Sparseness, conditional independence and unnormalized densities

Rasmus Waagepetersen

April 17, 2022

Outline:

1. conditional independence and sparseness
2. unnormalized densities

Conditional independence

Suppose X, Y, Z are random variables (or vectors). Then we define X and Y to be conditionally independent given Z if

$$p(x, y|z) = p(x|z)p(y|z)$$

The following statements are equivalent:

1. $p(x, y|z) = p(x|z)p(y|z)$
2. $p(x, y, z) = f(x, z)g(y, z)$ for some functions f and g
3. $p(x|y, z) = p(x|z)$
4. $p(y|x, z) = p(y|z)$

($p(\cdot)$ generic notation for (possibly conditional) probability densities)

Suppose $X \sim N(\mu, \Sigma)$ with precision matrix $Q = \Sigma^{-1}$.

Then X_i and X_j conditionally independent given $X_{-\{i,j\}} \Leftrightarrow$ $Q_{ij} = 0$. This follows from decomposition

$$(x-\mu)^{\mathsf{T}} Q (x-\mu) =$$
$$\left\{ (x_i - \mu_i)^2 Q_{ii} + 2 \sum_{k \neq i} (x_i - \mu_i)(x_k - \mu_k) Q_{ik} \right\} + \sum_{\substack{l,k: \\ l \neq i, k \neq i}} (x_l - \mu_l)(x_k - \mu_k) Q_{lk}$$

Note that x_i not in last term and $Q_{ij} = Q_{ji} = 0$ implies x_j not in first term. Thus we obtain factorization of density of X:

$$p(x) = f(x_i, x_{-\{i,j\}}) g(x_j, x_{-\{i,j\}})$$

In particular, if Q is sparse, a lot of X_i, X_j will be conditionally independent given the remaining variables.

Unnormalized density

To specify a probability density it is enough to specify an *unnormalized* density $h(\cdot)$:

$$f(x) \propto h(x) \Leftrightarrow f(x) = h(x)/c$$

where normalizing constant c uniquely determined by:

$$\int f(x)\mathrm{d}x = 1 \Leftrightarrow \int h(x)/c\mathrm{d}x = 1 \Leftrightarrow c = \int h(x)\mathrm{d}x$$

For example if X has density proportional to

$$h(x) = x^{a-1}\exp(-bx), \quad a, b > 0$$

we know that X has a Gamma distribution.

Conditional distribution of X_i

By previous slide

$$p(x_i|x_{-i}) \propto \exp(-\frac{1}{2}(x_i - \mu_i)^2 Q_{ii} - \sum_{k \neq i}(x_i - \mu_i)(x_k - \mu_k)Q_{ik})$$

Note for a normal distribution $Y \sim N(\xi, \sigma^2)$,

$$p(y) \propto \exp(-\frac{1}{2\sigma^2}y^2 + \frac{1}{\sigma^2}y\xi)$$

Comparing the two above equations we get

$$X_i|X_{-i} = x_{-i} \sim N(\mu_i - \frac{1}{Q_{ii}}\sum_{k \neq i} Q_{ik}(x_k - \mu_k), Q_{ii}^{-1})$$

Again we see that $Q_{ij} = Q_{ji} = 0 \Leftrightarrow X_i$ conditionally independent of X_j given $X_{-\{i,j\}}$.

Looking at bivariate distribution of (X_i, X_j) given $X_{-\{i,j\}}$ shows that the conditional (partial) correlation is

$$\mathbb{C}\mathrm{orr}[X_i, X_j|X_{-i,j}] = -Q_{ij}/\sqrt{Q_{ii}Q_{jj}}$$

Exercises

1. show the equivalence of 1.-4. on slide 3.
2. verify the factorization on slide 4 and the expression for the conditional distribution of X_i on slide 6.