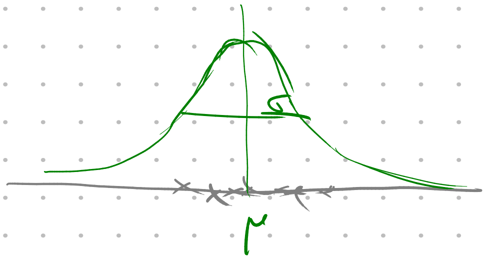
Байесовский вывод для гауссиана



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

$\tau = \frac{1}{\sigma^2}$

precision

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

$$\ln p(x_1 \dots x_n|\mu,\tau) = \sum \ln p(x_i|\mu,\tau) =$$

$$= \sum_{i=1}^{n}\left(\frac{1}{2}\ln\tau - \frac{1}{2}\ln 2\pi - \frac{\tau}{2}(x_i-\mu)^2\right)$$

1) $\underline{\tau = const}$

$$p(\mu) \times p(X|\mu) \propto p(\mu|X)$$

$$\underline{\ln p(\mu)} + \cancel{\frac{n}{2}\ln\tau} - \cancel{\frac{n}{2}\ln 2\pi} - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 = \underline{\ln p(\mu|X)} + const$$

$$p(\mu|\mu_0,\tau_0) = \mathcal{N}(\mu|\mu_0,\tau_0)$$

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

$$= const - \frac{1}{2}\left(\tau_0\mu^2 - 2\tau_0\mu\mu_0 + \tau\sum_{i=1}^{n}(\mu^2 - 2\mu x_i)\right) =$$

$$= const - \frac{1}{2}\left((\tau_0+n\tau)\mu^2 - 2\mu(\tau_0\mu_0 + \tau\sum x_i)\right) =$$

$$= const - \frac{\tau_0+n\tau}{2}\left(\mu - \frac{\tau_0\mu_0+\tau\cdot\sum x_i}{\tau_0+n\tau}\right)^2 \qquad \frac{1}{\sigma_n^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}$$

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

2) $\mu = const$

$$p(\tau) \times p(D|\tau) \propto p(\tau|D)$$

$$\ln p(\tau) + \underline{\frac{n}{2}\ln\tau} - \cancel{\frac{n}{2}\ln 2\pi} - \frac{\tau}{2}\sum_i (x_i-\mu)^2$$

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

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

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

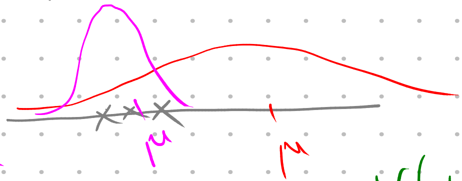
3) $p(\mu,\tau)=?$

$$p(\mu,\tau) \times p(D|\mu,\tau) \propto p(\mu,\tau|D)$$

$$\ln p(\mu,\tau) + \frac{n}{2}\ln\tau - \frac{n}{2}\ln 2\pi - \frac{\tau}{2}\sum_{i=1}^{n} (x_i-\mu)^2$$

$$p(\mu,\tau) = p(\mu)p(\tau) = N(\mu|\mu_0,\tau_0)\cdot Gamma(\tau|\alpha_0,\beta_0)$$

$$\ln p(\mu) + \ln p(\tau)$$



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

$$N(\mu|\mu', \lambda'\tau)$$

$$= Gamma(\tau|\alpha_0,\beta_0)\cdot N(\mu|\mu_0,\lambda_0\tau)$$

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

$$+ \frac{n}{2}\ln\tau - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 =$$

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

$$\lambda_0\mu^2 - 2\lambda_0\mu_0\mu + \lambda_0\mu_0^2 + \sum_i(\mu^2 - 2x_i\mu + x_i^2) =$$

$$= \mu^2(\lambda_0+n) - 2\mu\left(\lambda_0\mu_0+\sum x_i\right) + \lambda_0\mu_0^2 + \sum x_i^2 =$$

$$= (\lambda_0+n)\cdot\left(\mu - \frac{\lambda_0\mu_0+\sum x_i}{\lambda_0+n}\right)^2 + \lambda_0\mu_0^2 + \sum_i x_i^2 - \frac{(\lambda_0\mu_0+\sum x_i)^2}{\lambda_0+n}$$

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

$$- \tau\left(\beta_0 + \frac{1}{2}\left(\lambda_0\mu_0^2+\sum x_i^2 - \frac{(\lambda_0\mu_0+\sum x_i)^2}{\lambda_0+n}\right)\right)$$

$$\mathrm{Gamma}(\tau \mid \alpha_0, \beta_0) \cdot \mathcal{N}(\mu \mid \mu_0, \lambda_0 \tau) \quad \xrightarrow{D = \{x_1 \dots x_n\}}$$

$$\rightarrow \mathrm{Gamma}\left(\tau \mid \alpha_0 + \frac{n}{2}, \beta_0 + \frac{1}{2}(\dots)\right) \times$$

$$\times\ \mathcal{N}\left(\mu \mid \frac{\lambda_0\mu_0 + \sum x_i}{\lambda_0 + n}, (\lambda_0 + n)\tau\right)$$

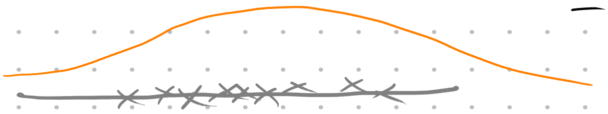
Σ^{-1}

$$p(\bar{x} \mid \bar{\mu}, \Lambda) = \sqrt{\frac{\det \Lambda}{(2\pi)^d}} \cdot e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T \Lambda (\bar{x}-\bar{\mu})}$$

4) $$p(x \mid x_1 \dots x_n) = \iint p(x \mid \mu, \tau)\, p(\mu, \tau \mid x_1 \dots x_n)\, d\mu\, d\tau =$$

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

$$= t\text{-Student}(x \mid \mu_{MAP}, n)$$



$$\ln p(x) = \ln p(x, z) - \ln p(z \mid x)$$

$$\ln p(x) = \underbrace{\int \ln \frac{p(x,z)}{q(z)}\, q(z)\, dz}_{} - \underbrace{\int \ln \frac{p(z \mid x)}{q(z)}\, q(z)\, dz}_{}$$

not calc   scarse

$$\ln p(x) = \underset{q\ \max}{L(q)} + \underset{q\ \min}{KL(q(z) \,\|\, p(z \mid x))}$$

$$q(z) = \prod q_i(z_i), \quad z_i \cap z_j = \emptyset$$

$$\ln q_i^*(z_i) = E_{q^*(z_{-i})}[\ln p(x, z)] + \text{const}$$

$$\mathcal{N}(\bar{x} \mid \bar{\mu}, \Lambda) \approx q(\bar{x}) = \underbrace{q_1(x_1)}_{}\underbrace{q_2(x_2)}_{}$$

Var. approx. for a Gaussian

$\mathrm{Gam}(\tau|\ldots)\,\mathcal{N}(\mu|\ldots,(\lambda\tau))$

$\shortparallel$

$D = \{x_1 \ldots x_n\}$

$$p(\tau,\mu) \times p(x_1 \ldots x_n|\tau,\mu) \propto p(\tau,\mu|D)$$

$$\mathrm{Gam}(\tau|\alpha_0,\beta_0)\,\mathcal{N}(\mu|\mu_0,\lambda_0\tau) \qquad \prod_i \mathcal{N}(x_i|\mu,\tau)$$

$$p(\tau,\mu|D) \approx q(\tau,\mu) = q_\mu(\mu)\cdot q_\tau(\tau)$$

$$\begin{cases} \ln q_\mu^*(\mu) = \mathbb{E}_\tau[\ln p(\tau,\mu,D)] + \text{const} \\ \ln q_\tau^*(\tau) = \mathbb{E}_\mu[\ln p(\tau,\mu,D)] + \text{const} \end{cases}$$

$$\ln p(\tau,\mu,D) = \text{const} + (\alpha_0 - 1)\ln\tau - \beta_0\tau + \frac{1}{2}\ln\tau - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 + \frac{n}{2}\ln\tau - \frac{\tau}{2}\sum_i (x_i-\mu)^2$$

$$\ln q_\mu^*(\mu) = \mathbb{E}_\tau\left[-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 - \frac{\tau}{2}\sum_i(x_i-\mu)^2\right] + \text{const}$$

$$= -\frac{\mathbb{E}_\tau[\tau]}{2}\cdot\left(\lambda_0(\mu-\mu_0)^2 + \sum_i(x_i-\mu)^2\right) + \text{const} =$$

$$\mu^2(\lambda_0+N) - 2\mu(\lambda_0\mu_0 + \Sigma x_i)$$

$$= -\frac{\mathbb{E}[\tau]}{2}(\lambda_0+n)\cdot\left(\mu - \frac{\lambda_0\mu_0+\Sigma x_i}{\lambda_0+n}\right)^2 + \text{const}$$

$$q_\mu^*(\mu) = \mathcal{N}\left(\frac{\lambda_0\mu_0+\Sigma x_i}{\lambda_0+n}, (\lambda_0+n)\cdot\mathbb{E}_{q^*}[\tau]\right)$$

$$\ln q_\tau^*(\tau) = \text{Const} + (\alpha_0-1)\ln\tau - \beta_0\tau + \frac{n+1}{2}\ln\tau - \frac{\tau}{2}\mathbb{E}_\mu\left[\lambda_0(\mu-\mu_0)^2 + \sum_i(x_i-\mu)^2\right]$$

$$q_\tau^*(\tau) = \mathrm{Gamma}\left(\tau \,\middle|\, \alpha_0 + \frac{n+1}{2}, \beta_0 + \frac{1}{2}\mathbb{E}_\mu\left[\lambda_0(\mu-\mu_0)^2 + \sum_i(x_i-\mu)^2\right]\right)$$

$p(\tau,\mu)$ — non-informative prior

$$\alpha_0 = \beta_0 = \mu_0 = \lambda_0 = 0$$

$$q_\mu^*(\mu) = \mathcal{N}\left(\mu \,\middle|\, \frac{1}{n}\sum x_i,\ n\cdot \mathbb{E}_\tau[\tau]\right)$$

$$\frac{n+1}{\mathbb{E}\tau}$$

$$q_\tau^*(\tau) = \mathrm{Gamma}\left(\tau \,\middle|\, \frac{n+1}{2},\ \frac{1}{2}\mathbb{E}_\mu\left[\sum_i (x_i-\mu)^2\right]\right)$$

$$\mathbb{E}_\tau[\tau] = \mathbb{E}_{\mathrm{Gamma}(\tau|\alpha,\beta)}[\tau] = \frac{\alpha}{\beta}$$

$$\mathbb{E}_{q_\tau}[\tau] = \frac{n+1/2}{1/2\ \mathbb{E}_\mu\left[\sum_i (x_i-\mu)^2\right]}$$

$$\frac{1}{\mathbb{E}\tau} = (\mathbb{E}_\mu[\mu])^2 + \mathrm{Var}[\mu]$$

$$\mathbb{E}_\mu[\mu] = \frac{1}{n}\sum x_i$$

$$\mathbb{E}_\mu\left[\sum_i (x_i-\mu)^2\right] = \sum_i x_i^2 - 2\sum x_i \cdot \mathbb{E}_\mu[\mu] + n\cdot \mathbb{E}_\mu[\mu^2] =$$

$$= \sum_i x_i^2 - 2\left(\sum x_i\right)\left(\frac{1}{n}\sum x_i\right) + n\left(\frac{1}{n^2}\left(\sum x_i\right)^2 + \frac{1}{n\cdot \mathbb{E}[\tau]}\right)$$

$$\mathbb{E}[\tau] = \frac{n+1}{\sum_i x_i^2 - \frac{2}{n}\left(\sum x_i\right)^2 + \frac{1}{n}\left(\sum x_i\right)^2 + \frac{1}{\mathbb{E}[\tau]}}$$

$$n+1 = 1 + \mathbb{E}[\tau]\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right)$$

$$\mathbb{E}[\tau] = \frac{n}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2}$$

$$\sigma^2 = \frac{1}{n}\sum_i \left(x_i - \frac{1}{n}\sum x_i\right)^2 = \frac{1}{n}\left(\sum x_i^2 - \frac{2}{n}\left(\sum x_i\right)^2 + \cancel{\frac{n}{n^2}\left(\sum x_i\right)^2}\right)$$

$$= \frac{1}{n}\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right)$$

∃kcn. cemeincho — Exponential family

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

natural parameters

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

1) Bernoulli:

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

$$p(x\mid\theta) = (1-\theta)\cdot e^{x\cdot\ln\frac{\theta}{1-\theta}}$$

log-odds

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

$$e^{\eta} = \frac{\theta}{1-\theta}$$

$$g(\eta) = \frac{1}{1+e^{\eta}}\qquad \theta = \frac{e^{\eta}}{1+e^{\eta}}$$

2) Gaussian

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

$$= \sqrt{\frac{\tau}{2\pi}}\, e^{-\frac{\tau}{2}x^2 + \tau\mu x - \frac{\tau}{2}\mu^2}$$

$$p(x\mid\mu,\tau) = \underbrace{\sqrt{\frac{\tau}{2\pi}}\cdot e^{-\frac{\tau}{2}\mu^2}}_{\text{"}g(\bar{\eta})}\cdot e^{\tau\mu x - \frac{\tau}{2}x^2}$$

$$h(x)=1\qquad \bar{\eta} = \begin{pmatrix}\tau\\ \tau\mu\end{pmatrix}$$

$$u(x) = \begin{pmatrix}-\frac{1}{2}x^2\\ x\end{pmatrix}$$

Conp. arp. poenp.:

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

$$\ln p(\bar{\eta}\mid\_) = \text{Const} + \nu\cdot\ln g(\bar{\eta}) + \nu\cdot(\bar{\eta}^T\bar{x})$$