1 REINFORCE VAE

The loss function for a VAE is the following:

$$L(\theta, \phi) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - KL(q_\phi(z|x)||p(z))$$

When using discrete latent variables we can't compute the gradient of the first term directly, but we can use REINFORCE [1] (also known as the score function estimator) to compute it, see [2] for more details:

$$\nabla \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z) \nabla_\phi \log q_\phi(z|x)]$$

In general we suppose that $q(z|x) = \prod_i q(z_i|x)$ and $p(z) = \prod_i p(z_i)$. Thus the KL can usually be computed analyticaly:

$$KL(q_\phi(z|x)||p(z)) = \sum_i KL(q(z_i|x)||p(z_i))$$

When $p(z)$ is the uniform distribution then:

$$KL(q_\phi(z|x)||p(z)) = -H(q_\phi(z|x)) + \log d$$

where H is the entropy and d is the dimension of the categorical distribution.

2 Gumbel-Softmax VAE [3, 4]

To sample from a categorical distribution we can use the gumbel-max trick [5]:

$$z = \arg\max_i (g_i + \log \pi_i)$$

where $g_i \sim Gumbel(0, 1)$ and π_i are the unnormalized probability.

$$p(z = i) = \frac{\pi_i}{\sum_k \pi_k}$$

To sample from a $Gumbel(0, 1)$ distribution, we can sample from a uniform distribution $u \sim Uniform(0, 1)$, then $g = -\log(-\log(u))$ comes from a gumbel distribution.
If we replace the argmax by a softmax we can use the reparametrization trick and backpropagate through the sampling procedure. Thus this estimator is biased, however if we add a parameter to control the temperature of the softmax we can control the bias-variance trade-off of the estimator. A unbiased version of the estimator was proposed in [6] where they combine REINFORCE with the gumbel-softmax trick.
In practice we can use a Straight-Through [7] variant of the estimator, where in the forward pass we use the argmax but we replace the argmax by a softmax in the backward pass (this works a little better in general).

3 VQ-VAE [8]

The state of the art approach for learning discrete representation with VAEs. Use a simple Straight-Through estimator, can be improved using the gumbel-softmax trick. The main idea is to learn a shared dictionary of embeddings, the VAE then learns a a latent representation which acts like a look up table where each latent represent the index of an embedding.

References

[1] Williams, Ronald J. "Simple statistical gradient-following algorithms for connectionist reinforcement learning." Machine learning 8.3-4 (1992): 229-256.

[2] http://blog.shakirm.com/2015/11/machine-learning-trick-of-the-day-5-log-derivative-trick/

[3] Jang, Eric, Shixiang Gu, and Ben Poole. "Categorical reparameterization with gumbel-softmax." arXiv preprint arXiv:1611.01144 (2016).

[4] Maddison, Chris J., Andriy Mnih, and Yee Whye Teh. "The concrete distribution: A continuous relaxation of discrete random variables." arXiv preprint arXiv:1611.00712 (2016).

[5] Maddison, Chris J., Daniel Tarlow, and Tom Minka. "A* sampling." Advances in Neural Information Processing Systems. 2014.

[6] Tucker, George, et al. "Rebar: Low-variance, unbiased gradient estimates for discrete latent variable models." Advances in Neural Information Processing Systems. 2017.

[7] Bengio, Yoshua, Nicholas Léonard, and Aaron Courville. "Estimating or propagating gradients through stochastic neurons for conditional computation." arXiv preprint arXiv:1308.3432 (2013).

[8] van den Oord, Aaron, and Oriol Vinyals. "Neural discrete representation learning." Advances in Neural Information Processing Systems. 2017.