Normalizing Flow history

General coupling layer Let $x \in \mathcal{X}$, I_1, I_2 a partition of $[\![1, D]\!]$ such that $d = |I_1|$ and m a function defined on $\mathbb{R}^d$, we can define $y = (y_{I_1}, y_{I_2})$ where:

$$y_{I_1} = x_{I_1}$$
$$y_{I_2} = g(x_{I_2}; m(x_{I_1}))$$

where $g : \mathbb{R}^{D-d} \times m(\mathbb{R}^d) \to \mathbb{R}^{D-d}$ is the *coupling law*, an invertible map with respect to its first argument given the second. The corresponding computational graph is shown Fig 2. If we consider $I_1 = [\![1, d]\!]$ and $I_2 = [\![d, D]\!]$, the Jacobian of this function is:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} I_d & 0 \\ \frac{\partial y_{I_2}}{\partial x_{I_1}} & \frac{\partial y_{I_2}}{\partial x_{I_2}} \end{bmatrix}$$

Where I_d is the identity matrix of size d. That means that $\det \frac{\partial y}{\partial x} = \det \frac{\partial y_{I_2}}{\partial x_{I_2}}$. Also, we observe we can invert the mapping using:

$$x_{I_1} = y_{I_1}$$
$$x_{I_2} = g^{-1}(y_{I_2}; m(y_{I_1}))$$

We call such a transformation a *coupling layer* with *coupling function* m.

Additive coupling layer For simplicity, we choose as coupling law an *additive coupling law* $g(a; b) = a + b$ so that by taking $a = x_{I_2}$ and $b = m(x_{I_1})$:

$$y_{I_2} = x_{I_2} + m(x_{I_1})$$
$$x_{I_2} = y_{I_2} - m(y_{I_1})$$

2014
Dinh, Laurent, David Krueger, and Yoshua Bengio. "Nice: Non-linear independent components estimation."

Normalizing Flow history

General coupling layer Let $x \in \mathcal{X}$, I_1, I_2 a partition of $[\![1, D]\!]$ such that $d = |I_1|$ and m a function defined on $\mathbb{R}^d$, we can define $y = (y_{I_1}, y_{I_2})$ where:

$$y_{I_1} = x_{I_1}$$
$$y_{I_2} = g(x_{I_2}; m(x_{I_1}))$$

where $g : \mathbb{R}^{D-d} \times m(\mathbb{R}^d) \to \mathbb{R}^{D-d}$ is the *coupling law*, an invertible map with respect to its first argument given the second. The corresponding computational graph is shown Fig 2. If we consider $I_1 = [\![1, d]\!]$ and $I_2 = [\![d, D]\!]$, the Jacobian of this function is:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} I_d & 0 \\ \frac{\partial y_{I_2}}{\partial x_{I_1}} & \frac{\partial y_{I_2}}{\partial x_{I_2}} \end{bmatrix}$$

Where I_d is the identity matrix of size d. That means that $\det \frac{\partial y}{\partial x} = \det \frac{\partial y_{I_2}}{\partial x_{I_2}}$. Also, we observe we can invert the mapping using:

$$x_{I_1} = y_{I_1}$$
$$x_{I_2} = g^{-1}(y_{I_2}; m(y_{I_1}))$$

We call such a transformation a *coupling layer* with *coupling function* m.

Additive coupling layer For simplicity, we choose as coupling law an *additive coupling law* $g(a; b) = a + b$ so that by taking $a = x_{I_2}$ and $b = m(x_{I_1})$:

$$y_{I_2} = x_{I_2} + m(x_{I_1})$$
$$x_{I_2} = y_{I_2} - m(y_{I_1})$$

2014
Dinh, Laurent, David Krueger, and Yoshua Bengio. "Nice: Non-linear independent components estimation."

$$y_{1:d} = x_{1:d} \tag{4}$$
$$y_{d+1:D} = x_{d+1:D} \odot \exp\left(s(x_{1:d})\right) + t(x_{1:d}), \tag{5}$$

where s and t stand for scale and translation, and are functions from $R^d \mapsto R^{D-d}$, and $\odot$ is the Hadamard product or element-wise product (see Figure 2(a)).

3.3 Properties

The Jacobian of this transformation is

$$\frac{\partial y}{\partial x^T} = \begin{bmatrix} \mathbb{I}_d & 0 \\ \frac{\partial y_{d+1:D}}{\partial x_{1:d}^T} & \operatorname{diag}\left(\exp\left[s\left(x_{1:d}\right)\right]\right) \end{bmatrix}, \tag{6}$$

2016
Dinh, Laurent, Jascha Sohl-Dickstein, and Samy Bengio. "Density estimation using real nvp."