2 **(EM for factor analysis)** (a) Derive the EM algorithm for estimating the model parameters of factor analysis model (with μ included in the model). Summarize the algorithm in iterative E and M steps.

(a) $Y_i = \mu + \Lambda X_i + W_i$, $X_i \sim N(0, I_k)$, $W_i \sim N(0, \psi)$, ψ diagonal

the log joint distribution is $\log p(Y, X | \mu, \Lambda, \psi) = \sum_{i=1}^{n} (\log p(Y_i | X_i, \mu, \Lambda, \psi) + \log p(X_i))$

E step: given μ, Λ, W_i, compute $p(X_i | Y_i) \sim N(m_i, V)$, $V = (I_k + (\Lambda^{(t)})^T (\psi^{(t)})^{-1} \Lambda^{(t)})^{-1}$, $m_i = V (\Lambda^{(t)})^T (\psi^{(t)})^{-1} (Y_i - \mu^{(t)})$

$\Rightarrow E[X_i] = m_i$, $E[X_i X_i^T] = V + m_i m_i^T$

M step: $\mu^{(t+1)} = \frac{1}{n} \sum_{i=1}^{n} Y_i$

$\Lambda^{(t+1)} = (\sum_{i=1}^{n} (Y_i - \mu) m_i^T)(\sum_{i=1}^{n} (V + m_i m_i^T))$

$\psi^{(t+1)} = \frac{1}{n} \text{diag} (\sum_{i=1}^{n} (Y_i - \mu - \Lambda^{(t+1)} - \Lambda^{(t+1)} m_i)(Y_i - \mu)^T)$