Gradients

$$\nabla_\phi f(w, \phi) = \nabla_\phi \sum_i \mathbb{E}_{q(t_i|x_i,\phi)} \log p(x_i \mid t_i, w)$$

Gradients

$$\nabla_\phi f(w, \phi) = \nabla_\phi \sum_i \int q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

Gradients

$$\nabla_\phi f(w, \phi) = \sum_i \int \nabla_\phi \, q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

Gradients

$$\nabla_\phi f(w, \phi) = \sum_i \int \nabla_\phi \, q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$\neq \sum_i \int q(t_i \mid x_i, \phi) \underbrace{\nabla_\phi \log p(x_i \mid t_i, w)}_{0} \, dt_i$$

Gradients

$$\nabla_\phi f(w, \phi) = \sum_i \int \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int \frac{q(t_i \mid x_i, \phi)}{q(t_i \mid x_i, \phi)} \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

Gradients

$$\nabla_\phi f(w, \phi) = \sum_i \int \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int \frac{q(t_i \mid x_i, \phi)}{q(t_i \mid x_i, \phi)} \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$\nabla \log g(\phi) = \frac{\nabla g(\phi)}{g(\phi)}$$

Gradients

$$\nabla_\phi f(w,\phi) = \sum_i \int \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int \frac{q(t_i \mid x_i, \phi)}{q(t_i \mid x_i, \phi)} \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int q(t_i \mid x_i, \phi) \nabla_\phi \log q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$\nabla \log g(\phi) = \frac{\nabla g(\phi)}{g(\phi)}$$

Gradients

$$\nabla_\phi f(w,\phi) = \sum_i \int \nabla_\phi \, q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int \frac{q(t_i \mid x_i, \phi)}{q(t_i \mid x_i, \phi)} \nabla_\phi q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int q(t_i \mid x_i, \phi) \nabla_\phi \log q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \mathbb{E}_{q(t_i \mid x_i, \phi)} \nabla_\phi \log q(t_i \mid x_i, \phi) \log p(x_i \mid t_i, w) dt_i$$

Log-derivative trick

Gradients

$$\nabla_\phi f(w,\phi) = \sum_i \int \nabla_\phi q(t_i \mid x_i,\phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int \frac{q(t_i \mid x_i,\phi)}{q(t_i \mid x_i,\phi)} \nabla_\phi q(t_i \mid x_i,\phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \int q(t_i \mid x_i,\phi) \nabla_\phi \log q(t_i \mid x_i,\phi) \log p(x_i \mid t_i, w) dt_i$$

$$= \sum_i \mathbb{E}_{q(t_i \mid x_i,\phi)} \nabla_\phi \log q(t_i \mid x_i,\phi) \log p(x_i \mid t_i, w) dt_i$$

Log-derivative trick

Like -1000000