

2.15

$$\begin{aligned}
H[\mathbf{x}] &= - \int p(\mathbf{x}) \ln(p(\mathbf{x})) d\mathbf{x} \\
&= - \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \ln(\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})) d\mathbf{x} \\
&= - \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \ln \left( \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\} \right) d\mathbf{x} \\
&= - \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \left( \ln \left( \frac{1}{(2\pi)^{D/2}} \right) + \ln \left( \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \right) - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) d\mathbf{x} \\
&= - \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \left( -\frac{D}{2} \ln 2\pi - \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) d\mathbf{x} \\
&= \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \left( \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) d\mathbf{x} \\
&= \int \frac{D}{2} \ln 2\pi \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&\quad + \int \frac{1}{2} \ln |\boldsymbol{\Sigma}| \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&\quad + \int \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&= \frac{D}{2} \ln 2\pi \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&\quad + \frac{1}{2} \ln |\boldsymbol{\Sigma}| \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&\quad + \frac{1}{2} \int (\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} \\
&= \frac{D}{2} \ln 2\pi (1) \\
&\quad + \frac{1}{2} \ln |\boldsymbol{\Sigma}| (1)
\end{aligned}$$

$$+\frac{1}{2} \int (\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

$$= \frac{D}{2} \ln 2\pi$$

$$+\frac{1}{2} \ln |\boldsymbol{\Sigma}|$$

$$+\frac{1}{2} \int (\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

$$= \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}|$$

$$+\frac{1}{2} \int \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

$$-\frac{1}{2} \int 2\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

$$+\frac{1}{2} \int \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

Using equation 318 from Matrix Cookbook, we get

$$\int \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} = \text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}) + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

This gives us:

$$= \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}|$$

$$+\frac{1}{2} (\text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}) + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})$$

$$-\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

$$+\frac{1}{2} \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \int \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}$$

$$= \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}|$$

$$\begin{aligned}
& + \frac{1}{2} (tr(\mathbf{\Sigma}^{-1} \mathbf{\Sigma}) + \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu}) \\
& \quad - \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} \\
& + \frac{1}{2} \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} (1)
\end{aligned}$$

$$\begin{aligned}
& = \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\mathbf{\Sigma}| + \frac{1}{2} tr(\mathbf{\Sigma}^{-1} \mathbf{\Sigma}) + \frac{1}{2} \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} + \frac{1}{2} \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} \\
& = \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\mathbf{\Sigma}| + \frac{1}{2} tr(\mathbf{\Sigma}^{-1} \mathbf{\Sigma}) \\
& = \frac{D}{2} \ln 2\pi + \frac{1}{2} \ln |\mathbf{\Sigma}| + \frac{1}{2} D \\
& = \frac{1}{2} \ln |\mathbf{\Sigma}| + \frac{D}{2} (1 + \ln(2\pi))
\end{aligned}$$