Appendix B: Bias and Variance Derivations

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

Appendix B.1: Parametric Dual Modeling

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

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

\[
\text{Bias}\left(\hat{y}^{(ewls)}\right) = \text{E}\left(\hat{y}^{(ewls)}\right) - \text{E}(y) \\
= \text{E}\left(H_{o}^{(ewls)}y\right) - X\beta - f \\
\text{where } H_{o}^{(ewls)} = X(X'V_{o}^{-1}(glm)X)^{-1}X'V_{o}^{-1}(glm) \\
= H_{o}^{(ewls)}(X\beta + f) - X\beta - f \\
= -(I - H_{o}^{(ewls)})(X\beta + f)\ast
\]

Regarding the notation above, the matrix \(V_{o}^{(glm)}\), which is a component of \(H_{o}^{(ewls)}\), is given by: \(V_{o}^{(glm)} = \text{diag}(\sigma_{1}^{2}(glm), \cdots, \sigma_{n}^{2}(glm))\) where \(\sigma_{i}^{2}(glm) = \exp\{z'\hat{\theta}(glm)\}\) and \(\hat{\theta}(glm)\) is obtained by regressing \(E(e^{2(ewls)})\) on \(\exp\{Z\theta\}\).

The derivation of the variance of \(\hat{y}^{(ewls)}\) proceeds as follows:
\[
\text{Var}(\hat{y}^{(ewls)}) = \text{Var}[H_0^{(ewls)}y]
\]
\[
= H_0^{(ewls)} \text{Var}(y) H_0^{(ewls)}
\]
\[
= H_0^{(ewls)} V H_0^{(ewls)}
\]

It is important to note that \( V = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2) \) where \( \sigma_i^2 \) is taken to be \( \exp\{z_i \theta\} + l_i \).

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

\[
\text{Bias} \left( \hat{\sigma}^{2(glm)} \right) = E\left( \hat{\sigma}^{2(glm)} \right) - E(\sigma^2)
\]
\[
= E\left[ \exp\left\{ Z\theta^{(glm)} \right\} \right] - \exp\{Z\theta\} - 1
\]

Approximating \( \exp\{Z\theta^{(glm)}\} \) by (A.1.10) in Appendix A.1, we can write the approximate bias as:

\[
\text{Bias} \left( \hat{\sigma}^{2(glm)} \right) \approx E\left[ \exp\left\{ Z\theta^{(glm)} \right\} + H^{(glm)} \left( e^{2(ewls)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) \right]
\]
\[
- \exp\{Z\theta\} - 1
\]
\[
= \left\{ \exp\left\{ Z\theta^{(glm)} \right\} - \left[ \exp\left\{ Z\theta \right\} - 1 \right] \right\}
\]
\[
+ H^{(glm)} \left( E\left( e^{2(ewls)} \right) - \exp\left\{ Z\theta^{(glm)} \right\} \right)
\]
\[
= \left\{ \exp\left\{ Z\theta^{(glm)} \right\} - \left[ \exp\left\{ Z\theta \right\} - 1 \right] \right\}
\]
\[
+ H^{(glm)} \left[ E\left( e_1^{2(ewls)} \right), \ldots, E\left( e_n^{2(ewls)} \right) \right]
\]

where \( E\left( e_i^{2(ewls)} \right) \) is given by:

From Appendix B.2 we have that \( E\left( e_i^{2(ewls)} \right) \) is given by:
\[
E(e_i^{2(ewls)}) = \left(1 - h_{ii}^{(ewls)}\right)^2 \sigma_i^2 + \sum_{j \neq i}^n h_{ji}^{2(ewls)} \sigma_j^2 + \left(1 - h_{ii}^{(ewls)}\right)f_i - \sum_{j \neq i}^n h_{ji}^{(ewls)} f_j \right)^2.
\]

In deriving the variance expression for \( \hat{\sigma}^2 \text{(glm)} \) recall that
\[
\text{Var}(\hat{\sigma}^2 \text{(glm)}) = \text{Var}\left[\exp\left(\mathbf{Z}\hat{\theta}^{\text{(glm)}}\right)\right].
\]

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:
\[
\text{Var}(\hat{\sigma}^2 \text{(glm)}) = \text{Var}\left[\exp\left(\mathbf{Z}\hat{\theta}^{\text{(glm)}}\right)\right] + \mathbf{H}^{\text{(glm)}}\left(\mathbf{e}^{2(ewls)} - \exp\left[\mathbf{Z}\hat{\theta}^{\text{(glm)}}\right]\right).
\]

Since \( \exp\left(\mathbf{Z}\hat{\theta}^{\text{(glm)}}\right) \) is a constant, we have that \( \text{Var}(\hat{\sigma}^2 \text{(glm)}) \) is approximated as:
\[
\text{Var}(\hat{\sigma}^2 \text{(glm)}) \approx \mathbf{H}^{\text{(glm)}} \mathbf{V}_{e^{2(ewls)}} \mathbf{H}^{\text{(glm)}}.
\]

where \( \mathbf{V}_{e^{2(ewls)}} \) is assumed to be a diagonal matrix with the \( \text{Var}(e_i^{2(ewls)}) \) as the \( i \)th diagonal element. From Appendix C.1 we have that:
\[
\text{Var}(e_i^{2(ewls)}) = 2 \left(1 - h_{ii}^{(ewls)}\right)^2 \sigma_i^2 + \sum_{j \neq i}^n h_{ji}^{2(ewls)} \sigma_j^2 + 4 \left(1 - h_{ii}^{(ewls)}\right)^2 \sigma_i^2 + \sum_{j \neq i}^n h_{ji}^{2(ewls)} \sigma_j^2 \left(1 - h_{ii}^{(ewls)}\right)f_i - \sum_{j \neq i}^n h_{ji}^{(ewls)} f_j \right)^2.
\]
Appendix B.2 Nonparametric, Difference-Based Dual Modeling

Consider the underlying means model $y = h(x) + g^{1/2}(z)\epsilon$ with $m(x) = X\beta + f$. We are assuming that $E(\epsilon) = 0$ and $\text{Var}(\epsilon) = V$, where $V = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. The variance is given by the model $\sigma^2 = g(z) = \exp\{Z\theta\} + 1$. In non-parametric, difference-based dual modeling, the means model fits are obtained by local linear regression and are written in matrix notation as $\hat{y}^{(llr)} = H^{(llr)}y$ where $H^{(llr)} = \begin{bmatrix} h_1^{(llr)} \\ \vdots \\ h_n^{(llr)} \end{bmatrix}$ and $h_i^{(llr)} = x_i' \left( X' W^{(llr)}(x_i) X \right)^{-1} X' W^{(llr)}(x_i)$. Recall from Section 2.B.4 that $W^{(llr)}(x_i)$ is a diagonal matrix consisting of the kernel weights associated with $x_i$. The estimated variances are given by $\hat{\sigma}^2^{(diff)} = H_{b_{\hat{\epsilon}}}^{(llr)} e^{2(\text{pseud})}$ where $e^{2(\text{pseud})}$ is the $n \times 1$ vector of squared pseudo residuals and $H_{b_{\hat{\epsilon}}}^{(llr)}$ is the local linear hat matrix used to smooth the set of squared pseudo residuals. The form of $H_{b_{\hat{\epsilon}}}^{(llr)}$ is the same as that of $H^{(llr)}$ above but the subscript "$b_{\hat{\epsilon}}"$ is used to denote that a different bandwidth is used in determining the kernel weights for $H_{b_{\hat{\epsilon}}}^{(llr)}$ than is used for the kernel weights of $H^{(llr)}$. The bias and variance expressions for $\hat{y}^{(llr)}$ and $\hat{\sigma}^2^{(diff)}$ are as follows:

$$\text{Bias}\left(\hat{y}^{(llr)}\right) = E\left(\hat{y}^{(llr)}\right) - E\left(y\right)$$

$$= E\left( H^{(llr)} y \right) - X\beta - f$$

$$= H^{(llr)} (X\beta + f) - X\beta - f$$

$$= -(I - H^{(llr)}) (X\beta + f)$$

$$\text{Var}\left(\hat{y}^{(llr)}\right) = \text{Var}\left[ H^{(llr)} y \right]$$

$$= H^{(llr)} \text{Var}(y) H^{(llr)}$$

$$= H^{(llr)} V H^{(llr)}$$

where $V = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. 

119
The bias expression for \( \hat{\sigma}^2 \text{(diff)} \) is derived as follows:

\[
\text{Bias} \left( \hat{\sigma}^2 \text{(diff)} \right) = \mathbb{E} \left( \hat{\sigma}^2 \text{(diff)} \right) - \mathbb{E}(\sigma^2) \\
= \mathbb{E} \left[ H_{b_\bar{e}}^{(llr)} \bar{e}^2 \text{(pseud)} \right] - \exp \left\{ \mathbf{Z} \theta \right\} - 1 \\
= H_{b_\bar{e}}^{(llr)} \mathbb{E}(\bar{e}^2 \text{(pseud)}) - \exp \left\{ \mathbf{Z} \theta \right\} - 1.
\]

From Müller and Stadmüller we have that \( \mathbb{E}(\bar{e}^2 \text{(pseud)}) = \sigma^2 = \exp \left\{ \mathbf{Z} \theta \right\} + 1 \). Substituting for \( \mathbb{E}(\bar{e}^2 \text{(pseud)}) \) into the bias expression we have:

\[
\text{Bias} \left( \hat{\sigma}^2 \text{(diff)} \right) \\
\phantom{= \mathbb{E} \left[ H_{b_\bar{e}}^{(llr)} \bar{e}^2 \text{(pseud)} \right] - \exp \left\{ \mathbf{Z} \theta \right\} - 1} = H_{b_\bar{e}}^{(llr)} \left( \exp \left\{ \mathbf{Z} \theta \right\} + 1 \right) - \exp \left\{ \mathbf{Z} \theta \right\} - 1 \\
\text{Bias} \left( \hat{\sigma}^2 \text{(diff)} \right) = - \left( \mathbf{I} - H_{b_\bar{e}}^{(llr)} \right) \exp \left\{ \mathbf{Z} \theta \right\} - 1^*.
\]

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

\[
\text{Var} \left( \hat{\sigma}^2 \text{(diff)} \right) = \text{Var} \left[ H_{b_\bar{e}}^{(llr)} \bar{e}^2 \text{(pseud)} \right] \\
= H_{b_\bar{e}}^{(llr)} \text{Var} \left( \bar{e}^2 \text{(pseud)} \right) H_{b_\bar{e}}^{(llr)} \\
= H_{b_\bar{e}}^{(llr)} V_{\bar{e}} H_{b_\bar{e}}^{(llr)} \\
\text{where } V_{\bar{e}} \text{ denotes the dispersion matrix of the } n \times 1 \text{ vector of squared pseudo residuals. }^*
Appendix B.3: Dual Model Robust Regression

Consider the underlying means model \( y = h(x) + g^{1/2}(z)\varepsilon \) with \( h(x) = X\beta + f \). We are assuming that \( E(\varepsilon) = 0 \) and \( \text{Var}(\varepsilon) = V \), where \( V = \text{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right) \). The variance is given by the model \( \sigma^2 = g(z) = \exp\left\{Z\theta\right\} + 1 \). The MMRR fitted values are \( \hat{y}^{(mmrr)} = H^{(mmrr)}y = \left[H^{(ewls)} + \lambda_\mu H_{b_\mu}^{(llr)}(I - H^{(ewls)})\right]y \), where \( H^{(ewls)} \) is the hat matrix used in parametrically fitting the mean and \( H_{b_\mu}^{(llr)} \) 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 \( \hat{\sigma}^2^{(vmrr)} = \exp\left\{Z\hat{\theta}^{(glm)}\right\} + \lambda_\sigma H_{b_\sigma}^{(llr)}\left(e^{2(mmm)} - \exp\left\{Z\hat{\theta}^{(mmrr)}\right\}\right) \) where \( H_{b_\sigma}^{(llr)} \) 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 \( \hat{y}^{(mmrr)} \) and \( \hat{\sigma}^2^{(vmrr)} \) are then as follows:

\[
\text{Bias}\left(\hat{y}^{(mmrr)}\right) = E\left(\hat{y}^{(mmrr)}\right) - E(y) = E\left(H^{(mmrr)}y\right) - X\beta - f
\]
\[
= H^{(mmrr)}(X\beta + f) - X\beta - f
\]
\[
= -\left(I - H^{(mmrr)}\right)\left(X\beta + f\right).
\]

\[
\text{Var}\left(\hat{y}^{(mmrr)}\right) = \text{Var}\left[H^{(mmrr)}y\right] = H^{(mmrr)}\text{Var}(y)H^{(mmrr)} = H^{(mmrr)}VH^{(mmrr)}.
\]

It is important to note that \( V = \text{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right) \) where \( \sigma_i^2 = \exp\left\{z_i\theta\right\} + l_i \). The matrix \( V_o \), however, which is a component of \( H^{(ewls)} \) in \( H^{(mmrr)} \), is given by:
\[
V_o = \text{diag}\left(\sigma_1^2^{(vmrr)}, \cdots, \sigma_n^2^{(vmrr)}\right) \text{ where } \sigma_i^2^{(vmrr)} = \exp\left\{z_i\theta^{(glm)}\right\} + \lambda_\sigma h_{i,b_\sigma}^{(llr)}\left(e^{2(mmm)} - \exp\left\{Z\theta^{(glm)}\right\}\right).
\]
and $h_{i,b\sigma}^{(llr)}$ is the $i$th row of $H_{b\sigma}^{(llr)}$. The value of $\theta^{(glm)}$ is obtained by regressing $E\left(e^{2(mmrr)}\right)$ on $\exp\left\{ z_i\theta \right\}$.

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

$$
\text{Bias}\left(\sigma^2^{(vmrr)}\right) = E\left(\sigma^2^{(vmrr)}\right) - E\left(\sigma^2\right) = E\left[ \exp\left\{ Z\theta^{(glm)} \right\} \right] + \lambda_{\sigma} H_{b\sigma}^{(llr)} \left( e^{2(mmrr)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) - \exp\left\{ Z\theta \right\} - 1.
$$

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

$$
\text{Bias}\left(\sigma^2^{(vmrr)}\right) = E\left[ \exp\left\{ Z\theta^{(glm)} \right\} \right] + H^{(glm)} \left( e^{2(mmrr)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) + \lambda_{\sigma} H_{b\sigma}^{(llr)} \left( e^{2(mmrr)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) - \exp\left\{ Z\theta \right\} - 1.
$$

Grouping terms we can write

$$
\text{Bias}\left(\sigma^2^{(vmrr)}\right) = \exp\left\{ Z\theta^{(glm)} \right\} - \left[ \exp\left\{ Z\theta \right\} + 1 \right] + H^{(glm)} \left( e^{2(mmrr)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) + \lambda_{\sigma} H_{b\sigma}^{(llr)} \left( e^{2(mmrr)} - \exp\left\{ Z\theta^{(glm)} \right\} \right) - \exp\left\{ Z\theta \right\} - 1.
$$
\[ \mathbf{H}^{(glm)} E \left( e^{2 \text{ (mmrr)}} - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) + \]

\[ \lambda_\sigma \mathbf{H}_b^{(llr)} E \left[ (\mathbf{I} - \mathbf{H}^{(glm)}) \left( e^{2 \text{ (mmrr)}} - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) \right] \]

\[ = \exp \{ \mathbf{Z} \theta^{(glm)} \} - \left[ \exp \{ \mathbf{Z} \theta \} + 1 \right] + \]

\[ \mathbf{H}^{(glm)} E \left( e^{2 \text{ (mmrr)}} - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) + \]

\[ \lambda_\sigma \mathbf{H}_b^{(llr)} (\mathbf{I} - \mathbf{H}^{(glm)}) E \left( e^{2 \text{ (mmrr)}} - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) \]

\[ = \exp \{ \mathbf{Z} \theta^{(glm)} \} - \left[ \exp \{ \mathbf{Z} \theta \} + 1 \right] + \]

\[ \mathbf{H}^{(vmrr)} E \left( e^{2 \text{ (mmrr)}} - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) \]

where \[ \mathbf{H}^{(vmrr)} = \mathbf{H}^{(glm)} + \lambda_\sigma \mathbf{H}_b^{(llr)} (\mathbf{I} - \mathbf{H}^{(glm)}) \].

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

\[ \text{Bias} (\sigma^2 \text{ (vmrr)}) = \exp \{ \mathbf{Z} \theta^{(glm)} \} - \left[ \exp \{ \mathbf{Z} \theta \} + 1 \right] + \]

\[ \mathbf{H}^{(vmrr)} \left( E \left( e^{2 \text{ (mmrr)}} \right) - \exp \{ \mathbf{Z} \theta^{(glm)} \} \right) \]

where \[ E \left( e^{2 \text{ (mmrr)}} \right) = \begin{bmatrix} E \left( e_1^{2 \text{ (mmrr)}} \right) \\ \vdots \\ E \left( e_n^{2 \text{ (mmrr)}} \right) \end{bmatrix} \].

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

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

\[ \left( \mathbf{1} - h_{ii}^{(mmrr)} \right) f_i - \sum_{j \neq i} h_{ji}^{(mmrr)} f_j \right]^2 \cdot \]
The derivation of the variance of $\sigma^2_{\text{vmrr}}$ proceeds as follows:

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

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

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

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

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

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

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

$$4 \left\{\begin{array}{l}
\left(1 - h_{ii}^{(\text{mmrr})}\right)^2 \sigma_i^2 + \sum_{j \neq i} h_{ji}^{2(\text{mmrr})} \sigma_j^2 \left[\left(1 - h_{ii}^{(\text{mmrr})}\right)f_i - \sum_{j \neq i} h_{ji}^{(\text{mmrr})} f_j\right]^2 \end{array}\right\}.$$
Appendix B.3.1 Bias and Variance of rµ

In this section, the bias and variance expressions for determining the optimal bandwidth bµ 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 \( \hat{r}_\mu = H_{b_\mu}^{(llr)} r_\mu \), where \( r_\mu = y - X\hat{\beta}^{(ewls)} \) and \( y = m(x) + g^{1/2}(z)\epsilon \) with \( m(x) = X\beta + f \). We are assuming that \( E(\epsilon) = 0 \) and \( \text{Var}(\epsilon) = V \), where \( V = \text{diag} \left( \sigma_1^2, \ldots, \sigma_n^2 \right) \). The derivation of the bias of \( \hat{r}_\mu \) is as follows:

\[
\begin{align*}
\text{Bias}(\hat{r}_\mu) &= E(\hat{r}_\mu) - E(r_\mu) \\
&= E(H_{b_\mu}^{(llr)} r_\mu) - E(r_\mu) \\
&= - \left( I - H_{b_\mu}^{(llr)} \right) \left[ E(y - X\hat{\beta}^{(ewls)}) \right] \\
&= - \left( I - H_{b_\mu}^{(llr)} \right) \left[ I - H^{(ewls)} \right] E(y) \\
&= - \left( I - H_{b_\mu}^{(llr)} \right) \left( I - H^{(ewls)} \right) \left( X\beta + f \right) \\
&= - \left( I - H_{b_\mu}^{(llr)} \right) \left( I - H^{(ewls)} \right) f
\end{align*}
\]

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

\[
\begin{align*}
\text{Var}(\hat{r}_\mu) &= \text{Var}(H_{b_\mu}^{(llr)} r_\mu) = H_{b_\mu}^{(llr)} \text{Var}(y - X\hat{\beta}^{(ewls)}) H_{b_\mu}^{(llr)'} \\
&= H_{b_\mu}^{(llr)} \text{Var}(y - X\hat{\beta}^{(ewls)}) H_{b_\mu}^{(llr)'} \\
&= H_{b_\mu}^{(llr)} \left( I - H^{(ewls)} \right) \text{Var}(y) \left( I - H^{(ewls)} \right)' H_{b_\mu}^{(llr)}' \\
&= H_{b_\mu}^{(llr)} \left( I - H^{(ewls)} \right) V \left( I - H^{(ewls)} \right)' H_{b_\mu}^{(llr)}'
\end{align*}
\]
Appendix B.3.2 Bias and Variance of $r_\sigma$

In this section, the bias and variance expressions for determining the optimal bandwidth $b_\sigma$ 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 $\hat{r}_\sigma = H^{(llr)}_{b_\sigma} r_\sigma$, where $r_\sigma = e^{2 \theta^{(mmrr)}} - \exp\{Z\hat{\theta}^{(mmrr)}\}$. The derivation of the bias of $\hat{r}_\sigma$ is given by:

$$
\text{Bias}(\hat{r}_\sigma) = E\left(\hat{r}_\sigma\right) - E\left( r_\sigma \right)
= E\left[H^{(llr)}_{b_\sigma} r_\sigma\right] - E\left( r_\sigma \right)
= H^{(llr)}_{b_\sigma} E\left( r_\sigma \right) - E\left( r_\sigma \right)
= - \left( I - H^{(llr)}_{b_\sigma} \right) E\left( r_\sigma \right)
= - \left( I - H^{(llr)}_{b_\sigma} \right) E\left( e^{2 \theta^{(mmrr)}} - \exp\{Z\hat{\theta}^{(glm)}\} \right).
$$

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

$$
\text{Bias}(\hat{r}_\sigma^{(vmrr)}) = - \left( I - H^{(llr)}_{b_\sigma} \right) E\left( e^{2 \theta^{(mmrr)}} - \exp\{Z\theta^{(glm)}\} \right).
$$

Regrouping terms the bias expression is written as:

$$
\text{Bias}(\hat{r}_\sigma^{(vmrr)}) = - \left( I - H^{(llr)}_{b_\sigma} \right) E\left[ I - H^{(glm)} \right] \left( e^{2 \theta^{(mmrr)}} - \exp\{Z\theta^{(glm)}\} \right) - H^{(glm)} \left( e^{2 \theta^{(mmrr)}} - \exp\{Z\theta^{(glm)}\} \right).
$$
\[
= - \left( \mathbf{I} - \mathbf{H}^{(llr)}_{b_\sigma} \right) \left( \mathbf{I} - \mathbf{H}^{(glm)} \right) \left( \mathbb{E}(e^{2(mmrr)}) - \exp\{\mathbf{Z}\theta^{(glm)}\} \right).
\]

From Appendix C.1 we have that \(\mathbb{E}(e^{2(mmrr)})\) is an \(n\times1\) vector in which the \(i^{th}\) element is given by:

\[
\mathbb{E}(e^{2(mmrr)})_i = \left[ (1 - h_{ii}^{(mmrr)})^2 \sigma_i^2 + \sum_{j \neq i}^n h_{ji}^{2(mmrr)} \sigma_j^2 \right] + \left[ (1 - h_{ii}^{(mmrr)}) f_i - \sum_{j \neq i}^n h_{ji}^{(mmrr)} f_j \right].
\]

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

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

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

\[
= \mathbf{H}^{(llr)}_{b_\sigma} \text{Var}\left[e^{2(mmrr)} - \exp\{\mathbf{Z}\hat{\theta}^{(glm)}\}\right] \mathbf{H}^{(llr)}_{b_\sigma}.
\]

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

\[
\text{Var}\left(\hat{r}_\sigma\right) = \mathbf{H}^{(llr)}_{b_\sigma} \text{Var}\left[e^{2(mmrr)} - \exp\{\mathbf{Z}\theta^{(glm)}\}\right] \mathbf{H}^{(llr)}_{b_\sigma}.
\]

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

\[ = H_{b\sigma}^{(llr)} (I - H^{(glm)}) \text{Var}(e^{2(\text{mmrr})}) (I - H^{(glm)})' H_{b\sigma}^{(llr)} \]

\[ = H_{b\sigma}^{(llr)} (I - H^{(glm)}) V_{e^2(\text{mmrr})} (I - H^{(glm)})' H_{b\sigma}^{(llr)}, \]

where \( V_{e^2(\text{mmrr})} = \text{diag}\left(\text{Var}(e^2_i)\right) \) and \( \text{Var}(e^2_i) \) was given in Appendix C.1 by:

\[
\text{Var}(e^2_i) = 2 \left( 1 - h_{ii}^{(\text{mmrr})} \right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} h_{ji}^{(\text{mmrr})} \sigma_j^2 + 4 \left\{ \left( 1 - h_{ii}^{(\text{mmrr})} \right)^2 \sigma_i^2 + \sum_{j \neq i}^{n} h_{ji}^{(\text{mmrr})} \sigma_j^2 \right\} \left( 1 - h_{ii}^{(\text{mmrr})} \right) f_i - \sum_{j \neq i}^{n} h_{ji}^{(\text{mmrr})} f_j \right\} \]
Appendix B.4: Nonparametric, Residual-Based Dual Modeling

Consider the underlying means model \( y = h(x) + g^{1/2}(z) \varepsilon \) with \( h(x) = X\beta + f \). We are assuming that \( \text{E}(\varepsilon) = 0 \) and \( \text{Var}(\varepsilon) = V \), where \( V = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2) \). The variance is given by the model \( \sigma^2 = g(z) = \exp\{Z\theta\} + 1 \). In non-parametric, residual-based dual modeling, the means model fits are obtained by local linear regression and are written in matrix notation as

\[
\hat{y}^{(llr)} = H_b^{(llr)}y
\]

where

\[
H_b^{(llr)} = \begin{bmatrix}
h_1^{(llr)} \\ \vdots \\ h_n^{(llr)}
\end{bmatrix}
\]

\[
h_i^{(llr)} = x_i'(X'W^{(llr)}(x_i)X)^{-1}X'W^{(llr)}(x_i)
\]

Recall from Section 2.B.4 that \( W^{(llr)}(x_i) \) 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

\[
\hat{\sigma}^2_{(res)} = H_{b_e}^{(llr)}e^{2(llr)}
\]

where \( e^{2(res)} \) is the \( n \times 1 \) vector of local linear squared residuals (residuals resulting from local linear fit to the mean). The matrix \( H_{b_e}^{(llr)} \) is the local linear hat matrix used to smooth the set of local linear squared residuals. The form of \( H_{b_e}^{(llr)} \) is the same as that of \( H_b^{(llr)} \) above but the subscript "\( b_e \)" is used to denote that a different bandwidth is used in determining the kernel weights for \( H_{b_e}^{(llr)} \) than is used for the kernel weights of \( H_b^{(llr)} \). The derivation of the bias of \( \hat{\sigma}^2_{(res)} \) is as follows:

\[
\text{Bias} \left( \hat{\sigma}^2_{(res)} \right) = \text{E} \left( \hat{\sigma}^2_{(res)} \right) - \text{E} \left( \sigma^2 \right)
\]

\[
= \text{E} \left[ H_{b_e}^{(llr)}e^{2(llr)} \right] - \exp\left\{ Z\theta \right\} - 1
\]

\[
= H_{b_e}^{(llr)}\text{E} \left( e^{2(llr)} \right) - \exp\left\{ Z\theta \right\} - 1
\]

where \( \text{E} \left( e^{2(llr)} \right) = \begin{bmatrix} \text{E} \left( e_1^{2(llr)} \right) \\ \vdots \\ \text{E} \left( e_n^{2(llr)} \right) \end{bmatrix} \).
and from Appendix C.2 we have that $E(e_i^{2(llr)})$ is given by:

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

$$\left[ (1 - h_{ii}^{(llr)}) (x_i \beta + f_i) - \sum_{j \neq i}^n h_{ji}^{(llr)} (x_j \beta + f_j) \right]^2.$$

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

$$\text{Var}\left(\hat{\sigma}^2(\text{res})\right) = \text{Var}\left[H_{bc}^{(llr)} e^{2(llr)}\right]$$

$$= H_{bc}^{(llr)} \text{Var}\left(e^{2(llr)}\right) H_{bc}^{(llr)}\text{}$$

$$= H_{bc}^{(llr)} V^{2 (llr)} H_{bc}^{(llr)}\text{}$$

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

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

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

$$\left[ (1 - h_{ii}^{(llr)}) (x_i \beta + f_i) - \sum_{j \neq i}^n h_{ji}^{(llr)} (x_j \beta + f_j) \right]^2.$$