EM for Ridge different parametrizations

(1.) $y \sim N(Xb, s^2)$ $Y \sim N(0, X^TXs_b^2 + s^2 I)$

$b \sim N(0, s_b^2 I)$ $= N(0, s_b^2(X^TX + \frac{s^2}{s_b^2}))$

$$\log p(y, b \mid s, s_b) = -\frac{p}{2}\log 2\pi s_b^2 - \frac{1}{2s_b^2}\|b\|_2^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\|y - Xb\|_2^2$$

$$= -\frac{p}{2}\log 2\pi s_b^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2}\left[\frac{y^Ty}{s^2} + b^T\left(\frac{X^TX}{s^2} + \frac{I}{s_b^2}\right)b - 2\frac{y^TXb}{s^2}\right]$$

$$= -\frac{p}{2}\log 2\pi s_b^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\left[y^Ty + b^T\left(X^TX + \frac{s^2}{s_b^2}I\right)b - 2y^TXb\right]$$

$b \mid y \sim N(\mu_1, \Sigma_1)$

$\Sigma_1^{-1} = \frac{1}{s^2}(X^TX + \frac{s^2}{s_b^2}I)$

$\Sigma_1^{-1}\mu_1 = \frac{X^Ty}{s^2}$

$\mu_1 = (X^TX + \frac{s^2}{s_b^2}I)^{-1}X^Ty \quad = \underset{p\times p}{V^{-1}}\ \underset{p\times n}{X^T}\ \underset{n\times 1}{y}$

$\Sigma_1 = s^2(X^TX + \frac{s^2}{s_b^2}I)^{-1} \quad = s^2V^{-1}$

$$E(\log p(y, b \mid s, s_b)) = -\frac{p}{2}\log 2\pi s_b^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\left[y^Ty + \mathrm{tr}\left(\left(X^TX + \frac{s^2}{s_b^2}I\right)\left[\mu_1\mu_1^T + \Sigma_1\right]\right) - 2y^TX\mu_1\right]$$

$$\frac{\partial}{\partial s^2} = -\frac{n}{2s^2} + \frac{1}{2(s^2)^2}\left[y^Ty + \mathrm{tr}\left((X^TX)(\mu_1\mu_1^T + \Sigma_1)\right) - 2y^TX\mu_1\right] \Rightarrow \hat{s}^2 = \frac{1}{n}\left[E\|y - Xb\|_2^2\right]$$

$$\frac{\partial}{\partial s_b^2} = -\frac{p}{2s_b^2} - \frac{1}{2(s_b^2)^2}E\|b_2\|^2 \Rightarrow \hat{s}_b^2 = \frac{1}{p}E\|b_2\|^2 = \frac{1}{p}\mathrm{tr}\left[\mu_1\mu_1^T + \Sigma_1\right]$$

---

(2) $y \sim N(s_bXb, s^2)$

$b \sim N(0, I)$

$$\log p(y, b \mid s, s_b) = \frac{p}{2}\log 2\pi - \frac{1}{2}\|b\|_2^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\|y - s_bXb\|_2^2$$

$b \mid y \sim N(\tilde{\mu}_1, \tilde{\Sigma}_1)$

$\tilde{\Sigma}_1^{-1} = \frac{1}{s^2}(s_b^2X^TX + s^2 I) = \frac{s_b^2}{s^2}(X^TX + \frac{s^2}{s_b^2}I)$

$\tilde{\mu}_1 = \frac{s_b}{s^2}(\frac{s_b^2}{s^2}X^TX + I)^{-1}X^Ty$

$= \frac{1}{s_b}(X^TX + \frac{s^2}{s_b^2}I)^{-1}X^Ty \left[= \frac{1}{s_b}\mu_1^{\text{above}}\right]$

$\tilde{\Sigma}_1^{-1} \tilde{\mu}_1 = \frac{X^T y\, s_b}{s^2} \quad = \left(\frac{s_b}{s^2} X^T X + I\right) \qquad \tilde{\Sigma}_1 = \frac{s^2}{s_b^2}\left(X^T X + \frac{s^2}{s_b^2} I\right)^{-1} \quad \left[= \frac{1}{s_b^2}\Sigma_1 \text{ as }\right]$

$\frac{\partial}{\partial s^2} = -\frac{n}{2s^2} + \frac{1}{2(s^2)^2} E\, \|y - s_b X b\|_2^2 \qquad \hat{s}^2 = \frac{1}{n} E\| y - s_b X b\|_2^2$

$\frac{\partial}{\partial s_b} = \frac{-1}{2s^2}\left[-2\mu_1 X^T y \quad + 2 s_b \; E\left[b^T X^T X b\right]\right] \Rightarrow \hat{s}_b = \frac{\mu_1 X^T y}{E(b^T X^T X b)}$ $= \mathrm{tr}(X^T X\, E(bb^T))$

---

(3) $y = s_y\left(h X b + \sqrt{1-h^2}\, E\right) \qquad e \sim N(0, I) \quad b \sim N(0, I) \qquad s^2 := s_y^2(1-h^2)$

$\log p(y, b \mid s_y, h^2) = \frac{p}{2}\log 2\pi - \frac{1}{2}\|b\|_2^2 - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\|y - s_y h X b\|_2^2$

$b \mid y \sim N(\tilde{\tilde{\mu}}_1, \tilde{\tilde{\Sigma}}_1) \qquad \tilde{\tilde{\mu}}_1 = \frac{1}{s_y h}\mu_1 \quad \tilde{\tilde{\Sigma}}_1 = \frac{1}{s_y^2 h^2}\Sigma_1$

$\log p(y, b \mid s_y, h^2) \;=\; -\frac{n}{2}\log s_y^2 - \frac{n}{2}\log(1-h^2) - \frac{1}{2s_y^2(1-h^2)}\Big(y^T y - 2 s_y h\, b^T X^T y$
$+ s_y^2 h^2 b^T X^T X b\Big)$

$= -\frac{n}{2}\log s_y^2 - \frac{n}{2}\log(1-h^2) - \frac{1}{2s_y^2(1-h^2)} y^T y + \frac{1}{2s_y}\frac{h}{1-h^2} b^T X^T y - \frac{1}{2}\frac{h^2}{1-h^2}\left(b^T X^T X b\right)$

$\frac{\partial}{\partial s_y} = -\frac{n}{s_y} + \frac{2y^T y}{2s_y^3(1-h^2)} - \frac{1}{2s_y^2}\frac{h}{1-h^2} b^T X^T y$

quadratic in s_y ...

$= 0$ at $s_y^2 =$

---

(4) $y \sim s(Xb + e) \qquad b \sim N\left(0, \frac{s_b^2}{s^2}\right)$

$\log p(y, b \mid s, s_b^2) = -\frac{p}{2}\log 2\pi \frac{s_b^2}{s^2} - \frac{s^2}{2s_b^2} b^T b - \frac{n}{2}\log 2\pi s^2$
$- \frac{1}{2s^2}(y - sXb)^T(y - sXb)$

$\frac{\partial}{\partial s_b^2} = -\frac{p}{2}\frac{1}{s_b^2} + \frac{s^2}{2s_b^4} E(b^T b) \qquad \Sigma_1^{-1}\mu_1 = \frac{1}{s} X^T y$

$\hat{s}_b^2 = \frac{1}{p} s^2 E(b^T b) \qquad \Sigma_1^{-1} = \left(X^T X + \frac{s^2}{s_b^2} I\right)$

$$\frac{\partial}{\partial s^2} = \frac{p}{2}\frac{1}{s^2} - \frac{n}{2}\frac{1}{s^2} - \frac{1}{2s_b^2} b^T b$$

$$\Sigma_1 = V$$

$$\mu_1 = \frac{1}{s} V^{-1} X^T y$$

$$+ \frac{1}{2s^4} y^T y \quad \ldots$$

quadratic in s^2.

---

5 $\quad y \sim s_b X b + e \qquad b \sim N(0, \lambda^2) \quad e \sim N(0, s^2)$

$$p(y, b \mid s, \lambda, s^2) = -\frac{p}{2}\log 2\pi\lambda^2 - \frac{0.5}{\lambda^2} b^T b - \frac{n}{2}\log 2\pi s^2 - \frac{1}{2s^2}\|y - X s_b b\|_2^2$$

$$p(b \mid y) \sim N(\mu_1, \Sigma_1)$$

$$\Sigma_1^{-1} = \frac{1}{s^2}\left(s_b^2 X^T X + \frac{s^2}{\lambda^2} I\right) = \frac{s_b^2}{s^2}\left(X^T X + \frac{s^2}{\lambda^2 s_b^2} I\right)$$

$$\Sigma_1^{-1}\mu_1 = \frac{s_b}{s^2} X^T y$$

$$\frac{\partial}{\partial \lambda^2} = -\frac{p}{2\lambda^2} + \frac{0.5}{\lambda^4} b^T b \Rightarrow \hat{\lambda}^2 = \frac{1}{p} E(b^T b) \quad \left(\text{same as } s_b \text{ in } 1\right)$$

$$\frac{\partial}{\partial s^2} = -\frac{n}{2s^2} + \frac{1}{2s^2} E\left(\|y - X s_b b\|_2^2\right)$$

$$= \hat{s}^2 = \frac{1}{n} E\left(\|y - X s_b b\|_2^2\right) \qquad [\text{as in } 2]$$

$$\frac{\partial}{\partial s_b^2} = \text{as in } 2 \Rightarrow s_b^2 = \frac{\mu_1 X^T y}{E(b^T X^T X b)}$$