We decided this derivation is too much work for a small part of an assignment. We provide the original question and the solution in this document.

This is a very good exercise to check your understanding of integration and Analysis, so we encourage you to try it yourself.

Todo: Compute the Entropy of the normal distribution.

$$\mathcal{N}(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})\,.$$

Solution:

The entropy of a Gaussian can be computed without performing any integrals directly:

$$\begin{aligned}
H\left[\mathcal{N}\right] &= -\int\limits_{-\infty}^{\infty} \log\left(N(x;\mu,\sigma^2)\right)\cdot N(x;\mu,\sigma^2)\,\mathrm{d}x \\
&= -\int\limits_{-\infty}^{\infty} \left[-\frac{(x-\mu)^2}{2\sigma^2} - \frac{1}{2}\log\left(2\pi\sigma^2\right)\right]\cdot N(x;\mu,\sigma^2)\,\mathrm{d}x \\
&= \frac{1}{2\sigma^2}\underbrace{\int\limits_{-\infty}^{\infty}(x-\mu)^2\cdot N(x;\mu,\sigma^2)\,\mathrm{d}x}_{=\sigma^2} + \frac{1}{2}\log\left(2\pi\sigma^2\right)\underbrace{\int\limits_{-\infty}^{\infty} N(x;\mu,\sigma^2)\,\mathrm{d}x}_{=1} \\
&= \frac{1}{2} + \frac{1}{2}\log\left(2\pi\sigma^2\right) = \frac{1}{2}\log\left(2\pi\mathrm{e}\sigma^2\right)
\end{aligned}$$

Todo: Find the Kullback-Leibler divergence between two Gaussian distributions.

$$D_{KL}\left(\mathcal{N}(\mu_1,\sigma_1^2)||\mathcal{N}(\mu_2,\sigma_2^2)\right)\,.$$

Solution:

Let's first simplify the expression for the Kullback-Leibler divergence, by introducing the notation:

$\mathcal{N}_1(x) = \mathcal{N}(x; \mu_1, \sigma_1^2)$ and $\mathcal{N}_2(x) = \mathcal{N}(x; \mu_2, \sigma_2^2)$:

$$
\begin{aligned}
D_{KL}\left(\mathcal{N}_1 || \mathcal{N}_2\right) &= \int_{-\infty}^{\infty} \mathcal{N}_1(x) \cdot \log\left(\frac{\mathcal{N}_1(x)}{\mathcal{N}_2(x)}\right) \mathrm{d}x \\
&= -H\left[\mathcal{N}_1\right] - \int_{-\infty}^{\infty} \log\left(\mathcal{N}_2(x)\right) \mathcal{N}_1(x) \,\mathrm{d}x \\
&= -\frac{\log\left(2\pi \mathrm{e} \sigma_1^2\right)}{2} - \int_{-\infty}^{\infty} \left(-\frac{\log(2\pi\sigma_2^2)}{2} - \frac{(x-\mu_2)^2}{2\sigma_2^2}\right) \cdot \mathcal{N}_1(x) \,\mathrm{d}x \\
&= -\frac{\log\left(2\pi \mathrm{e} \sigma_1^2\right)}{2} + \frac{\log(2\pi\sigma_2^2)}{2} + \int_{-\infty}^{\infty} \frac{(x-\mu_2)^2}{2\sigma_2^2} \cdot \mathcal{N}_1(x) \,\mathrm{d}x \\
&= \log\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{1}{2} + \int_{-\infty}^{\infty} \frac{(x-\mu_2)^2}{2\sigma_2^2} \cdot \mathcal{N}_1(x) \,\mathrm{d}x \qquad (1)
\end{aligned}
$$

For space reasons, we continue to compute only the last integral

$$
\begin{aligned}
\int_{-\infty}^{\infty} (x-\mu_2)^2 \cdot \mathcal{N}(x; \mu_1, \sigma_1^2) \,\mathrm{d}x &= \int_{-\infty}^{\infty} \underbrace{(x-\mu_1+\mu_1-\mu_2)^2}_{\substack{=(x-\mu_1)^2+(\mu_1-\mu_2)^2 \\ +2(x-\mu_1)(\mu_1-\mu_2)}} \cdot \mathcal{N}(x; \mu_1, \sigma_1^2) \,\mathrm{d}x \\
&= \underbrace{\int_{-\infty}^{\infty} (x-\mu_1)^2 \mathcal{N}(x; \mu_1, \sigma_1^2) \,\mathrm{d}x}_{=\sigma_1^2} \\
&\quad + (\mu_1-\mu_2)^2 \underbrace{\int_{-\infty}^{\infty} \mathcal{N}(x; \mu_1, \sigma_1^2) \,\mathrm{d}x}_{=1} \\
&\quad + 2(\mu_1-\mu_2) \underbrace{\int_{-\infty}^{\infty} (x-\mu_1) \mathcal{N}(x; \mu_1, \sigma_1^2) \,\mathrm{d}x}_{=0} \\
&= \sigma_1^2 + (\mu_1-\mu_2)^2
\end{aligned}
$$

Plugging this back into the previous equation yields

$$
D_{KL}\left(\mathcal{N}(\mu_1, \sigma_1^2) || \mathcal{N}(\mu_2, \sigma_2^2)\right) = \log\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{1}{2} + \frac{\sigma_1^2 + (\mu_1-\mu_2)^2}{2\sigma_2^2}.
$$

Note that the divergence is 0, if $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$.