Appendix: Variational Inference with Normalizing Flows

Danilo Jimenez Rezende DANILOR@GOOGLE.COM
Shakir Mohamed SHAKIR@GOOGLE.COM
Google DeepMind, London

A. Invertibility conditions

We describe the constraints required to have invertible maps for the planar and radial normalizing flows described in section 3.

A.1. Planar flows

Functions of the form (10) are not always invertible depending on the non-linearity and parameters chosen. When using $h(x) = \tanh(x)$, a sufficient condition for $f(\mathbf{z})$ to be invertible is that $\mathbf{w}^\top\mathbf{u} \geq -1$.

This can be seen by splitting $\mathbf{z}$ as a sum of a vector $\mathbf{z}_\perp$ perpendicular to $\mathbf{w}$ and a vector $\mathbf{z}_\parallel$, parallel to $\mathbf{w}$. Substituting $\mathbf{z} = \mathbf{z}_\perp + \mathbf{z}_\parallel$ into (10) gives

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

This equation can be solved for $\mathbf{z}_\perp$ given $\mathbf{z}_\parallel$ and $\mathbf{y} = f(\mathbf{z})$, having a unique solution

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

The parallel component can be further expanded as $\mathbf{z}_\parallel = \alpha\frac{\mathbf{w}}{||\mathbf{w}||^2}$, where $\alpha \in \mathbb{R}$. The equation that must be solved for α is derived by taking the dot product of (1) with $\mathbf{w}$, yielding the scalar equation

$$\mathbf{w}^T f(\mathbf{z}) = \alpha + \mathbf{w}^T\mathbf{u}h(\alpha + b). \quad (3)$$

A sufficient condition for (3) to be invertible w.r.t α is that its r.h.s $\alpha + \mathbf{w}^T\mathbf{u}h(\alpha + b)$ to be a non-decreasing function. This corresponds to the condition $1 + \mathbf{w}^T\mathbf{u}h'(\alpha + b) \geq 0 \equiv \mathbf{w}^T\mathbf{u} \geq -\frac{1}{h'(\alpha+b)}$. Since $0 \leq h'(\alpha + b) \leq 1$, it suffices to have $\mathbf{w}^T\mathbf{u} \geq -1$.

We enforce this constraint by taking an arbitrary vector $\mathbf{u}$ and modifying its component parallel to $\mathbf{w}$, producing a new vector $\hat{\mathbf{u}}$ such that $\mathbf{w}^\top\hat{\mathbf{u}} > -1$. The modified vector can be compactly written as $\hat{\mathbf{u}}(\mathbf{w}, \mathbf{u}) = \mathbf{u} + \left[m(\mathbf{w}^\top\mathbf{u}) - (\mathbf{w}^\top\mathbf{u})\right]\frac{\mathbf{w}}{||\mathbf{w}||^2}$, where the scalar function $m(x)$ is given by $m(x) = -1 + \log(1 + e^x)$.

A.2. Radial flows

Functions of the form (14) are not always invertible depending on the values of α and β. This can be seen by splitting the vector $\mathbf{z}$ as $\mathbf{z} = \mathbf{z}_0 + r\hat{\mathbf{z}}$, where $r = |\mathbf{z} - \mathbf{z}_0|$. Replacing this into (14) gives

$$f(\mathbf{z}) = \mathbf{z}_0 + r\hat{\mathbf{z}} + \beta\frac{r\hat{\mathbf{z}}}{\alpha + r}. \quad (4)$$

This equation can be uniquely solved for $\hat{\mathbf{z}}$ given r and $\mathbf{y} = f(\mathbf{z})$,

$$\hat{\mathbf{z}} = \frac{\mathbf{y} - \mathbf{z}_0}{r\left(1 + \frac{\beta}{\alpha+r}\right)}. \quad (5)$$

To obtain a scalar equation for the norm r, we can subtract both sides of (4) and take the norm of both sides. This gives

$$|y - \mathbf{z}_0| = r\left(1 + \frac{\beta}{\alpha + r}\right). \quad (6)$$

A sufficient condition for (6) to be invertible is for its r.h.s. $r\left(1 + \frac{\beta}{\alpha+r}\right)$ to be a non-decreasing function, which implies $\beta \geq -\frac{(r+\alpha)^2}{\alpha}$. Since $r \geq 0$, it suffices to impose $\beta \geq -\alpha$. This constraint is imposed by reparametrizing β as $\hat{\beta} = -\alpha + m(\beta)$, where $m(x) = -1 + \log(1 + e^x)$.