Neurons centered on a subset of positions

Model

Rotational Invariance Assumption:

$$f_k(\theta) = f_0(\theta - c_k)$$

After discretization:

$$f_k(\theta_m) = f_0(\theta_m - c_k) := f_{m-c_k}$$

Assume the centers are a regular subset with equal distance:

$$\begin{aligned}
\text{Distance between centers:} \quad & \delta \geq 1 \\
\text{Number of centers:} \quad & N = \frac{M}{\delta} \\
\text{Position of centers:} \quad & c_k = \delta k, k = 1, ..., N
\end{aligned}$$

Therefore $r_k|\theta_m$ satiesfies Poisson distribution:

$$p(r_k|\theta_m) = \frac{f_k(\theta_m)^{r_k}}{r_k!} e^{-f_k(\theta_m)} = \frac{f_{m-c_k}{}^{r_k}}{r_k!} e^{-f_{m-c_k}}$$

Rotational Invariance and Mutual Information

From rotational invariance, we show that

$$D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right) = D_{KL}\left(p(\mathbf{r}|\theta_{m+\delta})\|p(\mathbf{r})\right) \tag{1}$$

Proof.

$$\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta_m)} = \frac{\frac{1}{M}\sum_{i=1}^{M}\prod_{k=1}^{N}\frac{f_{i-c_k}{}^{r_k}}{r_k!}e^{-f_{i-c_k}}}{\prod_{k=1}^{N}\frac{f_{m-c_k}{}^{r_k}}{r_k!}e^{-f_{m-c_k}}} = \frac{1}{M}\sum_{i=1}^{M}\prod_{k=1}^{N}\left(\frac{f_{i-c_k}}{f_{m-c_k}}\right)^{r_k} e^{-(f_{i-c_k}-f_{m-c_k})}$$

$$\begin{aligned}
D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right) &= -E_{\mathbf{r}|\theta_i}\left[\ln\left(\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta_m)}\right)\right] \\
&= -\sum_{\mathbf{r}}\prod_{k=1}^{N}\frac{f_{m-c_k}{}^{r_k}}{r_k!}e^{-f_{m-c_k}}\ln\left(\frac{1}{M}\sum_{i=1}^{M}\prod_{k=1}^{N}\left(\frac{f_{i-c_k}}{f_{m-c_k}}\right)^{r_k} e^{-(f_{i-c_k}-f_{m-c_k})}\right) \\
&= -\sum_{\mathbf{r}} C_1(\mathbf{r})\prod_{k=1}^{N} f_{m-c_k}{}^{r_k} e^{-f_{m-c_k}} \ln\left(\prod_{k=1}^{N} f_{m-c_k}^{-r_k} e^{f_{m-c_k}} \cdot C_2(\mathbf{r})\right)
\end{aligned}$$

Where $C_1(\mathbf{r}) = \frac{1}{\prod_{k=1}^{M} r_k!}$, $C_2(\mathbf{r}) = \frac{1}{M}\sum_{i=1}^{M}\prod_{k=1}^{N} f_{i-c_k}^{r_k} e^{-f_{i-c_k}}$ does not depend on k. Since $c_k = k\delta$, $m + \delta - c_k = m - c_{k-1}$,

$$\begin{aligned}
D_{KL}\left(p(\mathbf{r}|\theta_{m+\delta})\|p(\mathbf{r})\right) &= -\sum_{\mathbf{r}} C_1(\mathbf{r})\prod_{k=1}^{N} f_{m-c_{k-1}}{}^{r_k} e^{-f_{m-c_{k-1}}} \ln\left(\prod_{k=1}^{N} f_{m-c_{k-1}}^{-r_k} e^{f_{m-c_{k-1}}} \cdot C_2(\mathbf{r})\right) \\
&= -\sum_{\mathbf{r}} C_1(\mathbf{r})\prod_{k=1}^{N} f_{m-c_k}{}^{r_{k+1}} e^{-f_{m-c_k}} \ln\left(\prod_{k=1}^{N} f_{m-c_k}^{-r_{k+1}} e^{f_{m-c_k}} \cdot C_2(\mathbf{r})\right)
\end{aligned}$$

Taking the cyclic permutation of $\mathbf{r}$ by 1 such that $\tilde{r}_k = r_{k+1}$, i.e. $\tilde{\mathbf{r}} = (r_2, r_3, \ldots, r_{M+1} = r_1)$. Therefore, since C_1 and C_2 are invariant under cyclic permutations of $\mathbf{r}$, we have

$$
\begin{aligned}
D_{KL}\left(p(\mathbf{r}|\theta_{m+\delta})\|p(\mathbf{r})\right) &= \sum_{\mathbf{r}} C_1(\mathbf{r}) \prod_{k=1}^{N} {f_{m-c_k}}^{r_{k+1}} e^{-f_{m-c_k}} \ln\left(\prod_{k=1}^{N} f_{m-c_k}^{-r_{k+1}} e^{f_{m-c_k}} \cdot C_2(\mathbf{r})\right) \\
&= \sum_{\tilde{\mathbf{r}}} C_1(\tilde{\mathbf{r}}_{-1}) \prod_{k=1}^{N} {f_{m-c_k}}^{\tilde{r}_k} e^{-f_{m-c_k}} \ln\left(\prod_{k=1}^{N} f_{m-c_k}^{-\tilde{r}_k} e^{f_{m-c_k}} \cdot C_2(\tilde{\mathbf{r}}_{\mathbf{-1}})\right) \\
&= \sum_{\tilde{\mathbf{r}}} C_1(\tilde{\mathbf{r}}) \prod_{k=1}^{N} {f_{m-c_k}}^{\tilde{r}_k} e^{-f_{m-c_k}} \ln\left(\prod_{k=1}^{N} f_{m-c_k}^{-\tilde{r}_k} e^{f_{m-c_k}} \cdot C_2(\tilde{\mathbf{r}})\right) \\
&= D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right)
\end{aligned}
$$

$\square$

A special case is when $\delta = 1$, our original model: $D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right) = D_{KL}\left(p(\mathbf{r}|\theta_0)\|p(\mathbf{r})\right)$ for all m. Now we derive the mutual information:

$$
\begin{aligned}
I(\mathbf{r};\theta) &= D_{KL}(p(\mathbf{r},\theta)||p(\mathbf{r})p(\theta)) \\
&= \int_\theta \int_{\mathbf{r}} p(\mathbf{r},\theta) \ln\left(\frac{p(\mathbf{r},\theta)}{p(\mathbf{r})p(\theta)}\right) d\mathbf{r}d\theta \\
&= \int_\theta \int_{\mathbf{r}} p(\mathbf{r}|\theta)p(\theta) \ln\left(\frac{p(\mathbf{r}|\theta)}{p(\mathbf{r})}\right) d\mathbf{r}d\theta \\
&= \frac{1}{M}\sum_{s=1}^{M} \int_{\mathbf{r}} p(\mathbf{r}|\theta_s) \ln\left(\frac{p(\mathbf{r}|\theta_s)}{p(\mathbf{r})}\right) d\mathbf{r} \\
&= \frac{1}{M}\sum_{m=1}^{M} D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right) \\
&= \frac{1}{\delta}\sum_{m=1}^{\delta} D_{KL}\left(p(\mathbf{r}|\theta_m)\|p(\mathbf{r})\right) \qquad (2) \\
&= -\frac{1}{\delta}\sum_{m=1}^{\delta} E_{\mathbf{r}|\theta_m}\left[\ln\left(\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta_m)}\right)\right] \qquad (3)
\end{aligned}
$$

1st Order Derivatives

For simplicity, we introduce the following notations:

$$
\begin{aligned}
P_m(\mathbf{r}) &:= P(\mathbf{r}|\theta_m) = \prod_{k=1}^{N} \frac{{f_{m-c_k}}^{r_k}}{r_k!} e^{-f_{m-c_k}} \qquad (4) \\
S_m(\mathbf{r}) &:= \frac{p(\mathbf{r})}{p(\mathbf{r}|\theta_m)} = \frac{1}{M}\sum_{i=1}^{M}\prod_{k=1}^{N}\left(\frac{f_{i-c_k}}{f_{m-c_k}}\right)^{r_k} e^{-(f_{i-c_k}-f_{m-c_k})} \qquad (5) \\
Q_m^j(\mathbf{r}) &:= \prod_{k=1}^{N}\left(\frac{f_{j-c_k}}{f_{m-c_k}}\right)^{r_k} e^{-(f_{j-c_k}-f_{m-c_k})} \qquad (6)
\end{aligned}
$$

Therefore,

$$S_m(\mathbf{r}) = \frac{1}{M}\sum_{j=1}^{M} Q_m^j(\mathbf{r}) \tag{7}$$

$$\begin{aligned}
I(\mathbf{r};\theta) &= -\frac{1}{M}\sum_{m=1}^{M} E_{\mathbf{r}|\theta_m}\left[\ln\left(\frac{p(\mathbf{r})}{p(\mathbf{r}|\theta_m)}\right)\right] \\
&= -\frac{1}{M}\sum_{m=1}^{M}\sum_{\mathbf{r}} P_m(\mathbf{r})\left[\ln S_m(\mathbf{r})\right] \qquad (8) \\
\frac{\partial I(\mathbf{r};\theta)}{\partial f_i} &= -\frac{1}{M}\sum_{m=1}^{M}\sum_{\mathbf{r}}\left[\frac{\partial P_m(\mathbf{r})}{\partial f_i}\ln S_m(\mathbf{r}) + \frac{P_m(\mathbf{r})}{S_m(\mathbf{r})}\frac{\partial S_m(\mathbf{r})}{\partial f_i}\right] \qquad (9)
\end{aligned}$$

Denote by $i \sim m : (i-m) \mod \delta = 0$.

First, we compute the partial derivatives of $P_m(\mathbf{r}) = \prod_{k=1}^{N}\frac{f_{m-c_k}{}^{r_k}}{r_k!}e^{-f_{m-c_k}}$: notice that it only contains certain f_i's such that $i = m - c_l = m - \delta l$ for some integer $l = 1, ..., N$:

$$\frac{\partial P_m}{\partial f_i} = 1_{\{i\sim m\}}\left(\frac{r_{\hat{m}}}{f_i} - 1\right)P_m \tag{10}$$

where $\hat{m} := \frac{m-i}{\delta}$ such that $m - \delta\hat{m} = i$.

Next, we show that the second term is zero:

$$\begin{aligned}
\frac{\partial S_m(\mathbf{r})}{\partial f_i} &= \frac{\partial}{\partial f_i}\left[\frac{1}{M}\sum_{j=1}^{M} Q_m^j\right] \\
&= \frac{1}{M}\sum_{j=1}^{M}\frac{\partial Q_m^j}{\partial f_i} \\
&= \frac{1}{M}\sum_{j=1}^{M}\frac{\partial}{\partial f_i}\left[\frac{\prod_{k=1}^{N} f_{j-\delta k}^{r_k}e^{-f_{j-\delta k}}}{\prod_{k=1}^{N} f_{m-\delta k}^{r_k}e^{-f_{m-\delta k}}}\right] \\
&= \frac{1}{M}\sum_{j=1}^{M}\left[1_{\{i\sim j\}}\left(\frac{r_{\hat{j}}}{f_i} - 1\right)Q_m^j + 1_{\{i\sim m\}}\left(-\frac{r_{\hat{m}}}{f_i} + 1\right)Q_m^j\right] \\
&= \frac{1}{M}\sum_{j=1}^{M}\left[1_{\{i\sim j\}}\left(\frac{r_{\hat{j}}}{f_i} - 1\right) - 1_{\{i\sim m\}}\left(\frac{r_{\hat{m}}}{f_i} - 1\right)\right]Q_m^j
\end{aligned}$$

Thus

$$
\begin{aligned}
f_i \cdot \sum_{\mathbf{r}} \frac{1}{M} \sum_{m=1}^{M} \frac{P_m(\mathbf{r})}{S_m(\mathbf{r})} \frac{\partial S_m(\mathbf{r})}{\partial f_i} &= f_i \cdot \sum_{\mathbf{r}} \frac{1}{M} \sum_{m=1}^{M} \frac{P_m(\mathbf{r})}{S_m(\mathbf{r})} \frac{\partial S_m(\mathbf{r})}{\partial f_i} \\
&= \frac{1}{M^2} \sum_{\mathbf{r}} \sum_{m=1}^{M} \frac{P_m(\mathbf{r})}{S_m(\mathbf{r})} \sum_{j=1}^{M} \left[1_{\{i\sim j\}} \left(r_{\hat{j}} - f_i \right) - 1_{\{i\sim m\}} \left(r_{\hat{m}} - f_i \right) \right] \frac{P_j(\mathbf{r})}{P_m(\mathbf{r})} \\
&= \frac{1}{M^2} \sum_{\mathbf{r}} \sum_{m=1}^{M} \frac{P_j(\mathbf{r})}{S_m(\mathbf{r})} \sum_{j=1}^{M} \left[1_{\{i\sim j\}} \left(r_{\hat{j}} - f_i \right) - 1_{\{i\sim m\}} \left(r_{\hat{m}} - f_i \right) \right] \\
&= \frac{1}{M} \sum_{\mathbf{r}} \sum_{m=1}^{M} \frac{\prod_{k=1}^{N} \frac{{f_{j-\delta k}}^{r_k}}{r_k!} e^{-f_{j-\delta k}}}{\sum_{l=1}^{M} \prod_{k=1}^{N} \left(\frac{f_{l-\delta k}}{f_{m-\delta k}} \right)^{r_k} e^{-(f_{l-\delta k} - f_{m-\delta k})}} \sum_{j=1}^{M} \left[1_{\{i\sim j\}} \left(r_{\hat{j}} - f_i \right) - 1_{\{i\sim m\}} \left(r_{\hat{m}} - f_i \right) \right] \\
&= \frac{1}{M} \sum_{\mathbf{r}} \frac{\sum_{m=1}^{M} \sum_{j=1}^{M} \prod_{k=1}^{N} {f_{m-\delta k}}^{r_k} e^{-f_{m-\delta k}} \cdot \prod_{k=1}^{N} f_{j-\delta k}^{r_k} e^{-f_{j-\delta k}} \cdot 1_{\{i\sim j\}} \left(r_{\hat{j}} - f_i \right)}{\prod_{k=1}^{N} r_k! \sum_{l=1}^{M} \prod_{k=1}^{N} {f_{l-\delta k}}^{r_k} e^{-f_{l-\delta k}}} \\
&\quad - \frac{1}{M} \sum_{\mathbf{r}} \frac{\sum_{m=1}^{M} \sum_{j=1}^{M} \prod_{k=1}^{N} {f_{m-\delta k}}^{r_k} e^{-f_{m-\delta k}} \cdot \prod_{k=1}^{N} f_{j-\delta k}^{r_k} e^{-f_{j-\delta k}} \cdot 1_{\{i\sim m\}} \left(r_{\hat{m}} - f_i \right)}{\prod_{k=1}^{N} r_k! \sum_{l=1}^{M} \prod_{k=1}^{N} {f_{l-\delta k}}^{r_k} e^{-f_{l-\delta k}}} \\
&= 0
\end{aligned}
$$

Thus

$$
\begin{aligned}
\frac{\partial I(\mathbf{r};\theta)}{\partial f_i} &= -\frac{1}{M} \sum_{m=1}^{M} \sum_{\mathbf{r}} \left[\frac{\partial P_m(\mathbf{r})}{\partial f_i} \ln S_m(\mathbf{r}) + \frac{P_m(\mathbf{r})}{S_m(\mathbf{r})} \frac{\partial S_m(\mathbf{r})}{\partial f_i} \right] \\
&= -\frac{1}{\delta} \sum_{m=1}^{\delta} \sum_{\mathbf{r}} \frac{\partial P_m(\mathbf{r})}{\partial f_i} \ln S_m(\mathbf{r}) (?) \\
&= -\frac{1}{\delta} \sum_{m=1}^{\delta} \sum_{\mathbf{r}} 1_{\{i\sim m\}} \left(\frac{r_{\hat{m}}}{f_i} - 1 \right) P_m(\mathbf{r}) \ln S_m(\mathbf{r}) \\
&= \frac{1}{\delta} \sum_{m=1}^{\delta} E \left[1_{\{i\sim m\}} \left(1 - \frac{r_{\hat{m}}}{f_i} \right) \ln S_m(\mathbf{r}) \right]
\end{aligned}
$$

Notes from code

δ: stepPop, c_k : center[k], M: numBin, N: numPop.

$$
\begin{aligned}
lrate_m(k,j) &:= \ln\left(\frac{f_{j-c_k}}{f_{m-c_k}}\right) \\
dexp_m(k,j) &:= \sum_{k=1}^{N} f_{m-c_k} - f_{j-c_k} \\
mexp_m(i) &:= \sum_k r_k lrate_m(k,i) + dexp_m(i) \\
Max_m &:= \max_i mexp_m(i) \\
\prod_{k=1}^{N}\left(\frac{f_{i-c_k}}{f_{m-c_k}}\right)^{r_k} e^{-(f_{i-c_k}-f_{m-c_k})} = Q_m(i) &= \exp mexp_m(i) \\
S_m(\mathbf{r}) &= \frac{1}{M}\sum_{i=1}^{M}\exp\left[mexp(i)\right] \\
&= \frac{1}{M}\sum_{i=1}^{M}\exp\left[mexp(i) - Max\right]\cdot e^{Max} \\
mean_m &:= \ln\left(S_m(\mathbf{r})\right) \text{ where } \mathbf{r} \sim P(\mathbf{r}|\theta_m) \\
lave_m &:= \widehat{mean}_m = \frac{\sum_{l=1}^{Niter} mean_m(l)}{Niter} \\
MIcond(m) &:= -lave_m \\
MIvalue &:= -\sum_{m=1}^{\delta} lave(m) \\
I(\mathbf{r};\theta) &\approx \frac{1}{\delta} MIvalue/\ln(2) \\
tmpgrad_m(i) &:= -E_{\mathbf{r}|\theta_m}\left[\frac{1}{S_m}\sum_{w=1}^{N}\left(\frac{r_w}{f_i}-1\right)Q_m(i+\delta w)\right] \\
MIgradcond(m,i) &:= -\partial_i L_m(?) \\
MIgrad(i) &:= \partial_i I
\end{aligned}
$$