Langevin Monte Carlo

Say we want to sample from $p(w \mid D)$

Start from w^0

For k = 1, …

$$w^{k+1} = w^k + \varepsilon \nabla \log p(w^k \mid D) + \eta^k,$$

$$= w^k + \varepsilon \nabla \left(\log p(w^k) + \sum_{i=1}^{N} \log p(y_i \mid x_i, w^k) \right) + \eta^k$$

Weight decay $-C\|w^k\|^2$

Usual cross entropy

$$\eta^k \sim \mathcal{N}(0, 2\varepsilon I)$$