Appendix B: Bias and Variance Derivations

** Note that all results here are for fixed bandwidths $\left(b_\mu \text{ and } b_\sigma\right)$ and fixed mixing parameters $\left(\lambda_\mu \text{ and } \lambda_\sigma\right)$.

Appendix B.1: Parametric Dual Modeling

Consider the underlying means model $\mathbf{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{g}^{1/2}(\mathbf{z})\boldsymbol{\varepsilon}$ with $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$. We are assuming that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where $\mathbf{V} = \mathrm{diag}\left\langle \sigma_1^2, \cdots, \sigma_n^2 \right\rangle$. The variance is given by the model $\boldsymbol{\sigma}^2 = \boldsymbol{g}(\boldsymbol{z}) = \exp\{\mathbf{Z}\boldsymbol{\theta}\} + \mathbf{l}$. In parametric dual modeling the means model fits are obtained by estimated weighted least squares and can be written in matrix notation as $\hat{\mathbf{y}}^{(\mathrm{ewls})} = \mathbf{H}^{(\mathrm{ewls})}\mathbf{y}$ where $\mathbf{H}^{(\mathrm{ewls})} = \mathbf{X}\left(\mathbf{X}'\hat{\mathbf{V}}^{-1(\mathrm{glm})}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1(\mathrm{glm})}$ and $\hat{\mathbf{V}}^{-1(\mathrm{glm})} = \mathrm{diag}\left\langle \hat{\boldsymbol{\sigma}}_1^{2(\mathrm{glm})}, \cdots, \hat{\boldsymbol{\sigma}}_n^{2(\mathrm{glm})} \right\rangle$. The estimated variances are given by $\hat{\sigma}_i^{2(\mathrm{glm})} = \exp\left(\mathbf{z}_i'\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right)$ where $\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}$ is obtained by regressing $\mathbf{e}^{2(\mathrm{ewls})}$ on the specified variance function $\exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\}$.

The bias expression for $\hat{\mathbf{y}}^{(\mathrm{ewls})}$ is derived as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{y}}^{(\mathrm{ewls})}\right) &= \mathrm{E}\left(\hat{\mathbf{y}}^{(\mathrm{ewls})}\right) - \mathrm{E}(\mathbf{y}) \\
&= \mathrm{E}\left(\mathbf{H}_o^{(\mathrm{ewls})}\mathbf{y}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&\text{where } \mathbf{H}_o^{(\mathrm{ewls})} = \mathbf{X}\left(\mathbf{X}'\mathbf{V}_o^{-1(\mathrm{glm})}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}_o^{-1(\mathrm{glm})} \\
&= \mathbf{H}_o^{(\mathrm{ewls})}\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&= -\left(\mathbf{I} - \mathbf{H}_o^{(\mathrm{ewls})}\right)\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}\right) \blacksquare
\end{aligned}
$$

Regarding the notation above, the matrix $\mathbf{V}_o^{(\mathrm{glm})}$, which is a component of $\mathbf{H}_o^{(\mathrm{ewls})}$, is given by: $\mathbf{V}_o^{(\mathrm{glm})} = \mathrm{diag}\left(\sigma_1^{2\,(\mathrm{glm})}, \cdots, \sigma_n^{2\,(\mathrm{glm})}\right)$ where $\sigma_i^{2\,(\mathrm{glm})} = \exp\left\{\mathbf{z}_i'\boldsymbol{\theta}^{(\mathrm{glm})}\right\}$ and $\boldsymbol{\theta}^{(\mathrm{glm})}$ is obtained by regressing $\mathrm{E}\left(\mathbf{e}^{2(\mathrm{ewls})}\right)$ on $\exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\}$.

The derivation of the variance of $\hat{\mathbf{y}}^{(\mathrm{ewls})}$ proceeds as follows:

$$\begin{aligned}\mathrm{Var}\left(\hat{\mathbf{y}}^{(\mathrm{ewls})}\right) &= \mathrm{Var}\left[\mathbf{H}_{\mathrm{o}}^{(\mathrm{ewls})}\mathbf{y}\right] \\ &= \mathbf{H}_{\mathrm{o}}^{(\mathrm{ewls})}\mathrm{Var}(\mathbf{y})\mathbf{H}_{\mathrm{o}}^{(\mathrm{ewls})} \\ &= \mathbf{H}_{\mathrm{o}}^{(\mathrm{ewls})}\mathbf{V}\,\mathbf{H}_{\mathrm{o}}^{(\mathrm{ewls})} \blacksquare\end{aligned}$$

It is important to note that $\mathbf{V} = \mathrm{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right)$ where σ_i^2 is taken to be $\exp\left\{\mathbf{z}_i^{'}\boldsymbol{\theta}\right\} + l_i$.

The bias expression for $\hat{\boldsymbol{\sigma}}^{2(\mathrm{glm})}$ is given as follows:

$$\begin{aligned}\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{glm})}\right) &= \mathrm{E}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{glm})}\right) - \mathrm{E}\left(\boldsymbol{\sigma}^2\right) \\ &= \mathrm{E}\left[\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}\right] - \exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} - \mathbf{l}\end{aligned}$$

Approximating $\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}$ by (A.1.10) in Appendix A.1, we can write the approximate bias as:

$$\begin{aligned}\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{glm})}\right) &\approx \mathrm{E}\left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} + \mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2\,(\mathrm{ewls})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\right] \\ &\quad - \exp\{\mathbf{Z}\boldsymbol{\theta}\} - \mathbf{l} \\ &= \left\{\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} - \mathbf{l}\right]\right\} \\ &\quad + \mathbf{H}^{(\mathrm{glm})}\mathrm{E}\left(\mathbf{e}^{2\,(\mathrm{ewls})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) \\ &= \left\{\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} - \mathbf{l}\right]\right\} \\ &\quad + \mathbf{H}^{(\mathrm{glm})}\left(\mathrm{E}\left(\mathbf{e}^{2\,(\mathrm{ewls})}\right) - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\end{aligned}$$

$$\text{where } \mathrm{E}\left(\mathbf{e}^{2(\mathrm{ewls})}\right) = \begin{bmatrix}\mathrm{E}\left(e_1^{2(\mathrm{ewls})}\right) \\ \vdots \\ \mathrm{E}\left(e_n^{2(\mathrm{ewls})}\right)\end{bmatrix}.$$

From Appendix B.2 we have that $\mathrm{E}\left(e_i^{2(\mathrm{ewls})}\right)$ is given by:

$$E\left(e_i^{2(\text{ewls})}\right) = \left[\left(1 - h_{ii}^{(\text{ewls})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\text{ewls})} \sigma_j^2\right] +$$

$$\left[\left(1 - h_{ii}^{(\text{ewls})}\right) f_i \;-\; \sum_{j \neq i}^{n} h_{ji}^{(\text{ewls})} f_j\right]^2 \blacksquare$$

In deriving the variance expression for $\hat{\sigma}^{2\,(\text{glm})}$ recall that

$$\text{Var}\left(\hat{\sigma}^{2\,(\text{glm})}\right) \;=\; \text{Var}\left[\exp\left\{\mathbf{Z}\hat{\theta}^{(\text{glm})}\right\}\right].$$

Approximating $\exp\left\{\mathbf{Z}\hat{\theta}^{(\text{glm})}\right\}$ by (A.1.10) in Appendix A.1, we can write the approximate variance as:

$$\text{Var}\left(\hat{\sigma}^{2\,(\text{glm})}\right) \approx \text{Var}\left[\exp\left\{\mathbf{Z}\theta^{(\text{glm})}\right\} \;+\; \mathbf{H}^{(\text{glm})}\left(\mathbf{e}^{2(\text{ewls})} \;-\; \exp\left\{\mathbf{Z}\theta^{(\text{glm})}\right\}\right)\right]$$

Since $\exp\left\{\mathbf{Z}\theta^{(\text{glm})}\right\}$ is a constant, we have that $\text{Var}\left(\hat{\sigma}^{2\,(\text{glm})}\right)$ is approximated as:

$$\text{Var}\left(\hat{\sigma}^{2\,(\text{glm})}\right) \approx \mathbf{H}^{(\text{glm})} \mathbf{V}_{e^{2\,(\text{ewls})}} \mathbf{H}^{(\text{glm})'}$$

where $\mathbf{V}_{e^{2\,(\text{ewls})}}$ is assumed to be a diagonal matrix with the $\text{Var}\left(e_i^{2(\text{ewls})}\right)$ as the i^{th} diagonal element. From Appendix C.1 we have that:

$$\text{Var}\left(e_i^{2(\text{ewls})}\right) = \;2\left[\left(1 - h_{ii}^{(\text{ewls})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\text{ewls})} \sigma_j^2\right]^2 +$$

$$4\left\{\left[\left(1 - h_{ii}^{(\text{ewls})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\text{ewls})} \sigma_j^2\right]\left[\left(1 - h_{ii}^{(\text{ewls})}\right) f_i \;-\; \sum_{j \neq i}^{n} h_{ji}^{(\text{ewls})} f_j\right]^2\right\} \blacksquare$$

Appendix B.2 Nonparametric, Difference-Based Dual Modeling

Consider the underlying means model $\mathbf{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{g}^{1/2}(\mathbf{z})\boldsymbol{\varepsilon}$ with $\boldsymbol{m}(\boldsymbol{x}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$. We are assuming that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where $\mathbf{V} = \mathrm{diag}\left\langle \sigma_1^2, \cdots, \sigma_n^2 \right\rangle$. The variance is given by the model $\boldsymbol{\sigma}^2 = \boldsymbol{g}(\boldsymbol{z}) = \exp\{\mathbf{Z}\boldsymbol{\theta}\} + \mathbf{l}$. In non-parametric, difference-based dual modeling, the means model fits are obtained by local linear regression and are written in matrix notation as $\hat{\mathbf{y}}^{(\mathrm{llr})} = \mathbf{H}^{(\mathrm{llr})}\mathbf{y}$ where $\mathbf{H}^{(\mathrm{llr})} = \begin{bmatrix} \mathbf{h}_1^{'(\mathrm{llr})} \\ \vdots \\ \mathbf{h}_n^{'(\mathrm{llr})} \end{bmatrix}$ and $\mathbf{h}_i^{'(\mathrm{llr})} = \mathbf{x}_\mathrm{i}^{'}\left(\mathbf{X}^{'}\mathbf{W}^{(\mathrm{llr})}(x_i)\mathbf{X}\right)^{-1}\mathbf{X}^{'}\mathbf{W}^{(\mathrm{llr})}(x_i)$. Recall from Section 2.B.4 that $\mathbf{W}^{(\mathrm{llr})}(x_i)$ is a diagonal matrix consisting of the kernel weights associated with x_i. The estimated variances are given by $\hat{\boldsymbol{\sigma}}^{2(\mathrm{diff})} = \mathbf{H}_{b_{\tilde{e}}}^{(\mathrm{llr})}\tilde{\mathbf{e}}^{2(\mathrm{pseud})}$ where $\tilde{\mathbf{e}}^{2(\mathrm{pseud})}$ is the n x 1 vector of squared pseudo residuals and $\mathbf{H}_{b_{\tilde{e}}}^{(\mathrm{llr})}$ is the local linear hat matrix used to smooth the set of squared pseudo residuals. The form of $\mathbf{H}_{b_{\tilde{e}}}^{(\mathrm{llr})}$ is the same as that of $\mathbf{H}^{(\mathrm{llr})}$ above but the subscript "$b_{\tilde{e}}$" is used to denote that a different bandwidth is used in determining the kernel weights for $\mathbf{H}_{b_{\tilde{e}}}^{(\mathrm{llr})}$ than is used for the kernel weights of $\mathbf{H}^{(\mathrm{llr})}$. The bias and variance expressions for $\hat{\mathbf{y}}^{(\mathrm{llr})}$ and $\hat{\boldsymbol{\sigma}}^{2(\mathrm{diff})}$ are as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{y}}^{(\mathrm{llr})}\right) &= \mathrm{E}\left(\hat{\mathbf{y}}^{(\mathrm{llr})}\right) - \mathrm{E}(\mathbf{y}) \\
&= \mathrm{E}\left(\mathbf{H}^{(\mathrm{llr})}\mathbf{y}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&= \mathbf{H}^{(\mathrm{llr})}(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&= -\left(\mathbf{I} - \mathbf{H}^{(\mathrm{llr})}\right)(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}) \ \blacksquare
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}\left(\hat{\mathbf{y}}^{(\mathrm{llr})}\right) &= \mathrm{Var}\left[\mathbf{H}^{(\mathrm{llr})}\mathbf{y}\right] \\
&= \mathbf{H}^{(\mathrm{llr})}\mathrm{Var}(\mathbf{y})\mathbf{H}^{(\mathrm{llr})} \\
&= \mathbf{H}^{(\mathrm{llr})}\mathbf{V}\,\mathbf{H}^{(\mathrm{llr})} \\
&\quad \text{where } \mathbf{V} = \mathrm{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right) \ \blacksquare
\end{aligned}
$$

The bias expression for $\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}$ is derived as follows:

$$\begin{aligned}
\text{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}\right) &= \text{E}\left(\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}\right) - \text{E}\left(\boldsymbol{\sigma}^{2}\right) \\
&= \text{E}\left[\mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\tilde{\mathbf{e}}^{2(\text{pseud})}\right] - \exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} - \mathbf{l} \\
&= \mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\text{E}\left(\tilde{\mathbf{e}}^{2(\text{pseud})}\right) - \exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} - \mathbf{l}.
\end{aligned}$$

From Müller and Stadtmüller we have that $\text{E}\left(\tilde{\mathbf{e}}^{2(\text{pseud})}\right) \approx \boldsymbol{\sigma}^{2} = \exp\{\mathbf{Z}\boldsymbol{\theta}\} + \mathbf{l}$. Substituting for $\text{E}\left(\tilde{\mathbf{e}}^{2(\text{pseud})}\right)$ into the bias expression we have:

$$\begin{aligned}
\text{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}\right) &\approx \mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\left(\exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} + \mathbf{l}\right) - \exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} - \mathbf{l} \\
\text{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}\right) &\approx -\left(\mathbf{I} - \mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\right)\left(\exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} + \mathbf{l}\right) \blacksquare
\end{aligned}$$

The variance expression for $\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}$ is derived as follows:

$$\begin{aligned}
\text{Var}\left(\hat{\boldsymbol{\sigma}}^{2\,(\text{diff})}\right) &= \text{Var}\left[\mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\tilde{\mathbf{e}}^{2(\text{pseud})}\right] \\
&= \mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\text{Var}\left(\tilde{\mathbf{e}}^{2(\text{pseud})}\right)\mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})'} \\
&= \mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})}\,\mathbf{V}_{\tilde{e}}\,\mathbf{H}_{b_{\tilde{e}}}^{(\text{llr})'}
\end{aligned}$$

where $\mathbf{V}_{\tilde{e}}$ denotes the dispersion matrix of the nx1 vector of squared pseudo residuals. ■

Appendix B.3: Dual Model Robust Regression

Consider the underlying means model $\mathbf{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{g}^{1/2}(\mathbf{z})\boldsymbol{\varepsilon}$ with $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$. We are assuming that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where $\mathbf{V} = \mathrm{diag}\left\langle \sigma_1^2, \cdots, \sigma_n^2 \right\rangle$. The variance is given by the model $\boldsymbol{\sigma}^2 = \boldsymbol{g}(\boldsymbol{z}) = \exp\left\{ \mathbf{Z}\boldsymbol{\theta} \right\} + \mathbf{l}$. The MMRR fitted values are $\hat{\mathbf{y}}^{(\mathrm{mmrr})} = \mathbf{H}^{(\mathrm{mmrr})}\mathbf{y} = \left[\mathbf{H}^{(\mathrm{ewls})} + \lambda_\mu \mathbf{H}_{b_\mu}^{(\mathrm{llr})}(\mathbf{I} - \mathbf{H}^{(\mathrm{ewls})}) \right]\mathbf{y}$, where $\mathbf{H}^{(\mathrm{ewls})}$ is the hat matrix used in parametrically fitting the mean and $\mathbf{H}_{b_\mu}^{(\mathrm{llr})}$ is the local linear hat matrix used in the local linear fit to the residuals from the parametric (EWLS) means fit. The VMRR fitted values are written as $\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})} = \exp\left\{ \mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})} \right\} + \lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{ \mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{mmrr})} \right\} \right)$ where $\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}$ is the local linear hat matrix used in the local linear fit to the residuals from the parametric (GLIM) fit. The bias and variance expressions for $\hat{\mathbf{y}}^{(\mathrm{mmrr})}$ and $\hat{\boldsymbol{\sigma}}^{2(\mathrm{vmrr})}$ are then as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{y}}^{(\mathrm{mmrr})}\right) &= \mathrm{E}\left(\hat{\mathbf{y}}^{(\mathrm{mmrr})}\right) - \mathrm{E}(\mathbf{y}) \\
&= \mathrm{E}\left(\mathbf{H}^{(\mathrm{mmrr})}\mathbf{y}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&= \mathbf{H}^{(\mathrm{mmrr})}\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{f} \\
&= -\left(\mathbf{I} - \mathbf{H}^{(\mathrm{mmrr})}\right)\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}\right) \blacksquare
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}\left(\hat{\mathbf{y}}^{(\mathrm{mmrr})}\right) &= \mathrm{Var}\left[\mathbf{H}^{(\mathrm{mmrr})}\mathbf{y}\right] \\
&= \mathbf{H}^{(\mathrm{mmrr})}\mathrm{Var}(\mathbf{y})\mathbf{H}^{'\,(\mathrm{mmrr})} \\
&= \mathbf{H}^{(\mathrm{mmrr})}\mathbf{V}\,\mathbf{H}^{'\,(\mathrm{mmrr})} \blacksquare
\end{aligned}
$$

It is important to note that $\mathbf{V} = \mathrm{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right)$ where σ_i^2 is written as $\exp\left\{\mathbf{z}_i^{'}\boldsymbol{\theta}\right\} + l_i$. The matrix $\mathbf{V}_\mathrm{o}$, however, which is a component of $\mathbf{H}^{(\mathrm{ewls})}$ in $\mathbf{H}^{(\mathrm{mmrr})}$, is given by : $\mathbf{V}_\mathrm{o} = \mathrm{diag}\left(\sigma_1^{2\,(\mathrm{vmrr})}, \cdots, \sigma_n^{2\,(\mathrm{vmrr})}\right)$ where

$$
\sigma_i^{2\,(\mathrm{vmrr})} = \exp\left\{\mathbf{z}_i^{'}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} + \lambda_\sigma \mathbf{h}_{i,b_\sigma}^{'(\mathrm{llr})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)
$$

and $\mathbf{h}_{i,b_\sigma}^{'(\mathrm{llr})}$ is the i^{th} row of $\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}$. The value of $\boldsymbol{\theta}^{(\mathrm{glm})}$ is obtained by regressing $\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})}\right)$ on $\exp\left\{\mathbf{z}_i^{'}\boldsymbol{\theta}\right\}$.

The expression for the bias of $\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}$ is derived as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) &= \mathrm{E}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) - \mathrm{E}\left(\boldsymbol{\sigma}^2\right) \\
&= \mathrm{E}\Big[\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\} + \\
&\qquad \lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}\right)\Big] \\
&\qquad - \exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} - \mathbf{l}.
\end{aligned}
$$

Approximating $\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}$ by (A.1.10) in Appendix A.1, the bias expression can now be written as

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) \approx \mathrm{E}\Big[&\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} + \mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) + \\
&\lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) - \\
&\lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\Big] \\
&- \exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} - \mathbf{l}.
\end{aligned}
$$

Grouping terms we can write

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) \approx{} &\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} + \mathbf{l}\right] + \\
&\mathbf{H}^{(\mathrm{glm})}\mathrm{E}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) + \\
&\lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) - \\
&\lambda_\sigma \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathbf{H}^{(\mathrm{glm})}\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)
\end{aligned}
$$

$$
= \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} + \mathbf{l}\right] +
$$

$$\mathbf{H}^{(glm)}\mathrm{E}\left(\mathbf{e}^{2\,(mmrr)} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right) +$$
$$\lambda_\sigma \mathbf{H}_{b_\sigma}^{(llr)}\,\mathrm{E}\left[\left(\mathbf{I} - \mathbf{H}^{(glm)}\right)\left(\mathbf{e}^{2(mmrr)} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right)\right]$$

$$= \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} + \mathbf{l}\right] +$$
$$\mathbf{H}^{(glm)}\mathrm{E}\left(\mathbf{e}^{2\,(mmrr)} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right) +$$
$$\lambda_\sigma \mathbf{H}_{b_\sigma}^{(llr)}\left(\mathbf{I} - \mathbf{H}^{(glm)}\right)\mathrm{E}\left(\mathbf{e}^{2(mmrr)} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right)$$

$$= \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} + \mathbf{l}\right] +$$
$$\mathbf{H}^{(vmrr)}\mathrm{E}\left(\mathbf{e}^{2(mmrr)} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right)$$
$$\text{where } \mathbf{H}^{(vmrr)} = \mathbf{H}^{(glm)} + \boldsymbol{\lambda}_{\boldsymbol{\sigma}}\mathbf{H}_{b_\sigma}^{(llr)}\left(\mathbf{I} - \mathbf{H}^{(glm)}\right).$$

Substituting for $\mathbf{H}^{(vmrr)}$ we have

$$\mathrm{Bias}\left(\hat{\sigma}^{2\,(vmrr)}\right) \approx \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\} - \left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}\right\} + \mathbf{l}\right] +$$
$$\mathbf{H}^{(vmrr)}\left(\mathrm{E}\left(\mathbf{e}^{2(mmrr)}\right) - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(glm)}\right\}\right)$$
$$\text{where } \mathrm{E}\left(\mathbf{e}^{2(mmrr)}\right) = \begin{bmatrix} \mathrm{E}\left(\mathrm{e}_1^{2(mmrr)}\right) \\ \vdots \\ \mathrm{E}\left(\mathrm{e}_n^{2(mmrr)}\right) \end{bmatrix}.$$

From Appendix C.1 we have that the expected value of the i^{th}, squared means model residual is given as:

$$\mathrm{E}\left(\mathrm{e}_i^{2(mmrr)}\right) = \left[\left(1 - \mathrm{h}_{ii}^{(mmrr)}\right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{2(mmrr)}\sigma_j^2\right] +$$
$$\left[\left(1 - \mathrm{h}_{ii}^{(mmrr)}\right)\mathrm{f}_i - \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{(mmrr)}\mathrm{f}_j\right]^2 \blacksquare$$

The derivation of the variance of $\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}$ proceeds as follows:

$$\mathrm{Var}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) = \mathrm{Var}\left[\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\} + \lambda_{\sigma}\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\ \mathbf{r}_{\sigma}\right]$$
$$= \mathrm{Var}\left[\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\} + \lambda_{\sigma}\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}\right)\right].$$

Approximating $\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}$ by (A.1.10) in Appendix A.1, the bias expression can now be written as

$$\mathrm{Var}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{vmrr})}\right) \approx \mathrm{Var}\left[\exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} + \mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) + \lambda_{\sigma}\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right) - \lambda_{\sigma}\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\,\mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\right]$$

$$= \mathrm{Var}\left[\mathbf{H}^{(\mathrm{vmrr})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\right]$$

$$= \mathbf{H}^{(\mathrm{vmrr})}\mathbf{V}_{\mathrm{e}^{2\,(\mathrm{mmrr})}}\mathbf{H}^{(\mathrm{vmrr})'}$$

where $\mathbf{V}_{\mathrm{e}^{2\,(\mathrm{mmrr})}}$ is assumed to be a diagonal matrix with the $\mathrm{Var}\left(\mathrm{e}_i^{2(\mathrm{mmrr})}\right)$ as the i^{th} diagonal element. Recall from Appendix C.1 that:

$$\mathrm{Var}\left(\mathrm{e}_i^{2(\mathrm{mmrr})}\right) = 2\left[\left(1-\mathrm{h}_{ii}^{(\mathrm{mmrr})}\right)^2\sigma_i^2 + \sum_{j\neq i}^{n}\mathrm{h}_{ji}^{2(\mathrm{mmrr})}\sigma_j^2\right]^2 +$$

$$4\left\{\left[\left(1-\mathrm{h}_{ii}^{(\mathrm{mmrr})}\right)^2\sigma_i^2 + \sum_{j\neq i}^{n}\mathrm{h}_{ji}^{2(\mathrm{mmrr})}\sigma_j^2\right]\left[\left(1-\mathrm{h}_{ii}^{(\mathrm{mmrr})}\right)\mathrm{f}_i - \sum_{j\neq i}^{n}\mathrm{h}_{ji}^{(\mathrm{mmrr})}\mathrm{f}_j\right]^2\right\} \blacksquare$$

Appendix B.3.1 Bias and Variance of $\mathbf{r}_\mu$

In this section, the bias and variance expressions for determining the optimal bandwidth b_μ for MMRR will be developed. This bandwidth is for the local linear fit to the residuals from the EWLS fit, which may be expressed as $\hat{\mathbf{r}}_\mu = \mathbf{H}_{b_\mu}^{(\text{llr})}\mathbf{r}_\mu$, where $\mathbf{r}_\mu = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^{(\text{ewls})}$ and $\mathbf{y} = \boldsymbol{m}(\boldsymbol{x}) + \boldsymbol{g}^{1/2}(\mathbf{z})\boldsymbol{\varepsilon}$ with $\boldsymbol{m}(\boldsymbol{x}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$. We are assuming that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where $\mathbf{V} = \mathrm{diag}\left\langle \sigma_1^2, \cdots, \sigma_n^2 \right\rangle$. The derivation of the bias of $\hat{\mathbf{r}}_\mu$ is as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{r}}_\mu\right) &= \mathrm{E}\left(\hat{\mathbf{r}}_\mu\right) - \mathrm{E}\left(\mathbf{r}_\mu\right) \\
&= \mathrm{E}\left(\mathbf{H}_{b_\mu}^{(\text{llr})}\mathbf{r}_\mu\right) - \mathrm{E}\left(\mathbf{r}_\mu\right) \\
&= \mathbf{H}_{b_\mu}^{(\text{llr})}\mathrm{E}\left(\mathbf{r}_\mu\right) - \mathrm{E}\left(\mathbf{r}_\mu\right) = -\left(\mathbf{I} - \mathbf{H}_{b_\mu}^{(\text{llr})}\right)\mathrm{E}\left(\mathbf{r}_\mu\right) \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\mu}^{(\text{llr})}\right)\left[\mathrm{E}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^{(\text{ewls})}\right)\right] \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\mu}^{(\text{llr})}\right)\left[\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\mathrm{E}(\mathbf{y})\right] \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\mu}^{(\text{llr})}\right)\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{f}\right) \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\mu}^{(\text{llr})}\right)\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\left(\mathbf{f}\right) \blacksquare
\end{aligned}
$$

The derivation of the variance of $\hat{\mathbf{r}}_\mu$ proceeds as follows:

$$
\begin{aligned}
\mathrm{Var}\left(\hat{\mathbf{r}}_\mu\right) &= \mathrm{Var}\left(\mathbf{H}_{b_\mu}^{(\text{llr})}\mathbf{r}_\mu\right) = \mathbf{H}_{b_\mu}^{(\text{llr})}\,\mathrm{Var}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^{(\text{ewls})}\right)\mathbf{H}_{b_\mu}^{(\text{llr})'} \\
&= \mathbf{H}_{b_\mu}^{(\text{llr})}\,\mathrm{Var}\left[\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\mathbf{y}\right]\mathbf{H}_{b_\mu}^{(\text{llr})'} \\
&= \mathbf{H}_{b_\mu}^{(\text{llr})}\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\mathrm{Var}(\mathbf{y})\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)'\mathbf{H}_{b_\mu}^{(\text{llr})'} \\
&= \mathbf{H}_{b_\mu}^{(\text{llr})}\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)\mathbf{V}\left(\mathbf{I} - \mathbf{H}^{(\text{ewls})}\right)'\mathbf{H}_{b_\mu}^{(\text{llr})'} \blacksquare
\end{aligned}
$$

Appendix B.3.2 Bias and Variance of $\mathbf{r}_\sigma$

In this section, the bias and variance expressions for determining the optimal bandwidth b_σ for VMRR will be developed. This bandwidth is for the local linear fit to the residuals from the parametric variance fit, which may be expressed as $\hat{\mathbf{r}}_\sigma = \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathbf{r}_\sigma$, where $\mathbf{r}_\sigma = \mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{mmrr})}\right\}$. The derivation of the bias of $\hat{\mathbf{r}}_\sigma$ is given by:

$$\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{r}}_\sigma\right) &= \mathrm{E}\left(\hat{\mathbf{r}}_\sigma\right) - \mathrm{E}\left(\mathbf{r}_\sigma\right) \\
&= \mathrm{E}\left[\mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathbf{r}_\sigma\right] - \mathrm{E}\left(\mathbf{r}_\sigma\right) \\
&= \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\mathrm{E}\left(\mathbf{r}_\sigma\right) - \mathrm{E}\left(\mathbf{r}_\sigma\right) \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\right)\mathrm{E}\left(\mathbf{r}_\sigma\right) \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\right)\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}\right).
\end{aligned}$$

Approximating $\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}$ by (A.1.10) in Appendix A.1, the bias expression can now be written as

$$\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{r}}_\sigma^{(\mathrm{vmrr})}\right) \approx -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\right)\mathrm{E}\Big[\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \\
\mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\Big].
\end{aligned}$$

Regrouping terms the bias expression is written as:

$$\begin{aligned}
\mathrm{Bias}\left(\hat{\mathbf{r}}_\sigma^{(\mathrm{vmrr})}\right) &\approx -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\right)\mathrm{E}\left[\left(\mathbf{I} - \mathbf{H}^{(\mathrm{glm})}\right)\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\right] \\
&= -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})}\right)\left(\mathbf{I} - \mathbf{H}^{(\mathrm{glm})}\right)\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)
\end{aligned}$$

$$= -\left(\mathbf{I} - \mathbf{H}_{b_\sigma}^{(\mathrm{llr})} \right)\left(\mathbf{I} - \mathbf{H}^{(\mathrm{glm})} \right)\left(\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})}\right) - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right).$$

From Appendix C.1 we have that $\mathrm{E}\left(\mathbf{e}^{2(\mathrm{mmrr})}\right)$ is an nx1 vector in which the i^{th} element is given by:

$$\mathrm{E}\left(\mathrm{e}_i^{2(\mathrm{mmrr})}\right) = \left[\left(1 - \mathrm{h}_{ii}^{(\mathrm{mmrr})}\right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{2(\mathrm{mmrr})} \sigma_j^2 \right] +$$

$$\left[\left(1 - \mathrm{h}_{ii}^{(\mathrm{mmrr})}\right)\mathrm{f}_i - \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{(\mathrm{mmrr})} \mathrm{f}_j \right] \blacksquare$$

The derivation of the variance of $\hat{\mathbf{r}}_\sigma$ is now given as follows:

$$\mathrm{Var}\left(\hat{\mathbf{r}}_\sigma\right) = \mathrm{Var}\left[\mathbf{H}_{b_\sigma}^{(\mathrm{llr})} \mathbf{r}_\sigma\right]$$

$$= \mathbf{H}_{b_\sigma}^{(\mathrm{llr})} \mathrm{Var}\left(\mathbf{r}_\sigma\right) \mathbf{H}_{b_\sigma}^{'\,(\mathrm{llr})}$$

$$= \mathbf{H}_{b_\sigma}^{(\mathrm{llr})} \mathrm{Var}\left[\mathbf{e}^{2\,(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}\right] \mathbf{H}_{b_\sigma}^{(\mathrm{llr})'}$$

Approximating $\exp\left\{\mathbf{Z}\hat{\boldsymbol{\theta}}^{(\mathrm{glm})}\right\}$ by (A.1.10) in Appendix A.1, the variance expression can now be written as

$$\mathrm{Var}\left(\hat{\mathbf{r}}_\sigma\right) \approx \mathbf{H}_{b_\sigma}^{(\mathrm{llr})} \mathrm{Var}\left[\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\} - \right.$$

$$\left. \mathbf{H}^{(\mathrm{glm})}\left(\mathbf{e}^{2(\mathrm{mmrr})} - \exp\left\{\mathbf{Z}\boldsymbol{\theta}^{(\mathrm{glm})}\right\}\right)\right] \mathbf{H}_{b_\sigma}^{(\mathrm{llr})'}$$

$$= \mathbf{H}_{b_\sigma}^{(\text{llr})} \operatorname{Var}\left[\left(\mathbf{I} - \mathbf{H}^{(\text{glm})}\right)\left(\mathbf{e}^{2\,(\text{mmrr})} - \exp\left\{\mathbf{Z\theta}^{(\text{glm})}\right\}\right)\right] \mathbf{H}_{b_\sigma}^{(\text{llr})'}$$

$$= \mathbf{H}_{b_\sigma}^{(\text{llr})} \left(\mathbf{I} - \mathbf{H}^{(\text{glm})}\right) \operatorname{Var}\left(\mathbf{e}^{2(\text{mmrr})}\right) \left(\mathbf{I} - \mathbf{H}^{(\text{glm})}\right)' \mathbf{H}_{b_\sigma}^{(\text{llr})'}$$

$$= \mathbf{H}_{b_\sigma}^{(\text{llr})} \left(\mathbf{I} - \mathbf{H}^{(\text{glm})}\right) \mathbf{V}_{\mathbf{e}^{2\,(\text{mmrr})}} \left(\mathbf{I} - \mathbf{H}^{(\text{glm})}\right)' \mathbf{H}_{b_\sigma}^{(\text{llr})'} ,$$

where $\mathbf{V}_{\mathbf{e}^{2\,(\text{mmrr})}} = \operatorname{diag}\left\langle \left[\operatorname{Var}\left(\mathrm{e}_i^{2(\text{mmrr})}\right)\right] \right\rangle$ and $\operatorname{Var}\left(\mathrm{e}_i^{2(\text{mmrr})}\right)$ was given in Appendix C.1 by:

$$\operatorname{Var}\left(\mathrm{e}_i^{2(\text{mmrr})}\right) = 2\left[\left(1 - \mathrm{h}_{ii}^{(\text{mmrr})}\right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{2(\text{mmrr})} \sigma_j^2\right]^2 +$$

$$4\left\{\left[\left(1 - \mathrm{h}_{ii}^{(\text{mmrr})}\right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{2(\text{mmrr})} \sigma_j^2\right]\left[\left(1 - \mathrm{h}_{ii}^{(\text{mmrr})}\right)\mathrm{f}_i - \sum_{j \neq i}^{n} \mathrm{h}_{ji}^{(\text{mmrr})} \mathrm{f}_j\right]^2\right\} \blacksquare$$

Appendix B.4: Nonparametric, Residual-Based Dual Modeling

Consider the underlying means model $\mathbf{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{g}^{1/2}(\mathbf{z})\boldsymbol{\varepsilon}$ with $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$. We are assuming that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where $\mathbf{V} = \mathrm{diag}\left\langle \sigma_1^2, \cdots, \sigma_n^2 \right\rangle$. The variance is given by the model $\boldsymbol{\sigma}^2 = \boldsymbol{g}(\boldsymbol{z}) = \exp\{\mathbf{Z\theta}\} + \mathbf{l}$. In non-parametric, residual-based dual modeling, the means model fits are obtained by local linear regression and are written in matrix notation as $\hat{\mathbf{y}}^{(\mathrm{llr})} = \mathbf{H}_b^{(\mathrm{llr})}\mathbf{y}$ where $\mathbf{H}_b^{(\mathrm{llr})} = \begin{bmatrix} \mathbf{h}_1^{'(\mathrm{llr})} \\ \vdots \\ \mathbf{h}_n^{'(\mathrm{llr})} \end{bmatrix}$ and $\mathbf{h}_i^{'(\mathrm{llr})} = \mathbf{x}_{\mathrm{i}}'\left(\mathbf{X}'\mathbf{W}^{(\mathrm{llr})}(x_i)\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}^{(\mathrm{llr})}(x_i)$. Recall from Section 2.B.4 that $\mathbf{W}^{(\mathrm{llr})}(x_i)$ is a diagonal matrix consisting of the kernel weights associated with x_i. Notice that this is the same estimate of the mean as the estimate provided by nonparametric, difference-based estimation. For this reason, the bias and variance expressions for the residual-based mean estimate can be referred to those developed in Appendix A:2. The estimated variances are given by $\hat{\boldsymbol{\sigma}}^{2(\mathrm{res})} = \mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}\mathbf{e}^{2(\mathrm{llr})}$ where $\mathbf{e}^{2(\mathrm{res})}$ is the n x 1 vector of local linear squared residuals (residuals resulting from local linear fit to the mean). The matrix $\mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}$ is the local linear hat matrix used to smooth the set of local linear squared residuals. The form of $\mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}$ is the same as that of $\mathbf{H}_b^{(\mathrm{llr})}$ above but the subscript "b_e" is used to denote that a different bandwidth is used in determining the kernel weights for $\mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}$ than is used for the kernel weights of $\mathbf{H}_b^{(\mathrm{llr})}$. The derivation of the bias of $\hat{\boldsymbol{\sigma}}^{2(\mathrm{res})}$ is as follows:

$$
\begin{aligned}
\mathrm{Bias}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{res})}\right) &= \mathrm{E}\left(\hat{\boldsymbol{\sigma}}^{2\,(\mathrm{res})}\right) - \mathrm{E}\left(\boldsymbol{\sigma}^2\right) \\
&= \mathrm{E}\left[\mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}\mathbf{e}^{2(\mathrm{llr})}\right] - \exp\left\{\mathbf{Z\theta}\right\} - \mathbf{l} \\
&= \mathbf{H}_{b_\mathrm{e}}^{(\mathrm{llr})}\mathrm{E}\left(\mathbf{e}^{2(\mathrm{llr})}\right) - \exp\left\{\mathbf{Z\theta}\right\} - \mathbf{l}
\end{aligned}
$$

$$
\text{where } \mathrm{E}\left(\mathbf{e}^{2(\mathrm{llr})}\right) = \begin{bmatrix} \mathrm{E}\left(\mathrm{e}_1^{2(\mathrm{llr})}\right) \\ \vdots \\ \mathrm{E}\left(\mathrm{e}_n^{2(\mathrm{llr})}\right) \end{bmatrix},
$$

and from Appendix C.2 we have that $\mathrm{E}\left(e_i^{2(\mathrm{llr})}\right)$ is given by:

$$\mathrm{E}\left(e_i^{2(\mathrm{llr})}\right) = \left[\left(1 - h_{ii}^{(\mathrm{llr})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\mathrm{llr})} \sigma_j^2\right] +$$

$$\left[\left(1 \;-\; h_{ii}^{(\mathrm{llr})}\right)\left(\mathbf{x}_i^{'}\boldsymbol{\beta} \;+\; f_i\right) \;-\; \sum_{j \neq i}^{n} h_{ji}^{(\mathrm{llr})}\left(\mathbf{x}_j^{'}\boldsymbol{\beta} \;+\; f_j\right)\right]^2 \blacksquare$$

The derivation of the variance of $\hat{\sigma}^{2(\mathrm{res})}$ is as follows:

$$\begin{aligned}\mathrm{Var}\left(\hat{\sigma}^{2\,(\mathrm{res})}\right) &= \mathrm{Var}\left[\mathbf{H}_{b_e}^{(\mathrm{llr})}\mathbf{e}^{2(\mathrm{llr})}\right] \\ &= \mathbf{H}_{b_e}^{(\mathrm{llr})}\mathrm{Var}\left(\mathbf{e}^{2(\mathrm{llr})}\right)\mathbf{H}_{b_e}^{(\mathrm{llr})'} \\ &= \mathbf{H}_{b_e}^{(\mathrm{llr})}\;\mathbf{V}_{e^{2\,(\mathrm{llr})}}\;\mathbf{H}_{b_e}^{(\mathrm{llr})'}\end{aligned}$$

where $\mathbf{V}_{e^{2\,(\mathrm{llr})}}$ is assumed to be a diagonal matrix with the $\mathrm{Var}\left(e_i^{2(\mathrm{llr})}\right)$ as the i^{th} diagonal element. From Appendix C.2 we have that:

$$\mathrm{Var}\left(e_i^{2(\mathrm{llr})}\right) = 2\left[\left(1 - h_{ii}^{(\mathrm{llr})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\mathrm{llr})} \sigma_j^2\right]^2 +$$

$$4\left\{\left[\left(1 - h_{ii}^{(\mathrm{mmrr})}\right)^2 \sigma_i^2 \;+\; \sum_{j \neq i}^{n} h_{ji}^{2(\mathrm{mmrr})} \sigma_j^2\right] \mathrm{x}\right.$$

$$\left.\left[\left(1 \;-\; h_{ii}^{(\mathrm{llr})}\right)\left(\mathbf{x}_i^{'}\boldsymbol{\beta} \;+\; f_i\right) \;-\; \sum_{j \neq i}^{n} h_{ji}^{(\mathrm{llr})}\left(\mathbf{x}_j^{'}\boldsymbol{\beta} \;+\; f_j\right)\right]^2 \right\} \blacksquare$$