

for number of training iterations **do**

for k steps **do**

- 在噪声先验分布为 $p_g(z)$ 的 m 个噪声样本 $\{z^{(1)}, \dots, z^{(m)}\}$ 中采一个 minibatch。
- 在数据分布为 $p_{data}(x)$ 的 m 个训练样本 $\{x^{(1)}, \dots, x^{(m)}\}$ 中采一个 minibatch。
- 通过随机梯度上升来更新判别器：

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log (1 - D(G(z^{(i)})))]$$

end for

- 在噪声先验分布为 $p_g(z)$ 的 m 个噪声样本 $\{z^{(1)}, \dots, z^{(m)}\}$ 中采一个 minibatch。
- 通过随机梯度下降来更新生成器：

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (1 - D(G(z^{(i)})))$$

end for

基于梯度的更新可以使用任何标准的基于梯度的学习准则。我们在实验中使用了动量准则。