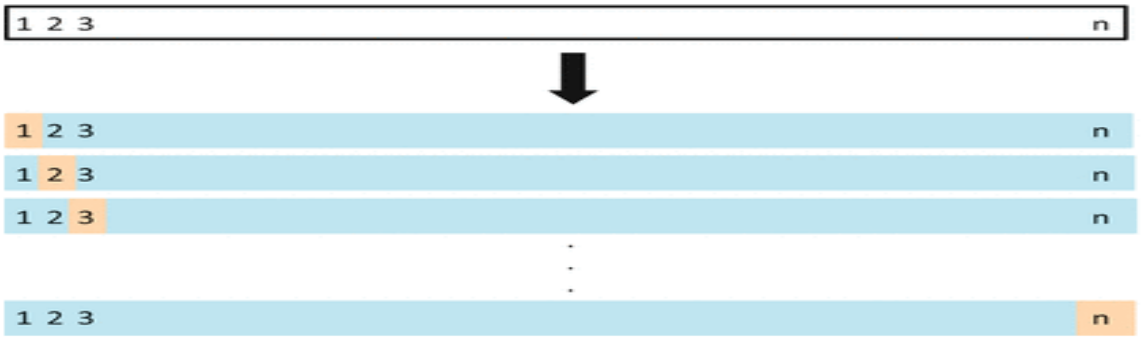
Leave One Out Cross Validation (LOOVC)

KOGOD SCHOOL of BUSINESS
AMERICAN UNIVERSITY • WASHINGTON, DC

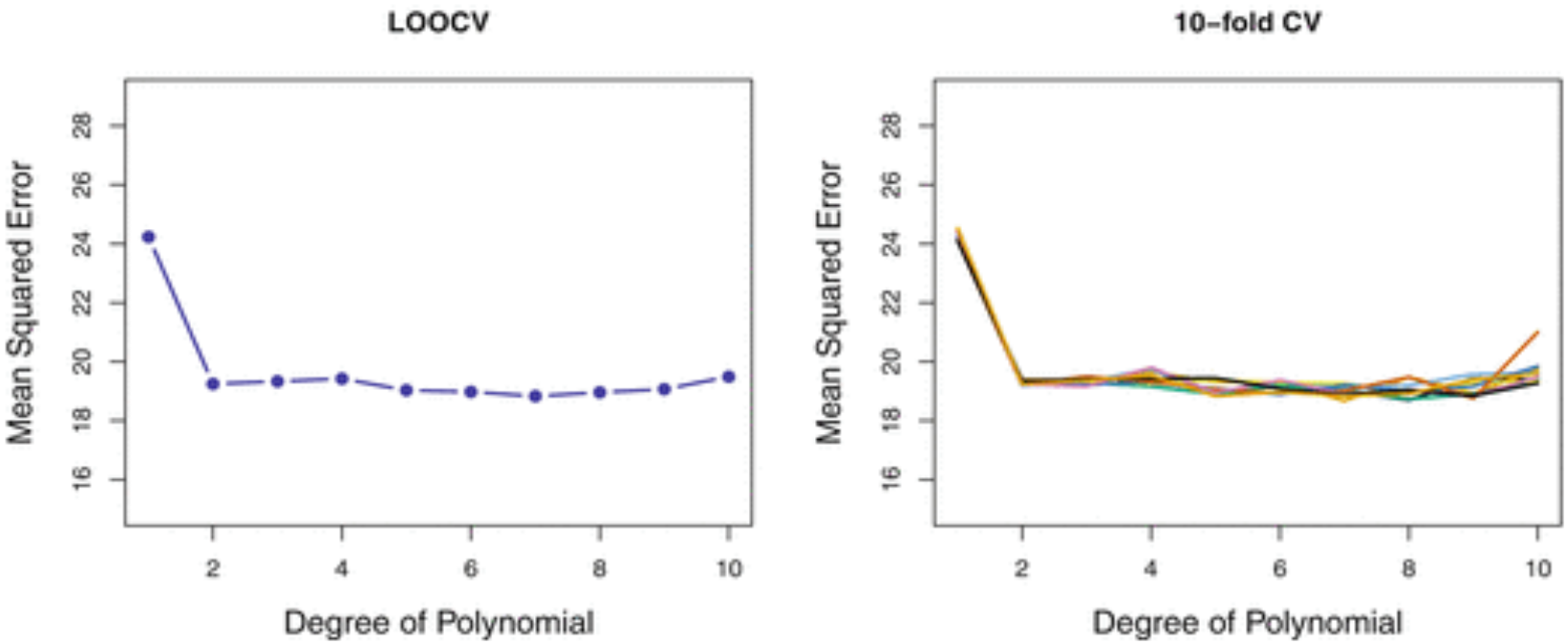
Leave One (or P) Out



- Leave One Out Cross Validation (**LOOCV**) is identical to **K-Fold** when **K** is **equal** to the number of observations **N**
- That is, the model is **fit** with **N-1** data points and the **MSE** calculated with the data point **left out**.
- The process is **repeated N times**, one for each data point
- This method is **computationally intensive** – you need to run one regression model estimation for each data point.
- But simulations show that it does **NOT outperform** the **K = 10** fold cross validation testing.
- **Leave P Out** Cross Validation is the same as LOOCV, but leaving P observations each time

LOOCV vs. 10-Fold CV

The example below was generated with the **"Auto"** data set in R, predicting **gas mileage** with **horsepower** (same model we showed before). The two graphs show that the LOOCV and the 10-Fold CV **performed similarly**, but the 10-Fold CV only requires 10 regression model estimations, whereas LOOCV requires one for each data point.



Tips

`cv.glm(){boot}` → `cv.glm()` function in the `{boot}` package to do cross-validation testing. Note that this function requires a `glm` object, not an `lm` object

`cv.mse=cv.glm(glm.fit, K=10)` → obtain MSE results from `glm` object and store in `cv.err` object; K=10 does a 10-Fold validation

`cv.err=cv.glm(glm.fit, K=10)$delta[1]` → The $delta[1] extracts the cross-validation error (first row [1]) from the cv.glm object

`cv.err.loocv=cv.glm(glm.fit)$delta[1]` → When the `K=` attribute is omitted, the `cv.glm()` function yields a Leave-One-Out cross validation (LOOCV)

KOGOD SCHOOL
of
BUSINESS