Interpretation of PACF

If you are trying to predict z_t using the earlier values

$$z_{t-1}, z_{t-2}, z_{t-3}, \cdots,$$

by using a regression model:

$$z_t = \beta_0 + \beta_1 z_{t-1} + \beta_2 z_{t-2} + \cdots + \beta_p z_{t-p} + \epsilon$$

then how far back in the past should you go? What value of p should you pick?

The PACF is meant to help you answer this question. If the PACF ϕ_{kk} is "small" for all lags $k > p$, then p is a reasonable choice.

PACF

The partial autocorrelation at lag k is (almost) the estimated regression coefficient for zlagk when doing the standard regression (OLS) of z on the variables zlag1, zlag2, ..., zlagk.

Compare the following with the PACF given earlier:

```
-------------------------------------------
Regression of z on zlag1 gives

Coefficients:
 (Intercept)      zlag1
    1.396133 0.7941462

-------------------------------------------
Regression of z on zlag1 and zlag2 gives

Coefficients:
 (Intercept)     zlag1       zlag2
    2.435215 1.384238 -0.7477757

--------------------------------------------
Regression of z on zlag1, zlag2 and zlag3 gives

Coefficients:
 (Intercept)     zlag1       zlag2     zlag3
    2.719594 1.295248 -0.5819785 -0.11964

--------------------------------------------
etc.
```