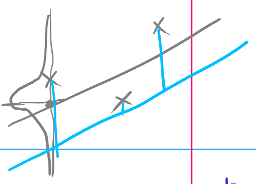


$$p(\theta|D) = \frac{\overbrace{p(\theta)}^{\text{prior}}\,\overbrace{p(D|\theta)}^{\text{likelihood}}}{p(D)}$$

$$p(D|\bar{w}) = p(\bar{y}|\bar{w}, X) = \prod_{n=1}^{N} p(y_n|\bar{w}, \bar{x}_n) = \prod_{n=1}^{N} \mathcal{N}(y_n|\bar{w}^T\bar{x}_n, \sigma^2).$$

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

$$p(\bar{w}) = \mathcal{N}(\bar{w}|\bar{\mu}_0, \Sigma_0)$$

$$\ln p(\bar{w}|D) = \text{const} - \frac{1}{2}(\bar{w}-\bar{\mu}_0)^T\Sigma_0^{-1}(\bar{w}-\bar{\mu}_0) - \frac{1}{2\sigma^2}\underbrace{\sum_{n=1}^{N}(y_n - \bar{w}^T\bar{x}_n)^2}_{} = (\bar{y}-X\bar{w})^T(\bar{y}-X\bar{w})$$

$$\begin{cases} \Sigma'^{-1} = \Sigma_0^{-1} + \frac{1}{\sigma^2}X^TX \\ \bar{\mu}' = \Sigma'\cdot\left(\Sigma_0^{-1}\bar{\mu}_0 + \frac{1}{\sigma^2}X^T\bar{y}\right) \end{cases}$$

$$D = \{(y_n, \bar{x}_n)\}_{n=1}^{N} \qquad p(\bar{w}), \qquad \bar{x}: \quad p(y|\bar{x}, D) = ?$$

$$p(y|\bar{x}, D) = \int p(y, \bar{w}|\bar{x}, D)\,d\bar{w} = \int \underbrace{p(y|\bar{w}, \bar{x})}_{\mathcal{N}(y|\bar{w}^T\bar{x}, \sigma^2)} \cdot \underbrace{p(\bar{w}|D)}_{\mathcal{N}(\bar{w}|\bar{\mu}_N, \Sigma_N)}\,d\bar{w} = ?(y)$$

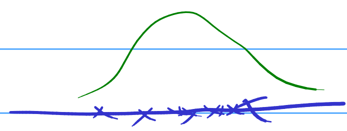
$$\frac{1}{\sqrt{2\pi a}}\int e^{-\frac{1}{2a}(x-b)^2}dx = 1$$

$$\ln p(y, \bar{w}|\bar{x}, D) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\bar{w}^T\bar{x} - y)^2 - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln\det\Sigma_N - \frac{1}{2}(\bar{w}-\bar{\mu}_N)^T\Sigma_N(\bar{w}-\bar{\mu}_N)$$

$$(\sqrt{2\pi})^d\sqrt{\det(\cdot)} = \int e^{\left[-\frac{1}{2}(\bar{w}-?)^T(\cdot)^{-1}(\bar{w}-?)\right]}d\bar{w}$$

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

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



$$D = \{x_n\}_{n=1}^{N}$$

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

① $\tau = \text{const}$

$$p(D|\mu) = \prod_n p(x_n|\mu) = \left(\cancel{\sqrt{\frac{\tau}{2\pi}}}\right)^N e^{-\frac{\tau}{2}\sum_{n=1}^{N}(x_n-\mu)^2}$$

$$\boxed{p(\mu) = \mathcal{N}(\mu|\mu_0, \tau_0) = \cancel{\sqrt{\frac{\tau_0}{2\pi}}}\, e^{-\frac{\tau_0}{2}(\mu-\mu_0)^2}}$$

$$\ln p(\mu|D) = \text{const} - \frac{\tau}{2}\sum_n(x_n-\mu)^2 - \frac{\tau_0}{2}(\mu-\mu_0)^2 =$$

$$-\frac{\tau}{2}\left(N\cdot\mu^2 - 2\mu\cdot\sum x_n + \sum x_n^2\right)$$

$$= \text{const} - \frac{\tau N}{2}\mu^2 + \mu\cdot\tau\cdot\sum x_n - \frac{\tau_0}{2}\mu^2 + \tau_0\mu_0\cdot\mu$$

$$= \text{const} - \frac{1}{2}\underline{(\tau_0 + \tau\cdot N)}\,\mu^2 + \mu\cdot\underline{(\tau_0\mu_0 + \tau\cdot\sum x_n)}$$

$$= \text{const} - \frac{\tau_0+\tau N}{2}\left(\mu - \frac{\tau_0\mu_0 + \tau\cdot\sum x_n}{\tau_0+\tau N}\right)^2$$

$$\begin{cases} \tau_N = \tau_0 + \tau\cdot N \\ \mu_N = \dfrac{\tau_0\mu_0 + \tau\cdot\sum x_n}{\tau_0 + \tau N} \end{cases}$$

(2) $\mu = \text{const}$, $\tau = ?$

$$p(D|\tau) = \left(\sqrt{\frac{\tau}{2\pi}}\right)^N \cdot e^{-\frac{\tau}{2}\cdot\sum(x_n-\mu)^2}$$

$$\ln p(D|\tau) = \text{const} + \frac{N}{2}\ln\tau - \tau\cdot\frac{1}{2}\sum(x_n-\mu)^2$$

$$\ln p(\tau) = \text{const} + (a_0-1)\ln\tau - b_0\tau$$

$$\boxed{p(\tau) \propto \tau^{a_0-1}\, e^{-b_0\tau}}$$

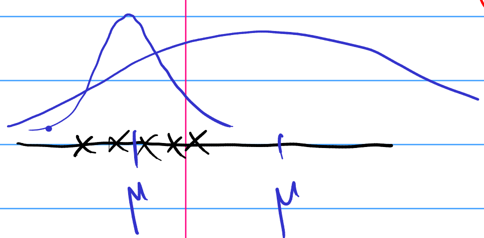
$$\ln p(\tau|D) = \text{const} + \frac{N}{2}\ln\tau - \frac{\tau}{2}\sum(\_)^2 + (a_0-1)\ln\tau - b_0\tau$$

$$a_N = a_0 + \frac{N}{2} \qquad b_N = b_0 + \frac{1}{2}\sum_{n=1}^{N}(x_n-\mu)^2$$

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

$$p(\mu,\tau) = \cancel{p(\mu)p(\tau)} = \mathcal{N}(\mu|\mu_0,\tau_0)\,\text{Gamma}(\tau|a_0,b_0)$$

$$\ln p(\mu,\tau) = \text{const} - \dots\mu\dots\mu^2\dots - \dots\ln\tau - \dots\tau \quad -\mu\tau \quad \mu\tau^2$$



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

$$p(\mu,\tau) = p(\tau)\,p(\mu|\tau) = \text{Gamma}(\tau|a_0,b_0)\,\mathcal{N}(\mu\,|\,\mu_0, D_0\tau)$$
$$= p(\mu,\tau\,|\,a_0,b_0,\mu_0,D_0)$$