

Gradients

$$\nabla_w f(w, \phi) = \nabla_w \sum_{i=1}^N \mathbb{E}_{q(t_i|x_i, \phi)} \log p(x_i \mid t_i, w)$$