1. It's given that $w \sim N(0, b^{-1}I)$.

Assume that $y \sim N(Aw, \sigma^2)$ where A is the design matrix and σ is the unknown variable (variance).

Let $\sigma^2 = a^{-1}$.

Input data

$$\because P(w|D) \propto P(D|w) \cdot P(w) \quad \text{(Bayes' rule)}$$

$$\propto e^{-\frac{1}{2\sigma^2}(Aw-y)^T(Aw-y)} \cdot e^{-\frac{1}{2} w^T (b^{-1}I)^{-1} w}$$

$$= e^{-\frac{a}{2}\left[(Aw-y)^T(Aw-y) + \frac{b}{a} w^T w\right]}$$

$$= e^{-\frac{a}{2}\left(w^T A^T A w - 2 w^T A^T y + y^T y + \frac{b}{a} w^T w\right)}$$

It's a scalar, so $(w^T A^T y)^T = y^T A w = w^T A^T y$

$$= e^{-\frac{1}{2}\left[w^T (aA^TA + bI) w - 2 a w^T A^T y + a y^T y\right]}$$

Let $\Lambda = aA^TA + bI$

$w^T \Lambda \mu = w^T a A^T y$

$\Rightarrow \mu = a \Lambda^{-1} A^T y$

對應的 quadratic form：$(\mathbf{w}-\boldsymbol{\mu})^T \Lambda (\mathbf{w}-\boldsymbol{\mu})$
$= \mathbf{w}^T \Lambda \mathbf{w} - 2\mathbf{w}^T \Lambda \boldsymbol{\mu} + \boldsymbol{\mu}^T \boldsymbol{\mu}$

$$\propto e^{-\frac{1}{2}(w^T \Lambda w - 2 w^T \Lambda \mu + \mu^T \mu)} + \text{constant}$$

normalization term

$$\therefore P(w|D) \sim N(\mu, \Lambda^{-1}) = N(a\Lambda^{-1}A^Ty, (aA^TA+bI)^{-1})$$

$$= N(a(aA^TA+bI)^{-1}A^Ty, (aA^TA+bI)^{-1})$$

$$= N((A^TA+\tfrac{b}{a}I)^{-1}A^Ty, (aA^TA+bI)^{-1})$$

which is also a multivariate gaussian distribution.

Appendix : Sequential Estimator

$$\mu_n = \frac{x_1 + x_2 + \dots + x_n}{n} \quad , \quad \mu_{n+1} = \frac{x_1 + x_2 + \dots + x_{n+1}}{n+1}$$

$$\Rightarrow \mu_{n+1} = \frac{n\mu_n + x_{n+1}}{n+1} = \frac{(n+1)\mu_n + x_{n+1} - \mu_n}{n+1} = \mu_n + \frac{x_{n+1} - \mu_n}{n+1}$$

M2：平方和

$$M2_{n+1} = \sum_{i=1}^{n+1} (x_i - \mu_{n+1})^2 = \sum_{i=1}^{n} (x_i - \mu_n + \mu_n - \mu_{n+1})^2 + (x_{n+1} - \mu_{n+1})^2$$

$$= \sum_{i=1}^{n} (x_i - \mu_n)^2 + \underbrace{\cancel{\sum_{i=1}^{n} 2 \cdot (x_i - \mu_n)(\mu_n - \mu_{n+1})}}_{=0} + \sum_{i=1}^{n} (\mu_n - \mu_{n+1})^2$$

$$+ (x_{n+1} - \mu_{n+1})^2$$

$$= M2_n + \sum_{i=1}^{n} (\mu_n - \mu_{n+1})^2 + (x_{n+1} - \mu_{n+1})^2$$

$$= M2_n + n \cdot \frac{(x_{n+1} - \mu_n)^2}{(n+1)^2} + \frac{n^2 (x_{n+1} - \mu_n)^2}{(n+1)^2}$$

$$= M2_n + \frac{n \cdot (x_{n+1} - \mu_n)^2}{n+1}$$

$$= M2_n + (x_{n+1} - \mu_n) \cdot \frac{n}{n+1} \cdot (x_{n+1} - \mu_n)$$

$$(n+1)\mu_{n+1} = (n+1)\mu_n + x_{n+1} - \mu_n$$

$$= M2_n + (x_{n+1} - \mu_n) \cdot \frac{n x_{n+1} - n\mu_n}{n+1}$$

$$= M2_n + (x_{n+1} - \mu_n) \cdot \frac{n x_{n+1} - [(n+1)\mu_{n+1} - x_{n+1}]}{n+1}$$

$$= M2_n + (x_{n+1} - \mu_n) \cdot (x_{n+1} - \mu_{n+1})$$

※