

$$\begin{aligned}
E(y_i) &= E(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_{n-1} x_{in-1} + \epsilon) \\
&= E\left(\sum_{j=0}^{n-1} \beta_j x_{ij}\right) + E(\epsilon) \\
&= \sum_{j=0}^{n-1} \beta_j x_{ij}
\end{aligned}$$

$$\begin{aligned}
V(y_i) &= V(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_{n-1} x_{in-1} + \epsilon) \\
&= V(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_{n-1} x_{in-1}) + V(\epsilon) \\
&= V(\epsilon) \\
&= \sigma^2
\end{aligned}$$

$$\begin{aligned}
E(\hat{\beta}) &= E((\mathbf{X}^T \mathbf{X})^{-1}) \mathbf{X}^T \mathbf{y} \\
&= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon})) \\
&= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{\beta}) + E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}) \\
&= E(\boldsymbol{\beta}) + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\boldsymbol{\epsilon}) \\
&= \boldsymbol{\beta}
\end{aligned}$$

$$\begin{aligned}
V(\hat{\beta}) &= V((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}) \\
&= V((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon})) \\
&= V((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{\beta}) + V((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}) \\
&= ((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) V(\boldsymbol{\epsilon}) ((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)^T \\
&= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X})^{-1} (\mathbf{X}^T)^{-1} \sigma^2 \\
&= (\mathbf{X}^T \mathbf{X})^{-1} \sigma^2
\end{aligned}$$

e)

$$\begin{aligned}
E[(\mathbf{y} - \tilde{\mathbf{y}})^2] &= E[(\mathbf{y} - E(\tilde{\mathbf{y}}) + E(\tilde{\mathbf{y}}) - \tilde{\mathbf{y}})^2] \\
&= E[(\mathbf{y} - E(\tilde{\mathbf{y}}))^2 + 2(\mathbf{y} - E(\tilde{\mathbf{y}}))(E(\tilde{\mathbf{y}}) - \tilde{\mathbf{y}}) + (\tilde{\mathbf{y}} - E(\tilde{\mathbf{y}}))^2] \\
&= E[(\mathbf{y} - E(\tilde{\mathbf{y}}))^2] + 2E[(\mathbf{y} - E(\tilde{\mathbf{y}}))(E(\tilde{\mathbf{y}}) - \tilde{\mathbf{y}})] + E[(\tilde{\mathbf{y}} - E(\tilde{\mathbf{y}}))^2]
\end{aligned}$$

$$\begin{aligned}
E[(\mathbf{y} - E(\tilde{\mathbf{y}}))^2] &= E[(f(x) + \epsilon - E(\tilde{\mathbf{y}}))^2] \\
&= E[(f(x) - \tilde{\mathbf{y}})^2 - 2(\epsilon(E(\tilde{\mathbf{y}}) - f(x))) + \epsilon^2] \\
&= E[(f(x) - \tilde{\mathbf{y}})^2] - 2E[\epsilon(E(\tilde{\mathbf{y}}) - f(x))] + E[\epsilon^2] \\
&= E[(\tilde{\mathbf{y}} - f(x))^2] - 2E[\epsilon]E[E(\tilde{\mathbf{y}}) - f(x)] + \sigma^2 \\
&= E[(\tilde{\mathbf{y}} - f(x))^2] + \sigma^2
\end{aligned}$$

$$\begin{aligned}
E[(\mathbf{y} - \tilde{\mathbf{y}})^2] &= (E[(\tilde{\mathbf{y}} - f(x))^2] + \sigma^2) + 2(E(\tilde{\mathbf{y}}) - \mathbf{y})E[(\tilde{\mathbf{y}} - E(\tilde{\mathbf{y}}))] + V(\tilde{\mathbf{y}}) \\
&= Bias[\tilde{\mathbf{y}}]^2 + 2(E(\tilde{\mathbf{y}}) - \mathbf{y})(E(\tilde{\mathbf{y}}) - E(\tilde{\mathbf{y}})) + V(\tilde{\mathbf{y}}) + \sigma^2 \\
&= Bias[\tilde{\mathbf{y}}]^2 + V(\tilde{\mathbf{y}}) + \sigma^2
\end{aligned}$$