

MATH 545: INTRODUCTION TO TIME SERIES ANALYSIS

WINTER 2025

PARTIAL AUTOCORRELATION

The partial autocorrelation function (PACF), $\alpha_X(h)$, for a $\{X_t\}$ can be regarded in two ways.

- We have by definition that

$$\alpha_X(h) = \text{Corr}[(X_t - \mathbb{E}[X_t | X_{(t+1):(t+h-1)}]), (X_{t+h} - \mathbb{E}[X_{t+h} | X_{(t+1):(t+h-1)}])] \quad h \geq 2 \quad (1)$$

with $\alpha_X(0) = 1$ and $\alpha_X(1) = \rho_X(1)$. This is a consequence of the general definition of partial correlation in multivariate statistics. The time-ordering of the X s leads to a specific implementation.

- We have from the Levinson-Durbin algorithm that $\alpha_X(h)$ is $\phi_{h,h}$, which is the last element of the optimal one-step ahead forecast vector obtained by solving

$$\begin{bmatrix} \gamma_X(0) & \gamma_X(1) & \gamma_X(2) & \cdots & \gamma_X(h-1) \\ \gamma_X(1) & \gamma_X(0) & \gamma_X(1) & \cdots & \gamma_X(h-2) \\ \gamma_X(2) & \gamma_X(1) & \gamma_X(0) & \cdots & \gamma_X(h-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_X(h-1) & \gamma_X(h-2) & \gamma_X(h-3) & \cdots & \gamma_X(0) \end{bmatrix} \begin{bmatrix} \phi_{h,1} \\ \phi_{h,2} \\ \phi_{h,3} \\ \vdots \\ \phi_{h,h} \end{bmatrix} = \begin{bmatrix} \gamma_X(1) \\ \gamma_X(2) \\ \gamma_X(3) \\ \vdots \\ \gamma_X(h) \end{bmatrix}. \quad (2)$$

To see that these definitions are equivalent if the conditional expectations in (1) are linear, first denote

$$\hat{X}_{n+1} = \sum_{i=1}^n \phi_{n,i} X_{n-i+1} \equiv \sum_{i=1}^n \hat{a}_i X_i \quad (3)$$

say, for zero mean stationary process $\{X_t\}$ as usual. Recall that the minimum mean square error (or OLS) criterion dictates that for $i = 1, \dots, n$,

$$\mathbb{E}[(X_{n+1} - \hat{X}_{n+1})X_{n-i+1}] = 0 \quad \therefore \quad \mathbb{E}[\hat{X}_{n+1}X_{n-i+1}] = \mathbb{E}[X_{n+1}X_{n-i+1}] \quad (4)$$

which, after stacking the results for $i = 1, \dots, n$, is identical to (2) with $h = n$, as

$$\mathbb{E}[\hat{X}_{n+1}X_{n-i+1}] = \mathbb{E}\left[\sum_{j=1}^n \phi_{n,j} X_{n-j+1} X_{n-i+1}\right] = \sum_{j=1}^n \phi_{n,j} \gamma_X(i-j).$$

Let

- $\mathcal{H}$ denote the space of linear combinations of $X_1, \dots, X_n$,
- $\mathcal{S}$ denote the space of linear combinations of $X_2, \dots, X_n$,

We aim to forecast X_1 and X_{n+1} based on $\mathcal{S}$

$$X_1 \mid \underbrace{X_2, X_3, \dots, X_{n-1}, X_n}_{\text{conditioning set}} \mid X_{n+1}$$

Let the optimal **forward** forecast of X_{n+1} based on $X_2, \dots, X_n$ be denoted

$$\hat{X}_{n+1}^{\mathcal{S}} = \sum_{i=2}^n s_i^{\mathcal{F}} X_i$$

Let the optimal **backward** forecast of X_1 based on $X_2, \dots, X_n$ be denoted

$$\hat{X}_1^{\mathcal{S}} = \sum_{i=2}^n s_i^{\mathcal{B}} X_i$$

with MSE optimal coefficients determined by the equivalent system to (4), that is, for $i = 2, \dots, n$

$$\mathbb{E} \left[(X_1 - \hat{X}_1^{\mathcal{S}}) X_i \right] = 0. \quad (5)$$

Let $\mathcal{T}$ denote the space of scalar multiples of $X_1 - \hat{X}_1^{\mathcal{S}}$.

Consider arbitrary $X \in \mathcal{H}$; we can write

$$X = \sum_{i=1}^n a_i X_i = \sum_{i=2}^n a_i X_i + a_1 X_1 = \sum_{i=2}^n a_i X_i + \left(a_1 X_1 - a_1 \sum_{i=2}^n s_i^{\mathcal{B}} X_i \right) + a_1 \sum_{i=2}^n s_i^{\mathcal{B}} X_i$$

so therefore

$$X = \sum_{i=2}^n (a_i - a_1 s_i^{\mathcal{B}}) X_i + a_1 \left(X_1 - \sum_{i=2}^n s_i^{\mathcal{B}} X_i \right) = \sum_{i=2}^n s_i^{\dagger} X_i + a_1 (X_1 - \hat{X}_1^{\mathcal{S}})$$

where for $i = 2, \dots, n$, $s_i^{\dagger} = (a_i - a_1 s_i^{\mathcal{B}})$. That is, an arbitrary $X \in \mathcal{H}$ can be decomposed

$$X = \text{element in } \mathcal{S} + \text{element in } \mathcal{T}$$

so that $\mathcal{H} = \mathcal{S} \oplus \mathcal{T}$. By (3), as $\hat{X}_{n+1} \in \mathcal{H}$, we have by equating coefficients that

$$\begin{aligned} \hat{X}_{n+1} &= \sum_{i=1}^n \hat{a}_i X_i = \sum_{i=2}^n s_i^{\mathcal{F}} X_i + c \left(X_1 - \sum_{i=2}^n s_i^{\mathcal{B}} X_i \right) \\ &= \hat{X}_{n+1}^{\mathcal{S}} + c(X_1 - \hat{X}_1^{\mathcal{S}}) \end{aligned} \quad (6)$$

for some c . By (5) which states that $(X_1 - \hat{X}_1^{\mathcal{S}})$ is uncorrelated with each of $X_2, \dots, X_n$, we have that $\mathcal{S}$ and $\mathcal{T}$ are **orthogonal** subspaces of $\mathcal{H}$: any element of $\mathcal{S}$ is uncorrelated with any element of $\mathcal{T}$.

- By stationarity, the vectors $(X_1, \dots, X_n)^{\top}$ and $(X_n, \dots, X_1)^{\top}$ have the **same** autocovariance structure, so optimally forecasting X_1 **backwards** using $X_2, \dots, X_n$ requires the **same** coefficients as forecasting X_n **forwards** using $X_1, \dots, X_{n-1}$.

That is, $s_i^{\mathcal{B}} = \phi_{n-1, i-1}$ for $i = 2, \dots, n$ and

$$\hat{X}_1^{\mathcal{S}} = \sum_{i=2}^n \phi_{n-1, i-1} X_i.$$

- Also by stationarity, the vectors $(X_1, \dots, X_n)^{\top}$ and $(X_2, \dots, X_{n+1})^{\top}$ have the **same** autocovariance structure, so comparing with (3), we have $s_i^{\mathcal{F}} = \phi_{n-1, n-i+1}$ for $i = 2, \dots, n$ and

$$\hat{X}_{n+1}^{\mathcal{S}} = \sum_{i=2}^n \phi_{n-1, n-i+1} X_i.$$

Then we may write that

$$\begin{aligned}
\hat{X}_{n+1} &= \sum_{i=2}^n \phi_{n-1,n-i+1} X_i + c \left(X_1 - \sum_{i=2}^n \phi_{n-1,i-1} X_i \right) \\
&= cX_1 + \sum_{i=2}^n (\phi_{n-1,n-i+1} - c\phi_{n-1,i-1}) X_i \\
&= cX_1 + \sum_{i=1}^{n-1} (\phi_{n-1,n-i} - c\phi_{n-1,i}) X_{i+1} & i \longrightarrow i-1 \\
&= cX_1 + \sum_{i=1}^{n-1} (\phi_{n-1,i} - c\phi_{n-1,n-i}) X_{n-i+1} & i \longrightarrow n-i
\end{aligned}$$

Thus by comparing with (3), we have the Levinson-Durbin recursion as $c = \phi_{n,n}$ and for $i = 1, \dots, n-1$,

$$\phi_{n,i} = \phi_{n-1,i} - \phi_{n,n}\phi_{n-1,n-i}.$$

Now from (6), we have

$$\hat{X}_{n+1} = \hat{X}_{n+1}^S + c(X_1 - \hat{X}_1^S)$$

so therefore, multiplying through by $(X_1 - \hat{X}_1^S)$ and taking expectations, we have

$$\mathbb{E}[\hat{X}_{n+1}(X_1 - \hat{X}_1^S)] = \mathbb{E}[\hat{X}_{n+1}^S(X_1 - \hat{X}_1^S)] + c\mathbb{E}[(X_1 - \hat{X}_1^S)^2]. \quad (7)$$

- For the left hand side of (7), we have

$$\mathbb{E}[\hat{X}_{n+1}(X_1 - \hat{X}_1^S)] = \mathbb{E}[(\hat{X}_{n+1} - X_{n+1})(X_1 - \hat{X}_1^S)] + \mathbb{E}[X_{n+1}(X_1 - \hat{X}_1^S)].$$

But by (4), $\mathbb{E}[(\hat{X}_{n+1} - X_{n+1})(X_1 - \hat{X}_1^S)] = 0$, as $X_1 - \hat{X}_1^S$ is a linear combination of $X_1, \dots, X_n$. Therefore

$$\mathbb{E}[\hat{X}_{n+1}(X_1 - \hat{X}_1^S)] = \mathbb{E}[X_{n+1}(X_1 - \hat{X}_1^S)].$$

- For the right hand side of (7), $\mathbb{E}[\hat{X}_{n+1}^S(X_1 - \hat{X}_1^S)] = 0$ by (5), as $\hat{X}_{n+1}^S$ is a linear combination of $X_2, \dots, X_n$.

Therefore from (7) we have

$$\begin{aligned}
c &= \frac{\mathbb{E}[X_{n+1}(X_1 - \hat{X}_1^S)]}{\mathbb{E}[(X_1 - \hat{X}_1^S)^2]} = \frac{\mathbb{E}[(X_{n+1} - \hat{X}_{n+1}^S + \hat{X}_{n+1}^S)(X_1 - \hat{X}_1^S)]}{\mathbb{E}[(X_1 - \hat{X}_1^S)^2]} \\
&= \frac{\mathbb{E}[(X_{n+1} - \hat{X}_{n+1}^S)(X_1 - \hat{X}_1^S)]}{\mathbb{E}[(X_1 - \hat{X}_1^S)^2]} + \frac{\mathbb{E}[\hat{X}_{n+1}^S(X_1 - \hat{X}_1^S)]}{\mathbb{E}[(X_1 - \hat{X}_1^S)^2]} \\
&= \frac{\mathbb{E}[(X_{n+1} - \hat{X}_{n+1}^S)(X_1 - \hat{X}_1^S)]}{\mathbb{E}[(X_1 - \hat{X}_1^S)^2]} + 0 \\
&= \frac{\text{Cov}[(X_{n+1} - \hat{X}_{n+1}^S), (X_1 - \hat{X}_1^S)]}{\text{Var}[(X_1 - \hat{X}_1^S)]}
\end{aligned}$$

as in this setting, under the linearity assumption

$$\hat{X}_{n+1}^S \equiv \mathbb{E}[X_{n+1}|X_{2:n}] \quad \hat{X}_1^S \equiv \mathbb{E}[X_1|X_{2:n}].$$

Also, by stationarity

$$\text{Var}[(X_1 - \hat{X}_1^S)] = \text{Var}[(X_{n+1} - \hat{X}_{n+1}^S)]$$

so therefore

$$\begin{aligned} c &= \frac{\text{Cov}[(X_{n+1} - \mathbb{E}[X_{n+1}|X_{2:n}]), (X_1 - \mathbb{E}[X_1|X_{2:n}])]}{\sqrt{\text{Var}[(X_{n+1} - \mathbb{E}[X_{n+1}|X_{2:n}])]\text{Var}[(X_1 - \mathbb{E}[X_1|X_{2:n}])]}} \\ &= \text{Corr}[(X_{n+1} - \mathbb{E}[X_{n+1}|X_{2:n}]), (X_1 - \mathbb{E}[X_1|X_{2:n}])] \end{aligned}$$

which demonstrates the equivalent definitions of the PACF, as this relationship holds for $n = 2, 3, \dots$

Example: AR(4) Process For the stationary AR(4) process

$$\Phi(B)X_t = Z_t$$

with $\Phi(\cdot)$ defined via the generating polynomial with factorization

$$\begin{aligned} \Phi(z) &= (1 - 0.1z)(1 + 0.2z)(1 - (0.8 + 0.4i)z)(1 - (0.8 - 0.4i)z) \\ &= (1 - 0.2z)(1 + 0.4z)(1 - 1.6z + 0.80z^2) \end{aligned}$$

```
library(polynom)
t1<-polynomial(c(1,-0.2))
t2<-polynomial(c(1,0.4))
t3<-polynomial(c(1,-1.6,0.80))
Phi<-(t1*t2)*t3
Phi

## 1 - 1.4*x + 0.4*x^2 + 0.288*x^3 - 0.064*x^4

round(1/polyroot(Phi),2)

## [1] 0.8-0.4i 0.8+0.4i -0.4+0.0i 0.2+0.0i

covec<--coef(Phi)
```

Thus the model takes the explicit form

$$X_t = 1.4X_{t-1} - 0.4X_{t-2} - 0.288X_{t-3} + 0.064X_{t-4} + Z_t$$

We can compute the ACF and the PACF using `ARMAacf` for $h = 0, 1, \dots, 10$.

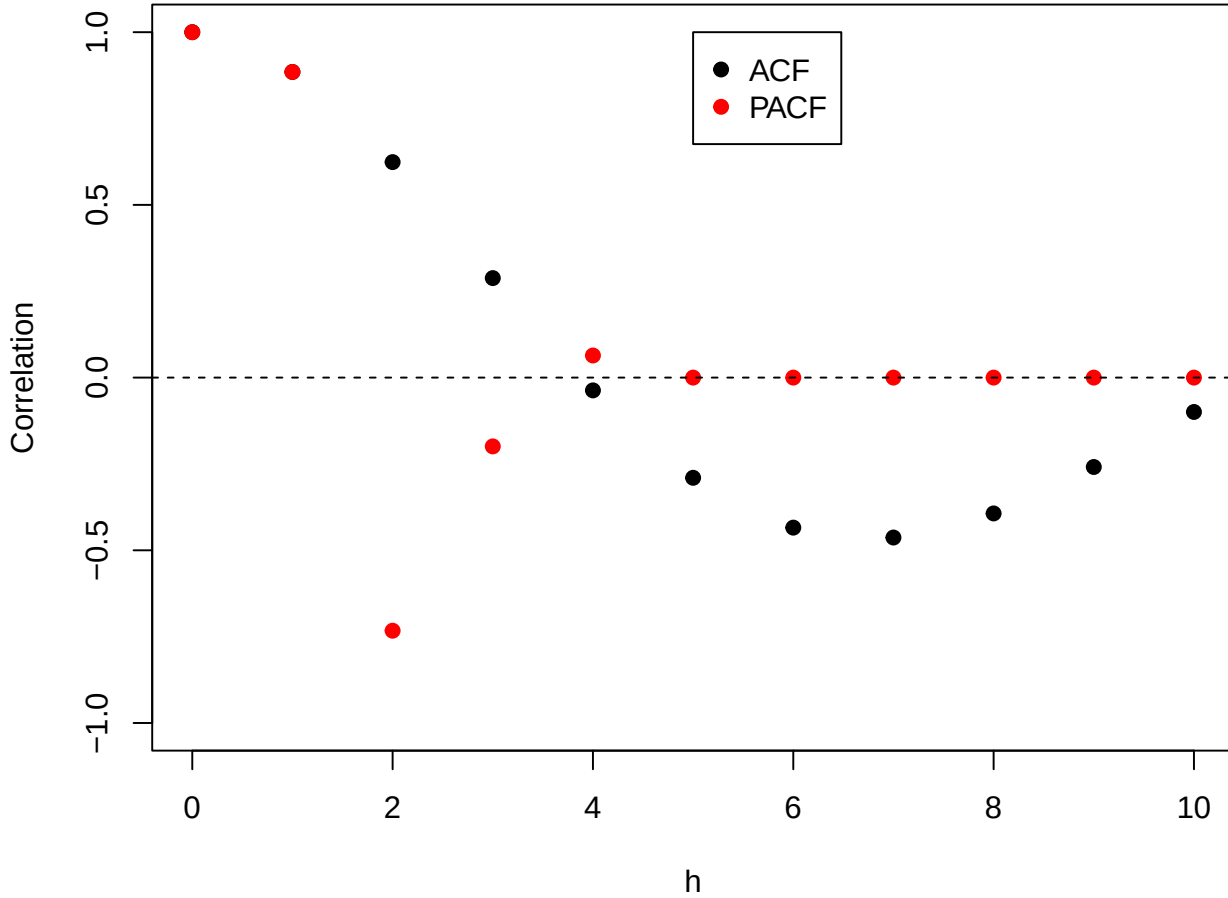
```
acf.vec<-ARMAacf(ar=covec[2:5],lag.max=10)
pacf.vec<-ARMAacf(ar=covec[2:5],lag.max=10,pacf=TRUE)
round(cbind(0:10,acf.vec,c(1,pacf.vec)),6)

##          acf.vec
## 0  0  1.000000  1.000000
## 1  1  0.884833  0.884833
## 2  2  0.623862 -0.732795
## 3  3  0.288102 -0.199216
## 4  4 -0.037033  0.064000
## 5  5 -0.290130  0.000000
## 6  6 -0.434416  0.000000
## 7  7 -0.463026  0.000000
## 8  8 -0.393282  0.000000
## 9  9 -0.258841  0.000000
## 10 10 -0.099516  0.000000
```

```

par(mar=c(4,4,1,2))
plot(0:10,acf.vec,pch=19, xlab='h', ylab='Correlation',ylim=range(-1,1))
points(0:10,c(1,pacf.vec),pch=19,col='red')
abline(h=0,lty=2)
legend(5,1,c('ACF','PACF'),pch=19,col=c('black','red'))

```



Note that for the AR(4) model, the PACF becomes identically zero beyond lag $h = 4$. This follows from the two facts

- (i) The AR(4) model takes the form

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4} + Z_t$$

so the forward prediction of X_{t+h} only depends on X_{t+h-1} , X_{t+h-2} , X_{t+h-3} and X_{t+h-4} , so for $h > 4$

$$X_{t+h} - \mathbb{E}[X_{t+h}|X_{t+1}, \dots, X_{t+h-1}] = X_t - (\phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4}) = Z_{t+h}$$

- (ii) The correlation structure of the reverse-time process is **identical** to that of the forward time process. Therefore for $h > 4$

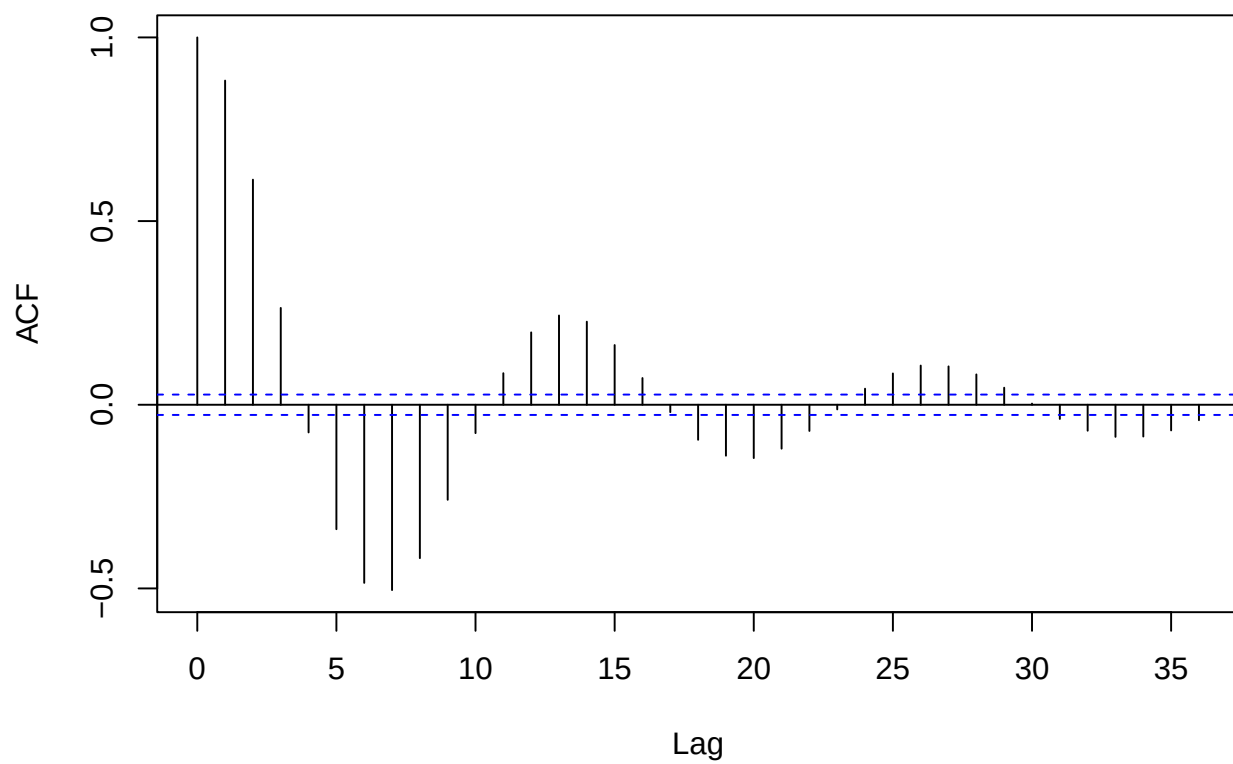
$$X_t - \mathbb{E}[X_t|X_{t+1}, \dots, X_{t+h-1}] = X_t - (\phi_1 X_{t+1} + \phi_2 X_{t+2} + \phi_3 X_{t+3} + \phi_4 X_{t+4}) = Z_t$$

Therefore for $h > 4$

$$X_t - \mathbb{E}[X_t|X_{t+1}, \dots, X_{t+h-1}] \quad \text{and} \quad X_{t+h} - \mathbb{E}[X_{t+h}|X_{t+1}, \dots, X_{t+h-1}]$$

are uncorrelated.

```
X<-arima.sim(n=5000,model=list(ar=covec[2:5]))  
par(mar=c(4,4,1,2))  
acf(X,main='')
```



```
pacf(X,main='')
```

