STAT 151A Lecture 26

Henry Liev

27 October 2025

Remark 0.1 (Diagnostics)
COVRATIO $= \frac{1}{(1-h_{ii})\left(\frac{n-p-2-e_i^{*2}}{n-p-1}\right)^{p+1}}$
Roughly, this answers to: "By what factor does size of confidence interval/region for β change when we delete observation i?"
If COVRATIO is > 1, CI is worse without i
If COVRATIO is < 1, CI is better without i
So far: diagnostics for individual points i
Another question: Is the NLM reasonable?
Use diagnostics after fitting model to check this
$Y = \mathbf{X}\beta + \vec{\varepsilon}, \vec{\varepsilon} \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_n)$
Linearity - Interaction plots, Residual plot should be centered at $y = 0$ across all x, if linearity is violated then use a transformation or add more ?? of x in your model
Homoskedastic - Residual vs fitted plot, spread around 0 should be fairly similar among all fitted values, not increasing or decreasing in spread, fix with a transformation or bootstrap or HLM
Normality - Q-Q plot, $\bar{e} = 0, \text{Var}(e) = \hat{\sigma}^2(\mathbb{I} - \mathbb{H})$, points should lie on $y = x$, fix by using HLM
Uncorrelatedness -