$y = Aw + \varepsilon$ $\qquad\qquad N(\mu, \sigma^2) \sim e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

We assume our likelihood $P(D|w) \sim N(Aw, a^{-1})$, where A is a design matrix, a is a precision

Also, we have prior $P(w) \sim N(0, b^{-1}I)$

$\therefore$ posterior $P(w|D) \propto N(Aw, a) \cdot N(0, b^{-1}I)$

$$\propto e^{-\frac{a}{2}(Aw-y)^T(Aw-y)} \cdot e^{-\frac{b}{2}(\vec{w}-\vec{0})^T I(\vec{w}-\vec{0})}$$

$$= e^{-\frac{a}{2}(Aw-y)^T(Aw-y) - \frac{b}{2}(\vec{w}-\vec{0})^T I(\vec{w}-\vec{0})} \quad \text{--- } ①$$

We know posterior is a normal distribution

if $P(w|D) \sim N(\mu, \Lambda^{-1})$, then

$$P(w|D) \propto e^{-\frac{1}{2}(w-\mu)^T \Lambda (w-\mu)}$$

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

for ①, we complete the square of power term

$$-\frac{1}{2}\left(a(Aw-y)^T(Aw-y) + b\,w^Tw\right)$$

$$= -\frac{1}{2}\left(a(w^TA^TAw - w^TA^Ty - y^TAw + y^Ty) + bw^Tw\right)$$

$$= -\frac{1}{2}\left(a\,w^TA^TAw - 2aw^TA^Ty + ay^Ty + w^T(bI)w\right)$$

$$= -\frac{1}{2}\left(w^T(aA^TA + bI)w - 2aw^TA^Ty + ay^Ty\right) \quad \text{--- } ③$$

if we compare ② and ③ we can find that

$$\begin{cases} \Lambda = aA^T A + bI \\ \Lambda \mu = aA^T y \end{cases}$$

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

$$= a\left(aA^T A + bI\right)^{-1} A^T y = \left(A^T A + \frac{b}{a} I\right)^{-1} A^T y$$

$$\sigma^2 = \Lambda^{-1} = \left(aA^T A + bI\right)^{-1}$$

#