1.

$$J(\beta) = RSS(\beta) + \phi(\beta)$$

$$RSS(\beta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad \text{with} \quad y_i = \sum_{j=0}^{d} x_{ij}\beta_j$$

$$\phi(\beta) = \beta^T R \beta = \sum_{i=1}^{n} \beta_i \sum_{j=1}^{n} R_{ij}\beta_j$$

$$0 = \frac{\partial J(\beta)}{\partial \beta_j} = -2\sum_{i=1}^{n} \left[(y_i - \hat{y}_i) x_{ij} + \frac{1}{2}\beta_i R_{ij} (\beta_j) \right]$$

$$= -2 \left[\underline{X}^T (y - \underline{X}\beta) + \frac{1}{2}\beta (R + R^T) \right]_j$$

$$0 = X^T (\underline{y} - \underline{X}\beta) + \frac{1}{2}\beta (\underline{R} + \underline{R}^T)$$

$$(\underline{X}^T \underline{X} + \frac{1}{2}\underline{R} + \frac{1}{2}\underline{R}^T)\beta = \underline{X}^T y$$

$$\beta_{opt} = (\underline{X}^T \underline{X} + \frac{1}{2}\underline{R} + \frac{1}{2}\underline{R}^T)^{-1} \underline{X}^T \underline{y}$$

2.

① $\prod_{i=1}^{n} p(y_i|\vec{x},\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y_i - e^{\theta}x_i)^2)$

② $loss = -\sum_{i=1}^{n} \ln p(y_i|x_i,\theta)$

$0 = \frac{\partial loss}{\partial \theta} = -\sum_{i=1}^{n} \frac{\partial \ln p(y_i|x_i,\theta)}{\partial \theta}$

$= -\sum_{i=1}^{n} \frac{1}{p(y_i|x_i,\theta)} \frac{\partial p(y_i|x_i,\theta)}{\partial \theta}$

$0 = \frac{\partial p}{\partial \theta} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{(-\frac{1}{2\sigma^2}(y_i - e^{\theta}x_i)^2)}$

$0 = -\sum \frac{1}{p(y_i|x_i,\theta)} \cdot \frac{\partial p}{\partial \theta} = \sum_{i=1}^{n} \frac{1}{\sigma^2}(y_i - e^{\theta}x_i)e^{\theta}x_i$

$0 = -\sum_{i=1}^{n} \frac{1}{\sigma^2}(y_i - e^{\theta}x_i)e^{\theta}x_i = 0$

$\Rightarrow e^{\theta}\sum_{i=1}^{n}(x_iy_i - e^{\theta}x_i^2) = 0$

$\Rightarrow \sum_{i=1}^{n} x_iy_i - e^{\theta}x_i^2 = 0$

$e^{\theta} = \sum_{i=1}^{n} \frac{x_iy_i}{x_i^2}$

(a)

3. $P_Y(y|\sigma^2)$

$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y-1)^2)$

$nll = -\sum_{i=1}^{n} \ln P_Y(y_i|\sigma)$

$\frac{nll}{\partial\sigma} = -\sum_{i=1}^{n} \frac{1}{p(y_i|\sigma)} \cdot \frac{\partial p(y_i|\sigma)}{\partial\sigma}$

$\frac{\partial p(y_i|\sigma)}{\partial\sigma} = -\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y_i-1)^2)$
$+ \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y_i-1)^2) \cdot (y_i-1)^2 \frac{1}{\sigma^3}$

$\frac{\partial nll}{\partial\sigma} = -\sum_{i=1}^{n}(-\frac{\sqrt{2\pi}}{\sqrt{2\pi\sigma^2}} + (y_i-1)^2 \frac{1}{\sigma^3}) = 0$

$= \sum_{i=1}^{n}(-\frac{1}{\sigma} + (y_i-1)^2 \frac{1}{\sigma^3}) = 0$

$\frac{1}{\sigma}\sum_{i=1}^{n}(-1 + (y_i-1)^2\frac{1}{\sigma^2}) = 0$

$\sum_{i=1}^{n} \frac{1}{\sigma^2}(y_i-1)^2) = n \quad \sigma^2 = \frac{\sum_{i=1}^{n}(y_i-1)^2}{n} \quad \sigma = \sqrt{\frac{\sum_{i=1}^{n}(y_i-1)^2}{n}}$

(b)

3. ② replace σ^2 with x

$p(y_i|\sigma^2) = \frac{1}{\sqrt{2\pi x}} \exp(-\frac{1}{2x}(y_i-1)^2)$

$nll = -\sum_{i=1}^{n} \ln p(y_i|x)$

$\frac{\partial nll}{\partial x} = -\sum_{i=1}^{n} \frac{1}{p(y_i|x)} \cdot \frac{\partial p}{\partial x}$

$\frac{\partial p}{\partial x} = -\frac{1}{2\sqrt{2\pi}} x^{-\frac{3}{2}} \exp(\ldots) + \frac{1}{\sqrt{2\pi x}} \exp(\ldots) \frac{1}{2} x^{-2}(y-1)^2$

$\frac{\partial nll}{\partial x} = -\sum_{i=1}^{n} -\frac{\sqrt{2\pi}}{2\sqrt{2\pi}} x^{-1} + \frac{1}{2} x^{-2}(y_i-1)^2 = 0$

$\frac{1}{2}x^{-1}\sum_{i=1}^{n}(-1 + x^{-1}(y_i-1)^2) = 0$

$\sum_{i=1}^{n}(x^{-1}(y_i-1)^2 - 1) = 0$

$\sum_{i=1}^{n} x^{-1}(y_i-1)^2 = n$

$x^{-1} = \frac{n}{\sum_{i=1}^{n}(y_i-1)^2} \Rightarrow x = \frac{\sum_{i=1}^{n}(y_i-1)^2}{n}$

so $\sigma^2 = \frac{\sum_{i=1}^{n}(y_i-1)^2}{n}$

It agrees with answer of part (a)