On Aggregate Dimension Reduction

Qin Wang, Xiangrong Yin, Bing Li and Zhihui Tang

The University of Alabama, University of Kentucky

Penn State University and PPDI

Supplementary Material

S1 Proof of Theorem 1

Let G be an open subset of $\Omega_{\mathbf{X}}$. Then, by part 1 of Proposition 11 in Yin et al. (2008), we have that $\mathcal{S}_{Y_G|\mathbf{X}_G} \subseteq \mathcal{S}_{Y|\mathbf{X}}$, which implies that $\text{span}\{\mathcal{S}_{Y_G|\mathbf{X}_G} : G \subseteq \Omega_{\mathbf{X}}\} \subseteq \mathcal{S}_{Y|\mathbf{X}}$.

By a result of Zhu and Zeng (2006), we have

$$\text{span}\left\{\partial h(y \mid \mathbf{x})/\partial \mathbf{x} : (\mathbf{x}, y) \in \Omega_{\mathbf{X}} \times \Omega_Y\right\} = \mathcal{S}_{Y|\mathbf{X}}. \quad \text{(S1.1)}$$

Apply the same result to $(\mathbf{X}_G, Y_G)$ to obtain

$$\text{span}\{\partial h(y \mid \mathbf{x})/\partial \mathbf{x} : (\mathbf{x}, y) \in G \times \Omega_Y\} = \mathcal{S}_{Y_G|\mathbf{X}_G}. \quad \text{(S1.2)}$$

Now let $(\mathbf{x}_0, y_0)$ be an arbitrary point in $(\Omega_{\mathbf{X}}, \Omega_Y)$, and let G be an open subset of $\Omega_{\mathbf{X}}$ that contains $\mathbf{x}_0$. Then, by part 3 of Proposition 1, $h_G(y \mid$

$\mathbf{x}) = h(y \mid \mathbf{x})$ for all $(\mathbf{x}, y) \in G \times \Omega_Y$. Therefore, $[\partial h_G(y \mid \mathbf{x})/\partial \mathbf{x}]_{\mathbf{x}_0, y_0} = [\partial h(y \mid \mathbf{x})/\partial \mathbf{x}]_{\mathbf{x}_0, y_0}$. Thus, by (S1.1) and (S1.2) we have

$$\mathcal{S}_{Y|\mathbf{X}} \subseteq \cup\{\mathcal{S}_{Y|\mathbf{X}_G} : G \subseteq \Omega_{\mathbf{X}}\} \subseteq \text{span}\{\mathcal{S}_{Y|\mathbf{X}_G} : G \subseteq \Omega_{\mathbf{X}}\}. \tag{S1.3}$$

Furthermore, by part 2 of Proposition 11 of Yin et al. (2008), there exists a compact set $K \subseteq \Omega_{\mathbf{X}}$ such that $\mathcal{S}_{Y_K|\mathbf{X}_K} = \mathcal{S}_{Y|\mathbf{X}}$, where $(\mathbf{X}_K, Y_K)$ is defined as $\mathbf{X}$ restricted on K. Since $\cup\{G : G \subseteq \Omega_{\mathbf{X}}\}$ forms an open cover of the compact set K, there is a finite subcover $\cup\{G_i : i = 1, \ldots, m\}$ of K. Hence by the same argument leading to (S1.3) we have $\mathcal{S}_{Y|\mathbf{X}} \subseteq \cup\{\mathcal{S}_{Y|\mathbf{X}_{G_i}} : i = 1, \ldots, m\}$, as desired. □

S2 Proof of Theorem 2

We have

$$E(\mathbf{X}_G \mid Y_G = y) = \int_G \mathbf{x}\, \frac{h(y \mid \mathbf{x}) p_G(\mathbf{x})}{g_G(y)} d\mathbf{x} = \frac{1}{g_G(y)} \int_G \mathbf{x}\, h(y \mid \mathbf{x}) p_G(\mathbf{x}) d\mathbf{x}. \tag{S2.1}$$

Let $\dot{h}(y \mid \mathbf{x})$ and $\ddot{h}(y \mid \mathbf{x})$ denote the first and second derivatives of h with respect to $\mathbf{x}$. By Taylor's theorem, for any $\mathbf{x} \in G$, there is a $\boldsymbol{\xi}$ with $\|\boldsymbol{\xi} - \boldsymbol{\mu}_G\| \leq \|G\|$ such that

$$h(y \mid \mathbf{x}) = h(y \mid \boldsymbol{\mu}_G) + \dot{h}^T(y \mid \boldsymbol{\mu}_G)(\mathbf{x} - \boldsymbol{\mu}_G) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_G)^T \ddot{h}(y \mid \boldsymbol{\xi})(\mathbf{x} - \boldsymbol{\mu}_G). \tag{S2.2}$$

In the meantime,

$$\begin{aligned} h(y \mid \boldsymbol{\mu}_{G} + \mathbf{P}_{\boldsymbol{\beta}_G}(\mathbf{x}-\boldsymbol{\mu}_{G})) =& h(y \mid \boldsymbol{\mu}_{G}) + \dot{h}^T(y \mid \boldsymbol{\mu}_{G})\mathbf{P}_{\boldsymbol{\beta}_G}(\mathbf{x}-\boldsymbol{\mu}_{G}) \\ &+ \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{G})^T\mathbf{P}_{\boldsymbol{\beta}_G}\ddot{h}(y \mid \boldsymbol{\xi})\mathbf{P}_{\boldsymbol{\beta}_G}(\mathbf{x}-\boldsymbol{\mu}_{G}). \end{aligned} \tag{S2.3}$$

However, by construction, it is easy to see that $\dot{h}(y \mid \boldsymbol{\mu}_{G}) \in \text{span}(\mathbf{H}_{G})$ almost everywhere in Ω_Y. Hence $\mathbf{P}_{\boldsymbol{\beta}_G}\dot{h}(y \mid \boldsymbol{\mu}_{G}) = \dot{h}(y \mid \boldsymbol{\mu}_{G})$. Because $\|\mathbf{x}-\boldsymbol{\mu}_{G}\| \leq \|G\|$ and the elements of $\ddot{h}(y \mid \boldsymbol{\xi})$ are bounded, the third terms on the right hand sides of (S2.2) and (S2.3) are of the order $O(\|G\|^2)$. Now subtract (S2.2) from (S2.3),

$$h(y|\mathbf{x}) = h(y|\boldsymbol{\mu}_{G} + \mathbf{P}_{\boldsymbol{\beta}_G}(\mathbf{x}-\boldsymbol{\mu}_{G})) + O(\|G\|^2) \quad \text{as} \quad \|G\| \to 0. \tag{S2.4}$$

Substitute (S2.2) into the right hand side of (S2.1), using the relations $E(\mathbf{X}_{G}-\boldsymbol{\mu}_{G}) = 0$ and $\text{Var}(\mathbf{X}_{G}) = \boldsymbol{\Sigma}_{G}$, to obtain

$$\begin{aligned} E(\mathbf{X}_{G}-\boldsymbol{\mu}_{G} \mid Y_{G} = y) =& \frac{1}{g_{G}(y)}\boldsymbol{\Sigma}_{G}\dot{h}(y \mid \boldsymbol{\mu}_{G}) \\ &+ \frac{1}{2g_{G}(y)}\int_{G}[(\mathbf{x}-\boldsymbol{\mu}_{G})(\mathbf{x}-\boldsymbol{\mu}_{G})^T\ddot{h}(y \mid \boldsymbol{\xi})(\mathbf{x}-\boldsymbol{\mu}_{G})]p_{G}(\mathbf{x})d\mathbf{x}. \end{aligned} \tag{S2.5}$$

Since $\|\mathbf{x}-\boldsymbol{\mu}_{G}\| \leq \|G\|$ and the components of $\ddot{h}(y \mid \boldsymbol{\xi})$ are bounded, the second term on the right is of the order $O(\|G\|^3)$. In other words,

$$E(\mathbf{X}_{G}-\boldsymbol{\mu}_{G} \mid Y_{G} = y) = \frac{1}{g_{G}(y)}\boldsymbol{\Sigma}_{G}\dot{h}(y \mid \boldsymbol{\mu}_{G}) + O(\|G\|^3).$$

Multiply both sides by $\boldsymbol{\Sigma}_{G}^{-1}$, keeping in mind that $\boldsymbol{\Sigma}_{G} = O(\|G\|^2)$, to obtain

$$\boldsymbol{\Sigma}_{G}^{-1}E(\mathbf{X}_{G} - \boldsymbol{\mu}_{G} \mid Y_{G} = y) = \frac{1}{g_{G}(y)}\dot{h}(y \mid \boldsymbol{\mu}_{G}) + O(\|G\|). \tag{S2.6}$$

Meanwhile, if we multiply both sides of the above equality by $\mathbf{P}_{\boldsymbol{\beta}_{G}}$, then, because $\dot{h}(y \mid \boldsymbol{\mu}_{G}) \in \text{span}(\boldsymbol{\beta}_{G})$ for almost every $y \in \Omega_Y$, we have

$$\mathbf{P}_{\boldsymbol{\beta}_{G}}\boldsymbol{\Sigma}_{G}^{-1}E(\mathbf{X}_{G} - \boldsymbol{\mu}_{G} \mid y) = \frac{1}{g_{G}(y)}\dot{h}(y \mid \boldsymbol{\mu}_{G}) + O(\|G\|). \tag{S2.7}$$

Now subtract (S2.7) from (S2.6) to prove (3.1). □

S3 Proof of Theorem 3

Let $\boldsymbol{\mu}_{G}^{*}$ be the center of G. Since p_{G} has bounded derivative, $p_{G}(\mathbf{x}) = p_{G}(\boldsymbol{\mu}_{G}^{*}) + O(\|G\|)$. Hence

$$\begin{aligned}\boldsymbol{\mu}_{G} &= \int_G (\mathbf{x} - \boldsymbol{\mu}_{G}^{*} + \boldsymbol{\mu}_{G}^{*})p_{G}(\mathbf{x})d\mathbf{x} \\ &= \boldsymbol{\mu}_{G}^{*} + \int_G (\mathbf{x} - \boldsymbol{\mu}_{G}^{*})[p_{G}(\boldsymbol{\mu}_{G}^{*}) + O(\|G\|)]d\mathbf{x} = \boldsymbol{\mu}_{G}^{*} + O(\|G\|^3).\end{aligned}$$

Hence the integral in the second term on the right hand side of (S2.5) is

$$\begin{aligned}&\int_G [(\mathbf{x} - \boldsymbol{\mu}_{G}^{*})(\mathbf{x} - \boldsymbol{\mu}_{G}^{*})^T\ddot{h}(y \mid \boldsymbol{\xi})(\mathbf{x} - \boldsymbol{\mu}_{G}^{*}) + O(\|G\|^5)][p_{G}(\boldsymbol{\mu}_{G}^{*}) + O(\|G\|)]d\mathbf{x} \\ &= \int_G [(\mathbf{x} - \boldsymbol{\mu}_{G}^{*})(\mathbf{x} - \boldsymbol{\mu}_{G}^{*})^T\ddot{h}(y \mid \boldsymbol{\xi})(\mathbf{x} - \boldsymbol{\mu}_{G}^{*})]p_{G}(\boldsymbol{\mu}_{G}^{*})d\mathbf{x} + O(\|G\|^5)\end{aligned}$$

However, the leading term on the right is also of the order $O(\|G\|^5)$, because

$$\begin{aligned}&\int_G [(\mathbf{x}-\boldsymbol{\mu}^*_G)(\mathbf{x}-\boldsymbol{\mu}^*_G)^T \ddot{h}(y \mid \boldsymbol{\xi})(\mathbf{x}-\boldsymbol{\mu}^*_G)] p_G(\boldsymbol{\mu}^*_G) d\mathbf{x} \\ &= p_G(\boldsymbol{\mu}^*_G) \int_G (\mathbf{x}-\boldsymbol{\mu}^*_G)(\mathbf{x}-\boldsymbol{\mu}^*_G)^T \ddot{h}(y \mid \boldsymbol{\mu}_G)(\mathbf{x}-\boldsymbol{\mu}^*_G) d\mathbf{x} + O(\|G\|^5),\end{aligned}$$

where the first term is 0 since G is an open ball. The rest of the proof is to the argument following (S2.5). □

Bibliography

Yin, X., B. Li, and R. D. Cook (2008). Successive direction extraction for estimating the central subspace in a multiple-index regression. *Journal of Multivariate Analysis 99*(8), 1733–1757.

Zhu, Y. and P. Zeng (2006). Fourier methods for estimating the central subspace and the central mean subspace in regression. *Journal of the American Statistical Association 101*, 1638–1651.