9.25

(9.70) より

$$\frac{\partial}{\partial\theta}\ln p(X|\theta) = \frac{\partial}{\partial\theta}\mathcal{L}(q,\theta) + \frac{\partial}{\partial\theta}KL(q\|p)$$

ここで

$$q(Z) = p(Z|X,\theta^{old})$$

を (9.72) に入れると

$$KL(q\|p) = -\sum_Z p(Z|X,\theta^{old})\ln\frac{p(Z|X,\theta)}{p(Z|X,\theta^{old})} = 0$$

なので

$$\frac{\partial}{\partial\theta}KL(q\|p) = -\sum_Z p(Z|X,\theta^{old})\frac{\frac{\partial}{\partial\theta}p(Z|X,\theta)}{p(Z|X,\theta)}$$

これより

$$\frac{\partial}{\partial\theta}KL(q\|p)\Big|_{\theta^{old}} = -\sum_Z \frac{\partial}{\partial\theta}\left[p(Z|X,\theta)\right]_{\theta^{old}} = -\frac{\partial}{\partial\theta}\left[\sum_Z p(Z|X,\theta)\right]_{\theta^{old}}$$

和と微分を換えてよい

$$= -\frac{\partial}{\partial\theta}1\Big|_{\theta^{old}} = 0$$

なので

$$\frac{\partial}{\partial\theta}\ln p(X|\theta)\Big|_{\theta^{old}} = \frac{\partial}{\partial\theta}\mathcal{L}(q,\theta)\Big|_{\theta^{old}} + \frac{\partial}{\partial\theta}KL(q\|p)\Big|_{\theta^{old}}$$

に入れて

$$\frac{\partial}{\partial\theta}\ln p(X|\theta)\Big|_{\theta^{old}} = \frac{\partial}{\partial\theta}\mathcal{L}(q,\theta)\Big|_{\theta^{old}}$$

を得る。