1 Change of Random Variables

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a bijective differentiable function. Let $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^n$ be random vectors such that $Y = f(X)$. The Jacobian matrix of f is defined by

$$J_f(x) = \frac{df}{dx} = \left[\frac{\partial f_i}{\partial x_j}\right]_{1 \le i,j \le n}$$

where f_i is the i-th coordinate function of f and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. The Jacobian matrix of the inverse function f^{-1} satisfies

$$J_{f^{-1}}(y) = \left(J_f(f^{-1}(y))\right)^{-1},$$
$$\det J_{f^{-1}}(y) = \frac{1}{\det J_f(f^{-1}(y))}$$

for any $y \in \mathbb{R}^n$.

Let p_X and p_Y be the joint probability density functions of X and Y. By using the change of variables, for any measurable subset A in $\mathbb{R}^n$,

$$\begin{aligned} P\{Y \in A\} &= P\{X \in f^{-1}(A)\} \\ &= \int \cdots \int_{x \in f^{-1}(A)} p_X(x)\, dx \\ &= \int \cdots \int_{y \in A} p_X(f^{-1}(y))\, |\det J_{f^{-1}}(y)|\, dy \\ &= \int \cdots \int_{y \in A} p_X(f^{-1}(y))\, \frac{1}{|\det J_f(f^{-1}(y))|}\, dy. \end{aligned}$$

Therefore, p_Y is given by

$$p_Y(y) = p_X(f^{-1}(y))\, \frac{1}{|\det J_f(f^{-1}(y))|}. \tag{1}$$

2 Change of Conditional Random Variables

Consider a conditional random vector $X_1|Y_1$ which follows a conditional probability density function

$$p_{X_1|Y_1}(x_1 \mid y_1) \stackrel{\text{def}}{=} \frac{p_{X_1,Y_1}(x_1, y_1)}{\int p_{X_1,Y_1}(x_1, y_1)\, dx_1}$$

where $X_1, x_1 \in \mathbb{R}^m$ and $Y_1, y_1 \in \mathbb{R}^n$. Let $f_Y : \mathbb{R}^m \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ be bijective differentiable transformations for any $Y \in \mathbb{R}^n$, and let $X_2 = f_{Y_1}(X_1)$ and $Y_2 = g(Y_1)$. Then, the transformation $h : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$ defined by

$$h(X_1, Y_1) = (f_{Y_1}(X_1), g(Y_1))$$

is also bijective and differentiable, whose inverse function is given by

$$h^{-1}(X_2, Y_2) = (f^{-1}_{g^{-1}(Y_2)}(X_2), g^{-1}(Y_2)).$$

The Jacobian matrix of h can be partitioned as

$$J_h(X_1, Y_1) = \begin{bmatrix} J_{f_{Y_1}}(X_1) & \frac{\partial f_{Y_1}}{\partial Y_1}(X_1) \\ 0 & J_g(Y_1) \end{bmatrix},$$

thus its Jacobian determinant is

$$|\det J_h(X_1, Y_1)| = |\det J_{f_{Y_1}}(X_1)| \cdot |\det J_g(Y_1)|. \tag{2}$$

The joint probability density function of X_2, Y_2 can be derived by applying (1) and (2):

$$\begin{aligned}
p_{X_2,Y_2}(x_2, y_2) &= p_{X_1,Y_1}(h^{-1}(x_2, y_2)) \frac{1}{|\det J_h(h^{-1}(x_2, y_2))|} \\
&= p_{X_1,Y_1}(f^{-1}_{g^{-1}(y_2)}(x_2), g^{-1}(y_2)) \frac{1}{|\det J_h(f^{-1}_{g^{-1}(y_2)}(x_2), g^{-1}(y_2))|} \\
&= p_{X_1,Y_1}(f^{-1}_{g^{-1}(y_2)}(x_2), g^{-1}(y_2)) \frac{1}{|\det J_{f_{g^{-1}(y_2)}}(f^{-1}_{g^{-1}(y_2)}(x_2))| \cdot |\det J_g(g^{-1}(y_2))|}.
\end{aligned}$$

Marginalising x_2 out gives the normalising constant:

$$\begin{aligned}
\int p_{X_2,Y_2}(x_2, y_2)\, dx_2 &= \frac{1}{|\det J_g(g^{-1}(y_2))|} \int p_{X_1,Y_1}(f^{-1}_{g^{-1}(y_2)}(x_2), g^{-1}(y_2)) \frac{1}{|\det J_{f_{g^{-1}(y_2)}}(f^{-1}_{g^{-1}(y_2)}(x_2))|}\, dx_2 \\
&= \frac{1}{|\det J_g(g^{-1}(y_2))|} \int p_{X_1,Y_1}(u, g^{-1}(y_2))\, du,
\end{aligned}$$

by using the change of variables with the substitution $u = f^{-1}_{g^{-1}(y_2)}(x_2)$. Now, combining these two results gives the conditional probability density function of $X_2|Y_2$:

$$\begin{aligned}
p_{X_2|Y_2}(x_2 \,|\, y_2) &= \frac{p_{X_2,Y_2}(x_2, y_2)}{\int p_{X_2,Y_2}(x_2, y_2)\, dx_2} \\
&= \frac{p_{X_1,Y_1}(f^{-1}_{g^{-1}(y_2)}(x_2), g^{-1}(y_2))}{\int p_{X_1,Y_1}(u, g^{-1}(y_2))\, du} \cdot \frac{1}{|\det J_{f_{g^{-1}(y_2)}}(f^{-1}_{g^{-1}(y_2)}(x_2))|} \\
&= p_{X_1|Y_1}(f^{-1}_{g^{-1}(y_2)}(x_2) \,|\, g^{-1}(y_2)) \cdot \frac{1}{|\det J_{f_{g^{-1}(y_2)}}(f^{-1}_{g^{-1}(y_2)}(x_2))|}. \qquad (3)
\end{aligned}$$

3 Reparameterisation

Let $X_1 \in \mathbb{R}^m$ and $Y_1 \in \mathbb{R}^n$ be random vectors. Our objective is to sample X_1 from the posterior density $p_{X_1|Y_1}(x_1 \,|\, y_1)$. To do this, we consider a bijective differentiable reparameterisation $h : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$ defined as above, and

$$(X_2, y_2) = h(X_1, y_1) = (f_{y_1}(X_1), g(y_1))$$

such that the posterior density $p_{X_2|Y_2}(x_2 \,|\, y_2)$ has a proper geometry. By doing so, we can easily sample X_2 from $p_{X_2|Y_2}(x_2 \,|\, y_2)$, and take $X_3 = f^{-1}_{g^{-1}(y_2)}(x_2)$. We claim that the random vector X_3 follows the desired posterior density $p_{X_1|Y_1}(x_1 \,|\, y_1)$.

Proof. By applying (1) and (3),

$$\begin{aligned}
p_{X_3}(x_3) &= p_{X_2|Y_2}(f_{g^{-1}(y_2)}(x_3) \,|\, y_2) \cdot \frac{1}{|\det J_{f^{-1}_{g^{-1}(y_2)}}(f_{g^{-1}(y_2)}(x_3))|} \\
&= p_{X_1|Y_1}(x_3 \,|\, g^{-1}(y_2)) \cdot \frac{1}{|\det J_{f_{g^{-1}(y_2)}}(x_3)|} \cdot |\det J_{f_{g^{-1}(y_2)}}(x_3)| \\
&= p_{X_1|Y_1}(x_3 \,|\, g^{-1}(y_2)) \\
&= p_{X_1|Y_1}(x_3 \,|\, y_1)
\end{aligned}$$

□