INVESTIGATIONS IN RELATIVE ENTROPY

1. Showing that one of the terms in $\partial_i I$ is zero

We are interested in a relative entropy I, which depends on the values f_i of the tuning curve at various points. We have that

$$I = E_{\mathbf{r}|\mathbf{s}}(\ln(S)),$$

where

$$S(\mathbf{r}) = \sum_{k=0}^{M-1} Q_s^k(\mathbf{r})$$

and

$$Q_s^k(\mathbf{r}) = \prod_{j=0}^{M-1} \left(\frac{f_{j-k}(s)}{f_j(s)}\right)^{r_j} = \prod_{j=0}^{M-1} f_j(s)^{r_{j+k}-r_j}.$$

From now on we suppress the value of s (might as well be $s = 0$, by rotational symmetry) and write $f_i := f_i(s)$. The expectation is taken relative to the Poisson probability distribution

$$P(\mathbf{r}) = \prod_i \frac{f_i^{r_i}}{r_i!} e^{-f_i}.$$

That is, our relative entropy is

$$I = \sum_{\mathbf{r}} P(\mathbf{r}) \ln(S(\mathbf{r})).$$

We then have

$$\partial_i I = \sum_{\mathbf{r}} P(\mathbf{r}) \frac{\partial_i S}{S} + \sum_{\mathbf{r}} \partial_i P(\mathbf{r}) \ln(S).$$

In this section we show that the first term (or rather, f_i times the first term) is zero. We compute:

$$f_i \partial_i S = \sum_{k=0}^{M-1} (r_{i+k} - r_i) \prod_j f_j^{r_{j+k}-r_j},$$

so

$$\frac{f_i \partial_i S}{S} = \frac{\sum_{k=0}^{M-1} (r_{i+k} - r_i) \prod_j f_j^{r_{j+k}}}{\sum_{k=0}^{M-1} \prod_j f_j^{r_{j+k}}},$$

where we have multiplied the numerator and denominator by $\prod_j f_j^{r_j}$. The extra factor of f_i is just a constant, and doesn't affect whether the total is zero or not. Now we take an

Date: February 9, 2018.

expectation. The probability of any particular value of $\mathbf{r}$ is proportional to $\dfrac{\prod_j f_j^{r_j}}{\prod_j r_j!}$, so our expectation is proportional to

$$\sum_{\mathbf{r}} \frac{\sum_k (f_{i+k} - f_i) \prod_j f_j^{r_j + r_{j+k}}}{\sum_k \prod_j f_j^{r_j+k} r_j!}.$$

The denominator is invariant under cyclic permutations of the entries of $\mathbf{r}$, so we can average the numerator over those cyclic permutations, getting

$$\frac{1}{M} \sum_{k=0}^{M-1} \sum_{\ell=0}^{M-1} (r_{i+\ell+k} - r_{i+\ell}) \prod_j f_j^{r_{j+\ell} + r_{j+\ell+k}}.$$

Every term in this expansion appears exactly twice, with opposite signs, so the sum is zero. For instance, the term that contains a factor of $f_i^{r_i + r_{m+i}}$ comes once from $\ell = 0$ and $k = m$, with coefficient $r_{i+m} - r_i$, and once from $\ell = m$ and $k = -m$, with coefficient $r_i - r_{i+m}$.

2. Computing the Hessian

We have shown that

$$\partial_i I = \sum (\partial_i P) \ln(S).$$

This implies that

$$\partial_{ij}^2 I = \sum (\partial_{ij}^2 P) \ln(S) + \sum (\partial_i P) \frac{\partial_j S}{S}.$$

We will investigate these terms one at a time.

The Poisson law has the convenient property that $\partial_i P(\mathbf{r}) = \left(\dfrac{r_i}{f_i} - 1 \right) P(\mathbf{r}) = P(\mathbf{r} - e_i) - P(\mathbf{r})$, where e_i is a unit vector in the ith direction. Similarly,

$$\partial_{ij}^2 P(\mathbf{r}) = P(\mathbf{r} - e_i - e_j) - P(\mathbf{r} - e_i) - P(\mathbf{r} - e_j) + P(\mathbf{r}),$$

so

$$\begin{aligned} \sum_{\mathbf{r}} \partial_{ij}^2 P(\mathbf{r}) \ln(S) &= \sum_{\mathbf{r}} [P(\mathbf{r} - e_i - e_j) - P(\mathbf{r} - e_i) - P(\mathbf{r} - e_j) + P(\mathbf{r})] \ln(S(\mathbf{r})) \\ &= \sum_{\mathbf{r}} P(\mathbf{r}) [\ln(S(\mathbf{r})) - \ln(S(\mathbf{r} + e_i)) - \ln(S(\mathbf{r} + e_j)) + \ln(S(\mathbf{r} + e_i + e_j))]. \end{aligned}$$

This term can be evaluated numerically by Monte Carlo.

Unfortunately, the second term is not zero. Instead, we have

$$\begin{aligned} f_i f_j (\partial_i P)(\partial_j S)/S &= \frac{P(\mathbf{r}))(r_i - f_i) \sum_k (r_{j+k} - r_j) \prod_\ell f_\ell^{r_{\ell+k}}}{\sum_k \prod_\ell f_\ell^{r_{\ell+k}}} \\ &= \frac{\sum_k r_i (r_{j+k} - r_j) \prod_\ell f_\ell^{r_{\ell+k} + r_\ell}}{\sum_k \prod_\ell f_\ell^{r_{k+\ell}} e^{f_\ell} r_\ell!}, \end{aligned}$$

where we used the results of the previous section to eliminate $f_i(r_j - r_{j+k})$. Unfortunately, the cyclic permutation trick of the previous section does not cancel $r_i(r_{j+k} - r_j)$ against

$r_i(r_j - r_{j+k})$. Rather, we wind up averaging $r_i(r_{j+k} - r_j)$ and $r_{i+k}(r_j - r_{j+k})$ to get an expression that is manifestly symmetric in i and j, namely

$$\sum_{\mathbf{r}} 2 f_i f_j (\partial_i P)(\partial_j S)/S = \sum_{\mathbf{r}} \frac{-\sum_k (r_i - r_{i+k})(r_j - r_{j+k}) \prod_\ell f_\ell^{r_{\ell+k}+r_\ell}}{\sum_k \prod_\ell f_\ell^{r_{k+\ell}} e^{f_\ell} r_\ell!}.$$

This can also be rewritten as the expectation of

$$\frac{-\sum_k (r_i - r_{i+k})(r_j - r_{j+k}) \prod_\ell f_\ell^{r_{\ell+k}}}{\sum_k \prod_\ell f_\ell^{r_{k+\ell}}}.$$

I don't see any way to simplify this. However, it doesn't look like it should be too hard to sample numerically.

I'll play around with some simple cases ($M = 2$ or 3) to see if I'm overlooking any simplification, but I doubt it.