Local Variable Selection and Parameter Estimation of Spatially Varying Coefficient Regression Models

Wesley Brooks

1. Asymptotics

1.1. Consistency

Theorem 1.1. *If* $h\sqrt{n}a_n \xrightarrow{p} 0$ *then* $\hat{\boldsymbol{\beta}}(\boldsymbol{s}) - \boldsymbol{\beta}(\boldsymbol{s}) - \frac{\kappa_2 h^2}{2\kappa_0}\{\boldsymbol{\beta}_{uu}(\boldsymbol{s}) + \boldsymbol{\beta}_{vv}(\boldsymbol{s})\} = O_p(n^{-1/2}h^{-1})$

Proof. The idea of the proof is to show that the objective being minimized achieves a unique minimum, which must be $\hat{\boldsymbol{\beta}}(\boldsymbol{s})$.

The order of convergence is $hn^{1/2}$ where $h = O(n^{-1/6})$.

To show: that for any ϵ, there is a sufficiently large constant C such that

$$\liminf_n P\left[\inf_{u \in \mathcal{R}: \|u\| \leqslant C} Q\left\{\boldsymbol{\beta}(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\right\} > Q\left\{\boldsymbol{\beta}(\boldsymbol{s})\right\}\right] > 1 - \epsilon$$

We show the result:

$$
\begin{aligned}
V_4^{(n)}(\boldsymbol{u}) &= Q\left\{\boldsymbol{\beta}(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\right\} - Q\left\{\boldsymbol{\beta}(\boldsymbol{s})\right\} \\
&= (1/2)\left[\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\left\{\boldsymbol{\beta}(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\right\}\right]^T \boldsymbol{W}(\boldsymbol{s})\left[\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\left\{\boldsymbol{\beta}(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\right\}\right] \\
&\quad + n\sum_{j=1}^{p}\lambda_j\|\boldsymbol{\beta}(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\| \\
&\quad - (1/2)\left\{\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\boldsymbol{\beta}(\boldsymbol{s})\right\}^T \boldsymbol{W}(\boldsymbol{s})\left\{\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\boldsymbol{\beta}(\boldsymbol{s})\right\} - n\sum_{j=1}^{p}\lambda_j\|\boldsymbol{\beta}(\boldsymbol{s})\| \\
&= (1/2)\boldsymbol{u}^T\left\{h^{-2}n^{-1}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{Z}(\boldsymbol{s})\right\}\boldsymbol{u} - \boldsymbol{u}^T\left[h^{-1}n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\left\{\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\boldsymbol{\beta}(\boldsymbol{s})\right\}\right] \\
&\quad + n\sum_{j=1}^{p}\lambda_j\|\boldsymbol{\beta}_j(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\| - n\sum_{j=1}^{p}\lambda_j\|\boldsymbol{\beta}_j(\boldsymbol{s})\| \\
&= (1/2)\boldsymbol{u}^T\left\{h^{-2}n^{-1}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{Z}(\boldsymbol{s})\right\}\boldsymbol{u} - \boldsymbol{u}^T\left[h^{-1}n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\left\{\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s})\boldsymbol{\beta}(\boldsymbol{s})\right\}\right] \\
&\quad + n\sum_{j=1}^{p}\lambda_j\|\boldsymbol{\beta}_j(\boldsymbol{s}) + h^{-1}n^{-1/2}\boldsymbol{u}\| - n\sum_{j=1}^{p_0}\lambda_j\|\boldsymbol{\beta}_j(\boldsymbol{s})\|
\end{aligned}
\tag{1}
$$

□

The final quantity in (1) is the sum of a quadratic term, a linear term, and a penalty term. We'll consider the terms of the sum in (1) separately.

Quadratic term.. By Lemma 2 of **?**, $\frac{1}{n}\boldsymbol{Z}^T(\boldsymbol{s}_i)\boldsymbol{W}(\boldsymbol{s}_i)\boldsymbol{Z}(\boldsymbol{s}_i) \xrightarrow{p} \Omega$, so the first term in (1) converges to $h^2\boldsymbol{u}^T\Omega\boldsymbol{u}$.

Linear term.. By a first-order Taylor expansion, we have that $\boldsymbol{\beta}(\boldsymbol{s}_i) = \boldsymbol{\beta}(\boldsymbol{s}) + \nabla\boldsymbol{\beta}(\boldsymbol{\xi}_i)(\boldsymbol{s}_i - \boldsymbol{s})$ where $\xi_i = \boldsymbol{s} + \theta(\boldsymbol{s}_i - \boldsymbol{s})$ and $\theta \in [0,1]$ for $i = 1, \ldots, n$. So

$$\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{s}_i)\boldsymbol{\beta}(\boldsymbol{s}_i) = \boldsymbol{m} + \boldsymbol{\varepsilon} - \boldsymbol{Z}(\boldsymbol{s}_i)\boldsymbol{\beta}(\boldsymbol{s}_i)$$

and so the linear term of (1) is

$$\boldsymbol{u}^T \left[n^{-1/2} \boldsymbol{Z}^T(\boldsymbol{s}) \boldsymbol{W}(\boldsymbol{s}) \left\{ \boldsymbol{m} + \boldsymbol{\varepsilon} - \boldsymbol{Z}(\boldsymbol{s}) \boldsymbol{\beta}(\boldsymbol{s}) \right\} \right]. \tag{2}$$

We wish to show that (2) is $O_p(1)$. Now, taking the three terms of the sum separately (derivations are in the appendix):

1.2. Third term

The first term is

$$h\, n^{-1/2} \boldsymbol{Z}^T(\boldsymbol{s}) \boldsymbol{W}(\boldsymbol{s}) \boldsymbol{Z}(\boldsymbol{s}) \boldsymbol{\beta}(\boldsymbol{s}) \tag{3}$$

The expectation of (3) is:

$$\begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0}_{p \times 2p} \\ \mathbf{0}_{2p \times p} & \mathbf{0}_{2p \times 2p} \end{pmatrix} \gamma(\boldsymbol{s}) f(\boldsymbol{s})$$

And the variance of (3) is:

$$\begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0}_{p \times 2p} \\ \mathbf{0}_{2p \times p} & \mathbf{0}_{2p \times 2p} \end{pmatrix} \gamma(\boldsymbol{s}) f(\boldsymbol{s})$$

2. Definitions

Let $\boldsymbol{A}$ be a matrix. Then $\{\boldsymbol{A}\}_j$ is the jth column as a column vector, and $\{\boldsymbol{A}\}_k^T$ be the kth row as a row vector.

3. Appendix: lemmas

Lemma 3.1. *If V is a random, symmetric $m \times m$ matrix with*

$$E(\{V\}_j) = \mu_j$$
$$E\left(\{V\}_j \{V\}_{j'}^T\right) - E(\{V\}_j)\, E\left(\{V\}_{j'}^T\right) = \Sigma_{jj'}$$

for $j, j' = 1, \ldots, m$, while U is a fixed $m \times m$ matrix, then

$$E\left(\{V\}_i^T U \{V\}_k\right) = \mu_i^T U \mu_k + tr(U\Sigma_{ik})$$

4. Appendix: proofs

Lemma 4.1. *The expectation and variance of*

$$n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{m} \tag{4}$$

are, respecively:

$$\begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}) f(\boldsymbol{s})$$

and:

$$\begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}) f(\boldsymbol{s})$$

Proof. Find the expectation and variance of the ith term in the sum $n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{m}$:

$$\begin{aligned}
&\{\boldsymbol{Z}^T(\boldsymbol{s})\}_i \; \{\boldsymbol{W}(\boldsymbol{s})\}_{ii} \; m(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s} - \boldsymbol{s}_i\|) \;\; \{\boldsymbol{Z}^T(\boldsymbol{s})\}_i \; \{\boldsymbol{Z}(\boldsymbol{s}_i)\}_i^T \; \boldsymbol{\gamma}(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s} - \boldsymbol{s}_i\|) \begin{pmatrix} X_1^2(\boldsymbol{s}_i) & \ldots & X_1(\boldsymbol{s}_i)X_p(\boldsymbol{s}_i) & \boldsymbol{0}_{1\times 2p} \\ \vdots & \ddots & \vdots & \vdots \\ X_1(\boldsymbol{s}_i)X_p(\boldsymbol{s}_i) & \ldots & X_p^2(\boldsymbol{s}_i) & \boldsymbol{0}_{1\times 2p} \\ \boldsymbol{0}_{2p\times 1} & \ldots & \boldsymbol{0}_{2p\times 1} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s} - \boldsymbol{s}_i\|) \begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i) & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}_i) \qquad (5)
\end{aligned}$$

So the expectation is:

$$\begin{aligned}
&\mathrm{E}[\;\{\boldsymbol{Z}^T(\boldsymbol{s})\}_i \; \{\boldsymbol{W}(\boldsymbol{s})\}_{ii} \; \{\boldsymbol{Z}(\boldsymbol{s}_i)\}_i^T \; \boldsymbol{\gamma}(\boldsymbol{s}_i)] \\
&\quad = \mathrm{E} \begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i) & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \int K_h(\|\boldsymbol{s} - \boldsymbol{t}\|)\boldsymbol{\gamma}(\boldsymbol{t}) f(\boldsymbol{t}) \partial \boldsymbol{t} \\
&\quad = \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}) f(\boldsymbol{s}) \qquad (6)
\end{aligned}$$

And the variance is:

$$
\begin{aligned}
&\mathrm{E}[\ \{\boldsymbol{Z}^T(\boldsymbol{s})\}_i\ \{\boldsymbol{W}(\boldsymbol{s})\}_{ii}\ \{\boldsymbol{Z}(\boldsymbol{s}_i)\}_i^T\ \boldsymbol{\gamma}(\boldsymbol{s}_i)\boldsymbol{\gamma}^T(\boldsymbol{s}_i)\ \{\boldsymbol{Z}^T(\boldsymbol{s}_i)\}_i\ \{\boldsymbol{W}(\boldsymbol{s})\}_{ii}\ \{\boldsymbol{Z}(\boldsymbol{s})\}_i^T\ ] \\
&\quad = \mathrm{E}\begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i) & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \int K_h(\|\boldsymbol{s}-\boldsymbol{t}\|)\boldsymbol{\gamma}(\boldsymbol{t})f(\boldsymbol{t})\partial\boldsymbol{t} \\
&\quad = \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{0}_{p\times 2p} \\ \boldsymbol{0}_{2p\times p} & \boldsymbol{0}_{2p\times 2p} \end{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s})f(\boldsymbol{s}) \qquad (7)
\end{aligned}
$$

□

4.1. Second term

The second term is

$$h\ n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{\varepsilon} \qquad (8)$$

The expectation of (3) is 0 and the variance of (3) is:

$$
\begin{pmatrix}
\boldsymbol{\Psi} & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_1 & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_2 \\
\boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_1 & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_1^2 & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_1(\boldsymbol{s}_i-\boldsymbol{s})_2 \\
\boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_2 & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_1(\boldsymbol{s}_i-\boldsymbol{s})_2 & \boldsymbol{\Psi}(\boldsymbol{s}_i-\boldsymbol{s})_2^2
\end{pmatrix} \times \sigma^2 \times f(\boldsymbol{s}) \times \nu_0
$$

Second term. Find the expectation and variance of the ith term in the sum $n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{Z}(\boldsymbol{s})\boldsymbol{\gamma}(\boldsymbol{s})$:

Third term. Find the expectation and variance of the ith term in the sum $n^{-1/2}\boldsymbol{Z}^T(\boldsymbol{s})\boldsymbol{W}(\boldsymbol{s})\boldsymbol{\varepsilon}$:

$$
\begin{aligned}
&\{\boldsymbol{Z}^T(\boldsymbol{s})\}_i \; \{\boldsymbol{W}(\boldsymbol{s})\}_{ii} \; \varepsilon(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s}-\boldsymbol{s}_i\|) \;\; \{\boldsymbol{Z}^T(\boldsymbol{s})\}_i \; \varepsilon(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s}-\boldsymbol{s}_i\|) \begin{pmatrix} X_1(\boldsymbol{s}_i) \\ \vdots \\ X_p(\boldsymbol{s}_i) \\ X_1(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_1 \\ \vdots \\ X_p(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_1 \\ X_1(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_2 \\ \vdots \\ X_p(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_2 \end{pmatrix} \varepsilon(\boldsymbol{s}_i) \\
&= K_h(\|\boldsymbol{s}-\boldsymbol{s}_i\|) \begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i) \\ \boldsymbol{X}(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_1 \\ \boldsymbol{X}(\boldsymbol{s}_i)(\boldsymbol{s}_i-\boldsymbol{s})_2 \end{pmatrix} \varepsilon(\boldsymbol{s}_i) \qquad (9)
\end{aligned}
$$

So the expectation is:

$$
\begin{aligned}
&\mathrm{E}[\ \{\boldsymbol{Z}^T(\boldsymbol{s})\}_i\ \{\boldsymbol{W}(\boldsymbol{s})\}_{ii}\ \varepsilon(\boldsymbol{s}_i)] \\
&= \mathrm{E}\begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i) \\ \boldsymbol{X}(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1 \\ \boldsymbol{X}(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_2 \end{pmatrix} \mathrm{E}\ \varepsilon(\boldsymbol{s}_i)\ \int K_h(\|\boldsymbol{s}-\boldsymbol{t}\|) f(\boldsymbol{t})\partial\boldsymbol{t} \\
&= \begin{pmatrix} \boldsymbol{\mu}(\boldsymbol{s}_i) \\ \boldsymbol{\mu}(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1 \\ \boldsymbol{\mu}(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_2 \end{pmatrix} \times 0 \times f(\boldsymbol{s}) \\
&= 0 \qquad (10)
\end{aligned}
$$

And the variance is:

$$
\begin{aligned}
&\mathrm{E}[\ \{\boldsymbol{Z}^T(\boldsymbol{s})\}_i\ \{\boldsymbol{W}(\boldsymbol{s})\}_{ii}\ \varepsilon^2(\boldsymbol{s}_i)\ \{\boldsymbol{W}(\boldsymbol{s})\}_{ii}\ \{\boldsymbol{Z}(\boldsymbol{s})\}_i^T\ ] \\
&= \mathrm{E}\begin{pmatrix} \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i) & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1 & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_2 \\ \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1 & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1^2 & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1(\boldsymbol{s}_i - \boldsymbol{s})_2 \\ \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_2 & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_1(\boldsymbol{s}_i - \boldsymbol{s})_2 & \boldsymbol{X}(\boldsymbol{s}_i)\boldsymbol{X}^T(\boldsymbol{s}_i)(\boldsymbol{s}_i - \boldsymbol{s})_2^2 \end{pmatrix} \\
&\quad \times E\ \varepsilon^2(\boldsymbol{s}_i)\ \int K_h^2(\|\boldsymbol{s}-\boldsymbol{t}\|) f(\boldsymbol{t})\partial\boldsymbol{t} \\
&= \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_1 & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_2 \\ \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_1 & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_1^2 & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_1(\boldsymbol{s}_i - \boldsymbol{s})_2 \\ \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_2 & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_1(\boldsymbol{s}_i - \boldsymbol{s})_2 & \boldsymbol{\Psi}(\boldsymbol{s}_i - \boldsymbol{s})_2^2 \end{pmatrix} \times \sigma^2 \times h^{-2} f(\boldsymbol{s})\nu_0 \\
&= 0 \qquad (11)
\end{aligned}
$$

Third term.. By assumption, $p_0\sqrt{n}a_n = O(\sqrt{n}a_n) = o_p(1)$.

So the quadratic term dominates the sum, implying that the difference $Q\left\{\boldsymbol{\beta}(\boldsymbol{s}_i)+n^{-1/2}\boldsymbol{u}\right\} > Q\left\{\boldsymbol{\beta}(\boldsymbol{s}_i)\right\}$ is positive, which proves the result.

5. References