Gradient updates for Rezende et al., 2014

Andrew Jones

January 17, 2020

Link to paper: `https://doi.org/10.3389/fncom.2014.00038`

Setup

- $u_i(t)$: membrane potential of neuron i at time t
- w_ij: synaptic weight from neuron i to neuron j
- $X_i(t)$: observed spike in neuron i at time t
- $\phi_i(t)$: evoked potential of neuron i at time t
- $\eta_i(t)$: adaptation potential of neuron i at time t
- $\rho_i(t)$: firing rate of neuron i at time t
- $\tau, \tau_{\text{adapt}}, \delta u, v$: neuron constants

Membrane potential update

This update is based on equation 1 in the paper.

$$
\begin{aligned}
u_i(t) &= \sum_j w_{ij}\phi_j(t) + \eta_i(t) \\
&= \sum_j w_{ij} \int_0^t \exp\left(-\frac{(t-s)}{\tau}\right) X_j(s)ds - \eta_0 \int_0^t \exp\left(-\frac{(t-s)}{\tau_{\text{adapt}}}\right) X_j(s)ds \\
&\approx \sum_j w_{ij} \sum_{k=0}^{t} \exp\left(-\frac{(t-k)}{\tau}\right) X_j(k)\Delta k - \eta_0 \sum_{k=0}^{t} \exp\left(-\frac{(t-k)}{\tau_{\text{adapt}}}\right) X_j(k)\Delta k \quad \text{(**discretize**)}
\end{aligned}
$$

M-network gradient updates

This update is based on equation 24 in the paper (and all associated equations).

We take the GLM's link function to be

$$g(u_i(t)) = \exp(u_i(t)).$$

The gradient of the log likelihood wrt w_{ij} is as follows:

$$
\begin{aligned}
\dot{w}_{ij}^{\mathcal{M}}(T) &\approx \int_0^T \frac{g'(u_i^{\mathcal{M}}(t))}{g(u_i^{\mathcal{M}}(t))} \left[X_i(t) - \rho_i^{\mathcal{M}}(t)\right] \phi_j(t)dt \\
&= \int_0^T \frac{\frac{\partial}{\partial u_i(t)} \left(\rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right)}{\rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]} \left[X_i(t) - \rho_i^{\mathcal{M}}(t)\right] \phi_j(t)dt \\
&= \int_0^T \frac{\rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]}{\rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]} \left[X_i(t) - \rho_i^{\mathcal{M}}(t)\right] \phi_j(t)dt \qquad (1) \\
&= \int_0^T \left[X_i(t) - \rho_i^{\mathcal{M}}(t)\right] \phi_j(t)dt \\
&= \int_0^T \left[X_i(t) - \rho_i^{\mathcal{M}}(t)\right] \phi_j(t)dt \\
&= \int_0^T \left[X_i(t) - \rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right] \int_0^t \exp\left(-\frac{(t-s)}{\tau}\right) X_j(s)dsdt \\
&\approx \sum_{t=0}^{T} \left[X_i(t) - \rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right] \sum_{s=0}^{t} \left[\exp\left(-\frac{(t-s)}{\tau}\right) X_j(s)\Delta s\right] \Delta t \quad \text{(**discretize**)} \qquad (2)
\end{aligned}
$$

In (1), we used the gradient of g with respect to $u_i(t)$:

$$
\frac{\partial}{\partial u_i(t)} \left(\rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right) = \rho_0 \exp\left[\frac{u_i(t)-v}{\Delta u}\right]
$$

Finally the update to w_{ij} is

$$
w_{ij}^{\mathcal{M}}(T) = w_{ij}^{\mathcal{M}}(T) + \mu^{\mathcal{M}} \dot{w}_{ij}^{\mathcal{M}}(T)
$$

Q-network gradient updates

We'd like to maximize the marginal likelihood of the visible neurons:

$$
p(X_V) = \int p(X_V, X_H) dX_H.
$$

The paper uses a variational approach to approximate $p(X_V)$. In particular, they posit a distribution q in order to approximate $p(X_H|X_V)$, and then minimize the divergence between p and q (i.e., fairly standard variational inference).

The KL-divergence is given by

$$
\begin{aligned}
KL(q,p) &= \int q(X_H|X_V)\log\frac{q(X_H|X_V)}{p(X_H|X_V)}dX_H \\
&= \int q(X_H|X_V)\log\frac{q(X_H|X_V)}{p(X_H,X_V)} + \log p(X_V)dX_H \\
&= \mathbb{E}_{q(X_H|X_V)}\left[\log q(X_H|X_V) - \log p(X_H,X_V)\right] + \log p(X_V)
\end{aligned}
$$

Using the fact that the KL divergence is always non-negative, we can obtain a lower bound on the negative log-likelihood:

$$
\begin{aligned}
&\mathbb{E}_{q(X_H|X_V)}\left[\log q(X_H|X_V) - \log p(X_H,X_V)\right] + \log p(X_V) \geq 0 \\
\implies &-\log p(X_V) \leq \mathbb{E}_{q(X_H|X_V)}\left[\log q(X_H|X_V) - \log p(X_H,X_V)\right]
\end{aligned}
$$

Thus, by minimizing this lower bound on the RHS (which is also known as the Helmholtz free energy), we can maximize the data log-likelihood.

This update is based on equation 25 in the paper (and all associated equations).

$$
\begin{aligned}
\dot{w}_{ij}^{\mathcal{Q}}(T) &= \hat{\mathcal{F}}\int_0^T dt\frac{g'(u_i^{\mathcal{Q}}(t))}{g(u_i^{\mathcal{Q}}(t))}\left[X_i(t) - \rho_i^{\mathcal{Q}}(t)\right]\phi_j(t) \\
&= \hat{\mathcal{F}}\int_0^T dt\left[X_i(t) - \rho_i^{\mathcal{Q}}(t)\right]\phi_j(t) \\
&= \hat{\mathcal{F}}\int_0^T \left[X_i(t) - \rho_0\exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right]\int_0^t \exp\left(-\frac{(t-s)}{\tau}\right)X_j(s)dsdt \\
&\approx \hat{\mathcal{F}}\sum_{t=0}^{T}\left[X_i(t) - \rho_0\exp\left[\frac{u_i(t)-v}{\Delta u}\right]\right]\Delta t\sum_{s=0}^{t}\exp\left(-\frac{(t-s)}{\tau}\right)X_j(s)\Delta s \quad \text{(**discretize**)}
\end{aligned}
$$

$\hat{\mathcal{F}}$ is defined as a point estimate of the free energy:

$$
\begin{aligned}
\hat{\mathcal{F}} &= \int_0^T \left[\sum_{i \in \mathcal{H}} \left[\log \rho_i^Q(t) X_i(t) - \rho_i^Q(t) \right] - \sum_{i \in \mathcal{V} \cup \mathcal{H}} \left[\log \rho_i^M(t) X_i(t) - \rho_i^M(t) \right] \right] dt \\
&= \int_0^T \left[\sum_{i \in \mathcal{H}} \left[\log \rho_i^Q(t) X_i(t) - \rho_i^Q(t) \right] - \sum_{i \in \mathcal{V} \cup \mathcal{H}} \left[\log \rho_i^M(t) X_i(t) - \rho_i^M(t) \right] \right] dt \\
&= \int_0^T \left[\sum_{i \in \mathcal{H}} \left[\log \left(\rho_0 \exp \left[\frac{u_i^{\mathcal{Q}}(t) - v}{\Delta u} \right] \right) X_i(t) - \rho_0 \exp \left[\frac{u_i^{\mathcal{Q}}(t) - v}{\Delta u} \right] \right] \right. \\
&\left. - \sum_{i \in \mathcal{V} \cup \mathcal{H}} \left[\log \left(\rho_0 \exp \left[\frac{u_i^{\mathcal{M}}(t) - v}{\Delta u} \right] \right) X_i(t) - \rho_0 \exp \left[\frac{u_i^{\mathcal{M}}(t) - v}{\Delta u} \right] \right] \right] dt && 1 \\
&\approx \sum_{t=0}^T \left[\sum_{i \in \mathcal{H}} \left[\log \left(\rho_0 \exp \left[\frac{u_i^{\mathcal{Q}}(t) - v}{\Delta u} \right] \right) X_i(t) - \rho_0 \exp \left[\frac{u_i^{\mathcal{Q}}(t) - v}{\Delta u} \right] \right] \right. \\
&\left. - \sum_{i \in \mathcal{V} \cup \mathcal{H}} \left[\log \left(\rho_0 \exp \left[\frac{u_i^{\mathcal{M}}(t) - v}{\Delta u} \right] \right) X_i(t) - \rho_0 \exp \left[\frac{u_i^{\mathcal{M}}(t) - v}{\Delta u} \right] \right] \right] \Delta t && (\text{**discretize**})
\end{aligned}
$$