Machine Learning for Graphs and Sequential Data Exercise Sheet 01

Normalizing Flows

Problem 1:

(a) We consider the following transformations:

- $f(\boldsymbol{z}) = \begin{bmatrix} 10z_1 + 1 \\ \cos(z_1)z_2 \\ \sin(z_1 z_2) \end{bmatrix}$ from $\mathbb{R}^3$ to $\mathbb{R} \times \mathbb{R} \times [-1, 1]$.
- $f(\boldsymbol{z}) = \begin{bmatrix} z_1^3 \\ e^{z_1} z_2^5 \\ e^{-z_1 - z_2} z_3^7 \end{bmatrix}$ from $\mathbb{R}^3$ to $\mathbb{R}^3$.

Are these transformations invertible ?

- By computating the Jacobian, we observe that it is triangular:

$$J_f = \begin{bmatrix} 10 & 0 & 0 \\ -z_2 \sin(z_1) & \cos(z_1) & 0 \\ z_2 \cos(z_1 z_2) & z_1 \cos(z_1 z_2) & 0 \end{bmatrix}$$

Consequently, its determinant is the product of the diagonal elements i.e. $\det(J_f) = 10 \times \cos(z_1) \times 0$. Since the determinant is 0 the tranformation is not invertible.

- We can compute the inverse of f by solving a system of (non-linear) equations.

$$\begin{cases} x_1 = z_1^3 \\ x_2 = e^{z_1} z_2^5 \\ x_3 = e^{-z_1 - z_2} z_3^7 \end{cases}$$

$$\begin{cases} z_1 = x_1^{\frac{1}{3}} \\ z_2 = \left(x_2 e^{-x_1^{\frac{1}{3}}} \right)^{\frac{1}{5}} \\ z_3 = \left(x_3 e^{x_1^{\frac{1}{3}} + \left(x_2 e^{-x_1^{\frac{1}{3}}} \right)^{\frac{1}{5}}} \right)^{\frac{1}{7}} \end{cases}$$

Thus f is invertible with inverse:

$$f^{-1}(\boldsymbol{x}) = \begin{bmatrix} x_1^{\frac{1}{3}} \\ \left(x_2 e^{-x_1^{\frac{1}{3}}}\right)^{\frac{1}{5}} \\ \left(x_3 e^{-x_1^{\frac{1}{3}} - \left(x_2 e^{-x_1^{\frac{1}{3}}}\right)^{\frac{1}{5}}}\right)^{\frac{1}{7}} \end{bmatrix}$$

(b) We consider the transformation $f(\boldsymbol{z}) = \begin{bmatrix} \sin(z_1) \\ \cos(z_2) \end{bmatrix}$ from $[a, b] \times [c, d]$ to $[-1, 1]^2$. Under what conditions on a, b, c, d is this tranformation invertible ?

case 1: There exist k such that $a, b \in [-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi]$ and k' such that $c, d \in [k'\pi, (k'+1)\pi]$. In this case, each of the element-wise transformation (i.e. sin and cos) are stricly monotonic on these intervals and are invertible. Hence, f is also invertible on this domain with inverse:

$$f^{-1}(x) = \begin{bmatrix} \arcsin(x_1) + k\pi \\ \arccos(x_2) + k'\pi \end{bmatrix}$$

case 2: There does not exist k such that $a, b \in [-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi]$. The function sinus is not invertible on such an interval meaning that we can find two points $z_1^{(1)} < z_1^{(2)}$ in $[a, b]$ such that $\sin(z_1^{(1)}) = \sin(z_1^{(2)})$. Consequently, we have for example $f\left(\begin{bmatrix} z_1^{(1)} \\ c \end{bmatrix}\right) = f\left(\begin{bmatrix} z_1^{(2)} \\ c \end{bmatrix}\right)$ and the transformation f is not invertible. We can apply a similar reasoning with cosinus if there does not exist k' such that $c, d \in [k'\pi, (k'+1)\pi]$.

(c) We consider the tranformation $f(\boldsymbol{z}) = \boldsymbol{A}\boldsymbol{z} + \boldsymbol{b}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$, where $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ and $\boldsymbol{b} \in \mathbb{R}^2$. Under what conditions on $\boldsymbol{A}$ and $\boldsymbol{b}$ is this tranformation invertible ?

The Jacobian determinant of a linear transformation f is:

$$\det J_f = \det A$$

We know the determinant of a matrix $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}$ in closed form:

$$\det A = a_{11}a_{22} - a_{21}a_{12}$$

The necessary and sufficient condition for f to be invertible is $a_{11}a_{22} - a_{21}a_{12} \neq 0$. There is no condition on b.

Problem 2: We consider the following forward tranformation $f(\boldsymbol{z}) = \begin{bmatrix} z_1 \\ e^{z_1} z_2 \\ |1+z_2|z_3 + \sin(z_1) \end{bmatrix} = \boldsymbol{x}$ from $\mathbb{R}^3$ to $\mathbb{R}^3$. We assume a uniform base distribution $p_1(\boldsymbol{z}) = U([0,2]^3)$. Evaluate the density $p_2(\boldsymbol{x})$ at the two points $\boldsymbol{x}^{(1)} = \begin{bmatrix} 1 \\ 2e \\ 3+\sin(1) \end{bmatrix}$ and $\boldsymbol{x}^{(2)} = \begin{bmatrix} 2 \\ e^2 \\ 6+\sin(2) \end{bmatrix}$.

We are asked for density estimation. Therefore, we first compute the reverse transformation by solving the system of (non-linear) equations:

$$\begin{cases} x_1 = z_1 \\ x_2 = e^{z_1} z_2 \\ x_3 = |1+z_2|z_3 + \sin(z_1) \end{cases}$$

$$\begin{cases} z_1 = x_1 \\ z_2 = x_2 e^{-x_1} \\ z_3 = \frac{x_3 - \sin(x_1)}{|x_2 e^{-x_1} + 1|} \end{cases}$$

Thus f is invertible with inverse:

$$f^{-1}(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ x_2 e^{-x_1} \\ \frac{x_3 - \sin(x_1)}{|x_2 e^{-x_1} + 1|} \end{bmatrix}$$

Second, we compute the Jacobian determinant. We remark that the Jacobian is triangular. Hence, we only need the diagonal coefficients to compute its determinant:

$$\det J_{f^{-1}(\boldsymbol{x})} = \begin{vmatrix} 1 & 0 & 0 \\ * & e^{-x_1} & 0 \\ * & * & \frac{1}{x_2 e^{-x_1} + 1} \end{vmatrix} = \frac{1}{|x_2 + e^{x_1}|}$$

Third, we compute the inverse of $\boldsymbol{x}^{(1)} = \begin{bmatrix} 1 \\ 2e \\ 3+\sin(1) \end{bmatrix}$ and $\boldsymbol{x}^{(2)} = \begin{bmatrix} 2 \\ e^2 \\ 6+\sin(2) \end{bmatrix}$:

$$f^{-1}(\boldsymbol{x}^{(1)}) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \boldsymbol{z}^{(1)}$$

$$f^{-1}(\boldsymbol{x}^{(2)}) = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \boldsymbol{z}^{(2)}$$

Finally, we use the change of variables formula. Since $f^{-1}(\boldsymbol{x}^{(1)}) \in [0,2]^3$, we have $p_1(f^{-1}(\boldsymbol{x}^{(1)})) = \frac{1}{2^3}$:

$$\begin{aligned} p_2(\boldsymbol{x}^{(1)}) &= p_1(f^{-1}(\boldsymbol{x}^{(1)}))|\det J_{f^{-1}(\boldsymbol{x}^{(1)})}| \\ &= \frac{1}{2^3}\frac{1}{3e} = \frac{1}{24e} \end{aligned}$$

Since $f^{-1}(\boldsymbol{x}^{(2)}) \notin [0,2]^2$, we have $p_1(f^{-1}(\boldsymbol{x}^{(2)})) = 0$:

$$\begin{aligned} p_2(\boldsymbol{x}^{(2)}) &= p_1(f^{-1}(\boldsymbol{x}^{(2)}))|\det J_{f^{-1}(\boldsymbol{x}^{(2)})}| \\ &= 0 \end{aligned}$$

Problem 3: We consider the following forward transformation $x = f(z) = \sum_{k=1}^{K} \sigma(kz)$ from $\mathbb{R}$ to $]0, K[$ with $\sigma(z) = \frac{1}{1+e^{-z}}$. We assume a Gaussian base distribution $p_1(z) = \mathcal{N}(0,1)$. We sampled one point from the base distribution $z^{(1)} = 0$. Compute the corresponding sample $x^{(1)}$ from the transformed distribution and evaluate its density $p_2(x^{(1)})$.

We are asked for sampling. We compute first the forward transformation of $z^{(1)}$:

$$f(z^{(1)}) = \sum_{k=1}^{K} \sigma(k \times 0) = \frac{K}{2}$$

To evaluate the density, we need first the Jacobian determinant:

$$\begin{aligned} \det J_{f(z)} &= \frac{\partial f(z)}{\partial z} \\ &= \sum_{k=1}^{K} \frac{\partial \sigma(kz)}{\partial z} \\ &= \sum_{k=1}^{K} \frac{ke^{-kz}}{(1+e^{-kz})^2} \\ &= \sum_{k=1}^{K} k\sigma(kz)(1-\sigma(kz)) \end{aligned}$$

Using the change of variables formula, we obtain:

$$\begin{aligned} p_2(x^{(1)}) &= p_1(z^{(1)})|\det J_{f(z^{(1)})}|^{-1} \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sum_{k=1}^{K} k\sigma(k \times 0)(1-\sigma(k \times 0))} \\ &= \frac{1}{\sqrt{2\pi}} \frac{4}{\sum_{k=1}^{K} k} \\ &= \frac{1}{\sqrt{2\pi}} \frac{8}{K(K+1)} \end{aligned}$$

Problem 4: We consider the forward transformation $x = f(z) = az + b$ from $\mathbb{R}$ to $\mathbb{R}$ where $a, b \in \mathbb{R}$ are learnable parameters. We assume a Gaussian base distribution $p_1(z) = \mathcal{N}(0, 1)$. We observed three points $x^{(1)} = 0, x^{(2)} = 1, x^{(3)} = 2$. Compute the maximum likelihood estimate of the parameters a, b.

The inverse of the transformation is $f^{-1}(x) = \frac{x-b}{a}$ and its Jacobian determinant $\det J_{f^{-1}(x^{(i)})} = \frac{1}{a}$. We compute the log-likelihood of the three points.

$$\begin{aligned}
\sum_{i=1}^{3} \log p_2(x^{(i)}) &= \sum_{i=1}^{3} \log p_1(f^{-1}(x^{(i)})) + \log \det J_{f^{-1}(x^{(i)})} \\
&= \sum_{i=1}^{3} -\frac{1}{2} \left(\frac{x^{(i)} - b}{a} \right)^2 - \log a - \log\sqrt{2\pi} \\
&= -\frac{1}{2a^2}(b^2 + (1-b)^2 + (2-b)^2) - 3\log a - 3\log\sqrt{2\pi} \\
&= -\frac{3b^2 - 6b + 5}{2a^2} - 3\log a - 3\log\sqrt{2\pi}
\end{aligned}$$

Computing the derivatives w.r.t. b and a and setting it to 0 gives:

$$\begin{cases} -\frac{1}{2a^2}(6b - 6) = 0 \\ \frac{1}{a^3}(3b^2 - 6b + 5) - \frac{3}{a} = 0 \end{cases}$$

Solving the system gives us two solutions $a = \frac{\sqrt{2}}{\sqrt{3}}, b = 1$ and $a = -\frac{\sqrt{2}}{\sqrt{3}}, b = 1$

.