

# Web-based Supplementary Material for “Local Adaptive Grouped Regularization and its Oracle Properties for Varying Coefficient Regression”

Wesley Brooks<sup>a</sup>, Jun Zhu<sup>b</sup>, Zudi Lu<sup>c</sup>

<sup>a</sup>*Department of Statistics, University of Wisconsin, Madison, WI 53706*

<sup>b</sup>*Department of Statistics and Department of Entomology, University of Wisconsin, Madison, WI 53706*

<sup>c</sup>*School of Mathematical Sciences, The University of Southampton, Highfield, Southampton UK*

---

---

## 1. Lemmas

### Lemma 1.

$$E \left[ \sum_{i=1}^n q_1 (\mathbf{Z}_i^T \boldsymbol{\zeta}(\mathbf{s}), Y_i) \mathbf{Z}_i K_h(\|\mathbf{s} - \mathbf{s}_i\|) \right] = \begin{pmatrix} 2^{-1} n^{1/2} h^3 f(\mathbf{s}) \kappa_2 \boldsymbol{\Gamma}(\mathbf{s}) (\nabla_{uu}^2 \boldsymbol{\beta}(\mathbf{s}) + \nabla_{vv}^2 \boldsymbol{\beta}(\mathbf{s}))^T \\ \mathbf{0}_{2p} \end{pmatrix} + o_p(h^2 \mathbf{1}_{3p})$$

and

$$\begin{aligned} \text{Var} \left[ \sum_{i=1}^n q_1 (\mathbf{Z}_i^T \boldsymbol{\zeta}(\mathbf{s}), Y_i) \mathbf{Z}_i K_h(\|\mathbf{s} - \mathbf{s}_i\|) \right] &= f(\mathbf{s}) \text{diag} \{ \nu_0, \nu_2, \nu_2 \} \otimes \boldsymbol{\Gamma}(\mathbf{s}) + o(1) \\ &= \Lambda + o(1) \end{aligned}$$

---

*Email addresses:* wrbrooks@uwalumni.com (Wesley Brooks), jzhu@stat.wisc.edu (Jun Zhu), Z.Lu@soton.ac.uk (Zudi Lu)

*Proof. Expectation:* For  $j = 1, \dots, p$ , by a Taylor expansion of  $\beta_j(\mathbf{s}_i)$  around  $\mathbf{s}$ ,

$$\beta_j(\mathbf{s}_i) = \beta_j(\mathbf{s}) + \nabla \beta_j(\mathbf{s})(\mathbf{s}_i - \mathbf{s}) + (\mathbf{s}_i - \mathbf{s})^T \{ \nabla^2 \beta_j(\mathbf{s}) \} (\mathbf{s}_i - \mathbf{s}) + o(h^2)$$

and thus, for  $\mathbf{x} \in \mathbb{R}^p$ ,

$$\mathbf{x}_i^T \boldsymbol{\beta}(\mathbf{s}_i) = \sum_{j=1}^p x_{ij} \left[ \beta_j(\mathbf{s}) + \nabla \beta_j(\mathbf{s})^T (\mathbf{s}_i - \mathbf{s}) + \tilde{\beta}_{ij}'' \right] + o(h^2).$$

Letting  $\mathbf{z}_i^T = \{(1, s_{i,1} - s_1, s_{i,2} - s_2) \otimes \mathbf{x}_i^T\}$  and  $\boldsymbol{\zeta}(\mathbf{s}) = (\boldsymbol{\beta}(\mathbf{s})^T, \nabla_u \boldsymbol{\beta}(\mathbf{s})^T, \nabla_v \boldsymbol{\beta}(\mathbf{s})^T)^T$ , we have that

$$\begin{aligned} \mathbf{x}_i^T \boldsymbol{\beta}(\mathbf{s}_i) - \mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}) &= \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_i'' + o(h^2) \\ &= O_p(h^2). \end{aligned}$$

By a Taylor expansion around  $\mathbf{x}^T \boldsymbol{\beta}(\mathbf{s}_i)$ , then,

$$\begin{aligned} q_1(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), \mu(\mathbf{s}_i, \mathbf{z}_i)) &= q_1(\mathbf{x}_i^T \boldsymbol{\beta}(\mathbf{s}_i), \mu(\mathbf{s}_i, \mathbf{z})) \\ &\quad - q_2(\mathbf{x}_i^T \boldsymbol{\beta}(\mathbf{s}_i), \mu(\mathbf{s}_i, \mathbf{z})) \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_i'' \\ &\quad + o(h^2). \end{aligned}$$

And by the definitions of  $q_1(\cdot)$  and  $q_2(\cdot)$ , we have that

$$q_1(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), \mu(\mathbf{s}_i, \mathbf{z}_i)) = \rho(\mathbf{s}_i, \mathbf{z}_i) \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_i'' + o(h^2).$$

Now the expectation of  $\Omega_n$  is

$$\begin{aligned} nE(\omega_i | \mathbf{Z}_i = \mathbf{z}_i, \mathbf{s}_i) &= (1/2) \alpha_n \mathbf{z}_i q_1(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), \mu(\mathbf{s}_i, \mathbf{z}_i)) K(h^{-1} \|\mathbf{s} - \mathbf{s}_i\|) \\ &= (1/2) \alpha_n h^2 \mathbf{z}_i \left\{ h^{-2} \rho(\mathbf{s}_i, \mathbf{z}_i) \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_i'' + o(\mathbf{1}_{3p}) \right\} K(h^{-1} \|\mathbf{s} - \mathbf{s}_i\|). \end{aligned}$$

To facilitate a change of variables, we observe that  $h^{-2}\tilde{\beta}_j'' = \left(\frac{\mathbf{s}_i - \mathbf{s}}{h}\right)^T \{\nabla^2 \beta_j(\mathbf{s})\} \left(\frac{\mathbf{s}_i - \mathbf{s}}{h}\right)$ . Thus,

$$E(\omega_i | \mathbf{s}_i) = (1/2) \alpha_n h^2 \left[ \begin{pmatrix} 1 \\ h^{-1}(s_{i,1} - s_1) \\ h^{-1}(s_{i,2} - s_2) \end{pmatrix} \otimes \left\{ \Gamma(\mathbf{s}_i) h^{-2} \tilde{\beta}_i'' \right\} + o(\mathbf{1}_{3p}) \right] K(h^{-1} \|\mathbf{s} - \mathbf{s}_i\|).$$

And, using the symmetry of the kernel function,

$$E(\omega_i) = (1/2) \alpha_n h^4 f(\mathbf{s}) \begin{pmatrix} \kappa_2 \\ h\kappa_3 \\ h\kappa_3 \end{pmatrix} \otimes [\Gamma(\mathbf{s}) \{ \nabla_{uu}^2 \boldsymbol{\beta}(\mathbf{s}) + \nabla_{vv}^2 \boldsymbol{\beta}(\mathbf{s}) \}] + o(h^2 \mathbf{1}_{3p})$$

where  $\{ \nabla_{uu}^2 \boldsymbol{\beta}(\mathbf{s}) + \nabla_{vv}^2 \boldsymbol{\beta}(\mathbf{s}) \} = (\nabla_{uu}^2 \beta_1(\mathbf{s}) + \nabla_{vv}^2 \beta_1(\mathbf{s}), \dots, \nabla_{uu}^2 \beta_p(\mathbf{s}) + \nabla_{vv}^2 \beta_p(\mathbf{s}))^T$ . Thus,

$$E(\Omega_n) = \begin{pmatrix} \alpha_n^{-1} 2^{-1} h^2 \kappa_2 f(\mathbf{s}) \Gamma(\mathbf{s}) (\nabla_{uu}^2 \boldsymbol{\beta}(\mathbf{s}) + \nabla_{vv}^2 \boldsymbol{\beta}(\mathbf{s}))^T \\ \mathbf{0}_{2p} \end{pmatrix} + o_p(h^2 \mathbf{1}_{3p})$$

**Variance:** By the previous result,  $E(\Omega_n) = O(h^2)$ . Thus,  $\text{var}(\Omega_n) \rightarrow E(\Omega_n^2)$ , and since the observations are independent,  $E(\Omega_n^2) = \sum_{i=1}^n E(\omega_i^2)$ . And, by Taylor expansion around  $\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i)$ ,

$$\begin{aligned} q_1^2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), Y_i) &= q_1^2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i), Y_i) \\ &\quad - q_1(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i), Y_i) q_2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i), Y_i) \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_i'' \\ &\quad + o(h^2). \end{aligned}$$

Since  $q_1(\cdot, \cdot)$  is the quasi-score function, it follows that

$$E(\omega_i^2 | \mathbf{Z}_i = \mathbf{z}_i, \mathbf{s}_i) = \alpha_n^2 \rho(\mathbf{s}_i, \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^T K(h^{-1} \|\mathbf{s} - \mathbf{s}_i\|) + o(h^2).$$

By the symmetry of the kernel function,

$$E(\omega_i^2) = n^{-1} f(\mathbf{s}) \text{diag}\{\nu_0, \nu_2, \nu_2\} \otimes \mathbf{\Gamma}(\mathbf{s}) + o(1).$$

Thus,

$$\text{Var}(\Omega_n) = f(\mathbf{s}) \text{diag}\{\nu_0, \nu_2, \nu_2\} \otimes \mathbf{\Gamma}(\mathbf{s}) + o(1).$$

□

**Lemma 2.**

$$\begin{aligned} E \left[ \sum_{i=1}^n q_2(\mathbf{Z}_i^T \boldsymbol{\zeta}(\mathbf{s}), Y_i) \mathbf{Z}_i \mathbf{Z}_i^T K_h(\|\mathbf{s} - \mathbf{s}_i\|) \right] &= -f(\mathbf{s}) \text{diag}\{\kappa_0, \kappa_2, \kappa_2\} \otimes \mathbf{\Gamma}(\mathbf{s}) + o(1) \\ &= -\Delta + o(1) \end{aligned}$$

and

$$\text{Var} \left\{ \left( \sum_{i=1}^n q_2(\mathbf{Z}_i^T \boldsymbol{\zeta}(\mathbf{s}), Y_i) \mathbf{Z}_i \mathbf{Z}_i^T K_h(\|\mathbf{s} - \mathbf{s}_i\|) \right)_{ij} \right\} = O(n^{-1} h^{-2})$$

*Proof. Expectation:* The approach is similar to the proof of Lemma 1. By the Taylor expansion of  $q_2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), \mu(\mathbf{s}_i, \mathbf{z}_i))$  around  $\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i)$ :

$$\begin{aligned} q_2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}), \mu(\mathbf{s}_i, \mathbf{z}_i)) &= q_2(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i), \mu(\mathbf{s}_i, \mathbf{z}_i)) + q_3(\mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i), \mu(\mathbf{s}_i, \mathbf{z}_i)) \{ \mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}) - \mathbf{z}_i^T \boldsymbol{\zeta}(\mathbf{s}_i) \} \\ &= -\rho(\mathbf{s}_i, \mathbf{z}_i) + o(1). \end{aligned}$$

And by the same arguments as before

$$\begin{aligned}
E(\delta_i | \mathbf{Z}_i = \mathbf{z}_i, \mathbf{s}_i) &= -\alpha_n^2 \rho(\mathbf{s}_i, \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^T K(h^{-1} \|\mathbf{s}_i - \mathbf{s}\|) \\
E(\delta_i | \mathbf{s}_i) &= -\alpha_n^2 \begin{pmatrix} 1 \\ h^{-1}(s_{i,1} - s_1) \\ h^{-1}(s_{i,2} - s_2) \end{pmatrix} \begin{pmatrix} 1 \\ h^{-1}(s_{i,1} - s_1) \\ h^{-1}(s_{i,2} - s_2) \end{pmatrix}^T \otimes \boldsymbol{\Gamma}(\mathbf{s}_i) K(h^{-1} \|\mathbf{s}_i - \mathbf{s}\|) \\
E(\delta_i) &= -nf(\mathbf{s}) \text{diag}\{\kappa_0, \kappa_2, \kappa_2\} \otimes \boldsymbol{\Gamma}(\mathbf{s}) + o(n^{-1})
\end{aligned}$$

Thus,

$$E(\Delta_n) = -f(\mathbf{s}) \text{diag}\{\kappa_0, \kappa_2, \kappa_2\} \otimes \boldsymbol{\Gamma}(\mathbf{s}) + o(1)$$

**Variance:** From the previous result, it follows that  $\{E(\delta_i)\}^2 = O(n^{-2})$ . By the definition of  $\delta_i$ ,

$$\begin{aligned}
E(\delta_i^2 | \mathbf{Z}_i = \mathbf{z}_i, \mathbf{s}_i) &= \\
&\alpha_n^4 \mathbf{z}_i^T \mathbf{z}_i q_2^2(\mathbf{s}_i, \mathbf{z}_i) \begin{pmatrix} 1 \\ h^{-1}(s_{i,1} - s_1) \\ h^{-1}(s_{i,2} - s_2) \end{pmatrix} \begin{pmatrix} 1 \\ h^{-1}(s_{i,1} - s_1) \\ h^{-1}(s_{i,2} - s_2) \end{pmatrix}^T \mathbf{z}_i \mathbf{z}_i^T K^2(h^{-1} \|\mathbf{s}_i - \mathbf{s}\|) + o(1)
\end{aligned}$$

And it follows that  $E(\delta_i^2) = O(n^{-1} \alpha_n^2)$ , and  $\text{Var}(\Delta_n) = O(\alpha_n^2)$ . □