Lin Mod M1 : 3B

MD

Standardized estimate of prediction error is $C_p = \dfrac{SSE(X_0)}{\hat{\sigma}^2_F} - (n - 2p_0)$

Q: If the simpler model is true and we replace $\hat{\sigma}^2_F$ by σ^2, what do you expect C_p should be?

Definitions: Know that $E[MSE] = \sigma^2$

$SSE[X_0]$ "residual sum of squares under simple model" $= (Y)'(M - M_0)(Y)$

$\hat{\sigma}^2_F$ "estimate of σ^2 under large, true model" $= E[MSE] = E\left[\dfrac{Y'(I-M)Y}{r(I-M)}\right]$

$$C_p = \frac{(Y)'(M-M_0)(Y)}{\underbrace{E\left[\frac{Y'(I-M)Y}{r(I-M)}\right]}_{=\sigma^2}} - (n - 2p_0)$$

$$= \frac{\sigma^2 tr(M_0) + (XB)'(M-M_0)(XB)}{\sigma^2} - (n-2p_0)$$

Use result from (A) to replace numerator.

$$\boxed{C_p = tr(M_0) - (n - 2p_0)}$$

Here: Assuming denominator is true σ^2, and that simple model is true. The latter implies $(XB)'(M-M_0)(XB) = 0$