

Lauren McCarthy - LM2221

HW 3

①

① a

$$y_i \sim \text{Normal}(x_i^T w, \tau^{-1})$$

$$w \sim \text{Normal}(0, \text{diag}(\alpha_1, \dots, \alpha_d)^{-1})$$

$$\alpha_k \sim \text{Gamma}(a_0, b_0)$$

$$\tau \sim \text{Gamma}(e_0, f_0)$$

$$\ln p = \sum_{i=1}^N \ln p(y_i | w, \tau, x, \alpha) + \ln p(\tau) + \ln p(w | \tau) + \sum_{k=1}^d \ln p(\alpha_k)$$

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

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

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

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

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

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

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

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

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

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

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

$$\underline{q(\alpha)} \propto \exp \left[\mathbb{E} \left\{ \frac{1}{2} \ln(\alpha_k) - \frac{1}{2} \alpha_k \omega_k^2 + (a_0 - 1) \ln \alpha_k - b_0 \alpha_k \right\} \right]$$

$$\exp \left[(a_0 + \frac{1}{2} - 1) \ln \alpha_k - (b_0 + \frac{1}{2} \mathbb{E}_{\omega}(\omega_k^2)) \alpha_k \right]$$

$$q(\alpha) \propto \text{Gamma}(a', b')$$

$$a' = a_0 + \frac{1}{2}$$

$$b' = b_0 + \frac{1}{2} \mathbb{E}_{\omega}(\omega_k^2)$$

$$\underline{q(\omega)} \propto \exp \left[\mathbb{E}_{\alpha, \tau} \left\{ -\frac{\gamma}{2} \sum_{i=1}^N (y_i - x_i^T \omega)^2 - \frac{1}{2} \sum_{k=1}^K \alpha_k \omega_k^2 \right\} \right]$$

$$\exp - \frac{1}{2} \mathbb{E}(\gamma) \sum_{i=1}^N (y_i^2 - y_i x_i^T \omega + \omega^T x_i x_i^T \omega) - \frac{1}{2} \omega^T \text{diag}(\omega) \omega$$

$$\exp - \frac{1}{2} \left[2 \mathbb{E}(\gamma) \left(\sum y_i x_i^T \right) \omega + \omega^T \left[\mathbb{E}(\gamma) \sum x_i x_i^T + \mathbb{E}\{\text{diag}(\omega)\} \right] \omega \right]$$

$$q(\omega) \propto N(\mu_{\omega}, \Sigma_{\omega}) \text{ where}$$

$$\Sigma_{\omega} = \mathbb{E}(\gamma) \sum x_i x_i^T + \mathbb{E}\{\text{diag}(\omega)\}^{-1}$$

$$\mu_{\omega} = \sum_{\omega} \mathbb{E}(\gamma) \sum_{i=1}^N y_i x_i$$

$$\mathbb{E}(\gamma) = \frac{e'}{f'} \quad \mathbb{E}(\alpha_k) = \frac{a'_k}{b'_k}$$

(b) Sudo Code

① Initialize params $a^0, b^0, e^0, f^0, \mu^0, \Sigma^0$

② For $t = 1, \dots, T$

Ⓐ update $q(\gamma) \rightarrow e' = e_0 + N/2$

$$f' = f_0 + \frac{1}{2} \sum_{i=1}^N (y_i - x_i^T \mu_{\omega})^2 + \sum_{i=1}^N x_i^T \Sigma_{\omega} x_i$$

Ⓑ update $q(\alpha) \rightarrow a' = a_0 + 1/2$

$$b' = b_0 + \frac{1}{2} \sum_{k=1}^K \Sigma_{k,k} + \mu_{\omega(k)}^2$$

Ⓒ update $q(\omega) \rightarrow \Sigma_{\omega} = \frac{e'}{f'} \sum x_i x_i^T + \mathbb{E}\{\text{diag}(\omega)\}^{-1}$

$$\mu_{\omega} = \sum_{\omega} \frac{e'}{f'} \sum_{i=1}^N y_i x_i$$

③ Compare $\mathcal{L}(a'_t, b'_t, \Sigma'_t, \mu'_t, e'_t, f'_t)$ when change is small it has converged.

①c objective function \mathcal{L}

$$\begin{aligned}\mathcal{L}(a, b, e, f, \mu, \Sigma) &= \mathbb{E} \{ \ln p(\gamma) \} + \mathbb{E} \{ \ln p(w | \alpha_1, \dots, \alpha_d) \} \\ &\quad + \sum_{k=1}^d \mathbb{E} \{ \ln p(\alpha_k) \} + \sum_{i=1}^N \mathbb{E} \{ \ln p(y_i | x_i, w, \gamma) \} \\ &\quad - \mathbb{E} \{ \ln q(\gamma) \} - \mathbb{E} \{ \ln q(w) \} - \sum_{k=1}^d \mathbb{E} \{ \ln q(\alpha_k) \}\end{aligned}$$

$$\begin{aligned}\mathbb{E} \{ \ln p(\gamma) \} &= e_0 \ln f_0 - \ln \Gamma(e_0) + (e_0 - 1) \mathbb{E} \{ \ln \gamma \} - f_0 \mathbb{E} \{ \gamma \} \\ &= e_0 \ln f_0 - \ln \Gamma(e_0) + (e_0 - 1) \underbrace{[\Psi(e_0) - \ln(f_0)]}_{\text{entropy}} - f_0 \frac{e_1}{f_1}\end{aligned}$$

$$\begin{aligned}\mathbb{E} \{ \ln p(w | \alpha_1, \dots, \alpha_d) \} &= -\frac{d}{2} \ln 2\pi + \frac{1}{2} \sum_{k=1}^d \mathbb{E} \{ \ln \alpha_k \} - \frac{1}{2} \text{trace}(\text{diag}[\mathbb{E}(\alpha_1), \dots, \mathbb{E}(\alpha_d)] \mathbb{E}(ww^T)) \\ &= -\frac{d}{2} \ln 2\pi + \frac{1}{2} \sum_{k=1}^d [\Psi(a_k) - \ln(b_k)] - \frac{1}{2} \text{trace} \left(\frac{a_{11}}{b_{11}}, \dots, \frac{a_{dd}}{b_{dd}} \right) (\Sigma + \mu \mu^T)\end{aligned}$$

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

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

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

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

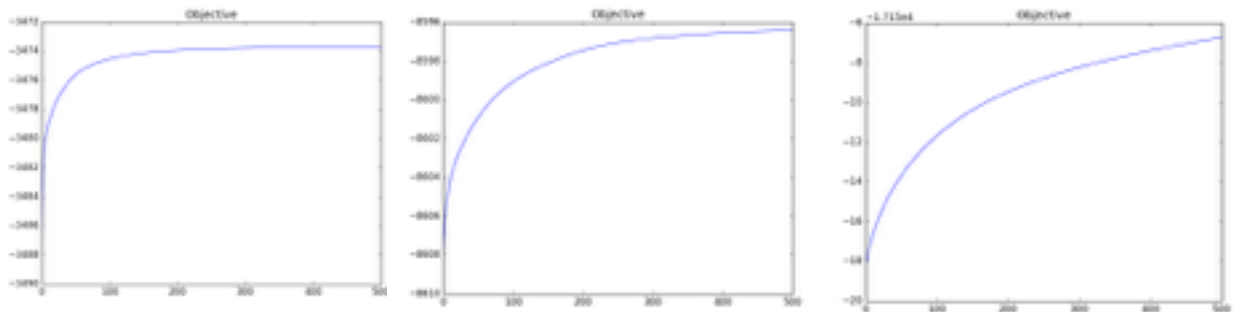
$$\mathbb{E} \{ \ln q(\alpha_k) \} = a_{1,k} \ln b_{1,k} + \ln \Gamma(a_{1,k}) + (1 - a_{1,k}) \Psi(a_{1,k})$$

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

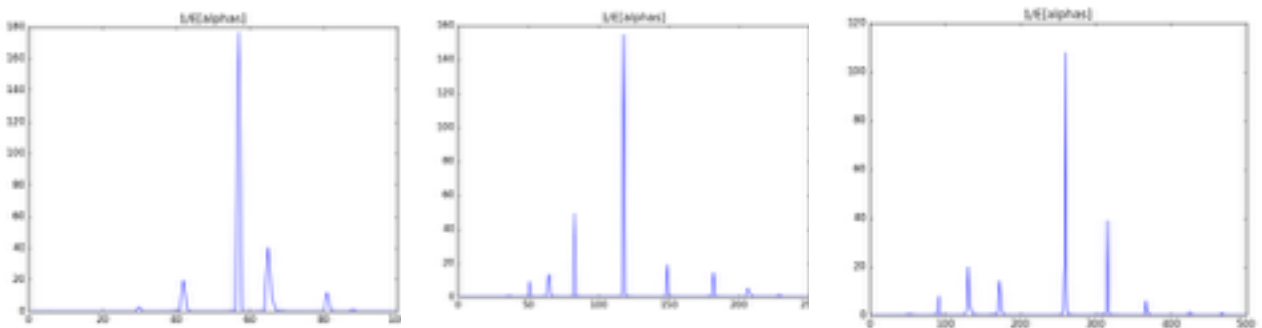
□

Problem 2:

a) Run your algorithm for 500 iterations and plot the variational objective function.



b) Using the final iteration, plot $1/E_q(\alpha)$ as a function of k .



c) Give the value of $1/E_q(\lambda)$ for the final iteration.

$1/E[\lambda]: 1.07981360709$
 $1/E[\lambda]: 0.899449415713$
 $1/E[\lambda]: 0.978100451976$

d) Final Plots

