Байесовский вывод для гауссиана

$$\cancel{y \sim \mathcal{N}(y \mid w^T x}, \sigma^2) \leftarrow \text{lin. regr.}$$

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

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

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

$$\mathcal{N}(x_1 \ldots x_n \mid \mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} e^{-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

① $\tau = \text{Const}$, $\mu = ?$

$$\log p(D \mid \mu) = \text{Const} - \frac{\tau}{2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$p(\mu) = \mathcal{N}(\mu \mid \mu_0, \tau_0) = \sqrt{\tau_0 / 2\pi}\; e^{-\tau_0/2\,(\mu-\mu_0)^2}$$

$$\log p(\mu \mid D) = \text{Const} - \frac{\tau}{2}\sum_{i=1}^{n}(x_i - \mu)^2 - \frac{\tau_0}{2}(\mu - \mu_0)^2 =$$

$$= \text{Const} - \frac{1}{2}\mu^2(\tau_0 + \tau \cdot n) + \mu\left(\tau_0\mu_0 + \tau \cdot \sum_i x_i\right) =$$

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

$$\tau_n = \tau_0 + \tau \cdot n \qquad \frac{1}{\sigma_n^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}$$

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

② $\mu = \text{Const}$, $\tau = ?$

$$\log p(D \mid \tau) = \text{Const} + \frac{n}{2}\log\tau - \frac{\tau}{2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$p(\tau \mid \alpha_0, \beta_0) = \boxed{\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}}\,\tau^{\alpha_0 - 1} e^{-\beta_0 \tau}$$

gamma distribution

$\tau \in (0, +\infty)$

$$\log p(\tau \mid x_1 \ldots x_n) = \text{Const} + \frac{n}{2}\log\tau - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + (\alpha_0 - 1)\tau - \beta_0\tau$$

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

③ $\mu, \tau = ?$ $\quad \log p(x_1 \dots x_n | \mu, \tau) = Const + \frac{n}{2}\log\tau - \frac{\tau}{2}\sum_{i=1}^{n} (x_i-\mu)^2$

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

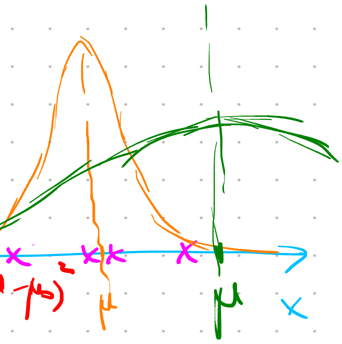
$$p(\mu,\tau) = \boxed{p(\mu)p(\tau)} = \mathcal{N}(\mu|\mu_0,\tau_0)\,Gam(\tau|\alpha_0,\beta_0)$$

$$\log p(\mu,\tau) = Const - \frac{\tau_0}{2}(\mu-\mu_0)^2 + (\alpha_0-1)\log\tau - \beta_0\tau$$

$Gam(\alpha_0,\beta_0)\,\mathcal{N}(\dots) \qquad \mathcal{N}(\mu_0, \frac{1}{\tau_0})\; Gam(--)$

$$p(\mu,\tau) = \underline{p(\tau)p(\mu|\tau)} = p(\mu)p(\tau|\mu)$$

$$\sqrt{\frac{\lambda_0\tau}{2\pi}}\, e^{-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2}$$



$$p(\mu,\tau) = Gam(\tau|\alpha_0,\beta_0)\cdot\mathcal{N}(\mu|\mu_0,\lambda_0\tau)$$

$$\log p(\mu,\tau) = Const + (\alpha_0-1)\log\tau - \beta_0\tau + \frac12\log\tau - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2$$

$$\log p(\mu,\tau|x_1\dots x_n) = Const + \frac n2\log\tau - \frac\tau2\sum(-)^2 + (\alpha_0-1)\log\tau - \beta_0\tau + \boxed{+\frac12\log\tau} - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 =$$

$$= Const + \overbrace{(\alpha_0+\frac n2 - 1)}^{\alpha_n}\log\tau + \frac12\log\tau - \frac\tau2\left(\lambda_0(\mu-\mu_0)^2 + \sum_i (x_i-\mu)^2\right) - \beta_0\tau$$

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

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

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

$$= Const + \overbrace{(\alpha_0+\frac n2-1)}^{\alpha_n}\log\tau + \frac12\log\tau - \frac{\overbrace{(\lambda_0+n)}^{\lambda_n}\tau}{2}\Big(\mu - \overbrace{\frac{\lambda_0\mu_0+\sum_i x_i}{\lambda_0+n}}^{\mu_n}\Big)^2$$

$$\underbrace{- \frac\tau2(\;\cdot\;) - \beta_0\tau}_{\beta_n}$$

$$p(\mu,\tau|x_1\dots x_n) = Gam(\tau|\alpha_n,\beta_n)\cdot\mathcal{N}(\mu|\mu_n,\lambda_n\tau), \text{ где}$$

$$\alpha_n = \alpha_0 + n/2 \qquad \beta_n = \beta_0 + \lambda_0\mu_0^2 + \sum_i x_i^2 - \frac{(\lambda_0\mu_0 + \sum_i x_i)^2}{\lambda_0 + n}$$

$$\mu_n = \frac{\lambda_0\mu_0 + \sum_i x_i}{\lambda_0 + n}, \qquad \lambda_n = \lambda_0 + n$$

(4) $$p(\mu \mid x_1 \ldots x_n) = \int_0^\infty p(\mu, \tau \mid x_1 \ldots x_n)\, d\tau =$$

$$= \int_0^\infty \mathrm{Gam}(\tau \mid \alpha_n, \beta_n) \cdot \mathcal{N}(\mu \mid \mu_n, \lambda_n \tau)\, d\tau =$$

$$= \mathrm{Const} \cdot \int_0^\infty \tau^{\alpha_n - 1} e^{-\beta_n \tau} \cdot \tau^{1/2} \cdot e^{-\frac{\lambda_n \tau}{2}(\mu - \mu_n)^2}\, d\tau$$

$$= \mathrm{Const} \int_0^\infty \tau^{(\alpha_n + \frac{1}{2}) - 1} \cdot e^{-\tau \cdot \left(\beta_n + \frac{\lambda_n}{2}(\mu - \mu_n)^2\right)}\, d\tau =$$

$$= \mathrm{Const} \cdot \frac{\Gamma(\alpha_n + 1/2)}{\left(\beta_n + \frac{\lambda_n}{2}(\mu - \mu_n)^2\right)^{\alpha_n + 1/2}} = p(\mu \mid x_1 \ldots x_n)$$

degrees of freedom

$-\frac{\nu+1}{2}$

Student's t-distribution $\quad p(x) = \mathrm{Const} \cdot \left(1 + \frac{x^2}{\nu}\right)$

(5) Predictive distribution

$$p(x \mid x_1 \ldots x_n) = \int_0^\infty \int_{-\infty}^\infty \mathcal{N}(x \mid \mu, \tau) \cdot \mathrm{Gam}(\tau \mid \alpha_n, \beta_n)\, \mathcal{N}(\mu \mid \mu_n, \lambda_n \tau)\, d\mu\, d\tau$$

$$\int \mathcal{N}(x \mid \mu)\, \mathcal{N}(\mu \mid \_)\, d\mu = \mathcal{N}(x \mid \mu', \underbrace{\tau'}_{\lambda'\tau})$$

$$= \int_0^\infty \mathcal{N}(x \mid \mu', \lambda'\tau) \cdot \mathrm{Gam}(\tau \mid \alpha_n, \beta_n)\, d\tau = \mathrm{Student}(x \mid \ldots)$$