Idea of Identifiability establishment: A Linear, Non-Gaussian Case

Kun Zhang, for 80-816/516

March 2025

1 Setting

Suppose the true model is

$$\mathbf{X} = \mathbf{A}\mathbf{S}, \tag{1}$$

where $S = (S_1, S_2, ..., S_d)^\intercal$ has independent components and $\mathbf{A}$ is the mixing matrix. Here let us assume at most one S_i is Gaussian. To show that this model is identifiable up to trivial permutation and scaling indeterminacies in the columns of $\mathbf{A}$ (or S_i), let us assume that $\mathbf{X}$ also admits the following model in which $\mathbf{Z} = (Z_1, Z_2, ..., Z_d)^\intercal$ has independent components:

$$\mathbf{X} = \mathbf{B}\mathbf{Z}. \tag{2}$$

Equations (1) and (2) imply that $\mathbf{Z}$ is a linear transformation of $\mathbf{S}$:

$$\mathbf{Z} = \mathbf{M}\mathbf{S}, \tag{3}$$

where $\mathbf{M} = \mathbf{B}^{-1}\mathbf{A}$. We now aim to show that $\mathbf{M}$ must be a generalized permuted matrix.

2 Basic Idea: Characterization of Independence

As mentioned multiple times in class, statistical independence is an essential property to be exploited to describe the data distribution in generative models. In a generative model, the joint distribution of the observed variables is specified by the model–what properties of the joint distribution guarantee independence?

Assume the density function is second-order differentiable (although it can be weakened. A commonly-used property used to guarantee independence between Z_1 and Z_2 is

$$\frac{\partial^2 \log p_{Z_1,Z_2}}{\partial z_1 \partial z_2} \equiv 0. \tag{4}$$

This is because $Z_1 \perp\!\!\!\perp Z_2$ if and only if

$$\log p_{Z_1,Z_2} = \log p_{Z_1} + \log p_{Z_2}.$$

Similarly, $Z_k \perp\!\!\!\perp Z_l \,|\, \mathbf{Z} \setminus \{Z_k, Z_l\}$ can be characterized by

$$\frac{\partial^2 \log p_{\mathbf{Z}}}{\partial z_k \partial z_l} \equiv 0. \tag{5}$$

3 Derivation

Let $\mathbf{Q} = \mathbf{M}^{-1}$, with entries q_{ij}. Thanks to the change of variables, Equation (3) implies

$$p_{\mathbf{Z}} = p_{\mathbf{S}}/|\mathbf{M}|,$$

where $|\mathbf{M}|$ is the absolute value of the determinant of $\mathbf{M}$. That is,

$$\log p_{\mathbf{Z}} = \log p_{\mathbf{S}} - \log|\mathbf{M}| = \sum_{i=1}^{d} \log p_{S_i} - \log|\mathbf{M}|, \tag{6}$$

because of the independence among S_i. Its partial derivative w.r.t. z_k is

$$\begin{aligned}
\frac{\partial \log p_{\mathbf{Z}}}{\partial z_k} &= \sum_{i=1}^{d} \frac{\partial \log p_{Si}}{\partial z_k} \\
&= \sum_{i=1}^{d} \frac{\partial \log p_{Si}}{\partial s_i} \cdot \frac{\partial s_i}{\partial z_k} \\
&= \sum_{i=1}^{d} q_{ik} \cdot \frac{\partial \log p_{Si}}{\partial s_i} \\
&= \sum_{i=1}^{d} q_{ik} \cdot \eta_i'(s_i),
\end{aligned} \tag{7}$$

where $\eta_i := \log p_{S_i}$. Moreover,

$$\begin{aligned}
\frac{\partial^2 \log p_{\mathbf{Z}}}{\partial z_k \partial z_l} &= \sum_{i=1}^{d} q_{ik} \cdot \eta_i''(s_i) \cdot \frac{\partial s_i}{\partial z_l}, \\
&= \sum_{i=1}^{d} q_{ik} q_{il} \cdot \eta_i''(s_i).
\end{aligned} \tag{8}$$

To enforce mutual independence among Z_i, let us first enforce a weak condition, which is $Z_k \perp\!\!\!\perp Z_l \,|\, \mathbf{Z} \setminus \{Z_k, Z_l\}$, which is equivalent to (5). Combining (5) and (8) gives

$$\sum_{i=1}^{d} q_{ik} q_{il} \cdot \eta_i''(s_i) \equiv 0. \tag{9}$$

We know that S_i is Gaussian if and only if $\eta_i''(s_i)$ is constant (and nonzero). Now suppose S_j is not Gaussian, then there exist at least two different values of S_j, denoted by $s_j^{(1)}$ and $s_j^{(2)}$, such that

$$\eta_j''(s_j^{(1)}) \neq \eta_j''(s_j^{(2)}). \tag{10}$$

Then (9) tells us

$$q_{jk}q_{jl} \cdot [\eta_j''(s_j^{(1)}) - \eta_j''(s_j^{(2)})] = 0,$$

which, together with (10), implies

$$q_{jk}q_{jl} = 0.$$

Now consider the case where one component of $\mathbf{S}$ is Gaussian. Denote by S_r this component. Since we already know $q_{jk}q_{jl} = 0$ for any $j \neq r$, Equation (9) implies

$$q_{rk}q_{rl} \cdot \eta_r''(s_r) = 0,$$

that is,

$$q_{rk}q_{rl} = 0.$$

In other words, we have shown that each column of $\mathbf{Q}$ has at most one nonzero entry. Further notice that $\mathbf{Q}$ is not singular, and we see that each column of $\mathbf{Q}$ has exactly one nonzero entry and that furthermore, each row of $\mathbf{Q}$ has exactly one nonzero entry. Hence $\mathbf{Q}$ is a generalized permutation matrix. So is $\mathbf{M}$, as its inverse.

By the way, since $\mathbf{M}$ is a generalized permutation matrix, Z_i, as a permuted and rescaled version of S_i, will be mutually independent, although in the proof above we only conditional independence $Z_k \perp\!\!\!\perp Z_l \,|\, \mathbf{A} \setminus \{Z_k, Zl\}$.