5.2.5.3 にて、式を追ってみる。

$\alpha_{\gamma_w}, \beta_{\gamma_w}$ が初期化できること、m, v が更新できること、尤度が $\mathcal{N}(w|0, \gamma^{-1})$ で表されることを前提とする。

そうした場合、q は (5.58) で表されるので、正規化定数は (5.64) の 1 行目で表される。

$$
\begin{aligned}
Z(\alpha_{\gamma_w}, \beta_{\gamma_w}) &= \int \mathcal{N}(w_{p,q}^{(r)}|0, \gamma_w^{-1}) q(W, \gamma_w, \gamma_y) dW d\gamma_w d\gamma_y \\
&= \int \mathcal{N}(w_{p,q}^{(r)}|0, \gamma_w^{-1}) \prod_i \prod_j \prod_l \mathcal{N}(w_{i,j}^{(l)}|m_{i,j}^{(l)}, v_{i,j}^{(l)}) Gam(\gamma_w|\alpha_{\gamma_w}, \beta_{\gamma_w}) Gam(\gamma_y|\alpha_{\gamma_y}, \beta_{\gamma_y}) \prod_i \prod_j \prod_l dw_{i,j}^{(l)} d\gamma_w d\gamma_y \\
&= \int \mathcal{N}(w_{p,q}^{(r)}|0, \gamma_w^{-1}) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) Gam(\gamma_w|\alpha_{\gamma_w}, \beta_{\gamma_w}) dw_{p,q}^{(r)} d\gamma_w
\end{aligned} \tag{1}
$$

となり、(5.61) の式が求まる。(5.59),(5.60) については尤度が決まれば、同じように求められる。

この式に対して、更に計算を進めると、(3.62) 式の導出過程 (3.48) を考えて、

$$
\begin{aligned}
Z(\alpha_{\gamma_w}, \beta_{\gamma_w}) &= \int \mathcal{N}(w_{p,q}^{(r)}|0, \gamma_w^{-1}) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) Gam(\gamma_w|\alpha_{\gamma_w}, \beta_{\gamma_w}) dw_{p,q}^{(r)} d\gamma_w \\
&= \int (\int \mathcal{N}(w_{p,q}^{(r)}|0, \gamma_w^{-1}) Gam(\gamma_w|\alpha_{\gamma_w}, \beta_{\gamma_w}) d\gamma_w) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) dw_{p,q}^{(r)} \\
&= \int St(w_{p,q}^{(r)}|0, \frac{\alpha_{\gamma_w}}{\beta_{\gamma_w}}, 2\alpha_{\gamma_w}) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) dw_{p,q}^{(r)}
\end{aligned} \tag{2}
$$

ここで、(3.64),(3.65) を考慮すると

$$
\mathbb{E}_{St(w_{p,q}^{(r)}|0, \frac{\alpha_{\gamma_w}}{\beta_{\gamma_w}}, 2\alpha_{\gamma_w})} = 0 \tag{3}
$$

$$
\mathbb{V}_{St(w_{p,q}^{(r)}|0, \frac{\alpha_{\gamma_w}}{\beta_{\gamma_w}}, 2\alpha_{\gamma_w})} = \frac{2\alpha_{\gamma_w}}{\frac{\alpha_{\gamma_w}}{\beta_{\gamma_w}}(2\alpha_{\gamma_w} - 2)} = \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w} - 1} \tag{4}
$$

よって、平均分散が一致するようにガウス分布で近似すると、

$$
\begin{aligned}
Z(\alpha_{\gamma_w}, \beta_{\gamma_w}) &= \int St(w_{p,q}^{(r)}|0, \frac{\alpha_{\gamma_w}}{\beta_{\gamma_w}}, 2\alpha_{\gamma_w}) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) dw_{p,q}^{(r)} \\
&\approx \int \mathcal{N}(w_{p,q}^{(r)}|0, \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w} - 1}) \mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)}) dw_{p,q}^{(r)}
\end{aligned} \tag{5}
$$

ここで以下の式を考える。

$$
\begin{aligned}
\mathcal{N}(w|0, \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w} - 1}) \mathcal{N}(w|m, v) &\equiv \mathcal{N}(w|0, a) \mathcal{N}(w|m, v) = \frac{1}{2\pi\sqrt{av}} exp(-\frac{w^2}{2a} - \frac{(w-m)^2}{2v}) \\
&= \frac{1}{2\pi\sqrt{av}} exp(-\frac{(v+a)w^2 - 2amw + am^2}{2av}) = \frac{1}{2\pi\sqrt{av}} exp(-\frac{(v+a)(w - \frac{am}{v+a})^2 + am^2 - \frac{a^2m^2}{v+a}}{2av}) \\
&= \frac{1}{2\pi\sqrt{av}} exp(-\frac{v+a}{2av}(w - \frac{am}{v+a})^2 + \frac{m^2}{2(v+a)}) = \frac{1}{\sqrt{2\pi\frac{av}{v+a}}} exp(-\frac{1}{2}\frac{v+a}{av}(w - \frac{am}{v+a})^2) \frac{1}{\sqrt{2\pi(v+a)}} exp(-\frac{1}{2}\frac{m^2}{(v+a)}) \\
&= \mathcal{N}(w|\frac{am}{v+a}, \frac{av}{v+a}) \mathcal{N}(m|0, v+a) = \mathcal{N}(w|\frac{\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}m}{v + \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}}, \frac{\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}v}{v + \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}}) \mathcal{N}(m|0, v + \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w} - 1})
\end{aligned} \tag{6}
$$

より一般的には”ガウス分布 積”などで検索すると多くのサイトで言及されている。

この式を使って Z を考える。

$$
\begin{aligned}
Z(\alpha_{\gamma_w}, \beta_{\gamma_w}) &= \int \mathcal{N}(w_{p,q}^{(r)}|0, \frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1})\mathcal{N}(w_{p,q}^{(r)}|m_{p,q}^{(r)}, v_{p,q}^{(r)})dw_{p,q}^{(r)} \\
= \int \mathcal{N}(w_{p,q}^{(r)}|\frac{\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}m_{p,q}^{(r)}}{v_{p,q}^{(r)}+\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}}, \frac{\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}v_{p,q}^{(r)}}{v_{p,q}^{(r)}+\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1}})&\mathcal{N}(m_{p,q}^{(r)}|0, v_{p,q}^{(r)}+\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1})dw_{p,q}^{(r)} \\
&= \mathcal{N}(m_{p,q}^{(r)}|0, v_{p,q}^{(r)}+\frac{\beta_{\gamma_w}}{\alpha_{\gamma_w}-1})
\end{aligned}
\tag{7}
$$

γ_w の更新に関して考えると、$q_{new}(W, \gamma_w, \gamma_y) \propto p(w|\gamma_w)q_{old}(W, \gamma_w, \gamma_y)$ になるので、記載されているように Z は w のときと同様に計算できる。