$$p(x,\theta) = p(x|\theta)\,p(\theta) \qquad X = (x_1, \ldots, x_n)$$

$$p(\theta|X) = \frac{p(X|\theta)\,p(\theta)}{\int p(X|\theta)\,p(\theta)\,d\theta} = \frac{\prod_{i=1}^{n} p(x_i|\theta)\,p(\theta)}{\int \prod_{i=1}^{n} p(x_i|\theta)\,p(\theta)\,d\theta} \qquad \varepsilon$$

$$\theta_{MP} = \arg\max_{\theta} p(\theta|X) = \arg\max_{\theta} \prod_{i=1}^{n} p(x_i|\theta)\,p(\theta) =$$

$$= \arg\max_{\theta}\left[\sum_{i=1}^{n} \log p(x_i|\theta) + \log p(\theta)\right]$$ — Poor man's Bayes

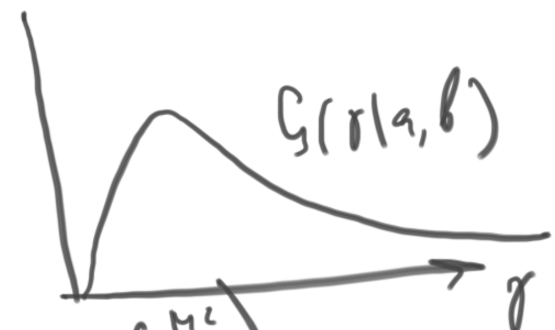
$$\left.\begin{array}{l} p(x|\theta) \in \mathcal{A}(\theta) \\ p(\theta) \in \mathcal{B}(\alpha) \end{array}\right\} \text{Conjugate} \iff \underline{p(\theta|X) \in \mathcal{B}(\alpha')}$$

$$\alpha \longrightarrow \alpha'$$

$$\begin{cases} p(x|\mu) = \mathcal{N}(x|\mu,1) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x-\mu)^2\right) \\ p(\mu) = \mathcal{N}(\mu|m,s^2) \end{cases}$$

$$\begin{cases} p(x|\gamma) = \mathcal{N}(x|0,\gamma^{-1}) = \sqrt{\frac{\gamma}{2\pi}}\exp\left(-\frac{\gamma}{2}x^2\right) & \gamma > 0 \\ p(\gamma) = \mathcal{G}(\gamma|a,b) = \frac{b^a}{\Gamma(a)}\gamma^{a-1}\exp(-b\gamma) & a,b > 0 \end{cases}$$

$$\boxed{\begin{aligned} p(x|\mu,\gamma) &= \mathcal{N}(x|\mu,\gamma^{-1}) \;\; (\triangleq) \\ p(\mu,\gamma) &= p(\mu)\,p(\gamma) \end{aligned}}$$



$$(\triangleq)\;\; \sqrt{\frac{\gamma}{2\pi}}\exp\left(-\frac{\gamma}{2}(x-\mu)^2\right) = \sqrt{\frac{\gamma}{2\pi}}\exp\left(-\frac{\gamma}{2}x^2 + \gamma\mu x - \gamma\frac{\mu^2}{2}\right)$$

$$p(\theta|X) = \frac{p(X|\theta)\,p(\theta)}{\int p(X|\theta)\,p(\theta)\,d\theta} \approx q(\theta) = \arg\min_{q \in Q} D\big(p(\theta|X), q(\theta)\big)$$

Intractable

$$KL\big(q(\theta)\,\|\,p(\theta|X)\big) \stackrel{def}{=} \int q(\theta) \log \frac{q(\theta)}{p(\theta|X)}\,d\theta \geq 0$$

$$= \iff q(\theta) = p(\theta|X)$$

$$\underbrace{\log p(x)}_{\text{не зависит от } q} = \underbrace{\int q(\theta) \log \frac{p(x,\theta)}{q(\theta)}\,d\theta}_{\mathcal{L}(q)} + \underbrace{\int q(\theta) \log \frac{q(\theta)}{p(\theta|X)}\,d\theta}_{KL(q(\theta)\,\|\,p(\theta|X))}$$

$$q(\theta) = \arg\max_{q \in Q} \mathcal{L}(q)$$

1) Mean-field $\quad q(\theta) = \prod_{j=1}^{m} q_j(\theta_j) \quad \theta = \bigcup_{j=1}^{m} \theta_j$

2) Parametric VI $\quad q(\theta) = q(\theta \mid \lambda)$

Conditional Conjugacy : $\theta = \theta_1 \cup \ldots \cup \theta_m$

$p(x \mid \theta_{\neg j}, \theta_j) \quad " \quad p(\theta_j)$ - conjugate

---

$$\log q_j(\theta_j) = \underbrace{E_{\theta_{\neg j}} \log p(x, \theta)}_{\text{tractable}} + \text{Const}$$

$$p(x|\mu,\gamma) = \mathcal{N}(x|\mu,\gamma^{-1})$$

$$p(\mu,\gamma) = p(\mu)\,p(\gamma) = \mathcal{N}(\mu|m,\beta^{-1})\,\mathcal{G}(\gamma|a,b)$$

$$X = (x_1 \ldots x_n) \qquad p(\mu,\gamma|X) \approx q(\mu)\,q(\gamma)$$

можно вынести за $\mathbb{E}$

$$\log q(\mu) = \mathbb{E}_{q(\gamma)} \log p(X,\mu,\gamma) = \mathbb{E}_\gamma\Big[\log p(X|\mu,\gamma) + \log p(\mu) + \log p(\gamma)\Big] =$$

Const

не зависит от μ

$$= \text{Const} \underbrace{- \frac{d}{2}\log 2\pi - \frac{\beta}{2}(\mu - m)^2}_{\log p(\mu)} + \mathbb{E}_\gamma\Big[\frac{n}{2}\log\gamma - \frac{n}{2}\log 2\pi -$$

$$- \frac{\gamma}{2}\sum_{i=1}^{n}(x_i - \mu)^2\Big] = \text{Const} - \frac{\beta}{2}(\mu - m)^2 - \frac{1}{2}\mathbb{E}\gamma \cdot \sum_{i=1}^{n}(x_i - \mu)^2 =$$

$$= \text{Const} - \frac{\beta}{2}\mu^2 + \beta\mu m - \frac{\beta}{2}m^2 - \frac{1}{2}\mathbb{E}\gamma\Big(\sum_i x_i^2 - 2\mu\sum x_i + n\mu^2\Big) =$$

$$= -\mu^2\Big(+\frac{\beta}{2} + n\mathbb{E}\gamma\Big) + \mu\Big(\beta m + \mathbb{E}\gamma\sum x_i\Big) + \text{Const}$$

$$q(\mu) = \mathcal{N}(\mu|m',\beta')$$

$$m' = \frac{\beta m + \sum x_i \, \mathbb{E}\gamma}{\beta + n\,\mathbb{E}\gamma} \qquad \beta' = \beta + n\,\mathbb{E}\gamma \qquad \mathbb{E}\gamma = \frac{a'}{b'} \qquad p(\gamma) = \frac{b^a}{\Gamma(a)} \gamma^{a-1} \exp(-b\gamma)$$

$$\log q(\gamma) = \mathbb{E}_{q(\mu)} \log p(X, \mu, \gamma) = \mathbb{E}_{q(\mu)}\left[\log p(X \mid \mu, \gamma) + \log p(\mu) + \log p(\gamma)\right]$$

Const

no depend. of μ

$$= a \log b - \log \Gamma(a) + (a-1)\log\gamma - b\gamma + \text{Const} +$$

$$+ \mathbb{E}_{q(\mu)} \sum_{i=1}^{n} \left[\frac{1}{2}\log\gamma - \frac{1}{2}\log 2\pi - \frac{\gamma}{2}(x_i - \mu)^2\right] =$$

$$p(X \mid \mu, \gamma) = \prod_{i=1}^{n} \sqrt{\frac{\gamma}{2\pi}} \exp\left(-\frac{\gamma}{2}(x_i - \mu)^2\right)$$

$$= (a + n - 1)\log\gamma - \gamma\left[b + \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}_{\mu}(x_i - \mu)^2\right] + \text{Const}$$

$$\boxed{q(\gamma) = G(\gamma \mid a', b')}$$

$$a' = a + n$$

$$b' = b + \frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - 2\,\mathbb{E}\mu \sum_{i=1}^{n} x_i + \underbrace{\mathbb{E}\mu^2}_{(m')^2 + (\beta')^{-1}}\right)$$

$p(X|\theta) \to \max_\theta$

$$p(X|\theta) = \int p(X,z|\theta)\,dz = $$
$$= \int p(x|z,\theta)\,p(z|\theta)\,dz$$

EM-algorithm

E-step $\quad q(z) = \arg\max_q \mathcal{L}(q,\theta) = \arg\min_{q\in Q} KL(q(z)\,\|\,p(z|X,\theta)) =$

$$= p(z|X,\theta) \approx \prod_{j=1}^{m} q_j(z_j)$$

M-step $\quad \theta = \arg\max_\theta \mathcal{L}(q,\theta) = \arg\max \mathbb{E}_{q(z)} \log p(X,z|\theta)$

$$p(X,z,\theta) = p(X,z|\theta)\,p(\theta) = p(X|z,\theta)\,p(z|\theta)\,p(\theta)$$

$$\theta_{MP} = \arg\max_\theta p(\theta|X) = \arg\max_\theta p(X|\theta)\,p(\theta) = \arg\max_\theta\Big[\underbrace{\log p(X|\theta)}_{\mathcal{L}(q,\theta)} + \log p(\theta)\Big]$$

$\log p(X|\theta) \geq \mathcal{L}(q,\theta)$

$\uparrow \ \uparrow$

M'-step $\quad \theta = \arg\max_\theta \Big[\mathbb{E}_{q(z)} \log p(X,z|\theta) + \log p(\theta)\Big]$

$$p(X,Z,\theta) = p(X|Z,\theta)\,p(Z|\theta)\,p(\theta)$$

Наблюдаем $X = (x_1 \dots x_n)$

Свойство модели	Алгоритм	Вид приближения
Полное сопряжение относит. Z, θ	Байесовский вывод	$p(Z,\theta\|X)$
Условное сопряжение на $Z\|\theta$ и на $\theta\|Z$	Mean-field	$q(Z)\,q(\theta) = \arg\min KL(\dots)$; $q(\theta) \in \Delta(\theta)$
Условное сопряж. на $Z\|\theta$	EM'	$p(Z\|X,\theta)\,\delta(\theta-\theta_{MP})$
Условное сопряж. на $\theta\|Z$	M'E	$\delta(Z-Z_{MP})\,p(\theta\|X,Z)$
Условное сопряж. на $z_j\|Z_{\setminus j},\theta$	MF EM'	$\prod_{j=1}^{m} q_j(z_j)\,\delta(\theta-\theta_{MP})$
Нет сопряженности	Байес для бедных	$\delta(Z-Z_{MP})\,\delta(\theta-\theta_{MP})$