Gumbel-Softmax and Reparametrization

Paper 1

Paper 2

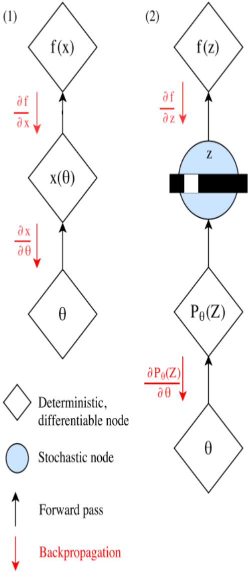
Paper 1: THE CONCRETE DISTRIBUTION: A CONTINUOUS RELAXATION OF DISCRETE RANDOM VARIABLES

Chris J. Maddison[1,2], Andriy Mnih[1], Yee Whye Teh[1]
1: DeepMind, London, United Kingdom
2: University of Oxford, Oxford, United Kingdom

Paper 2: CATEGORICAL REPARAMETERIZATION WITH GUMBEL-SOFTMAX

Eric Jang[1], Shixiang Gu[2], Ben Poole[3]
1: Google Brain
2: University of Cambridge
3: Stanford University

Stochastic Nodes



Optimizing Stochastic Computation Graphs

$$\nabla_\theta L(\theta,\phi) = \nabla_\theta \mathbb{E}_{X\sim p_\phi(x)}[f_\theta(X)] = \mathbb{E}_{X\sim p_\phi(x)}[\nabla_\theta f_\theta(X)]$$

Figure: Objective

$$\nabla_\theta L(\theta,\phi) \simeq \frac{1}{S}\sum_{s=1}^{S} \nabla_\theta f_\theta(X^s),$$

Figure: wrt θ

$$\nabla_\phi L(\theta,\phi) = \nabla_\phi \int p_\phi(x) f_\theta(x)\,\mathrm{d}x = \int f_\theta(x) \nabla_\phi p_\phi(x)\,\mathrm{d}x,$$

Figure: wrt *phi*

Related: Score function estimators

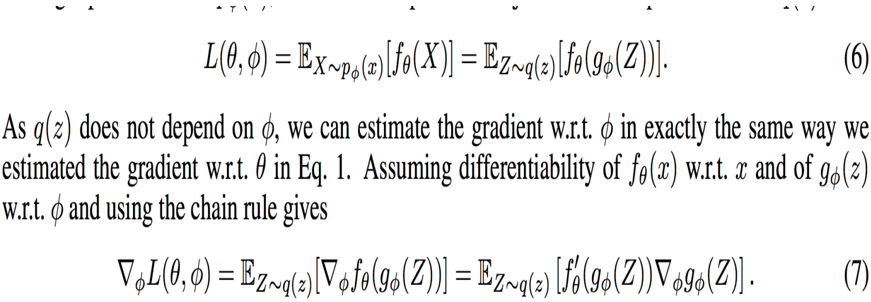
$$\nabla_\phi L(\theta, \phi) = \mathbb{E}_{X \sim p_\phi(x)} \left[f_\theta(X) \nabla_\phi \log p_\phi(X) \right].$$

Estimating this expectation using naive Monte Carlo gives the estimator

$$\nabla_\phi L(\theta, \phi) \simeq \frac{1}{S} \sum_{s=1}^{S} f_\theta(X^s) \nabla_\phi \log p_\phi(X^s),$$

Figure: REINFORCE

Related: Reparametrization



$$L(\theta,\phi)=\mathbb{E}_{X\sim p_\phi(x)}[f_\theta(X)]=\mathbb{E}_{Z\sim q(z)}[f_\theta(g_\phi(Z))]. \tag{6}$$

As $q(z)$ does not depend on ϕ, we can estimate the gradient w.r.t. ϕ in exactly the same way we estimated the gradient w.r.t. θ in Eq. 1. Assuming differentiability of $f_\theta(x)$ w.r.t. x and of $g_\phi(z)$ w.r.t. ϕ and using the chain rule gives

$$\nabla_\phi L(\theta,\phi)=\mathbb{E}_{Z\sim q(z)}[\nabla_\phi f_\theta(g_\phi(Z))]=\mathbb{E}_{Z\sim q(z)}\left[f'_\theta(g_\phi(Z))\nabla_\phi g_\phi(Z)\right]. \tag{7}$$

Figure: REPARAMETRIZATION

Gumbel-Softmax : Reparametrization + continuous

$$z = \mathtt{one\_hot}\left(\arg\max_i \left[g_i + \log \pi_i\right]\right)$$

Figure: REPARAMETRIZATION

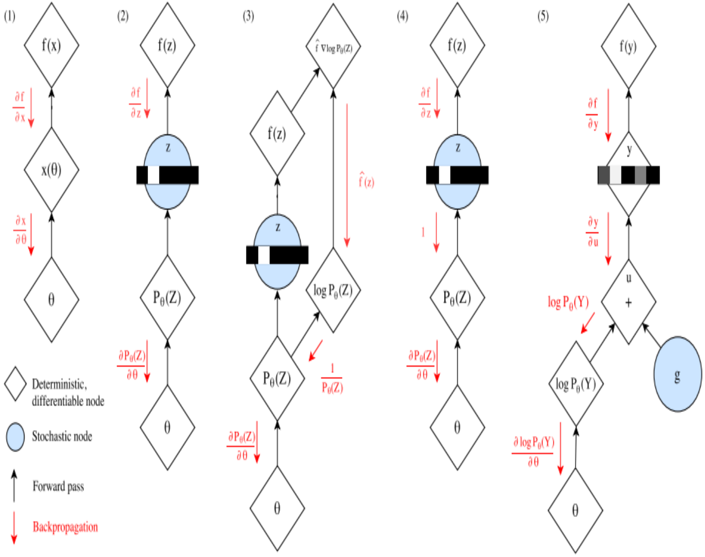
where

$$y_i = \frac{\exp((\log(\pi_i) + g_i)/\tau)}{\sum_{j=1}^{k} \exp((\log(\pi_j) + g_j)/\tau)} \qquad \text{for } i = 1, ..., k.$$

The density of the Gumbel-Softmax distribution (derived in Appendix B) is:

$$p_{\pi,\tau}(y_1, ..., y_k) = \Gamma(k)\tau^{k-1}\left(\sum_{i=1}^{k} \pi_i / y_i^{\tau}\right)^{-k} \prod_{i=1}^{k} \left(\pi_i / y_i^{\tau+1}\right)$$

Everything together



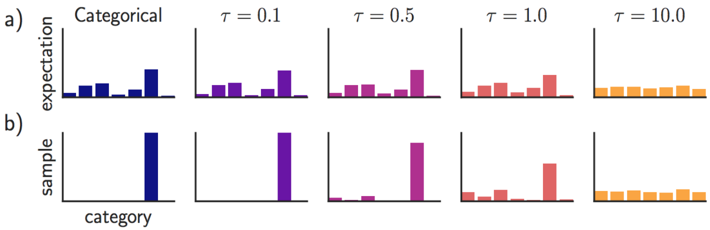


Figure 1: The Gumbel-Softmax distribution interpolates between discrete one-hot-encoded categorical distributions and continuous categorical densities. (a) For low temperatures ($\tau = 0.1, \tau = 0.5$), the expected value of a Gumbel-Softmax random variable approaches the expected value of a categorical random variable with the same logits. As the temperature increases ($\tau = 1.0, \tau = 10.0$), the expected value converges to a uniform distribution over the categories. (b) Samples from Gumbel-Softmax distributions are identical to samples from a categorical distribution as $\tau \to 0$. At higher temperatures, Gumbel-Softmax samples are no longer one-hot, and become uniform as $\tau \to \infty$.

Figure: REPARAMETRIZATION