Lauren McCarthy - LM2221

HW 3

(1) (a)

$$y_i \sim \text{Normal}(x_i^T w, \tau^{-1})$$
$$w \sim \text{Normal}(0, \text{diag}(\alpha_1, \dots \alpha_d)^{-1})$$
$$\alpha_k \overset{iid}{\sim} \text{Gamma}(a_0, b_0)$$
$$\tau \sim \text{Gamma}(e_0, f_0)$$

$$\ln p = \sum_{i=1}^{N} \ln p(y_i | w, \tau, x, \alpha) + \ln p(\tau) + \ln p(w|\tau) + \sum_{i=1}^{N} \ln p(\alpha_1, \dots, \alpha_n)$$

$$\sum_{i=1}^{N} \ln p(y_i | w, \tau, x, \alpha) = \sum_{i=1}^{N} \tfrac{1}{2} \ln\left(\frac{\tau}{2\pi}\right) - \frac{\tau}{2}(y_i - x_i^T w)^2$$

$$\ln w = \tfrac{1}{2} \sum_{k=1}^{d} \ln(\alpha_k) - \frac{d}{2} \ln(2\pi) - \tfrac{1}{2} \sum_{k=1}^{d} \alpha_k w_k^2$$

$$\ln(\tau) = e_0 \ln f_0 - \ln\Gamma(e_0) + (e_0 - 1)\ln\tau - f_0\tau$$

$$\ln(\alpha) = \sum_{k=1}^{d} a_0 \ln b_0 - \ln\Gamma(a_0) + (a_0 - 1)\ln\alpha_k - b_0\alpha_k$$

$$\underline{\underline{q(\tau)}} \propto \exp\left[\frac{N}{2}\ln\tau - \frac{\tau}{2}\mathbb{E}_w\left\{\sum_{i=1}^{N}(y_i - x_i^T w)^2\right\} + (e_0 - 1)\ln\tau - f_0\tau\right]$$

$$q(\tau) \propto \text{Gamma}(e', f') \qquad e' = e_0 + \frac{N}{2}$$

$$f' = f_0 + \tfrac{1}{2}\underbrace{\mathbb{E}_w\left\{\sum_{i=1}^{N}(y_i - x_i^T w)^2\right\}}$$

$$\mathbb{E}_w\left\{\sum_{i=1}^{N}(y_i - x_i^T w)^2\right\}$$

$$\mathbb{E}_w\left\{\sum_{i=1}^{N}(y_i^2 - 2y_i x_i^T w + x_i^T w w^T x_i)\right\}$$

$$\sum_{i=1}^{N} y_i^2 - 2y_i x_i^T \mu_w + \sum_{i=1}^{N} x_i^T \underbrace{\mathbb{E}_w(ww^T)}_{\mu_w\mu_w^T} x_i$$

$$\sum_{i=1}^{N}(y_i - x_i^T\mu_w)^2 + \sum_{i=1}^{N} x_i^T \Sigma_w x_i$$

$q(\alpha) \propto \exp\left[\mathbb{E}\{\tfrac{1}{2}\ln(\alpha_K) - \tfrac{1}{2}\alpha_K w_K^2 + (a_0-1)\ln\alpha_K - b_0\alpha_K\}\right]$

$\exp\left[(a_0+\tfrac{1}{2}-1)\ln\alpha_K - (b_0+\tfrac{1}{2}w_K^2)\alpha_K\right.$

$q(\alpha) \propto \text{Gamma}(a', b')$ $\quad a' = a_0 + \tfrac{1}{2}$

$b' = b_0 + \tfrac{1}{2}\underbrace{\mathbb{E}_w(w_K^2)}_{\Sigma_{K,K} + \mu_{w(K)}^2}$

$q(w) \propto \exp\left[\mathbb{E}_{\alpha,\tau}\{-\tfrac{\tau}{2}\sum_{i=1}^{N}(y_i - x_i^T w)^2 - \tfrac{1}{2}\sum_{K=1}^{d}\alpha_K w_K^2\}\right]$

$\exp -\tfrac{1}{2}\mathbb{E}(\tau)\sum_{i=1}^{N}(y_i^2 - y_i x_i^T w + w^T x_i x_i^T w) - \tfrac{1}{2}w^T \text{diag}\, w$

$\exp -\tfrac{1}{2}\left[2\mathbb{E}(\tau)(\sum y_i x_i^T)w + w^T \left[\mathbb{E}(\tau)\sum x_i x_i^T + \mathbb{E}\{\text{diag}(\alpha)\}w\right]\right.$

$q(w) \propto N(\mu_w, \Sigma_w)$ where $\quad \Sigma_w = \mathbb{E}(\tau)\sum x_i x_i^T + \mathbb{E}\{\text{diag}(\alpha)\}^{-1}$

$\mu_w = \Sigma_w \mathbb{E}(\tau)\sum_{i=1}^{N} y_i x_i$

$\mathbb{E}(\tau) = \frac{e'}{f'} \quad \mathbb{E}(\alpha_K) = \frac{a'_K}{b'_K}$

(b) Sudo Code

1. Initalize params $a^0, b^0, e^0, f^0, \mu^0, \Sigma^0$
2. FOR $t = 1 \ldots T$
   - (a) update $q(\tau) \rightarrow e' = e_0 + N/2$
     $f' = f_0 + \tfrac{1}{2}\sum_{i=1}^{N}(y_i - x_i^T\mu_w)^2 + \sum_{i=1}^{N} x_i^T \Sigma_w x_i$
   - (b) update $q(\alpha) \rightarrow a' = a_0 + \tfrac{1}{2}$
     $b' = b_0 + \tfrac{1}{2}\Sigma_{K,K} + \mu_{w(K)}^2$
   - (c) update $q(w) \rightarrow \Sigma_w = \frac{e'}{f'}\sum x_i x_i^T + \mathbb{E}\{\text{diag}(\alpha)\}^{-1}$
     $\mu_w = \Sigma_w \frac{e'}{f'}\sum_{i=1}^{N} y_i x_i$
3. Compare $\mathcal{L}(a'_t, b'_t, \Sigma'_t, \mu'_t, e'_t, f'_t)$ when change is small it has converged.

1 c objective function $\mathcal{L}$

$$\mathcal{L}(a_1, b_1, e_1, f_1, \mu_1, \Sigma_1) = \mathbb{E}\{\ln p(\lambda)\} + \mathbb{E}\{\ln p(w|\alpha_1, \dots \alpha_d)\} + \sum_{k=1}^{d} \mathbb{E}\{\ln p(\alpha_k)\} + \sum_{i=1}^{N} \mathbb{E}\{\ln p(y_i|x_i, w, \lambda)\} - \mathbb{E}\{\ln q(\lambda)\} - \mathbb{E}\{\ln q(w)\} - \sum_{k=1}^{d} \mathbb{E}\{\ln q(\alpha_k)\}$$

$$\mathbb{E}\{\ln p(\lambda)\} = e_0 \ln f_0 - \ln\Gamma(e_0) + (e_0 - 1)\,\mathbb{E}\{\ln\lambda\} - f_0\,\mathbb{E}\{\lambda\}$$
$$= e_0 \ln f_0 - \ln\Gamma(e_0) + (e_0-1)\underbrace{[\psi(e_1) - \ln(f_1)]}_{\text{entropy}} - f_0 \frac{e_1}{f_1}$$

$$\mathbb{E}\{\ln p(w|\alpha_1, \dots \alpha_d)\} = -\frac{d}{2}\ln 2\pi + \frac{1}{2}\sum_{k=1}^{d} \mathbb{E}\{\ln \alpha_k\} - \frac{1}{2}\operatorname{trace}(\operatorname{diag}[\mathbb{E}(\alpha_1), \dots \mathbb{E}(\alpha_d)]\,\mathbb{E}(ww^T))$$
$$= -\frac{d}{2}\ln 2\pi + \frac{1}{2}\sum_{k=1}^{d}[\psi(a_{1k}) - \ln(b_k)] - \frac{1}{2}\operatorname{trace}\left(\frac{a_{11}}{b_{11}}, \dots \frac{a_{dd}}{b_{dd}}\right)(\Sigma_1 + \mu\mu^T)$$

$$\mathbb{E}\{\ln p(\alpha_k)\} = a_{0,k}\ln b_{0,k} - \ln\Gamma(a_{0,k}) + (a_{0,k}-1)[\psi(a_{1,k}) - \ln b_{1,k}] - b_{0,k}\frac{a_{1,k}}{b_{1,k}}$$

$$\mathbb{E}\{\ln p(y_i|x_i, w, \lambda)\} = -\frac{1}{2}\ln(2\pi) + \frac{\mathbb{E}\{\ln\lambda\}}{2} - \frac{\mathbb{E}\{\lambda\}}{2}\,\mathbb{E}\left\{\sum_{i=1}^{N}(y_i - x_i^T w)^2\right\}$$
$$= -\frac{1}{2}\ln(2\pi) + \frac{1}{2}[\psi(e_1) - \ln(f_1)] - \frac{1}{2}\frac{e_1}{f_1}\left(\sum_{i=1}^{N}(y_i - x_i^T w)^2 + x_i^T \Sigma x_i\right)$$

$$\mathbb{E}\{\ln q(\lambda)\} = e_1 \ln f_1 + \ln\Gamma(e_1) + (1-e_1)\psi(e_1)$$

$$\mathbb{E}\{\ln q(w)\} = \frac{N}{2} + \frac{N}{2}\ln(2\pi) + \frac{1}{2}\ln\Sigma$$

$$\mathbb{E}\{\ln q(\alpha_k)\} = a_{1,k}\ln b_{1,k} + \ln\Gamma(a_{1,k}) + (1-a_{1k})\psi(a_{1k})$$

Plugging in these values to the equation above we get the objective.

□