CS6202: Problem Set 1

Question 1

Inverse-transform sampling is based on the probability integral transform which states the following—if X is continuous random variable (RV) with the cumulative distribution function (CDF) F_X, then the random variable Y has the uniform distribution on $[0,1]$. Verify this statement by giving an example in the context of normalizing flows (based on the slides on "Transformation of a Random Variable" from the presentation). Specifically, define a RV X with probability density $p(x)$ and the smooth, invertible transformation function $f = F_X$. Derive the analytic density of $Y = f(X)$ and show that it is equal to $\mathcal{U}(0,1)$, either using the CDF technique or the general change-of-variable formula.

Solution 1

Let $p(x) = \lambda e^{-\lambda x}$ which is the exponential distribution where $\lambda > 0$ and support $x \in [0, \infty)$. The CDF of the exponential distribution is given by $F_X(x) = 1 - e^{-\lambda x}$. Define $f = F_X$. Now, if $Y = f(X)$, we can see that the support of $p(y)$ is $[0,1]$ by plugging in the support of $p(x)$ in f, i.e., $f(0) = 1 - e^{-\lambda 0} = 0$ and $\lim_{x\to\infty} 1 - e^{-\lambda x} = 1$.

Now, we show that the density function of Y is 1, i.e., it's a uniform continuous distribution on $[0,1]$.

- Using the CDF technique:

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(1 - e^{-\lambda X} \leq y) \\
&= P(e^{-\lambda X} \geq -y + 1) \\
&= P(-\lambda X \geq \log(1-y)) \\
&= P\left(X \leq -\frac{\log(1-y)}{\lambda}\right) \\
&= F_X\left(-\frac{\log(1-y)}{\lambda}\right)
\end{aligned}$$

  Now, $p(y) = F'_Y(y)$

$$\begin{aligned}
p(y) &= \frac{d}{dy}\left\{F_Y(y)\right\} \\
&= \frac{d}{dy}\left\{F_X\left(-\frac{\log(1-y)}{\lambda}\right)\right\} \\
&= \frac{d}{dy}\left\{\int_0^{-\frac{\log(1-y)}{\lambda}} \lambda e^{-\lambda x} dx\right\} \\
&= \frac{d}{dy}\left\{\left[\frac{\lambda e^{-\lambda x}}{-\lambda}\right]_0^{-\frac{\log(1-y)}{\lambda}}\right\} \\
&= \frac{d}{dy}\left\{\left[-e^{-\lambda \cdot -\frac{\log(1-y)}{\lambda}} - [-e^{-\lambda \cdot 0}]\right]\right\} \\
&= \frac{d}{dy}\left\{[y - 1 - [-1]]\right\} = \frac{d}{dy}\left\{y\right\} = 1
\end{aligned}$$

- Using change-of-variable formula:

  The change-of-variable formula says that

$$
\begin{aligned}
p(y) &= p(x)\left|\frac{df}{dx}\right|^{-1} \\
&= \lambda e^{-\lambda x}\left|\frac{d}{dx}\left\{1-e^{-\lambda x}\right\}\right|^{-1} \\
&= \lambda e^{-\lambda x}\left|\lambda e^{-\lambda x}\right|^{-1} \\
&= \frac{\lambda e^{-\lambda x}}{\lambda e^{-\lambda x}} = 1
\end{aligned}
$$

Question 2

The transformation $f(\mathbf{z}) = \mathbf{z} + \mathbf{u}h(\mathbf{w}^\top\mathbf{z} + b)$ used in planar flows is not always invertible. What are the constraints on $\mathbf{u}$, $\mathbf{w}$, and b (if any) for the invertiblity of the function? Derive the constraints for the case when $z, u, w \in \mathbb{R}$ and then extend to the case when $\mathbf{z}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^d$. Assume that the fuction h is the logistic function $h(x) = \frac{1}{1+e^{-x}}$.

Solution 2

When $z, u, w \in \mathbb{R}$, we have $f(z) = z + u \cdot h(wz + b)$. A sufficient condition for the invertibility of $f(z)$ is that it is an increasing function, i.e., $f'(z) > 0$.

$$
\begin{aligned}
f'(z) &> 0 \\
1 + uw \cdot h'(wz + b) &> 0 \\
uw &> -\frac{1}{h'(wz + b)}
\end{aligned}
$$

The function $h'(x) = \frac{e^x}{(e^x+1)^2}$ takes its maximum value at $x = 0$ which is 0.25. Therefore, it is sufficient that $uw > -4$ for the function f to be invertible.

Let's extend this to the case when $\mathbf{z}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^d$. We begin by splitting $\mathbf{z}$ into two components $\mathbf{z}_\perp$ (perpendicular to $\mathbf{w}$) and $\mathbf{z}_\parallel$ (parallel to $\mathbf{w}$). Note that $\because$ $\mathbf{z}_\parallel$ is parallel to $\mathbf{w}$, it can also be written as $\alpha\frac{\mathbf{w}}{\|\mathbf{w}\|^2}$. With this is mind,

$$
f(\mathbf{z}) = \mathbf{z}_\perp + \mathbf{z}_\parallel + \mathbf{u}h(\mathbf{w}^\top\mathbf{z}_\parallel + b) \tag{1}
$$

Let $\mathbf{y} = f(\mathbf{z})$; now, $\mathbf{z}_\perp$ can be uniquely computed as follows (given that we have $\mathbf{z}_\parallel$)

$$
\mathbf{z}_\perp = \mathbf{y} - \mathbf{z}_\parallel - \mathbf{u}h(\mathbf{w}^\top\mathbf{z}_\parallel + b) \tag{2}
$$

Substituting $\mathbf{z}_\parallel = \alpha\frac{\mathbf{w}}{\|\mathbf{w}\|^2}$ in Eq. (1)

$$
f(\mathbf{z}) = \mathbf{z}_\perp + \alpha\frac{\mathbf{w}}{\|\mathbf{w}\|^2} + \mathbf{u}h(\alpha + b) \tag{3}
$$

$$
\mathbf{w}^\top f(\mathbf{z}) = \alpha + \mathbf{w}^\top\mathbf{u}h(\alpha + b) \qquad \text{(dot product with } \mathbf{w}\text{)} \tag{4}
$$

We need to solve the scalar equation in Eq. (4) to solve for α. For a unique solution, the RHS of Eq. (4) must be an increasing function, i.e.,

$$
\begin{aligned}
1 + \mathbf{w}^\top\mathbf{u}h'(\alpha + b) &> 0 \\
\mathbf{w}^\top\mathbf{u} &> -\frac{1}{h'(\alpha + b)}
\end{aligned}
$$

Again, as before, $\max h'(x) = 0.25$ at $x = 0$. Therefore, it suffices that $\mathbf{w}^\top \mathbf{u} > -4$.

Question 3

Implement a `Bijector` in Tensorflow Probability for the sigmoid activation function.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Solution 3

Forward function:

$$y = \frac{1}{1 + e^{-x}}$$

Inverse function:

$$x = -\log\left(\frac{1}{y} - 1\right)$$

Inverse Jacobian's diagonal element:

$$\left.\frac{\partial \sigma^{-1}}{\partial y}\right|_{ii} = \frac{1}{y_i - y_i^2}$$