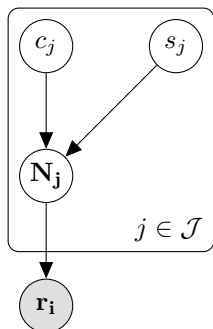


Divisive Normalization Notes

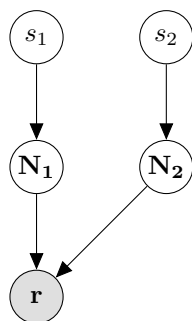


$$s_1 \sim \text{Unif}(-60, 60)$$

$$s_2 \sim \text{Unif}(s_1, 60)$$

$N_{1,2}$ is the response of neurons to grating 1, 2
(Poisson neurons with Gaussian tuning curves)

r is response of neurons ($\sum N$)



$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$

$$\mathbf{r}_{\text{hid}} = \phi(\mathbf{W}\mathbf{r} + \mathbf{b})$$

$$\mathbf{r} = \text{Poisson}(g_1 \mathbf{f}(s_1) + g_2 \mathbf{f}(s_2))$$

$$\hat{s}_1$$

$$\hat{s}_2$$

$f_i(s)$ is tuning curve of neuron i

Tuning curve assumptions:

Tuning curves cover the space so $\sum_i f_i(s_j)$ is independent of s_j

The tuning curves are von Mises: $f_i(s_j) = e^{\kappa(\cos(s_j - s_i^{\text{pref}}) - 1)}$

N_{ij} is the response of neuron i to grating j

$P(N|c, s) = \prod_i \text{Poisson}(N_{ij}; f_i(s_j)c_j)$

$P(c) = \text{Gamma}(c_j; \alpha, \beta)$

$P(s) = \text{Unif}(0, 2\pi)$

$$\begin{aligned} P(\mathbf{c}, \mathbf{s}|\mathbf{r}) &= \sum_{\mathbf{N}} P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) = \sum_{\mathbf{N}} P(\mathbf{r}|\mathbf{N})P(\mathbf{N}|\mathbf{c}, \mathbf{s})P(\mathbf{c})P(\mathbf{s}) \\ &\propto \prod_i (\delta(r_i - \sum_j N_{ij}) \prod_j \frac{(f_i(s_j)c_j)^{N_{ij}} e^{-f_i(s_j)c_j}}{N_{ij}!}) (\prod_j \frac{\beta^\alpha}{\Gamma(\alpha)} c_j^{\alpha-1} e^{-\beta c_j}) \end{aligned} \quad (1)$$

For the variational inference, we need

$$\begin{aligned} \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) &= \sum_i [\log(\delta(r_i - \sum_j N_{ij})) - \sum_j [-\log N_{ij}! + N_{ij} \log(f_i(s_j)c_j) - f_i(s_j)c_j]] \\ &\quad + \sum_j [(\alpha - 1) \log c_j - \beta c_j - \log(\beta^{-\alpha} \Gamma(\alpha))] \end{aligned} \quad (2)$$

Then we approximate this using a factorized distribution:

$$Q(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) = Q(\mathbf{N}|\mathbf{r})Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r}) \quad (3)$$

Then

$$\begin{aligned} \log Q(\mathbf{N}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\ &= \sum_i [\log(\delta(r_i - \sum_j N_{ij})) - \sum_j [N_{ij} \langle \log(f_i(s_j))c_j \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} - \log N_{ij}!]] \end{aligned} \quad (4)$$

$$\begin{aligned}
\log Q(\mathbf{c}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\
&= \sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \log c_j - \langle f_i(s_j) \rangle_{Q(\mathbf{s}|\mathbf{r})} c_j] \\
&\quad + \sum_j [(\alpha - 1) \log c_j - \beta c_j - \log(\beta^{-\alpha} \Gamma(\alpha))] \\
&= \sum_{ij} [(\alpha + \langle N_{ij} \rangle_{Q(\mathbf{c}|\mathbf{r})} - 1) \log c_j - (\beta + \langle f_i(s_j) \rangle_{Q(\mathbf{s}|\mathbf{r})}) c_j] \\
&\quad - \sum_j \log(\beta^{-\alpha} \Gamma(\alpha))
\end{aligned} \tag{5}$$

$$\begin{aligned}
\log Q(\mathbf{s}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{c}|\mathbf{r})} \\
&= \sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \log(f_i(s_j)) - f_i(s_j) \langle c_j \rangle_{Q(\mathbf{c}|\mathbf{r})}] = \sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \log(f_i(s_j))]
\end{aligned} \tag{6}$$

Last equality is because of space covering assumption

So $Q(\mathbf{N}|\mathbf{r})$ is a multinomial, $Q(\mathbf{c}|\mathbf{r})$ is a Gamma and $Q(\mathbf{s}|\mathbf{r})$ is von Mises

$$Q(\mathbf{N}|\mathbf{r}) = \prod_i \delta(r_i - \sum_j N_{ij}) r_i! \prod_j \frac{\langle (f_i(s_j) c_j) \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})}^{N_{ij}}}{N_{ij}!} \tag{7}$$

$$Q(\mathbf{c}|\mathbf{r}) = \prod_j \text{Gamma}(c_j | \alpha + \langle N_{ij} \rangle_{Q(\mathbf{c}|\mathbf{r})}, \beta + \langle f_i(s_j) \rangle_{Q(\mathbf{s}|\mathbf{r})}) \tag{8}$$

$$\begin{aligned}
Q(\mathbf{s}|\mathbf{r}) &\propto \prod_j e^{\kappa(\sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \cos(s_j - s_i^{\text{pref}}])} \\
&= \prod_j e^{\kappa(\sum_j \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} (\cos s_j \cos s_i^{\text{pref}} + \sin s_j \sin s_i^{\text{pref}}))} \\
&= \prod_j e^{\kappa(\sum_j (\cos s_j \sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \cos s_i^{\text{pref}} + \sin s_j \sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \sin s_i^{\text{pref}}))} \\
&= \prod_j e^{\kappa(\cos s_j \cos(\sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} s_i^{\text{pref}}) + \sin s_j \sin(\sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} s_i^{\text{pref}}))} \\
&= \prod_j e^{\tilde{\kappa} \cos(s_j - \hat{s})}
\end{aligned} \tag{9}$$

$$\text{where } \hat{s} = \sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} s_i^{\text{pref}}$$

Calculations with: $P(r|c, s) = \prod_i \text{Poisson}(r_i; \phi(\sum_j f_i(s_j)c_j))$

$$P(\mathbf{c}, \mathbf{s}|\mathbf{r}) = P(\mathbf{r}|\mathbf{c}, \mathbf{s})P(\mathbf{c})P(\mathbf{s})$$

$$\propto \prod_i \frac{(\phi(\sum_j (f_i(s_j)c_j)))^{r_i} e^{-\phi(\sum_j (f_i(s_j)c_j))}}{r_i!} \left(\prod_j \frac{\beta^\alpha}{\Gamma(\alpha)} c_j^{\alpha-1} e^{\beta c_j} \right) \quad (10)$$

$$\log P(\mathbf{c}, \mathbf{s}|\mathbf{r}) = \sum_i [-\log r_i! + r_i \log(\phi(\sum_j (f_i(s_j)c_j))) - (\phi(\sum_j (f_i(s_j)c_j)))]$$

$$+ \sum_j [(\alpha - 1) \log c_j - \beta c_j - \log(\beta^{-\alpha} \Gamma(\alpha))]$$

(11)

$$Q(\mathbf{c}, \mathbf{s}|\mathbf{r}) = Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r}) \quad (12)$$

$$\log Q(\mathbf{c}|\mathbf{r}) = \langle \log P(\mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{s}|\mathbf{r})}$$

$$= \sum_i [r_i \langle \log(\phi(\sum_j (f_i(s_j)c_j))) \rangle_{Q(\mathbf{s}|\mathbf{r})} - \langle \phi(\sum_j (f_i(s_j)c_j)) \rangle_{Q(\mathbf{s}|\mathbf{r})}]$$

$$+ \sum_j [(\alpha - 1) \log c_j - \beta c_j - \log(\beta^{-\alpha} \Gamma(\alpha))]$$

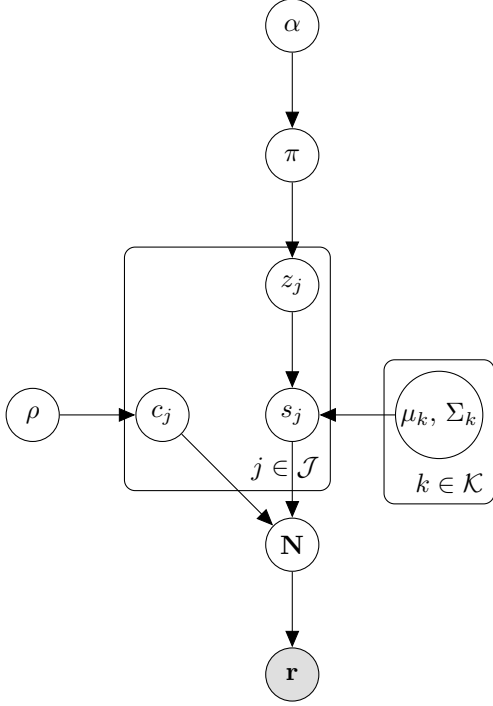
(13)

$$\log Q(\mathbf{s}|\mathbf{r}) = \langle \log P(\mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{c}|\mathbf{r})}$$

$$= \sum_i [r_i \langle \log(\phi(\sum_j (f_i(s_j)c_j))) \rangle_{Q(\mathbf{c}|\mathbf{r})} - \langle \log(\phi(\sum_j (f_i(s_j)c_j))) c_j \rangle_{Q(\mathbf{c}|\mathbf{r})}]$$

(14)

For later: consider with relative contrast (Dirichlet?)



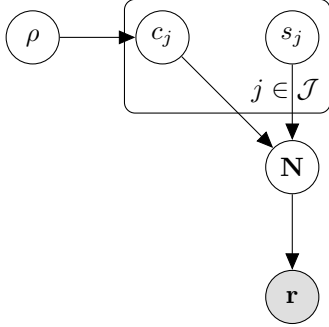
$$\begin{aligned}
 s_j | z_j &\sim \mathcal{N}(\mu_{z_j}, \Sigma_{z_j}) \\
 N_{ij} | \{s_j\} &\sim \text{Poisson}(c_j f_i(s_j)) \\
 r_i | N_{ij} &\sim \delta(r_i - z_j N_{ij}) \\
 \pi &\sim \text{Dir}(\alpha) \\
 c &\sim \text{Dir}(\rho)
 \end{aligned}$$

$$\begin{aligned}
 \log Q(z) &= \langle \log P(z, s, c, \mu, \Sigma, N, r) \rangle_{Q(s)Q(N)Q(c)Q(\mu, \Sigma)} \\
 &= \langle \log(\text{Mult}(z; \pi) \text{Dir}(\pi; \alpha) \prod_{j=1}^J \mathcal{N}(s_j; \mu_{z_j}, \Sigma_{z_j}) \text{Dir}(c, \rho) \dots \\
 &\quad \prod_{j=1}^J ((\delta(r_i - z_j N_{ij}) \text{Poisson}(c_j f_i(s_j))) \dots \\
 &\quad \prod_{k=1}^K \mathcal{N}(\mu_k; m_k, (\beta_0 \Sigma_k)) \text{Wi}(\Sigma_k^{-1}; \mathbf{L}_0, \nu_0)) \rangle_{Q(s)Q(N)Q(c)Q(\mu, \Sigma)}
 \end{aligned} \tag{15}$$

$$\begin{aligned}
\log Q(s) &= \langle \log \mathcal{N}(\mu_{z_j}, \Sigma_{z_j}) \rangle_{Q(z)Q(\mu, \Sigma)} + \sum_{i=1}^I \langle N_{ij} \log f_i(s_j) \rangle_{Q(N)} \\
&= -(\langle (s_j - \mu_j)^T \Sigma_{z_j}^{-1} (s_j - \mu_j) \rangle_{Q(z)Q(\mu, \Sigma)} + \sum_{i=1}^I \langle N_{ij} \rangle_{Q(N)} \log f_i(s_j)) \\
&= -(\sum_{k=1}^K Q(z_j = k) \langle (s_j - \mu_j)^T \Sigma_{z_j}^{-1} (s_j - \mu_j) \rangle_{Q(\mu, \Sigma)} - \langle N_{ij} \rangle_{Q(N)} \frac{\| (s_j - s_i^{\text{pref}}) \|^2}{L})
\end{aligned} \tag{16}$$

$$\begin{aligned}
\log Q(c) &= \langle \sum_j (\rho_j - 1) \log c_j \rangle + \sum_{ij} \langle N_{ij} \log c_j \rangle \\
&= \sum_j [\log c_j] [\rho_j - 1 + \sum_i \langle N_{ij} \rangle]
\end{aligned} \tag{17}$$

Redoing above with this graphical model:



s_j is stimulus j

c_j is contrast of grating j

r_i is response of neuron i

$f_i(s)$ is tuning curve of neuron i

Tuning curve assumptions:

Tuning curves cover the space so $\sum_i f_i(s_j)$ is independent of s_j

The tuning curves are von Mises: $f_i(s_j) = e^{\kappa(\cos(s_j - s_i^{\text{pref}}) - 1)}$

N_{ij} is the response of neuron i to grating j

$N_{ij} | \{s_j\} \sim \text{Poisson}(c_j f_i(s_j))$

$r_i | N_{ij} \sim \delta(r_i - z_j N_{ij})$

$$\begin{aligned}
\pi &\sim \text{Dir}(\alpha) \\
c &\sim \text{Dir}(\rho) \\
s &\sim \text{Unif}(0, \pi)
\end{aligned}$$

$$\begin{aligned}
P(\mathbf{c}, \mathbf{s}|\mathbf{r}) &= \sum_N P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) = \sum_N P(\mathbf{r}|\mathbf{N})P(\mathbf{N}|\mathbf{c}, \mathbf{s})P(\mathbf{c})P(\mathbf{s}) \\
&\propto \prod_i (\delta(r_i - \sum_j N_{ij}) \prod_j \frac{(f_i(s_j)c_j)^{N_{ij}} e^{-f_i(s_j)c_j}}{N_{ij}!}) (\prod_j \frac{1}{\beta(\rho)} c_j^{\rho-1}) \quad (18)
\end{aligned}$$

For the variational inference, we need

$$\begin{aligned}
\log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) &= \sum_i [\log(\delta(r_i - \sum_j N_{ij}))] - \sum_j [-\log N_{ij}! + N_{ij} \log(f_i(s_j)c_j)] \\
&\quad + \sum_j [(\rho_j - 1) \log c_j] \quad (19)
\end{aligned}$$

Then we approximate this using a factorized distribution:

$$Q(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) = Q(\mathbf{N}|\mathbf{r})Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r}) \quad (20)$$

$$\begin{aligned}
\log Q(\mathbf{N}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\
&= \sum_i [\log(\delta(r_i - \sum_j N_{ij}))] - \sum_j [N_{ij} \langle \log(f_i(s_j))c_j \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} - \log N_{ij}!] \quad (21)
\end{aligned}$$

$$\begin{aligned}
\log Q(\mathbf{c}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\
&= \sum_j [\log c_j (\rho_j - 1 + \sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})})] \quad (22)
\end{aligned}$$

$$\begin{aligned}
\log Q(\mathbf{s}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) \rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{c}|\mathbf{r})} \\
&= \sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \log(f_i(s_j)) - f_i(s_j) \langle c_j \rangle_{Q(\mathbf{c}|\mathbf{r})}] = \sum_{ij} [\langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})} \log(f_i(s_j))]
\end{aligned} \tag{23}$$

So $Q(\mathbf{N}|\mathbf{r})$ is multinomial, $Q(\mathbf{c}|\mathbf{r})$ is Dirichlet and $Q(\mathbf{s}|\mathbf{r})$ is von Mises

$$Q(\mathbf{N}|\mathbf{r}) = \prod_i \delta(r_i - \sum_j N_{ij}) r_i! \prod_j \frac{\langle (f_i(s_j) c_j) \rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})}^{N_{ij}}}{N_{ij}!} \tag{24}$$

$$Q(\mathbf{c}|\mathbf{r}) = \prod_j \text{Dir}(c_j; \rho_j - 1 + \sum_i \langle N_{ij} \rangle_{Q(\mathbf{N}|\mathbf{r})}) \tag{25}$$