X|\bar\theta^{(m)}) = \log \int p(X,Z|\theta)\,dZ - \log p(X|\bar\theta^{(m)})$$

$$= \log \int p(Z|X,\bar\theta^{(m)}) \cdot \frac{p(X,Z|\bar\theta)}{p(Z|X,\bar\theta^{(m)})}\,dZ - \log p(X|\bar\theta^{(m)})$$

$$= \log E_{p(Z|X,\bar\theta^{(m)})}\left[\frac{p(X,Z|\bar\theta)}{p(Z|X,\bar\theta^{(m)})}\right] - \log p(X|\bar\theta^{(m)}) \geq$$

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y)$$

$$f(E_p y) \geq E_p[f(y)]$$

Jensen's inequality

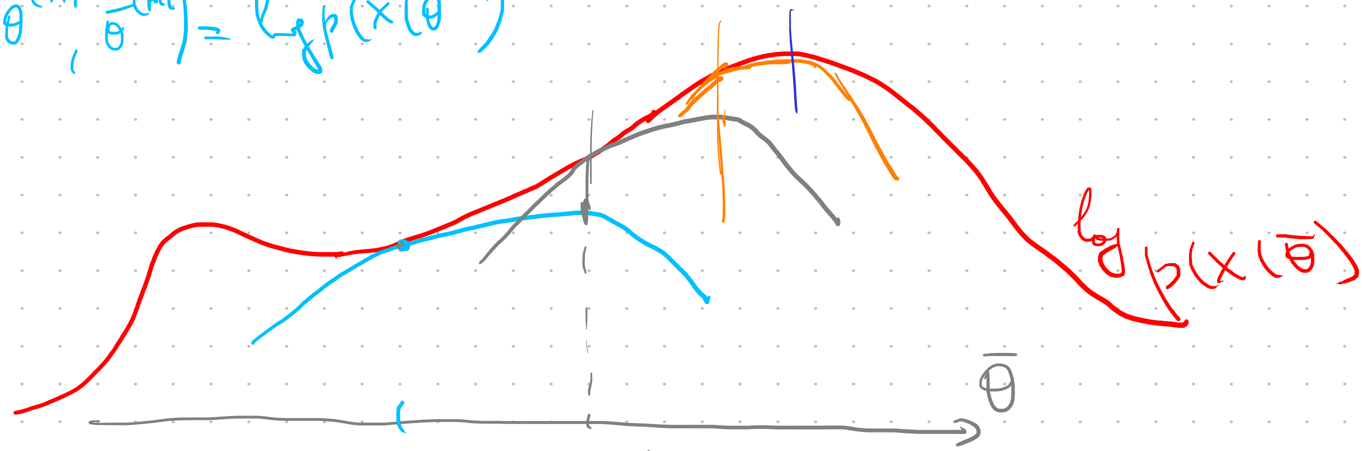
$$\geq E_{p(Z|X,\bar\theta^{(m)})}\left[\log \frac{p(X,Z|\bar\theta)}{p(Z|X,\bar\theta^{(m)})}\right] - \log p(X|\bar\theta^{(m)}) =$$

$$= \int p(z|x,\bar\theta^{(m)}) \log \frac{p(x,z|\bar\theta)}{\underbrace{p(z|x,\bar\theta^{(m)})\,p(x|\bar\theta^{(m)})}_{= p(x,z|\bar\theta^{(m)})}} \, dz$$

$$\log p(x|\theta) - \log p(x|\bar\theta^{(m)}) \geq \mathbb{E}_{p(z|x,\bar\theta^{(m)})}\left[\log \frac{p(x,z|\bar\theta)}{p(x,z|\bar\theta^{(m)})}\right]$$

$$\log p(x|\bar\theta) \geq L(\bar\theta, \bar\theta^{(m)}) = \log p(x|\bar\theta^{(m)}) + \mathbb{E}_{..}[\,..\,]$$

$$L(\bar\theta^{(m)}, \bar\theta^{(m)}) = \log p(x|\bar\theta^{(m)})$$



$\bar\theta^{(m)}$ $\quad$ $\bar\theta^{(m+1)} = \arg\max_{\bar\theta} L(\bar\theta, \bar\theta^{(m)})$

$$L(\bar\theta,\bar\theta^{(m)}) = \underbrace{\log p(x|\bar\theta^{(m)})}_{Const(\bar\theta)} + \underbrace{\mathbb{E}_{p(z|x,\bar\theta^{(m)})}\left[\log p(x,z|\bar\theta)\right.}_{Q(\bar\theta,\bar\theta^{(m)})} - \left.\underbrace{\log p(x,z|\bar\theta^{(m)})}_{Const(\bar\theta)}\right]$$

Mixture model: $\quad z_n = (0 \_ 1 \_ 0)\,, \quad p(x_n|z_n,\bar\theta) = \sum \pi_k \, p(\bar x_n|\bar\theta_k)$

$$p(x,z|\bar\theta) = \prod_n p(\bar z_n|\bar\theta)\, p(\bar x_n|\bar z_n,\bar\theta) = \prod_n \prod_k \left(\pi_k \, p(\bar x_n|\bar\theta_k)\right)^{z_{nk}}$$

max - M-step

$$Q(\bar\theta,\bar\theta^{(m)}) = \mathbb{E}_{p(z|x,\bar\theta^{(m)})}\left[\log p(x,z|\bar\theta)\right] =$$ E-step

$$= \mathbb{E}_z\left[\sum_n \sum_k z_{nk}\left(\log\pi_k + \log p(\bar x_n|\bar\theta_k)\right)\right] = \sum_n \sum_k \mathbb{E}[z_{nk}]\left(\log\pi_k + \log p(\bar x_n|\bar\theta_k)\right)$$

для μ_k, Σ_k: $\sum_n \mathbb{E}[z_{nk}] \cdot \log \mathcal{N}(\bar{x}_n \mid \bar{\mu}_k, \Sigma_k) \xrightarrow{\mu_k, \Sigma_k} \max$

$$\sum_n \mathbb{E}[z_{nk}] \cdot \left(\text{Const} - \frac{1}{2} \log \det \Sigma_k - \frac{1}{2} (\bar{x}_n - \bar{\mu}_k)^T \Sigma_k^{-1} (\bar{x}_n - \bar{\mu}_k) \right)$$

5 Дискретные примеры

доминантный / рецессивный: + / −

(Ceppellini et al., 1955)

θ

$q = p(+)$

$1 - q = p(-)$

a	$++$	$\rightarrow$ фенотип A
c	$--$	$\rightarrow$ фенотип B
b	$+-$, $-+$	$\rightarrow$ фенотип A

Дано: X: $\boxed{(a+b), \; c}$

$p(--\mid B) = 1$

$p(++) = q^2$

$p(--) = (1-q)^2$

$p(+-, -+) = 2q(1-q)$

$$p(++\mid A) = \frac{p(++)}{p(++) + p(+-,-+)} = \frac{q^2}{q^2 + 2q(1-q)} = \frac{q}{2-q}$$

$q^{(0)}$: $\mathbb{E}_{q^{(0)}}[a] = (a+b) \cdot p(++\mid A) = (a+b) \cdot \frac{q^{(0)}}{2 - q^{(0)}}$

E-шаг: $\mathbb{E}_{q^{(0)}}[b] = (a+b) - \mathbb{E}[a] = (a+b) \cdot \frac{2(1-q^{(0)})}{2 - q^{(0)}}$

M-шаг: $q^{(1)} := \mathbb{E}\left[\frac{2a+b}{n}\right] = \frac{a+b}{n} \cdot \frac{2q^{(0)} + 2(1-q^{(0)})}{2 - q^{(0)}} = \frac{a+b}{n} \cdot \frac{2}{2 - q^{(0)}}$

$$q^{(m+1)} = \frac{a+b}{n} \cdot \frac{2}{2 - q^{(m)}}$$

String clustering

M

ACCAGGTAGC

ACGAGCTAGC

$\{A, C, G, T\}$ $\qquad \{a_1, \dots, a_L\}$

$k = 1, \dots, K$, $\pi_k = p(C_k)$

$C_k = \{ p(x_i = a_l \mid \bar{x} \in C_k) \}$

$$p(\bar{x}\mid C_k) = \prod_{i=1}^{M} p(x_i\mid C_k) = \prod_{i=1}^{M}\prod_{l=1}^{L} p(x_i = a_l\mid C_k)^{[x_i = a_l]} =$$

$$\theta_{kil}$$

$$= \prod_{i=1}^{M}\prod_{l=1}^{L} \theta_{kil}^{[x_i = a_l]} \qquad \forall k,i \quad \sum_l \theta_{kil} = 1$$

latent variables

$$\boxed{z_{nk} = [\bar{x}_n \in C_k]}$$

$$p(X, Z\mid\bar{\theta}) = \prod_{n=1}^{N}\prod_{k}\left(\pi_k\, p(\bar{x}_n\mid C_k)\right)^{z_{nk}} = \prod_{n=1}^{N}\prod_{k}\left[\pi_k \prod_{i=1}^{M}\prod_{l=1}^{L}\theta_{kil}^{[x_i = a_l]}\right]^{z_{nk}}$$

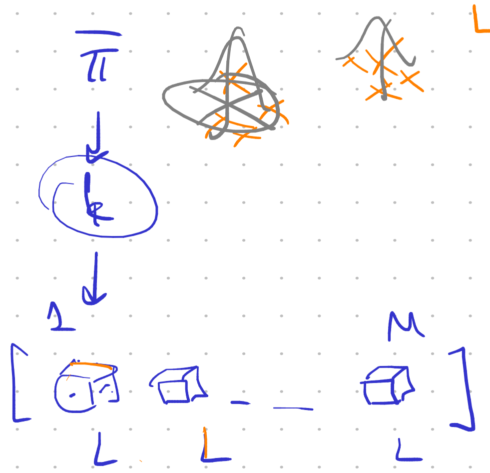
E-war: $$\mathbb{E}[z_{nk}] = \frac{\pi_k \prod_i\prod_l \theta_{kil}^{[x_i = a_l]}}{\sum_s \pi_s \prod_i\prod_l \theta_{sil}^{[x_i = a_l]}}$$

M-war: $$\mathbb{E}_Z[\log p(X, Z\mid\bar{\theta})] = \sum_{n=1}^{N}\sum_{k=1}^{K}\mathbb{E}[z_{nk}]\left(\log\pi_k + \sum_i\sum_l [x_{ni} = a_l]\log\theta_{kil}\right)$$

$$\pi, \theta \to \max$$

$$= \sum_{k=1}^{K}\left(\sum_n \mathbb{E}[z_{nk}]\log\pi_k\right) +$$

$$+ \sum_{k=1}^{K}\sum_{i=1}^{M}\left[\sum_{l=1}^{L}\left(\sum_{n=1}^{N}[x_{ni} = a_l]\cdot\mathbb{E}[z_{nk}]\right)\log\theta_{kil}\right]$$



$$\theta_{kil}^{(m+1)} = \frac{\sum_n [x_{ni} = a_l]\,\mathbb{E}[z_{nk}]}{\underbrace{\sum_{l'}\sum_n [x_{ni} = a_{l'}]\,\mathbb{E}[z_{nk}]}_{= \sum_n \mathbb{E}[z_{nk}]}}$$

$$\boxed{\theta_{kil}} = p(x_i = a_l \mid \bar{x} \in C_k)$$

6 EM extensions

$$\mathbb{E}\left[\log p(X, Z \mid \theta)\right] \to \max \qquad D = \{\bar{x}_n\}_{n=1}^N$$

$$D' = \{(\bar{x}_n, \bar{z}_n)\}_{n=1}^N$$

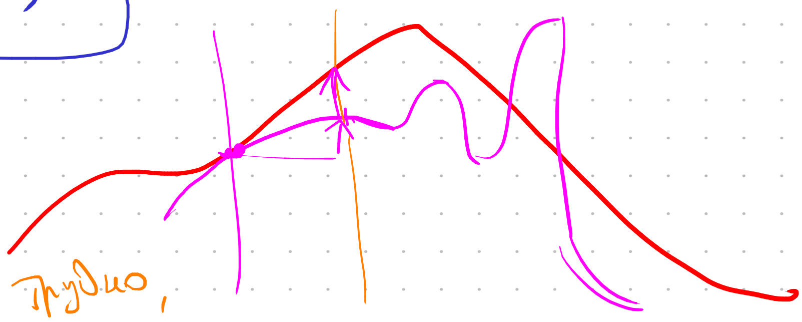
$D' = \{y_n\}_{n=1}^N \qquad p(y \mid \theta) \to \max$ — легко

$p(x \mid \theta) \to \max$ — трудно

$$\boxed{p(\bar{x}_n \mid \bar{y}_n, \bar{\theta}) = p(\bar{x}_n \mid \bar{y}_n)}$$ — удобно

$$\boxed{\bar{x}_n = f(\bar{y}_n)}$$

$\theta^{(m+1)}$



Generalized EM:

если $\bar{\theta}^{(m+1)} = \arg\max Q(\bar{\theta}, \bar{\theta}^{(m)})$ трудно,

то можно: $\bar{\theta}^{(m+1)}: \; Q(\bar{\theta}^{(m+1)}, \bar{\theta}^{(m)}) > Q(\bar{\theta}^{(m)}, \bar{\theta}^{(m)})$

$D' = \{(\bar{x}_n, \bar{z}_n)\}$

$$Q(\bar{\theta}, \bar{\theta}^{(m)}) = \mathbb{E}_Z\left[\log \prod_n p(\bar{x}_n, \bar{z}_n \mid \bar{\theta})\right] =$$

$$= \sum_{n=1}^N \mathbb{E}_{\bar{z}_n \mid \bar{x}_n, \bar{\theta}^{(m)}}\left[\log p(\bar{x}_n, \bar{z}_n \mid \bar{\theta})\right]$$

7 Randomized extensions

$$Q(\bar{\theta}, \bar{\theta}^{(m)}) = \mathbb{E}_{z \sim \underbrace{p(z \mid x, \bar{\theta}^{(m)})}_{\text{сэмплы}}}^{\text{sampling}}\left[\log p(x, z \mid \bar{\theta})\right]$$

$$z_1^{(m)}, z_2^{(m)}, \ldots, z_R^{(m)} \sim p(z \mid x, \bar{\theta}^{(m)}) \qquad R=1: \text{ Stochastic EM}$$

$$Q(\bar{\theta}, \bar{\theta}^{(m)}) \approx \frac{1}{R} \sum_{r=1}^R \log p(x, z_r^{(m)} \mid \bar{\theta}) \xrightarrow[\bar{\theta}]{} \max \qquad \text{Monte Carlo EM}$$

8 Priors

$$\log p(x|\bar\theta) \xrightarrow{\bar\theta} \max$$

θ_{MAP}: $\log p(\bar\theta|x) \to \max$?

$$= \log p(x|\bar\theta) + \log p(\bar\theta) + \text{Const} \xrightarrow{\bar\theta} \max$$

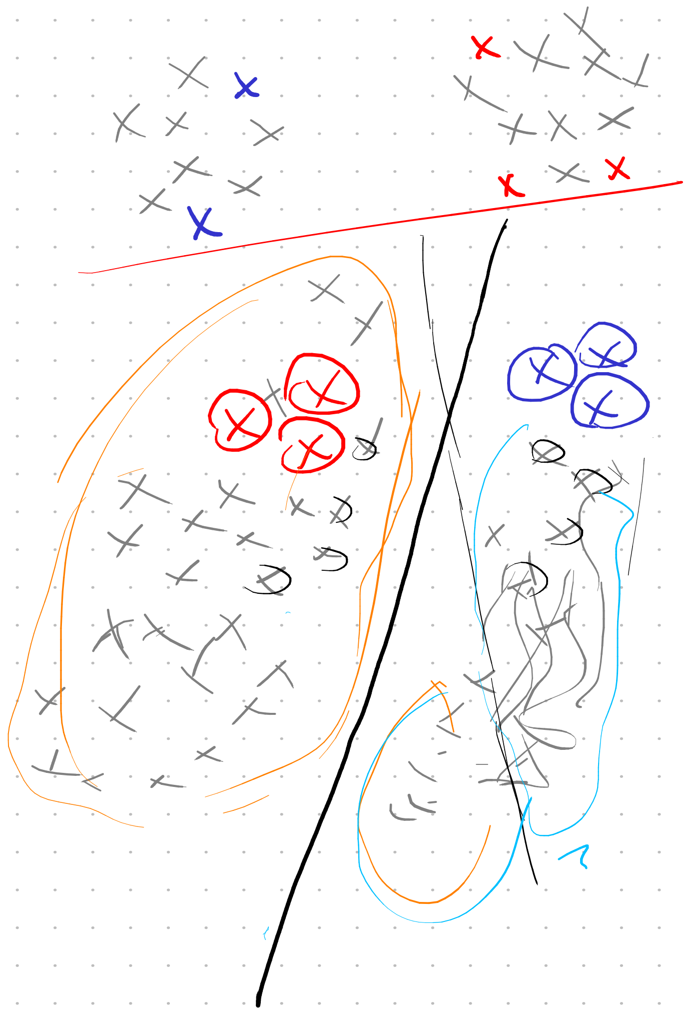
$$L(\bar\theta, \bar\theta^{(m)}) \le \log p(x|\bar\theta)$$

$$L(\bar\theta, \bar\theta^{(m)}) + \log p(\bar\theta) + \text{const} \le \log p(x|\bar\theta) + \log p(\bar\theta) + \text{const} = \log p(\bar\theta|x)$$

$$\boxed{\theta^{(m+1)} = \underset{\theta}{\arg\max}\left(Q(\bar\theta, \bar\theta^{(m)}) + \log p(\bar\theta)\right)}$$

5 Semi-supervised learning

$$Q(\bar\theta, \bar\theta^{(m)}) = E_z\left[\log p(x, z|\bar\theta)\right]$$



pseudolabels

(10) Bradley-Terry models

$1, \dots, n$ $\quad i \succ j,\ j \succ i$ $\quad i \longrightarrow \gamma_i$

если $\gamma_i > \gamma_j$, ∞ $\quad \underline{p(i \succ j) > 1/2}$

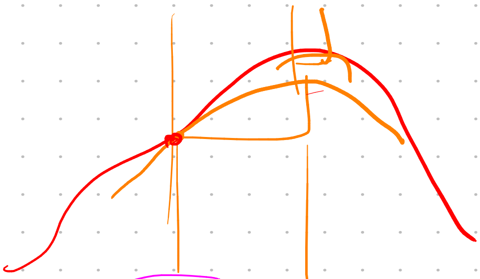
$$\boxed{p(i \succ j) = \frac{\gamma_i}{\gamma_i + \gamma_j} \qquad p(j \succ i) = \frac{\gamma_j}{\gamma_i + \gamma_j}}$$

мед.

$$D = \{(i,j)\} = \{w_{ij}\}_{i,j=1}^{n}$$

$$p(D \mid \bar\gamma) = \prod_{(i,j) \in D} \frac{\gamma_i}{\gamma_i + \gamma_j} = \prod_{i=1}^{n} \prod_{j=1}^{n} \left(\frac{\gamma_i}{\gamma_i + \gamma_j} \right)^{w_{ij}} \xrightarrow[\bar\gamma]{} \max$$

$$\log p(D \mid \bar\gamma) = \sum_i \sum_j w_{ij} \left(\log \gamma_i - \log(\gamma_i + \gamma_j) \right) \xrightarrow[\bar\gamma]{} \max$$

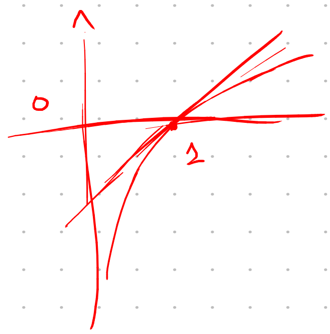


$\boxed{\bar\theta^{(m)}}$ $\quad \bar\theta^{(m+1)}$

MM-algorithms

minorization - maximization

$$\underline{\log x \le x - 1}$$



$$\log \frac{\gamma_i + \gamma_j}{\gamma_i^{(m)} + \gamma_j^{(m)}} \le \frac{\gamma_i + \gamma_j}{\gamma_i^{(m)} + \gamma_j^{(m)}} - 1$$

$$\log \gamma_i - \log(\gamma_i + \gamma_j) \ge$$

$$\ge \log \gamma_i - \frac{\gamma_i + \gamma_j}{\gamma_i^{(m)} + \gamma_j^{(m)}} + 1 + \log(\gamma_i^{(m)} + \gamma_j^{(m)})$$

$$\log p(D \mid \bar\gamma) \ge \sum_i \sum_j w_{ij} \left(\log \gamma_i - \frac{\gamma_i + \gamma_j}{\gamma_i^{(m)} + \gamma_j^{(m)}} + \text{const} \right) \xrightarrow[\bar\gamma]{} \max$$

$$\frac{\partial L}{\partial \gamma_k} = \sum_i \sum_j w_{ij} \left(\frac{\partial \log \gamma_i}{\partial \gamma_k} - \frac{\partial (\gamma_i + \gamma_j)}{\partial \gamma_k} \cdot \frac{1}{\gamma_i^{(m)} + \gamma_j^{(m)}} \right) = 0$$

$$\sum_{j=1}^{n} w_{kj}\left(\frac{1}{\gamma_k} - \frac{1}{\gamma_k^{(m)}+\gamma_j^{(m)}}\right) - \sum_{i=1}^{n} w_{ik}\cdot\frac{1}{\gamma_i^{(m)}+\gamma_k^{(m)}} = 0$$

Wins$_k$ Games$_{ki}$

$$\frac{\sum_j w_{kj}}{\gamma_k} = \sum_{i=1}^{n}\frac{w_{kj}+w_{ik}}{\gamma_i^{(m)}+\gamma_k^{(m)}}$$

$$\boxed{\gamma_k^{(m+1)} = \frac{\text{Wins}_k}{\sum_{i=1}^{n}\frac{\text{Games}_{ki}}{\gamma_i^{(m)}+\gamma_k^{(m)}}}}$$

Ehunsa $\theta > 1$

$$p(i>j) = \frac{\gamma_i}{\gamma_i+\theta\cdot\gamma_j}$$

$$p(j>i) = \frac{\gamma_j}{\gamma_j+\theta\cdot\gamma_i}$$

$$p(i=j) = 1 - \frac{\gamma_i}{\gamma_i+\theta\gamma_j} - \frac{\gamma_j}{\gamma_j+\theta\gamma_i}$$

$\log(\gamma_i+\theta\gamma_j)$

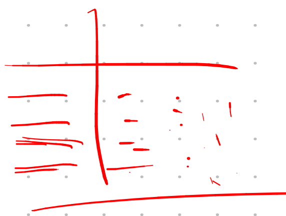
$\log(\gamma_j+\theta\gamma_i)$

$a > 1$

$$p(i>j,\ i\ \text{[illegible]}) = \frac{a\cdot\gamma_i}{a\cdot\gamma_i+\gamma_j}$$

[illegible]

$$\frac{\gamma_i}{\gamma_i+a\cdot\gamma_j}$$



$p(\pi_{1\ldots4})$