Question:
Considering The graphical model $z \rightarrow x$ where z and x are two random variables. And the joint distribution of x, z can be given as follows:

$$P(x, z) = P(z)P(x|z)$$

Where z is a latent variable. If we assume $p(z) = N(z|\mathbf{0}, \boldsymbol{I})$ and $x = \boldsymbol{W}\boldsymbol{z} + \boldsymbol{\mu} + \epsilon$. The dimension of each sample is D, $\boldsymbol{\mu}$ is the mean vector of random variable x and $\epsilon \sim N(\mathbf{0}, \sigma^2\boldsymbol{I})$. Denote $\boldsymbol{X}$ as data matrix. Try to derive the maximum likelihood form of variable x.
solution:
If $P(x, z) = P(z)P(x|z)$, then we have $p(x) = \int p(x|z)p(z)dz$. And we can get:

$$\begin{aligned}
E[x] &= \int_{-\infty}^{+\infty} x f_X(x)dx = \int_{-\infty}^{+\infty} x[\int_{-\infty}^{+\infty} f_Z(z)f_{X|Z}(x|z)dz]dx \\
&= \int_{-\infty}^{+\infty} \{[\int_{-\infty}^{+\infty} x f_{(X|Z)}(x|z)dx]\}f_Z(z)dz \\
&= E_z[E_x[x|z]] = E_z[\boldsymbol{W}\boldsymbol{z} + \mu]
\end{aligned}$$

$$\begin{aligned}
Var[x] &= E_z\{E_x(x^2|z) - E_x^2(x|z)\} \\
&= E[x^2] - E_z[E_x^2(x|z)] + E_z[E_x^2[x|z]] - E_z^2(E_x[x|z]) \\
&= E_z[Var_x[x|z]] + Var_z[E_x[x|z]] \\
&= E_z[\sigma^2\boldsymbol{I}] + Var_z[\boldsymbol{W}\boldsymbol{z} + \mu] = \sigma^2\boldsymbol{I} + \boldsymbol{W}\boldsymbol{W}^T
\end{aligned}$$

Thus, $p(x) = N(x|\mu, \boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I})$, which lead to $P(\boldsymbol{X}|\mu, \boldsymbol{W}, \sigma^2) = \prod_{n=1}^{N} p(x_n|\mu, \boldsymbol{W}, \sigma^2)$. The log likelihood function can be written as:

$$\begin{aligned}
logp(\boldsymbol{X}|\mu, \boldsymbol{W}, \sigma^2) &= \sum_{n=1}^{N} p(x_n|\mu, \boldsymbol{W}, \sigma^2) \\
&= -\frac{ND}{2}log(2\pi) - \frac{N}{2}log|\boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I}| \\
&- \frac{1}{2}\sum_{n=1}^{n}(x_n - \mu)^T(\boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I}(x_n - \mu) \\
&= -\frac{N}{2}\{Dlog(2\pi) + log|\boldsymbol{C}| + Tr(\boldsymbol{C}^{-1}\boldsymbol{S})\}
\end{aligned}$$

where $\boldsymbol{C} = \boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I}$, $\boldsymbol{S} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu)(x_n - \mu)^T$