Lab notebook: Smoothing orientation maps

Carl Smith

April 27, 2012 - May 1, 2012

Notes

Tuesday, May 1, 2012

Talked to Liam today. He suggests that I do a quick experiment with synthetic data to show some results on the 2D orientation map at the uptown talk next Monday. In particular I am to 1) generate smooth 2D fields of orientations from the low-rank model using Gibbs sampling on the prior, 2) generate von Mises observations at each location and 3) estimate the orientations given the observations by block-wise Gibbs sampling. I should produce some figure like a colormap or a vector field depicting the orientation over the region.

 Another matter is non-grid approaches. How should neurons be connected by these sum-of-separables potentials when they are randomly placed throughout some region?

Monday, April 30, 2012

Spoke with Kamiar for a while over lunch. In his setup, the observations Y are drawn from a Poisson distribution parametrized by the angle X:

$$y_t \sim Poiss(a \cos x_t + b \sin x_t).$$

Also, I need to confirm that the block-wise Gibbs sampling approach is valid. Also, Kamiar points me to (Macke et al., 2009) for information on the observation distribution.

Saturday, April 28, 2012

A graphical model of pinwheel cells

Figure 1 is a graphical model of orientation selectivity in a grid of pinwheel cells. (To ask Kamiar: what topology is the most appropriate?) The following edge potentials serve to smooth the orientation maps, coupling adjacent nodes to keep their values similar:

$$p(o_{i,j}|o_{i-1,j}) \propto \sum_{k=0}^{R} e^{\kappa \cos\left(o_{i-1,j} - \frac{2\pi k}{R+1}\right)} e^{\kappa \cos\left(o_{i,j} - \frac{2\pi k}{R+1}\right)}$$

$$p(o_{i,j}|o_{i,j-1}) \propto \sum_{k=0}^{R} e^{\kappa \cos\left(o_{i,j-1} - \frac{2\pi k}{R+1}\right)} e^{\kappa \cos\left(o_{i,j} - \frac{2\pi k}{R+1}\right)}$$
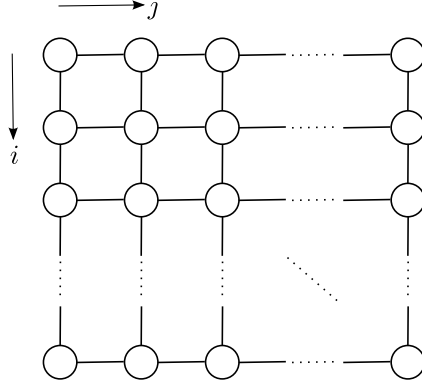
Figure 1: A graphical model of orientation selectivity, omitting observed nodes and their dependencies.

We can suppose that each node has a flat prior $p(o_{i,j})$, as we do here for simplicity, or we can suppose von Mises priors about set modes. (I'm being a bit fast and loose about whether this is a Markov random field or a Bayesian net, dealing with edge potentials as if they were transition probabilities. Hopefully the meaning is clear and I don't think there's anything illogical.) Throughout, the concentration parameter κ is assumed to be fixed. With the observations included the graphical model may be depicted as in Figure 2, where the small black circles denote the observations which have probability

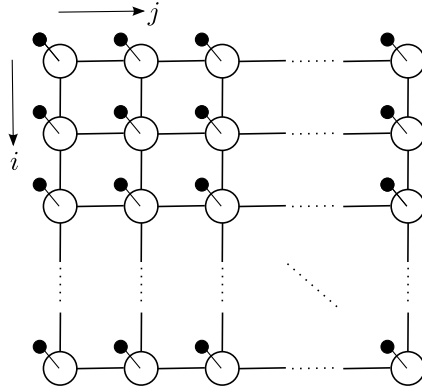$$p(z_{i,j}|o_{i,j}) \propto e^{\kappa \cos(z_{i,j}-o_{i,j})}.$$

Figure 2: A graphical model of orientation selectivity, including observed nodes and their dependencies.

The von Mises-Fisher distribution

On the unit circle, the von Mises distribution can be parametrized by the "mean" angle ϕ where the mode of the density resides:

$$f(x|\phi,\kappa) = \frac{e^{\kappa \cos(x-\phi)}}{2\pi I_0(\kappa)}$$

where $I_0(x)$ is the order 0 modified Bessel function. (Assume throughout that κ is fixed.) Instead, however, consider the unit vector $\mathbf{x}$ from the circle's center to the point on the perimeter an angle x away from the x-axis. Similarly define μ to be the vector with tail at the circle's center and point on the perimeter making the angle ϕ with the x-axis. This we may write

$$f(\mathbf{x}|\mu, \kappa) = \frac{e^{\kappa \mu^T \mathbf{x}}}{2\pi I_0(\kappa)}.$$

This is more generally an expression for the p-dimensional von Mises-Fisher distribution on the surface of the p-sphere, where $\mathbf{x}$ and μ are p-vectors and where $(2\pi I_0(\kappa))^{-1}$ is replaced by the more general normalization constant

$$C_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)}.$$

The conditional distribution of a column

Define $\mathbf{r}_k$ to be the unit vector that makes an angle of $\frac{2\pi k}{R+1}$ with the x-axis. Define $\mathbf{o}_{i,j}$ to be the vector that makes an angle of $o_{i,j}$ with the x-axis. Define $\mathbf{z}_{i,j}$ to be the unit vector that makes an angle of $z_{i,j}$ with the x-axis. Suppose that was have marginal estimates for every node and that we wish to block-wise draw a sample from the jth column. Without loss of generality, assume $1 < j < N$ where N is the number of columns and M the number of rows. The conditional distribution of the jth column given the rest of the nodes and given the observations is

$$\begin{aligned}
P_j &\equiv p(O_{1:M,j}|O\backslash O_{1:M,j}, Z) \\
&= p(O_{1:M,j}|O_{1:M,j-1}, O_{1:M,j+1}, Z_{1:M,j}) \\
&= p(o_{1,j}|o_{1,j-1}, o_{1,j+1}, z_{1,j}) \prod_{i=2}^{M} p(o_{i,j}|o_{i-1,j}, o_{i,j-1}, o_{i,j+1}, z_{i,j}).
\end{aligned}$$

This first factor is

$$\begin{aligned}
p(o_{1,j}|o_{1,j-1}, o_{1,j+1}, z_{1,j}) &\propto p(o_{1,j-1})p(o_{1,j}|o_{1,j-1})p(o_{1,j+1}|o_{1,j})p(z_{1,j}|o_{1,j}) \\
&\propto p(o_{1,j}|o_{1,j-1})p(o_{1,j+1}|o_{1,j})p(z_{1,j}|o_{1,j}) \\
&\propto \sum_{k=0}^{R} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{1,j-1}} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{1,j}} \sum_{k'=0}^{R} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{1,j}} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{1,j+1}} \cdot e^{\kappa \mathbf{z}_{1,j}^T \mathbf{o}_{1,j}}.
\end{aligned}$$

The second factor is

$$\begin{aligned}
p(o_{i,j}|o_{i-1,j}, o_{i,j-1}, o_{i,j+1}, z_{i,j}) &\propto p(o_{i,j-1})p(o_{i-1,j})p(o_{i,j}|o_{i,j-1})p(o_{i,j+1}|o_{i,j})p(z_{i,j}|o_{i,j}) \\
&\propto p(o_{i,j}|o_{i,j-1})p(o_{i,j+1}|o_{i,j})p(z_{i,j}|o_{i,j}) \\
&\propto \sum_{k=0}^{R} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{i,j-1}} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{i,j}} \sum_{k'=0}^{R} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{i,j}} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{i,j+1}} \cdot e^{\kappa \mathbf{z}_{i,j}^T \mathbf{o}_{i,j}}.
\end{aligned}$$

Putting these together, the conditional distribution is

$$P_j \propto \sum_{k=0}^{R} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{1,j-1}} e^{\kappa \mathbf{r}_k^T \mathbf{o}_{1,j}} \sum_{k'=0}^{R} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{1,j}} e^{\kappa \mathbf{r}_{k'}^T \mathbf{o}_{1,j+1}} \cdot e^{\kappa \mathbf{z}_{1,j}^T \mathbf{o}_{1,j}} \times$$

$$\prod_{i=2}^{T} \sum_{k''=0}^{R} e^{\kappa \mathbf{r}_{k''}^T \mathbf{o}_{i,j-1}} e^{\kappa \mathbf{r}_{k''}^T \mathbf{o}_{i,j}} \sum_{k'''=0}^{R} e^{\kappa \mathbf{r}_{k'''}^T \mathbf{o}_{i,j}} e^{\kappa \mathbf{r}_{k'''}^T \mathbf{o}_{i,j+1}}.$$

Sampling a column block-wise by forward backward

Along column j, starting from the top ($i = 1$), we compute forward variables each corresponding to a particular value of the auxiliary value k_i that connects $o_{i,j}$ to $o_{i+1,j}$. It is helpful to have the joint factor graph representation in Figure 3 in mind.

$$A_1^{(k_1)} = \int do_{1j} p(o_{1j}) p(k_1|o_{1j}) p(z_{1j}|o_{1j})$$

$$\propto \int do_{1j} \left(\sum_{l_1=0}^{R} e^{\kappa \mathbf{r}_{l_1}^T \mathbf{o}_{1,j-1}} e^{\kappa \mathbf{r}_{l_1}^T \mathbf{o}_{1j}} \sum_{m_1=0}^{R} e^{\kappa \mathbf{r}_{m_1}^T \mathbf{o}_{1j}} e^{\kappa \mathbf{r}_{m_1}^T \mathbf{o}_{1,j+1}} \right) e^{\kappa \mathbf{r}_{k_1}^T \mathbf{o}_{1j}} e^{\kappa \mathbf{z}_{1j}^T \mathbf{o}_{1j}}$$

$$= \sum_{l_1=0}^{R} \sum_{m_1=0}^{R} e^{\kappa(\mathbf{r}_{l_1}^T \mathbf{o}_{1,j-1} + \mathbf{r}_{m_1}^T \mathbf{o}_{1,j+1})} \int do_{1j} e^{\kappa(\mathbf{r}_{l_1} + \mathbf{r}_{m_1} + \mathbf{r}_{k_1} + \mathbf{z}_{1j})^T \mathbf{o}_{1j}}$$

$$= \sum_{l_1=0}^{R} \sum_{m_1=0}^{R} e^{\kappa(\mathbf{r}_{l_1}^T \mathbf{o}_{1,j-1} + \mathbf{r}_{m_1}^T \mathbf{o}_{1,j+1})} \frac{1}{C_p(\kappa \cdot ||\mathbf{r}_{l_1} + \mathbf{r}_{m_1} + \mathbf{r}_{k_1} + \mathbf{z}_{1j}||)}$$

$$A_i^{(k_i)} = \sum_{k_{i-1}=0}^{R} A_{i-1}^{(k_{i-1})} \int do_{i,j} p(o_{i,j}|k_{i-1}) p(k_i|o_{i,j}) p(z_{i,j}|o_{i,j})$$

$$\propto \sum_{k_{i-1}=0}^{R} A_{i-1}^{(k_{i-1})} \int do_{i,j} \left(\sum_{l_i=0}^{R} e^{\kappa \mathbf{r}_{l_i}^T \mathbf{o}_{i,j-1}} e^{\kappa \mathbf{r}_{l_i}^T \mathbf{o}_{i,j}} \sum_{m_i=0}^{R} e^{\kappa \mathbf{r}_{m_i}^T \mathbf{o}_{i,j}} e^{\kappa \mathbf{r}_{m_i}^T \mathbf{o}_{i,j+1}} \right) e^{\kappa \mathbf{r}_{k_{i-1}}^T \mathbf{o}_{i,j}} e^{\kappa \mathbf{r}_{k_i}^T \mathbf{o}_{i,j}}$$

$$= \sum_{k_{i-1}=0}^{R} A_{i-1}^{(k_{i-1})} \sum_{l_i=0}^{R} \sum_{m_i=0}^{R} e^{\kappa \left(\mathbf{r}_{l_i}^T \mathbf{o}_{i,j-1} + \mathbf{r}_{m_i}^T \mathbf{o}_{i,j+1} \right)} \int do_{i,j} e^{\kappa(\mathbf{r}_{l_i} + \mathbf{r}_{m_i} + \mathbf{r}_{k_{i-1}} + \mathbf{r}_{k_i} + \mathbf{z}_{i,j})^T \mathbf{o}_{i,j}}$$

$$= \sum_{k_{i-1}=0}^{R} A_{i-1}^{(k_{i-1})} \sum_{l_i=0}^{R} \sum_{m_i=0}^{R} e^{\kappa \left(\mathbf{r}_{l_i}^T \mathbf{o}_{i,j-1} + \mathbf{r}_{m_i}^T \mathbf{o}_{i,j+1} \right)} \frac{1}{C_p(\kappa \cdot ||\mathbf{r}_{l_i} + \mathbf{r}_{m_i} + \mathbf{r}_{k_{i-1}} + \mathbf{r}_{k_i} + \mathbf{z}_{i,j}||)}$$

Similarly we may derive backward variables:

$$B_M^{(k_{M-1})} = \int do_{M,j} p(o_{M,j}|k_{M-1,j}) p(z_{M,j}|o_{M,j})$$

$$= \sum_{l_M=0}^{R} \sum_{m_M=0}^{R} e^{\kappa \left(\mathbf{r}_{l_M}^T \mathbf{o}_{M,j-1} + \mathbf{r}_{m_M}^T \mathbf{o}_{M,j+1} \right)} \frac{1}{C_p(\kappa \cdot ||\mathbf{r}_{l_M} + \mathbf{r}_{m_M} + \mathbf{r}_{k_{M-1}} + \mathbf{z}_{M,j}||)}$$

$$B_i^{k_{i-1}} = \sum_{k_i=0}^{R} B_{i+1}^{k_i} \int do_{i,j} p(o_{i,j}|k_{i-1}) p(k_i|o_{i,j}) p(z_{i,j}|o_{i,j})$$
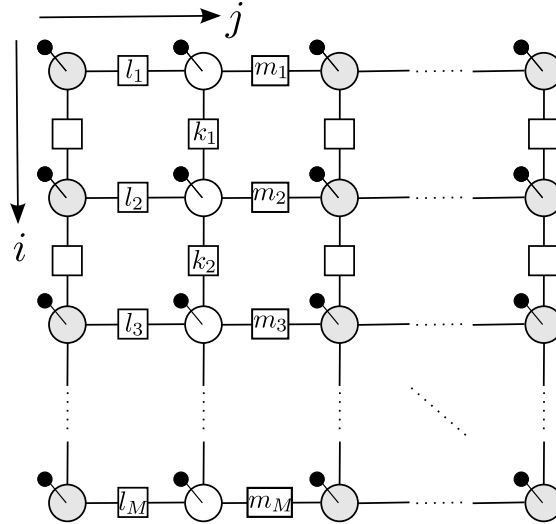
Figure 3: A graphical model of orientation selectivity, including observed nodes and auxiliary variables and their dependencies, drawing attention to a forward backward run on the second column from the left. The shaded nodes' values are held as constants, as are the observations. At each step of the forward recursion, to each possible value of k_i corresponds a forward variable $A_i^{(k_i)}$.

$$= \sum_{k_i=0}^{R} B_{i+1}^{k_i} \sum_{l_i=0}^{R} \sum_{m_i=0}^{R} e^{\kappa\left(\mathbf{r}_{l_i}^T \mathbf{o}_{i,j-1} + \mathbf{r}_{m_i}^T \mathbf{o}_{i,j+1}\right)} \frac{1}{C_p(\kappa \cdot ||\mathbf{r}_{l_i} + \mathbf{r}_{m_i} + \mathbf{r}_{k_{i-1}} + \mathbf{r}_{k_i} + \mathbf{z}_{i,j}||)}$$

DERIVE EXPRESSION FOR SINGLETON MARGINAL IN TERMS OF FORWARD AND BACKWARD VARIABLES.

Friday, April 27, 2012

Liam had an idea for using the low-rank models to smooth orientation maps on cortical surface. From an email:

> imagine you observe a bunch of neurons on the cortical surface. each neuron is labeled by its position on the surface. we know that in many cases nearby neurons have similar tuning properties. so if we want to estimate the tuning properties of a bunch of these neurons that we observe simultaneously, it makes sense to share information across neurons by doing some smoothing.
>
> in V1, for example, the preferred orientation is usually a smooth function of position. so we might introduce the following simple hierarchical model:
>
> let $o(x, y)$ denote the preferred orientation of the neuron at position (x, y). we observe some information $z(x, y)$ about each neuron's orientation. so we can write down

$$p(O, Z) = p(O)p(Z|O)$$
$$p(Z|O) = \prod_{x,y} p(z(x, y)|o(x, y)).$$

5

now we need a smoothing prior for $p(O)$. one convenient choice is to use a spatial prior of the low-rank form discussed in (Smith et al., 2012).

$$p(O) = \prod_{x,y} \phi[o(x,y), o(x, y+1)]\phi[o(x,y), o(x+1, y)]$$

here the term $\phi[o(x,y), o(x, y+1)]$ penalizes roughness in the vertical direction, and $\phi[o(x,y), o(x+1, y)]$ penalizes the horizontal direction.

if $\phi(\cdot)$ has the low-rank form discussed in the paper, then inference in this model can proceed via an efficient blockwise gibbs approach: choose a row (or column) of $\{o(x,y)\}_{x=i}$, and draw an exact sample from

$$p(o(x,y)_{x=i} | z(x,y)_{x=i}, o(x,y)_{x=i-1}, o(x,y)_{x=i+1}),$$

using the fast method in (Smith et al., 2012). iterating across rows and columns gives samples from $P(O|Y)$.

Liam has in mind the kind of data described in (Ohki et al., 2006).

References

Macke, J., Gerwinn, S., White, L., Kaschube, M., and Bethge, M. (2009). Bayesian estimation of orientation preference maps. In Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C. K. I., and Culotta, A., editors, Advances in Neural Information Processing Systems 22, pages 1195–1203.

Ohki, K., Chung, S., Kara, P., Hubener, M., Bonhoeffer, T., and Reid, R. C. (2006). Highly ordered arrangement of single neurons in orientation pinwheels. Nature, 442(24):925–928.

Smith, C., Wood, F., and Paninski, L. (2012). Low rank continuous-space graphical models. In Proc. 15th Intl. Conf. on Artificial Intelligence and Statistics (AISTATS), pages 1064–1072.