1.2 $\theta = (\theta_0, \dots \theta_i, \dots \theta_p)^T$

$$P(\theta) = \frac{1}{\sqrt{|2\pi \Sigma_0|}} \exp\left(-\frac{1}{2}(\theta-\mu_0)^T \Sigma_0^{-1} (\theta-\mu_0)\right)$$

Σ_0 prior covariance

μ_0 prior mean

set $\Sigma_0 = I$, $\mu_0 = 0$

$$\rightarrow p(\theta) = \frac{1}{\sqrt{|2\pi I|}} \exp\left(-\frac{1}{2}\theta^T\theta\right)$$

a) $$P(\theta|y) = \frac{P(y|\theta)\,P(\theta)}{P(y)} \propto P(y|\theta)\,p(\theta)$$

$$P(y|\theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \prod_{i=1}^N \exp\left(-\frac{(y_i - (\theta_0 + \theta_1 x_i + \dots + \theta_p x_i^p))^2}{2\sigma^2}\right)$$

$$P(\theta|y) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \prod_{i=1}^N \exp\left(-\frac{(y_i - (\theta_0 + \theta_1 x_i + \dots + \theta_p x_i^p))^2}{2\sigma^2}\right) \frac{1}{\sqrt{|2\pi I|}} \exp\left(-\frac{1}{2}\theta^T\theta\right)$$

$$\underline{\mathrm{Log}(P(\theta|y))} = -N\log(\sqrt{2\pi}\sigma) - \sum_{i=1}^N \frac{(y_i - (\theta_0 + \theta_1 x_i + \dots + \theta_p x_i^p))^2}{2\sigma^2} - \frac{1}{2}\theta^T\theta - \log\left(\sqrt{|2\pi I|}\right)$$

b) derive MAP estimate $\rightarrow \frac{\partial \mathrm{Log}(\dots)}{\partial\theta} \overset{!}{=} 0$

$$\frac{\partial \log(P(\theta|y))}{\partial\theta} = \frac{\partial}{\partial\theta}\Big(\overbrace{-N\log(\sqrt{2\pi}\sigma)}^{\text{vanish}} - \frac{1}{2}\theta^T\theta - \overbrace{\log\sqrt{|2\pi\Sigma_0|}}^{\text{vanish}} - \sum_{i=1}^N \frac{(y_i - (\theta_0 + x_i\theta_1 + \dots + \theta_p x_i^p))^2}{2\sigma^2}\Big)$$

$y \in \mathbb{R}^N$

$\theta \in \mathbb{R}^p$

$X \in \mathbb{R}^{N \times p}$

$$= \frac{\partial}{\partial\theta}\left(-\frac{1}{2}\theta^T\theta - \frac{1}{2\sigma^2}\|y - X\theta\|^2\right)$$

$$= -\theta - \frac{1}{2\sigma^2}\frac{\partial}{\partial\theta}\left((y-X\theta)^T(y-X\theta)\right)$$

$$= -\theta - \frac{1}{2\sigma^2}\frac{\partial}{\partial\theta}\left(y^Ty - y^TX\theta - \theta^TX^Ty + \theta^T\underbrace{X^TX}_{A}\theta\right)$$

$$= -\theta - \frac{1}{2\sigma^2}(-2X^Ty + 2A\theta) = -\theta - \frac{1}{2\sigma^2}(-2X^Ty + 2X^TX\theta)$$

$\overset{!}{=} 0$

$$-\theta - \frac{1}{\sigma^2}\left(-X^T y + X^T X \theta\right) = 0$$

$$-\theta\left(I + \frac{1}{\sigma^2} X^T X\right) + \frac{1}{\sigma^2} X^T y = 0$$

$$\Rightarrow \boxed{\hat{\theta} = \left(X^T X + \underbrace{\sigma^2 I}_{\text{shift of } \sigma^2}\right)^{-1} X^T y}$$