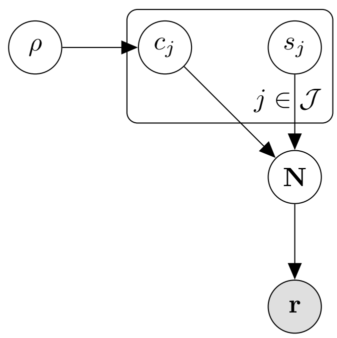
Divisive Normalization Notes

We consider the problem of demixing stimuli that are mixed due to neurons having large receptive fields or wide tuning curves.

0.1 Generative model



s_j is the j^{th} stimulus, $\mathbf{s} \sim \text{Unif}(0, \pi)$
c_j is the contrast of the j^{th} stimulus, $\mathbf{c} \sim \text{Dir}(\rho)$

$f_i(s)$ is the tuning curve of the i^{th} neuron in response to a single stimulus s
Tuning curve assumptions:

- Tuning curves cover the space so $\sum_i f_i(s)$ is independent of s
- The tuning curves are Gaussian: $f_i(s_j) \sim \mathcal{N}(s_i^{\text{pref}}, \sigma_{\text{tc}}^2)$

N_{ij} is the spike count of the i^{th} neuron elicited by the j^{th} stimulus
$N_{ij}|s_j \sim \text{Poisson}(c_j f_i(s_j))$
r_i is the total spike count of the i^{th} neuron in response to all stimuli
$r_i = \sum_j N_{ij}$

0.2 Inference

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

$$\begin{aligned} P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) &= P(\mathbf{r}|\mathbf{N})P(\mathbf{N}|\mathbf{c}, \mathbf{s})P(\mathbf{c})P(\mathbf{s}) \\ &\propto \prod_i \left[\delta\Big(r_i - \sum_j N_{ij}\Big)\Big(\prod_j \frac{(f_i(s_j)c_j)^{N_{ij}}}{N_{ij}!}\Big)\right] \prod_j \frac{1}{\beta(\rho)} c_j^{\rho_j - 1} \end{aligned} \tag{1}$$

For the variational inference, we need

$$\log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r}) = \sum_i \Big[\log \delta\Big(r_i - \sum_j N_{ij}\Big) + \sum_j \big[-\log N_{ij}! + N_{ij}\log(f_i(s_j)c_j)\big]\Big] \\ + \sum_j (\rho_j - 1)\log c_j \tag{2}$$

0.3 Variational approximation

Then we approximate this using a factorized distribution:

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

$$\begin{aligned}
\log Q(\mathbf{N}|\mathbf{r}) &= \sum_i \log Q(\mathbf{N}_i|\mathbf{r}_i) \\
\log Q(\mathbf{N}_i|\mathbf{r}_i) &= \langle \log P(\mathbf{N_i}, \mathbf{c}, \mathbf{s}|\mathbf{r_i})\rangle_{Q(\mathbf{c}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\
&= \log(\delta\Big(r_i - \sum_j N_{ij}\Big) + \sum_j N_{ij}\big(\langle \log c_j\rangle + \langle \log f_i(s_j)\rangle\big) - \sum_j \log N_{ij}! \\
\log Q(\mathbf{c}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r})\rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{s}|\mathbf{r})} \\
&= \sum_j [\log(c_j\rho_j - 1) + \sum_i \langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}] \\
\log Q(\mathbf{s}|\mathbf{r}) &= \langle \log P(\mathbf{N}, \mathbf{c}, \mathbf{s}|\mathbf{r})\rangle_{Q(\mathbf{N}|\mathbf{r})Q(\mathbf{c}|\mathbf{r})} \\
&= \sum_{ij} [\langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}\log(f_i(s_j)) - f_i(s_j)\langle c_j\rangle_{Q(\mathbf{c}|\mathbf{r})}] \\
&= \sum_{ij} \langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}\log(f_i(s_j))
\end{aligned} \tag{4}$$

So we now know

$$\begin{aligned}
Q(\mathbf{c}|\mathbf{r}) &\sim \mathrm{Dir}(\boldsymbol{\alpha}) \\
\alpha_j &= \rho_j - 1 + \sum_i \langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}
\end{aligned} \tag{5}$$

and

$$
\begin{aligned}
Q(s_j|r) &\propto \frac{1}{2\sigma_{\text{tc}}^2}\sum_{i=1}^{I}\langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}(s_j - s_i^{\text{pref}})^2 \\
&\sim \mathcal{N}(\mu_j, \tau_j) \\
\tau_j &= \sum_{i=1}^{I}\frac{\langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}}{2\sigma_{\text{tc}}^2} \\
\mu_j &= \frac{1}{\tau_j}\sum_{i=1}^{I} s_i^{\text{pref}}\frac{\langle N_{ij}\rangle_{Q(\mathbf{N}|\mathbf{r})}}{2\sigma_{\text{tc}}^2}
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
Q(N_{ij}|r) &\sim \text{Mult}(r_i, p_{ij}) \\
p_{ij} &= \frac{e^{\langle\log c_j\rangle + \langle\log f_i(s_j)\rangle}}{\sum_{j=1}^{J} e^{\langle\log c_j\rangle + \langle\log f_i(s_j)\rangle}}
\end{aligned} \tag{7}
$$

From this we can compute:

$$
\langle\log c_j\rangle_{Q(c_j)} = \Psi(\alpha_j) - \Psi(\sum_{j=1}^{J}\alpha_j) \tag{8}
$$

and

$$
\begin{aligned}
\langle\log f_i(s_j)\rangle_{Q(s_j)} &= \frac{-\langle(s_j - s_i^{\text{pref}})^2\rangle_{Q(s_j)}}{2\sigma_{\text{tc}}^2} \\
&= \frac{-[(s_i^{\text{pref}} - \mu_j)^2 + \tau_j^{-1}]}{2\sigma_{\text{tc}}^2}
\end{aligned} \tag{9}
$$

Therefore:

$$
p_{ij} = \frac{e^{\Psi(\alpha_j) - \Psi(\sum_{j=1}^{J}\alpha_j) - \frac{(s_i^{\text{pref}} - \mu_j)^2 + \tau_j^{-1}}{2\sigma_{\text{tc}}^2}}}{\sum_{j=1}^{J} e^{\Psi(\alpha_j) - \Psi(\sum_{j=1}^{J}\alpha_j) - \frac{(s_i^{\text{pref}} - \mu_j)^2 + \tau_j^{-1}}{2\sigma_{\text{tc}}^2}}} \tag{10}
$$

$$
\alpha_j = \rho_j - 1 + \sum_i r_i p_{ij} \tag{11}
$$

$$
\tau_j = \sum_{i=1}^{I}\frac{r_i p_{ij}}{2\sigma_{\text{tc}}^2} \tag{12}
$$

$$
\mu_j = \frac{1}{\tau_j}\sum_{i=1}^{I} s_i^{\text{pref}}\frac{r_i p_{ij}}{2\sigma_{\text{tc}}^2} \tag{13}
$$

We actually want the natural parameters so

$$\eta_j = \mu_j \tau_j = \sum_{i=1}^{I} s_i^{\mathrm{pref}} \frac{r_i p_{ij}}{2\sigma_{\mathrm{tc}}^2} \tag{14}$$

Also we don't want complex division in the neural computations so:

$$\begin{aligned} F_i(\mu_j, \alpha_j, \tau_j) &= e^{\Psi(\alpha_j) - \Psi(\sum_{j=1}^{J} \alpha_j) - \frac{(s_i^{\mathrm{pref}} - \frac{\eta_j}{\tau_j})^2 + \tau_j^{-1}}{2\sigma_{\mathrm{tc}}^2}} \\ p_{ij} &= F_i(\mu_j, \alpha_j, \tau_j)\pi_i \\ \pi_i &= \frac{1}{\sum_{j=1}^{J} F_i(\mu_j, \alpha_j, \tau_j)} \end{aligned} \tag{15}$$

And now we have update equations:

$$\begin{aligned} \frac{d\eta_j}{dt} &= -\eta_j + \frac{1}{2\sigma_{\mathrm{tc}}^2} \sum_{i=1}^{I} s_i^{\mathrm{pref}} r_i p_{ij} \\ \frac{d\tau_j}{dt} &= -\tau_j + \frac{1}{2\sigma_{\mathrm{tc}}^2} \sum_{i=1}^{I} r_i p_{ij} \\ \frac{dp_{ij}}{dt} &= -p_{ij} + F_i(\mu_j, \alpha_j, \tau_j)\pi_i \\ \frac{d\pi_i}{dt} &= 1 - \pi_i \sum_{j=1}^{J} F_i(\mu_j, \alpha_j, \tau_j) \\ \frac{d\alpha_j}{dt} &= -\alpha_j + \rho_j - 1 + \sum_{i}^{I} r_i p_{ij} \end{aligned} \tag{16}$$