

$$\begin{aligned}
 & E_{g(\pi)}[\ln p(\pi)] \\
 &= E_{g(\pi)}\left[\ln C(\alpha_0) \prod_{k=1}^K \pi_k^{\alpha_0-1}\right] \xrightarrow{(10.39)} \\
 &= E_{g(\pi)}\left[\ln C(\alpha_0) + \sum_k (\alpha_0-1) \ln \pi_k\right] \\
 &= \ln C(\alpha_0) + \sum_k (\alpha_0-1) E_{g(\pi)}[\ln \pi_k] \\
 &= \ln C(\alpha_0) + \sum_k (\alpha_0-1) \ln \tilde{\pi}_k \quad \dots (10.73)
 \end{aligned}$$

を得る。

$$\begin{aligned}
 & E_{g(M, \Lambda)}[\ln p(M, \Lambda)] \\
 &= E_{g(M|\Lambda)g(\Lambda)}[\ln p(M|\Lambda)p(\Lambda)] \xleftarrow{\text{乘法原理}} \\
 &= E_{g(M|\Lambda)g(\Lambda)}[\ln p(M|\Lambda)] + E_{g(M|\Lambda)g(\Lambda)}[\ln p(\Lambda)] \\
 &= E_{g(M|\Lambda)g(\Lambda)}[\ln \prod_k \pi_k N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1})] + E_{g(\Lambda)}[\ln \prod_k \pi_k^{\frac{1}{2}} p(\Lambda_k)] \xleftarrow{E_{g(M|\Lambda)g(\Lambda)}[\ln \pi_k] = \underbrace{E_{g(M|\Lambda)}[1]}_{1} E_{g(\Lambda)}[\ln p(\Lambda)]} \\
 &\quad \xleftarrow{(10.40)} \quad \Lambda_1, \Lambda_2, \dots, \Lambda_K \text{ 互いに独立} \quad \pi_k = \pi \Lambda_k / \prod_{j \neq k} \Lambda_j \dots \\
 &= E_{g(M|\Lambda)g(\Lambda)}\left[\sum_k \ln N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1})\right] + E_{g(\Lambda)}\left[\sum_k \ln \pi_k^{\frac{1}{2}} p(\Lambda_k)\right] \\
 &= E_{g(M|\Lambda)g(\Lambda)}\left[\sum_k \ln N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1})\right] + E_{g(\Lambda)}\left[\sum_k \ln p(\Lambda_k)\right] \\
 &= \sum_k \left[E_{g(M|\Lambda)g(\Lambda)}\left[\ln N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1})\right] + E_{g(\Lambda)}\left[\ln p(\Lambda_k)\right]\right] \\
 &= \sum_k \left[\frac{D}{2} \ln \left(\frac{\beta_k}{2\pi} \right) + \frac{1}{2} \ln \tilde{\Lambda}_k - \frac{\beta_k}{2} \left\{ D \beta_k^{-1} + \nu_k (m_k - m_0)^T W_k (m_k - m_0) \right\} \right. \xleftarrow{\text{得る}} \\
 &\quad \left. + \ln B(W_0, \nu_0) + \frac{\nu_0 - D - 1}{2} \ln \tilde{\Lambda}_k - \frac{1}{2} \nu_k \text{Tr}(W_0^{-1} W_k) \right] \\
 &= \frac{1}{2} \sum_k \left\{ D \ln \left(\frac{\beta_k}{2\pi} \right) + \ln \tilde{\Lambda}_k - \frac{D \beta_k}{\beta_k} - \beta_k \nu_k (m_k - m_0)^T W_k (m_k - m_0) \right\} \\
 &\quad + K \ln B(W_0, \nu_0) + \frac{\nu_0 - D - 1}{2} \sum_k \ln \tilde{\Lambda}_k - \frac{1}{2} \sum_k \nu_k \text{Tr}(W_0^{-1} W_k) \quad \dots (10.74)
 \end{aligned}$$

を得る。

$$\begin{aligned}
 & E_{g(Z)}[\ln g(Z)] \\
 &= E_{g(Z)}\left[\ln \prod_k r_{nk}^{z_{nk}}\right] \xleftarrow{(10.48)} \\
 &= \sum_n \sum_k E_{g(Z)}[z_{nk} \ln r_{nk}] \\
 &= \sum_n \sum_k \ln r_{nk} E_{g(Z)}[z_{nk}] \\
 &= \sum_n \sum_k r_{nk} \ln r_{nk} \quad \dots (10.75)
 \end{aligned}$$

を得る。

$$\begin{aligned}
 & E_{g(Z)}[\ln r_{nk}| \Lambda_k] \xleftarrow{\Lambda_1, \Lambda_2, \dots, \Lambda_K} \\
 &= E_{g(\Lambda_k)}[\ln p(\Lambda_k)] \\
 &= E_{g(\Lambda_k)}[\ln \pi_k^{\frac{1}{2}} p(\Lambda_k)] \\
 &= E_{g(\Lambda_k)}[\ln \pi_k^{\frac{1}{2}} W_k(m_k, \Lambda_k)] \xleftarrow{(10.40)} \\
 &= E_{g(\Lambda_k)}[\ln B(W_k, \nu_k) + \frac{\nu_k - D - 1}{2} \ln |\Lambda_k| - \frac{1}{2} \text{Tr}(W_k^{-1} \Lambda_k)] \\
 &= \ln B(W_0, \nu_0) + \frac{\nu_0 - D - 1}{2} \ln |\Lambda_k| - \frac{1}{2} \nu_k \text{Tr}(W_0^{-1} \Lambda_k) \\
 &= E_{g(\Lambda_k)}[\ln B(W_k, \nu_k)] \xleftarrow{(10.45)} \\
 &= E_{g(\Lambda_k)}[\text{Tr}(W_k^{-1} \Lambda_k)] \\
 &= \text{Tr}(E_{g(\Lambda_k)}[W_k^{-1} \Lambda_k]) \\
 &= \text{Tr}(W_k^{-1} E_{g(\Lambda_k)}[\Lambda_k]) \\
 &= \text{Tr}(W_k^{-1} \nu_k W_k) \xleftarrow{-(10.59), (10.50)} \\
 &= \nu_k \text{Tr}(W_k^{-1} W_k)
 \end{aligned}$$

$$\begin{aligned}
 & E_{g(Z)}[z_{nk}] \xleftarrow{2, 2, \dots, \text{得る}} \\
 &= E_{g(\Lambda_k)}[z_{nk}] \\
 &= E_{g(\Lambda_k)}[1] \dots E_{g(\Lambda_k)}[z_{nk}] \dots \\
 &= E_{g(\Lambda_k)}[z_{nk}] \xleftarrow{(10.48)} \\
 &= \sum_n \prod_{j \neq n} z_{jk} \tilde{z}_{nk}
 \end{aligned}$$

$$= r_{nk}$$

$$\begin{aligned}
& E_{q(\pi)}[\ln q(\pi)] \\
&= E_{q(\pi)}[\ln \text{Dir}(\pi | \alpha)] \quad \xrightarrow{(10.57)} \\
&= E_{q(\pi)}[\ln C(\alpha) \prod_k \pi_k^{\alpha_k - 1}] \\
&= \ln C(\alpha) + E_{q(\pi)}\left[\sum_k (\alpha_k - 1) \ln \pi_k\right] \\
&= \ln C(\alpha) + \sum_k (\alpha_k - 1) E_{q(\pi)}[\ln \pi_k] \\
&= \ln C(\alpha) + \sum_k (\alpha_k - 1) \ln \tilde{\pi}_k \quad \cdots (10.76) \\
&\text{を得る。}
\end{aligned}$$

$$\begin{aligned}
& E_{q(\mu, \Lambda)}[\ln q(\mu, \Lambda)] \\
&= E_{q(\mu, \Lambda)}\left[\ln \prod_k N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1}) W(\Lambda_k | W_k, V_k)\right] \quad \xrightarrow{(10.59)} \\
&= E_{q(\mu, \Lambda)}\left[\sum_k \left\{ \ln N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1}) + \ln W(\Lambda_k | W_k, V_k) \right\} \right] \\
&= \sum_k \left\{ E_{q(\mu_k)}\left[\ln N(\mu_k | m_k, (\beta_k \Lambda_k)^{-1})\right] + E_{q(\Lambda_k)}\left[\ln W(\Lambda_k | W_k, V_k)\right] \right\} \\
&= \sum_k \left\{ E_{q(\mu_k)}\left[-\frac{D}{2} \ln(2\pi) + \frac{1}{2} \ln |\beta_k \Lambda_k| - \frac{1}{2} (\mu_k - m_k)^T (\beta_k \Lambda_k) (\mu_k - m_k) \right] \right. \\
&\quad \left. + E_{q(\Lambda_k)}\left[\ln B(W_k, V_k) + \frac{V_k - D - 1}{2} \ln |\Lambda_k| - \frac{1}{2} \text{Tr}(W_k^T \Lambda_k)\right] \right\} \\
&= \sum_k \left\{ -\frac{D}{2} \ln(2\pi) + \frac{D}{2} \ln \beta_k + \frac{1}{2} E_{q(\mu_k)}[\ln |\Lambda_k|] - \frac{\beta_k}{2} E_{q(\mu_k, \Lambda_k)}[(\mu_k - m_k)^T \Lambda_k (\mu_k - m_k)] \right. \\
&\quad \left. + \ln B(W_k, V_k) + \frac{V_k - D - 1}{2} E_{q(\Lambda_k)}[\ln |\Lambda_k|] - \frac{1}{2} \text{Tr}(W_k^T E_{q(\mu_k)}[\Lambda_k]) \right\} \\
&= \sum_k \left\{ \frac{D}{2} \ln \left(\frac{\beta_k}{2\pi}\right) + \frac{1}{2} \ln \tilde{\Lambda}_k - \frac{\beta_k}{2} D \beta_k^{-1} \right. \quad \xrightarrow{(10.65)} \\
&\quad \left. + \ln B(W_k, V_k) + \frac{V_k - D - 1}{2} E_{q(\Lambda_k)}[\ln |\Lambda_k|] - \frac{1}{2} V_k D \right\} \\
&= \sum_k \left\{ \frac{D}{2} \ln \left(\frac{\beta_k}{2\pi}\right) + \frac{1}{2} \ln \tilde{\Lambda}_k - \frac{D}{2} - H[\Lambda_k] \right\} \quad \cdots (10.77) \quad \xrightarrow{(10.72)} \\
&\text{を得る。}
\end{aligned}$$

$E_{q(\mu_k)}[\ln |\Lambda_k|]$
 $= E_{q(\mu_k)}[\ln |\Lambda_k|]$
 $= E_{q(\mu_k)[1]} E_{q(\mu_k)[2]} \cdots E_{q(\mu_k)[D]} [\ln |\Lambda_k|] \cdots$
 $= E_{q(\mu_k)}[\ln |\Lambda_k|]$

 $E_{q(\Lambda_k)}[(\mu_k - m_k)^T \Lambda_k (\mu_k - m_k)]$
 $= E_{q(\Lambda_k)}[1] E_{q(\Lambda_k)}[2] \cdots E_{q(\Lambda_k)}[D] [\Lambda_k (\mu_k - m_k)^T (\mu_k - m_k)] \cdots$
 $= E_{q(\Lambda_k)}[(\mu_k - m_k)^T \Lambda_k (\mu_k - m_k)]$

 $E_{q(\mu_k)}[\ln |\Lambda_k|]$
 $= E_{q(\mu_k)}[\ln |\Lambda_k|]$
 $= \text{Tr}(W_k^T E_{q(\mu_k)}[\Lambda_k])$

 $E_{q(\mu_k, \Lambda_k)}[(\mu_k - m_k)^T \Lambda_k (\mu_k - m_k)]$
 $= D \beta_k^{-1} + \frac{1}{2} (m_k - m_k)^T \Lambda_k (m_k - m_k)$
 $= D \beta_k^{-1}$

 $\text{Tr}(W_k^T E_{q(\Lambda_k)}[\Lambda_k])$
 $= \text{Tr}(W_k^T W_k V_k) \cdots (10.65)$
 $= V_k D$