DISCRETIZING GAUSSIAN MODELS

DUSTIN CARTWRIGHT

Let Σ be a positive semidefinite matrix with non-zero diagonal entries, and G_Σ the corresponding (possibly singular) Gaussian distribution on n random variables with mean 0. In this report, we study the discretization formed by taking just the sign of each of the variables, i.e. the model on sign patterns $s \in \{\pm 1\}^n$ whose distribution is

$$p_s = P_\Sigma(s_1 x_1 > 0, \ldots, s_n x_n > 0)$$

The goal is to describe as explicitly as possible the possible probability distributions on sign patterns which can arise in this way. This is done completely for n at most 3, but for $n = 4$, these possible probability distributions are defined by a non-algebraic equation.

First note that we can rescale each of the random variables by its variance to get a correlation matrix whose (i, j) entry is $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$. The set of all positive semidefinite matrices with ones on the diagonal is called an elliptope and we will denote it E_n. From now on Σ will always denote a point in the elliptope E_n.

By the symmetry of the Gaussian distribution, the discretization will always satisfy $p_s = p_{-s}$. Thus, we will regard such a distribution as lying within the $(2^{n-1} - 1)$-dimensional simplex $\Delta^{2^{n-1}-1}$.

The Cholesky decomposition of a postive semidefinite allows us to reinterpret the discretization in terms of the spherical volume of the cones cut out by a configuration of oriented hyperplanes. Factoring Σ as AA^T for a real matrix A tells us that the distribution on x is the same as that coming from the transformation $x = Ax'$ where the entries of x' are independently distributed samples from the standard Gaussian distribution. Thus $x_i = 0$ corresponds to the hyperplane defined by the ith column of A (which is non-zero by our assumption on Σ), and so the sign pattern of a vector x is determined by where it lies in the complement of these hyperplanes. Since the probability distribution on x' is rotationally symmetric, it suffices to compute the relative volume of the intersection of any of these components with an $(n-1)$-dimensional sphere centered at the origin.

For $n = 2$, this allows us to understand the discretization map explicitly. Let $r = \Sigma_{12}$ be the off-diagonal entry of the matrix. Then the Cholesky decomposition gives us two hyperplanes which lie at an angle of $\pi/2 + \arcsin r$. Rescaling, we get:

$$p_{++} = p_{--} = \tfrac{1}{4} + \tfrac{1}{2\pi} \arcsin r$$
$$p_{+-} = p_{-+} = \tfrac{1}{4} - \tfrac{1}{2\pi} \arcsin r$$

In particular, the probabilities are not algebraic functions of the entries of Σ. Nonetheless, this parametrization tells us that the resulting probability distributions are exactly all such distributions which satisfy the above noted symmetry $p_s = p_{-s}$.

Date: December 10, 2008.

From this case we get the following:

Proposition 1. *Discretization induces a homeomorphism of E_n onto its image.*

Proof. We can write down a continuous inverse:

$$r_{ij} = \sin\left(2\pi\left(\sum_{s\in\{\pm1\}^n : s_i=s_j=+1} p_s - \tfrac{1}{4}\right)\right)$$

The summation amounts to marginalizing all the variables except i and j. Then the formula follows from the 2-variable situation analyzed above. □

Now we turn to the situation with 3 variables. Here, we have

$$p_{s_1s_2s_3} = \frac{1}{8} + \frac{s_1s_2}{4\pi}\arcsin r_{12} + \frac{s_1s_3}{4\pi}\arcsin r_{13} + \frac{s_2s_3}{4\pi}\arcsin r_{23} \tag{1}$$

for any $s_i = \pm1$. This formula can be proved from the Gauss-Bonnet theorem, but it can also be derived from the 2-variable case. Marginalizing any of the 3 variables gives us the three equations

$$\begin{aligned} p_{+++} + p_{++-} &= \frac{1}{4} + \frac{1}{2\pi}\arcsin r_{12} \\ p_{+++} + p_{+-+} &= \frac{1}{4} + \frac{1}{2\pi}\arcsin r_{13} \\ p_{+++} + p_{-++} &= \frac{1}{4} + \frac{1}{2\pi}\arcsin r_{23} \end{aligned}$$

Taking these three equations together with the constraint that the sum of the probabilities equals 1, we can solve and get (1). As in the 2-variable case, the image is surjective:

Proposition 2. *Discretization is a homeomorphism from E_3 to Δ_3.*

Proof. By Proposition 1, it suffices to show that the map is surjective.

First, we claim that the boundary of E_3 surjects onto the boundary of Δ_3, i.e. probability distributions with at least one 0. Such probability distributions are obtained from configurations of hyperplanes which intersect in a line and thus divide $\mathbb{R}^3$ into at most 6 cones, i.e. at most 3 pairs of opposite cones. However, by rotating the hyperplanes around their common line of intersection, we can obtain all possible probability distributions on these 3 cones, thus proving the claim.

Now, we use algebraic topology to show that the map is surjective. If there were a point of Δ_3 not in the image, then regarding Δ_3 topologically as a 3-dimensional disk, we can project away from that point to get a map to the boundary. Composing with the discretization map gives us a map from E_3, which, again, is topologically a disk to the boundary of Δ^3, which is homeomorphic to the boundary of E_3 by the claim and Proposition 1. However, this is impossible because there is no retract from a disk to its boundary. □

We can say a little bit more about this map. The preimages of the vertices of Δ^3 are the four rank 1 matrices on the boundary of the elliptope. The preimages of the edges are the six families of matrices:

$$\begin{bmatrix} 1 & \pm1 & a \\ \pm1 & 1 & \pm a \\ a & \pm a & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & a & \pm1 \\ a & 1 & \pm a \\ \pm1 & \pm a & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & a & \pm a \\ a & 1 & \pm1 \\ \pm a & \pm1 & 1 \end{bmatrix}$$

where $-1 \leq a \leq 1$ and $\pm$ denotes the same sign within each matrix. These are exactly the matrices in the elliptope where one of the 2×2 principal minors vanishes. In this case, the non-negativity of the determinant forces the relationship between the other two off-diagonal entries.

Finally, we come to the case of $n = 4$, in which case the image is no longer semialgebraic:

Theorem 3. *The image of E^4 in Δ^7 is an analytic hypersurface, but there is no non-trivial polynomial which vanishes on it.*

Proof. We restrict our attention to matrices in the elliptope of the form

$$\begin{bmatrix} 1 & r_{12} & & \\ r_{12} & 1 & r_{23} & \\ & r_{23} & 1 & r_{34} \\ & & r_{34} & 1 \end{bmatrix}$$

The corresponding distributions are known as orthoschemes in the statistics literature. The three zero entries tell us that the image of these matrices lies in a codimension 3 linear space. From Equation (2.2) in [Ab], we have the following, attributed to van der Vaart:

$$p_{s_1 s_2 s_3 s_4} = \frac{1}{16} + \frac{1}{8\pi}(s_1 s_2 a_{12} + s_2 s_3 a_{23} + s_3 s_4 a_{34}) + \frac{1}{4\pi^2}\left(s_1 s_2 s_3 s_4 \int_0^{a_{12}} \int_0^{a_{34}} \frac{d\alpha\, d\beta}{\sqrt{1 - \frac{\sin^2 a_{23}}{\cos^2 \alpha \cos^2 \beta}}} \right)$$

An affine linear change of coordinates gives us the variables a_{12}, a_{23}, a_{34}, together with the final double integral, which we will denote f. It will suffice to show that there is no polynomial relation between these coordinates.

If g were such a relation, then differentiating g with respect to a_{12} gives us another relation involving $\frac{\partial f}{\partial a_{12}}$:

$$\frac{\partial g}{\partial f}\frac{\partial f}{\partial a_{12}} + \frac{\partial g}{\partial a_{12}} = 0$$

Eliminating f from these two relations gives us a non-trivial relation between the a_{ij} and $\frac{\partial f}{\partial a_{12}}$. Repeating variations of this argument, we get a relation among the a_{ij} and

$$1 - \left(\frac{\partial^2 f}{\partial a_{12}\partial a_{34}}\right)^{-2} = \frac{\sin^2 a_{23}}{\cos^2 a_{12} \cos^2 a_{34}}$$

which is impossible since sin and cos are not algebraic functions. $\square$

References

[Ab] Abrahmson, I. G. Orthant probabilities for the quadrivariate normal distribution. *Ann. Math. Statistic.* **35**: 4 (1964). 1685–1703.