Sheet 3 Exercise 5

Kai

30 November 2016

The correct model: $\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$
$X = [\tilde{X}, x_p]$; $x_p = [x_{1p}, \ldots, x_{np}]'$; $\beta_p = [\beta_{p-1}^{'}, \beta_p]'$
Forgetting the p-th covariate: $\mathbf{y} = \tilde{\mathbf{X}}\boldsymbol{\beta_{p-1}} + \boldsymbol{\varepsilon}$; $\hat{\beta}_{p-1} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'y$

$$
\begin{aligned}
E[\hat{\beta}_{p-1}] &= E[(\tilde{X}'\tilde{X})^{-1}\tilde{X}'y] \\
&= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'E[y] \\
&= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'[\tilde{X}, x_p]\begin{bmatrix}\beta_{p-1} \\ \beta_p\end{bmatrix} + E[\varepsilon] \\
&= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'[\tilde{X}\beta_{p-1} + x_p\beta_p] \\
&= \beta_{p-1} + (\tilde{X}'\tilde{X})^{-1}\tilde{X}'x_p\beta_p \\
&\rightarrow bias: \underbrace{[\tilde{X}'\tilde{X}]^{-1}\tilde{X}'x_p}_{\delta}\beta_p
\end{aligned} \tag{1}
$$

δ is the least-squares estimator of the regression $x_p = \tilde{X}\gamma + \tilde{\varepsilon}$
If δ is approx. 0, the bias is approx. 0 $\rightarrow$ uncorrelated.
If β_p is approx. 0, the bias is also approx 0 $\rightarrow$ the covariate has no influence to the model.