Reparameterization Derivations

Clément Corbeau–Izorche

1 Beta distribution

Let us consider a Kumaraswamy distribution with parameters $(a, b) > 0$ whose density function is given by:

$$\text{Kumaraswamy}(x \mid a, b) = f(x \mid a, b) = abx^{a-1}(1 - x^a)^{b-1} \quad \text{for} \quad x \in [0, 1]$$

The associated cumulative distribution function is:

$$F(x \mid a, b) = \int_0^x f(t \mid a, b)dt = \Big[-(1 - x^a)^b \Big]_0^x = 1 - (1 - x^a)^b$$

To perform inverse sampling, we write x in terms of u:

$$u = F(x \mid a, b) = 1 - (1 - x^a)^b$$

$$\text{then,} \quad x = [1 - (1 - u)^{\frac{1}{b}}]^{\frac{1}{a}} = F^{-1}(u)$$

Therefore, by sampling $u \sim \mathcal{N}(0, I)$, x values will be samples from Kumaraswamy(a, b).

2 Dirichlet distribution

Let us consider a Dirichlet distribution with parameters $\alpha = (\alpha_1, ..., \alpha_K)$, whose density function is given by:

$$\text{Dirichlet}(\theta \mid \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{K} \theta_i^{\alpha_i - 1}$$

Srivastava Sutton, 2017 [1] re-writed this distribution in the softmax basis as following:

$$P(\theta(\mathbf{h}) \mid \alpha) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \prod_k \theta_k^{\alpha_k} g(\mathbf{1}^T \mathbf{h})$$

where $\theta = \sigma(\mathbf{h})$, with σ the softmax function.

Using Laplace approximation from Blei & Laffeirti, 2006 [2], this can be approximated by a multivariate normal with parameters μ and Σ such that:

$$\mu_k = \log \alpha_k - \frac{1}{K} \sum_i \alpha_i$$

and Σ is diagonal with:

$$\Sigma_{k,k} = \frac{1}{\alpha_k} \left(1 - \frac{2}{K}\right) + \frac{1}{K^2} \sum_i \frac{1}{\alpha_i}$$

Finally, the Dirichlet distribution can be approximated in the simplex basis with a logistic normal distribution $\mathcal{LN}(\mu, \Sigma)$.

Given that, samples from the approximation of the Dirichlet can be obtained by using the reparameterization trick for normal distributions:

$$\theta = e^{\mu + \Sigma\epsilon}$$

where $\epsilon \sim \mathcal{N}(0, I)$.

References

[1] Akash Srivastava Charles Sutton. Autoencoding variational inference for topic models. *ICLR*, 2017.

[2] Philipp Hennig David H Stern Ralf Herbrich Thore Graepel. Kernel topic models. *AISTATS*, 2012.