c) $\hat{f}(x) = R_x \hat{a}$

Local Polynomial Regression

$= R_x (R_x^T W R_x)^{-1} R_x^T W Y$

$R_x = \begin{pmatrix} (x_1-x)^0 & \cdots & (x_1-x)^0 \\ \vdots & & \\ (x_n-x)^0 & & (x_n-x)^0 \end{pmatrix}$

$R_x^T W R_x = \begin{pmatrix} \vdots & x_1-x \\ \vdots & \\ 1 & x_n-x \end{pmatrix} \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{pmatrix} \begin{pmatrix} 1 & \cdots & \cdot \\ x_1-x & \cdots & x_n-x \end{pmatrix}$

$R_x = \begin{pmatrix} \vdots & x_1-x \\ \vdots & \\ 1 & x_n-x \end{pmatrix}$

$= \begin{pmatrix} w_1 + \cdots + w_n & \sum w_i (x_i - x) \\ \sum w_i (x_i - x) & \sum w_i (x_i - x)^2 \end{pmatrix}$

$= \frac{1}{h} \begin{pmatrix} s_0 & s_1 \\ s_1 & s_2 \end{pmatrix}$

$(R_x^T W R_x)^{-1} = h \cdot \frac{1}{s_0 s_2 - s_1^2} \begin{pmatrix} s_2 & -s_1 \\ -s_1 & s_0 \end{pmatrix}$

$\begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \end{pmatrix} = \hat{a} = h \cdot \frac{1}{s_0 s_2 - s_1^2} \begin{pmatrix} s_2 & -s_1 \\ -s_1 & s_0 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ x_1-x & \cdots & x_n-x \end{pmatrix} \begin{pmatrix} w_1 y_1 \\ \vdots \\ w_n y_n \end{pmatrix}$

$\hat{a}_0 = h \cdot \frac{1}{s_0 s_2 - s_1^2} \left(\sum w_i s_2 y_i - \sum w_i s_1 y_i (x_i - x) \right)$

$\hat{f}(x) = \frac{\sum w_i^* y_i}{\sum w_i^*} = \sum w_i$

$Y_i = f(x_i) + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2)$

d) $\hat{a} = [\underbrace{R(x)^T K(x) R(x)}_{A}]^{-1} \underbrace{R(x) K(x)}_{B} Y$

- Assume residuals have constant variance σ^2
- Derive mean + variance of sampling distribution for the local polynomial estimate $\hat{f}(x)$ @ an arbitrary pt, x

$\hat{f}(x) = \hat{a}_0 = e_1^T \hat{a} \qquad e_1 = [1\ 0\ 0\ 0 \cdots 0]$

$\hat{a} = A^{-1} B Y \qquad E(\hat{a}) = A^{-1} B E(Y) = A^{-1} B \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$

$\mathrm{Var}(\hat{a}) = A^{-1} B \,\mathrm{Var}(Y) B^T A^{-1T} = \sigma^2 A^{-1} B B^T A^{-1T}$

$E(\hat{f}(x)) = e_1^T A^{-1} B E(Y)$

$\mathrm{Var}(\hat{f}(x)) = \sigma^2 e_1^T A^{-1} B B^T A^{-1} e_1$

e) $\mathrm{Cov}(Y) = \sigma^2 I \qquad$ Recall $y_i = f(x_i) + \varepsilon_i$

$E(Y) = \mu$

Define $r = y - Hy$

$E(\|r\|_2^2) = E[(y - Hy)^T (y - Hy)] = E[y^T y - 2y^T H y + y^T H^T H y]$

$= tr(I\sigma^2 I) + \mu^T \mu - 2(tr(H\sigma^2 I) + \mu^T H \mu) + tr(H^T H \sigma^2 I) + \mu^T H^T H \mu$

$= \sigma^2 n + \mu^T \mu - 2\sigma^2 tr(H) - 2\mu^T H \mu + \sigma^2 tr(H^T H) + \mu^T H^T H \mu$

$$E(\sigma^2) = \sigma^2 + \frac{-2\mu^T H \mu + \mu^T H H \mu}{n - 2\,tr(H) + tr(H^T H)}$$