

# On the Gumbel-Softmax Trick for Inference of Discrete Variables

## Abstract

The re-parameterization trick enables optimizing stochastic computation graphs via gradient descent. The essence of the trick is to re-factor each stochastic node into a differentiable function of its parameters and a random variable with fixed distribution. After refactoring, the gradients of the loss propagated by the chain rule through the graph are low variance unbiased estimators of the gradients of the expected loss. While many continuous random variables have such re-parameterizations, discrete random variables lack useful re-parameterizations due to the discontinuous nature of discrete states. As a remedy, the concrete distribution serves as a continuous relaxation of discrete random variables.

## 1 Reparameterization Trick

The law of the unconscious statistician states (LOTUS) states that for a random variable  $\epsilon$  with pdf  $f_X$  and a measurable function  $g$ , it holds:

$$\mathbb{E}(g(\epsilon)) = \int g(X) dF_\epsilon(x). \quad (1)$$

In other words, to compute the expectation of  $z = g(\epsilon)$  we only need to know the mapping  $g$  and the distribution of  $\epsilon$ , but we do not need the explicit distribution of  $z$

$$\mathbb{E}_{\epsilon \sim p(\epsilon)}(g(\epsilon)) = \mathbb{E}_{z \sim p(z)}(z). \quad (2)$$

Now, suppose  $z$  has a distribution that depends on a parameter  $\phi$ , i.e.  $z \sim p_\phi(z)$ . Moreover, assume one can express  $z = g(\epsilon, \phi)$  for known function  $g$  and a certain noise distribution, e.g.  $\epsilon \sim \mathcal{N}(0, 1)$ . Then LOTUS states that for any measurable function  $f$ :

$$\mathbb{E}_{z \sim p_\phi(z)}(f(z)) = \mathbb{E}_{\epsilon \sim p(\epsilon)}(f(g(\epsilon, \phi))). \quad (3)$$

In black-box variational inference formulations we encounter the gradient of some expectation with respect to a parameter  $\phi$  and use the following equality:

$$\nabla_\phi \mathbb{E}_{z \sim p(z)} = \nabla_\phi \mathbb{E}_{\epsilon \sim p(\epsilon)}(f(g(\epsilon, \phi))) = \mathbb{E}_{\epsilon \sim p(\epsilon)}(\nabla f(g(\epsilon, \phi))). \quad (4)$$

Further, we have conveniently expressed  $z$  so that expectations of functions of  $z$  can be expressed as integrals with respect to a density that does not depend on the parameter. Therefore, we can exchange the expectation and gradient.

Finally, the re-parameterization gives rise to an unbiased estimate of the above gradient via MCMC:

$$\nabla_{\phi} \mathbb{E}_{z \sim p(z)} \approx \frac{1}{N} \sum_{i=1}^N \nabla f(g(\epsilon_i, \phi)). \quad (5)$$

For reasons not yet completely understood, empirically it is seen this re-parameterization based estimate of the gradient exhibits much less variance than competing estimators.

## 2 Variational Inference

### References