FINM33165 Homework 1

Sarah Adilijiang

Polynomial Regression

(a)

$$p(D|M_n) = \int p(D,w|M_n)dw = \int \frac{p(D,w,M_n)}{p(M_n)}dw = \int \frac{p(D|w,M_n)\ p(w,M_n)}{p(M_n)}dw$$

$$= \int p(D|w,M_n)\ \frac{p(w,M_n)}{p(M_n)}\ dw = \int p(D|w,M_n)\ p(w|M_n)\ dw$$

(b)

The design matrix:

$$X_n = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N-1} & x_{N-1}^2 & \cdots & x_{N-1}^n \end{bmatrix}$$

So:

$$y = X_n w + \sigma^2 N(0, I_N)$$

where

$$y = (y_0, y_1, \cdots, y_{N-1})^T, \quad w = (w_0, w_1, \cdots, w_n)^T$$

Since Σ_p is an $n \times n$ positive definite covariance matrix, so matrix Σ_p is symmetric and all its eigenvalues are positive. And the inverse matrix Σ_p^{-1} is also symmetric. And $\Sigma_p^{-1}\Sigma_p = \Sigma_p\Sigma_p^{-1} = I$.

Then the matrix $A = \left(\frac{1}{\sigma^2}X_n^T X_n + \Sigma_p^{-1}\right)^{-1}$ is also a symmetric matirx, i.e. $A = A^{-1}$. And note that $A^{-1}A = AA^{-1} = I$.

Thus the right side of the equation is:

$$(w-\mu)^T A^{-1}(w-\mu) + \frac{1}{\sigma^2}\left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y\right)$$

$$= \left(w - A(\frac{1}{\sigma^2}X_n^T y)\right)^T A^{-1}\left(w - A(\frac{1}{\sigma^2}X_n^T y)\right) + \frac{1}{\sigma^2}\left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y\right)$$

$$= \left(w^T - \frac{1}{\sigma^2}y^T X_n A\right) A^{-1}\left(w - \frac{1}{\sigma^2}A X_n^T y\right) + \frac{1}{\sigma^2}\left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y\right)$$

$$= w^T A^{-1} w - \frac{1}{\sigma^2}y^T X_n w - \frac{1}{\sigma^2}w^T X_n^T y + \frac{1}{\sigma^4}y^T X_n A X_n^T y + \frac{1}{\sigma^2}y^T y - \frac{1}{\sigma^4}y^T X_n A^T X_n^T y$$

$$= w^T A^{-1} w - \frac{1}{\sigma^2}y^T X_n w - \frac{1}{\sigma^2}w^T X_n^T y + \frac{1}{\sigma^2}y^T y$$

$$= w^T\left(\frac{1}{\sigma^2}X_n^T X_n + \Sigma_p^{-1}\right) w - \frac{1}{\sigma^2}y^T X_n w - \frac{1}{\sigma^2}w^T X_n^T y + \frac{1}{\sigma^2}y^T y$$

$$= \frac{1}{\sigma^2} w^T X_n^T X_n w + w^T \Sigma_p^{-1} w - \frac{1}{\sigma^2} y^T X_n w - \frac{1}{\sigma^2} w^T X_n^T y + \frac{1}{\sigma^2} y^T y$$

$$= \frac{1}{\sigma^2}\left(w^T X_n^T X_n w - y^T X_n w - w^T X_n^T y + y^T y\right) + w^T \Sigma_p^{-1} w$$

$$= \frac{1}{\sigma^2}(y^T y - 2w^T X_n^T y + w^T X_n^T X_n w) + w^T \Sigma_p^{-1} w$$

Note that $w^T X_n^T y$ is a scalar, so $w^T X_n^T y = (w^T X_n^T y)^T = y^T X_n w$.

Since the left side of the equation is:

$$\frac{1}{\sigma^2}(y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w$$

$$= \frac{1}{\sigma^2}(y^T - w^T X_n^T)(y - X_n w) + w^T \Sigma_p^{-1} w$$

$$= \frac{1}{\sigma^2}(y^T y - 2w^T X_n^T y + w^T X_n^T X_n w) + w^T \Sigma_p^{-1} w$$

Therefore, we have proved that:

$$\frac{1}{\sigma^2}(y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w \;=\; (w-\mu)^T A^{-1} (w-\mu) + \frac{1}{\sigma^2}\left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y\right)$$

(c)

The left side of equation is:

$$N(y; X_n w, \sigma^2 I_N)\; N(w; 0, \Sigma_p)$$

$$= \frac{1}{2\pi^{N/2}} \frac{1}{(\det\sigma^2 I_N)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(y - X_n w)^T (y - X_n w)\right\} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det\Sigma_p)^{1/2}} \exp\left\{-\frac{1}{2} w^T \Sigma_p^{-1} w\right\}$$

$$= \frac{1}{2\pi^{N/2}} \frac{1}{(\det\sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det\Sigma_p)^{1/2}} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}(y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w\right]\right\}$$

From (b), we get that:

$$\frac{1}{\sigma^2}(y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w \;=\; (w-\mu)^T A^{-1} (w-\mu) + \frac{1}{\sigma^2}\left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y\right)$$

where

$$\mu = A(\frac{1}{\sigma^2} X_n^T y) \quad \text{and} \quad A = \left(\frac{1}{\sigma^2} X_n^T X_n + \Sigma_p^{-1}\right)^{-1}$$

So the right side of equation is:

$$\frac{1}{2\pi^{N/2}} \frac{1}{(\det\sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det\Sigma_p)^{1/2}} \; (2\pi)^{(n+1)/2} (\det A)^{1/2}$$

$$\times \exp\left\{-\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T\right) y\right\} \; N(w; A(\frac{1}{\sigma^2} X_n^T y), A)$$

$$= \quad \frac{1}{2\pi^{N/2}} \frac{1}{(\det \sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det \Sigma_p)^{1/2}}$$

$$\times \exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\} \exp\left\{ -\frac{1}{2} \left[(w-\mu)^T A^{-1} (w-\mu) \right] \right\}$$

Since:

$$\exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\} \exp\left\{ -\frac{1}{2} \left[(w-\mu)^T A^{-1} (w-\mu) \right] \right\}$$

$$= \exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y - \frac{1}{2} \left[\frac{1}{\sigma^2} (y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w - \frac{1}{\sigma^2} \left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y \right) \right] \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2} (y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w \right] \right\}$$

$$\times \exp\left\{ -\frac{1}{2} \left[y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y - \frac{1}{\sigma^2} \left(y^T y - \frac{1}{\sigma^2} y^T X_n A^T X_n^T y \right) \right] \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2} (y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w \right] \right\}$$

Thus, the right side equals to:

$$\frac{1}{2\pi^{N/2}} \frac{1}{(\det \sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det \Sigma_p)^{1/2}} \exp\left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2} (y - X_n w)^T (y - X_n w) + w^T \Sigma_p^{-1} w \right] \right\}$$

which is the same as the left side.

Therefore, the equation has been proved.

(d)

$$\int N(y; X_n w, \sigma^2 I_N) \; N(w; 0, \Sigma_p)$$

$$= \frac{1}{2\pi^{N/2}} \frac{1}{(\det \sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det \Sigma_p)^{1/2}} \; (2\pi)^{(n+1)/2} (\det A)^{1/2}$$

$$\exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\} \int N(w; A(\frac{1}{\sigma^2} X_n^T y), A)$$

$$= \frac{1}{2\pi^{N/2}} \frac{1}{(\det \sigma^2 I_N)^{1/2}} \frac{1}{2\pi^{(n+1)/2}} \frac{1}{(\det \Sigma_p)^{1/2}} \; (2\pi)^{(n+1)/2} (\det A)^{1/2} \; \exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\}$$

$$= \frac{(\det A)^{1/2}}{(2\pi)^{N/2} (\sigma^{2N})^{1/2} (\det \Sigma_p)^{1/2}} \; \exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\}$$

$$= \frac{\sqrt{\det A}}{(2\pi)^{N/2} \sigma^N \sqrt{\det \Sigma_p}} \; \exp\left\{ -\frac{1}{2} y^T \left(\frac{1}{\sigma^2} I_N - \frac{1}{\sigma^4} X_n A^T X_n^T \right) y \right\}$$

(e)

This is down in the jupyter notebook file.