RISHI SHAH CS4501-SLGM HW5

① $p(\theta|y,\sigma^2) \propto p(y|\theta,\sigma^2)\, p(\theta|\sigma^2)$

$= \exp(-\frac{1}{2\sigma^2}(X\theta-y)^T(X\theta-y))\, \exp(-\frac{1}{2}(m_0-\theta)^T S_0^{-1}(m_0-\theta))$

we expect $\theta|y,\sigma^2$ to be Normally distributed

$\Rightarrow$

$$(\theta-\hat{\theta})^T k^{-1}(\theta-\hat{\theta}) = \frac{1}{\sigma^2}(X\theta-y)^T(X\theta-y) + (m_0-\theta)^T S_0^{-1}(m_0-\theta)$$

$$\theta^T k^{-1}\theta - 2\theta^T k^{-1}\hat{\theta} + \hat{\theta}^T k^{-1}\hat{\theta} = \theta^T\left(\frac{X^TX}{\sigma^2}\right)\theta - 2\theta^T\left(\frac{X^Ty}{\sigma^2}\right) + \frac{y^Ty}{\sigma^2} + m_0^T S_0^{-1} m_0 - 2\theta^T S_0^{-1} m_0 + \theta^T S_0^{-1}\theta$$

Rearranging and dropping irrelevant constants:

$$\theta^T k^{-1}\theta - 2\theta^T k^{-1}\hat{\theta} = \theta^T\left(\frac{X^TX}{\sigma^2} + S_0^{-1}\right)\theta - 2\theta^T\left(\frac{X^Ty}{\sigma^2} + S_0^{-1}m_0\right)$$

Let $k^{-1} = \frac{X^TX}{\sigma^2} + S_0^{-1}$, Then,

$$-2\theta^T k^{-1}\hat{\theta} = -2\theta^T\left(\frac{X^Ty}{\sigma^2} + S_0^{-1}m_0\right)$$

$$\Rightarrow k^{-1}\hat{\theta} = \frac{X^Ty}{\sigma^2} + S_0^{-1}m_0$$

$$\Rightarrow \hat{\theta} = k\left(\frac{X^Ty}{\sigma^2} + S_0^{-1}m_0\right)$$

So, $\theta|y,\sigma^2 \sim N(\hat{\theta}, k\sigma^2)$

where $\hat{\theta} = k\left(\frac{X^Ty}{\sigma^2} + S_0^{-1}m_0\right)$

and $K = \left(\frac{X^TX}{\sigma^2} + S_0^{-1}\right)^{-1}$

② a)

$$\left(Q\begin{pmatrix}a_1 & & 0\\ & \ddots & \\ 0 & & a_n\end{pmatrix}Q^T\right)^{-1} = \left(Q^T\right)^{-1}\begin{pmatrix}a_1 & & 0\\ & \ddots & \\ 0 & & a_n\end{pmatrix}Q^{-1}$$

Q is orthogonal

$$= Q\begin{pmatrix}a_1 & & 0\\ & \ddots & \\ 0 & & a_n\end{pmatrix}^{-1}Q^T = Q\begin{pmatrix}\frac{1}{a_1} & & 0\\ & \ddots & \\ 0 & & \frac{1}{a_n}\end{pmatrix}Q^T$$

diagonal matrix

b)

$$\hat{y} = X\hat{\theta}^{LS} = X(X^TX)^{-1}X^Ty = (U\Sigma V^T)\left((U\Sigma V^T)^T(U\Sigma V^T)\right)^{-1}(U\Sigma V^T)^Ty$$

$$= U\Sigma V^T(V\Sigma^TU^TU\Sigma V^T)^{-1}(V\Sigma^TU^T)y$$

$$= U\Sigma V^T(V\Sigma^T\Sigma V^T)^{-1}V\Sigma^TU^Ty = U\Sigma V^T\left((V^T)^{-1}\Sigma^{-1}\Sigma^{-1}V^{-1}\right)V\Sigma U^Ty$$

$$= U\Sigma V^TV\Sigma^{-1}\Sigma^{-1}V^TV\Sigma U^Ty = U\Sigma\Sigma^{-1}\Sigma^{-1}\Sigma U^Ty$$

$$= \underline{UU^Ty}$$

c)

$$\hat{y} = X\hat{\theta}^{RR} = X(X^TX+\lambda I)^{-1}X^Ty = U\Sigma V^T\left((U\Sigma V^T)^T(U\Sigma V^T)+\lambda I\right)(U\Sigma V^T)^Ty$$

$$= U\Sigma V^T(V\Sigma^TU^TU\Sigma V^T+\lambda I)^{-1}V\Sigma^TU^Ty$$

$$= U\Sigma V^T(V\Sigma\Sigma V^T+V\lambda IV^T)^{-1}V\Sigma U^Ty$$

$$= U\Sigma V^T(V(\Sigma\Sigma+\lambda I)V^T)^{-1}V\Sigma U^Ty = U\Sigma V^T(V(\Sigma\Sigma+\lambda I)^{-1}V^T)V\Sigma U^Ty$$

$$= U\Sigma(\Sigma\Sigma+\lambda I)^{-1}\Sigma U^Ty = U\Sigma\begin{pmatrix}\sigma_1^2+\lambda & & 0\\ & \ddots & \\ 0 & & \sigma_m^2+\lambda\end{pmatrix}^{-1}\Sigma U^Ty$$

$$= U\Sigma\begin{pmatrix}\frac{1}{\sigma_1^2+\lambda} & & 0\\ & \ddots & \\ 0 & & \frac{1}{\sigma_m^2+\lambda}\end{pmatrix}\Sigma U^Ty = \underline{U\begin{pmatrix}\frac{\sigma_1^2}{\sigma_1^2+\lambda} & & 0\\ & \ddots & \\ 0 & & \frac{\sigma_m^2}{\sigma_m^2+\lambda}\end{pmatrix}U^Ty}$$