Ch3. ARMA(p, q) models

1. Define ARMA(p, q) model
2. ACVF/ACF
3. PACF (Partial Autocorrelation Function) $\approx$ ACF

We will determine the order of ARMA(p,q) model

Motivation

Nonstationary data
↓
remove trend / seasonality
↓
stationary data
↓
test of randomness
IID / Not IID
Not IID ↓ ARMA(p,q)

- In Chapter 2, we learned that linear process

  determined by few parameters

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

  provides general framework to study stationary TS.
- We also learned that sample average and SACVF/SACF provides reasonable estimates of stationary TS.
- However, SACVF

  "method of moment estimator"

  works well with small h →

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h} - \overline{X})(X_t - \overline{X})$$

  performs badly for large h in finite samples.
- Therefore, we will consider some "parametric" modeling of linear process known as ARMA(p, q). That is, coefficients $\{\psi_j\}$ will be fully determined by $(p + q)$ parameters.

ARMA(p, q) process

$*$ ARMA(1,1) => $X_t - \phi_1 X_{t-1} = Z_t + \theta_1 Z_{t-1}$

$\{X_t\}$ is an ARMA(p, q) process if $\{X_t\}$ is **stationary** and

AR(p) MA(q)

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q},$$

where $\{Z_t\} \sim WN(0, \sigma^2)$. $(\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma^2) \leftarrow$ desired parameters

Compact notation using backshift operator:

$$\boxed{\phi(B)X_t = \theta(B)Z_t}$$

$$\phi(B) = 1 - \phi_1 B - \ldots \phi_p B^p$$

$$\theta(B) = 1 + \theta B + \ldots + \theta_q B^q$$

$$B^j X_t = X_{t-j}$$

Restriction on ARMA coefficients

We will impose some restrictions on coefficients to achieve:

- Stationarity (existence and uniqueness of solution)
- Causality (only depends on past values)

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \sum_{j=0}^{\infty} |\psi_j| < \infty$$

- Invertibility (useful in forecasting)

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad \sum_{j=0}^{\infty} |\pi_j| < \infty$$

assumptions for forecasting

- Identifiability (modelling perspective)

– assumption for model uniqueness

Stationarity

- Recall Proposition 2.2.1 saying linear filter of stationary TS is again stationary TS
- For ARMA(p, q) series

$$\phi(B)X_t = \theta(B)\overbrace{Z_t}^{\text{already stationary TS (because it's WN}} \Rightarrow X_t = \phi(B)^{-1}\underbrace{\overbrace{\theta(B)}^{q^{th}\text{ polynomial} \to \text{finite sum} \to \text{absolutely summable}}Z_t}_{\text{stationary process}}$$

  Since $\theta(B)$ is finite filter, $\theta(B)Z_t$ is again stationary process.
- Thus, stationarity is determined by $\phi(B)^{-1}$.
- Suppose that $\phi(z) = 0$ have p-roots (may duplicate, but in the complex-field, fundamental theorem of algebra ensures that), say $\alpha_1, \ldots, \alpha_p$.

$$\phi(B) = \underbrace{(1 - \alpha_1^{-1}B)}_{\substack{0 \\ \Rightarrow B = \alpha_1}}\underbrace{(1 - \alpha_2^{-1}B)}_{\substack{0 \\ \Rightarrow B = \alpha_2}}\cdots\underbrace{(1 - \alpha_p^{-1}B)}_{\substack{0 \\ \Rightarrow B = \alpha_p}} \quad \text{factorization}$$

Stationarity

$\divideontimes \ \left(1-\frac{B}{\alpha_j}\right)^{-1} = \frac{1}{1-\frac{B}{\alpha_j}} = \sum_{k=0}^{\infty}\left(\frac{B}{\alpha_j}\right)^k$

- If $|1/\alpha_j| < 1$, $j = 1, \ldots, p$, then let $\left|\frac{B}{\alpha_j}\right| < 1$

  $\Rightarrow \alpha_j$ are roots $\phi(B)$

$$\phi(B)^{-1} = \prod_{j=1}^{p}\left(1-\frac{B}{\alpha_j}\right)^{-1} = \prod_{j=1}^{p}\left(\sum_{k=0}^{\infty}\left(\frac{B}{\alpha_j}\right)^k\right)$$

geometric sum

$$= \prod_{j=1}^{p}\left(\sum_{k=0}^{\infty}\left(\frac{1}{\alpha_j}\right)^k B^k\right) < \infty.$$

- If $|1/\alpha_j| > 1$, $j = 1, \ldots, p$, then

$$\phi(B)^{-1} = \prod_{j=1}^{p}\left\{\frac{-B}{\alpha_j}\left(1-\frac{\alpha_j}{B}\right)\right\}^{-1} = \prod_{j=1}^{p}\left(\frac{\alpha_j}{-B}\right)\left(1-\frac{\alpha_j}{B}\right)^{-1}$$

$\divideontimes \ \sum_{n}^{\infty} r^n = \frac{1}{1-r}, \ |r|<1$

$$= \prod_{j=1}^{p}\left(\frac{\alpha_j}{-B}\right)\sum_{k=0}^{\infty}\left(\frac{\alpha_j}{B}\right)^k = \prod_{j=1}^{p}(-\alpha_j)\sum_{k=0}^{\infty}\alpha_j^k B^{-(k+1)} < \infty$$

However, linear process will depends on future values.

Stationarity

unit root case

- If $|\alpha_j| = 1$ for some $j = 1, \ldots, p$, then it is still possible to write it as

$$\left(1 - \frac{B}{\alpha_j}\right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{B}{\alpha_j}\right)^k = \sum_{k=0}^{\infty} \left(\frac{1}{\alpha_j}\right)^k B^k,$$

hence only depends on the past but it diverges.

ARMA(p, q) has unique stationary solution if and only if
not unit root

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \neq 0 \text{ for all } |z| = 1$$

In short, no roots on the unit circle!
$\Rightarrow$ stationary solutions

Causality

- Causality means that X_t only depends on past values. Since

$$X_t = \phi(B)^{-1}\theta(B)Z_t$$

and as argued above we have that

ARMA(p, q) is causal if

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \neq 0 \text{ for all } |z| \leq 1$$

"no roots inside the unit root" ==

In short, roots are outside unit circle $\left|\frac{1}{a_j}\right| < 1$

- Note that 1 is included.

Invertibility

all the features will depend on $\theta(B)^{-1}$

It is already given that it is finite order so

Note that

$$Z_t = \theta(B)^{-1}\{\phi(B) X_t\}$$

and arguing similarly as above gives that

ARMA(p, q) is invertible, that is,

$$Z_t = \theta(B)^{-1}\phi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad \sum_{j=0}^{\infty} |\pi_j| < \infty$$

if

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q \neq 0 \text{ for all } |z| \leq 1$$

Again, roots are outside unit circle

Identifiability

"If the parameters are same, then the values are same"

- Statistical model P_θ (distribution) is identifiable if

$$P_{\theta_1} = P_{\theta_2} \iff \theta_1 = \theta_2$$

- For ARMA(p, q) model it corresponds to characteristic polynomials $\phi(z)$ and $\theta(z)$ has no common roots.
- Indeed. If $\phi(z)$ and $\theta(z)$ has common root, say s^*, then

$$\phi(z) = (1 - z/s^*)\phi_1(B), \quad \theta(z) = (1 - z/s^*)\theta_1(z)$$

$$\phi(B)X_t = \theta(B)Z_t \Rightarrow \phi_1(B)X_t = \theta_1(B)Z_t.$$

Therefore, it actually reduces to ARMA$(p-1, q-1)$.

Causal, invertible and stationary ARMA(p, q) process

ARMA(p, q) process has a causal, invertible and stationary solution if

$\phi(z)$ has roots outside unit curcle - Stationary & Causality

$\theta(z)$ has roots outside unit curcle - invertibility

$\phi(z)$ and $\theta(z)$ has no common roots

conditions for identifiability

You must be able to know whether a given ARMA(p, q) is a stationary / causal / invertible process. Then, calculate coefficients ψ_j and π_j if they are causal and invertible, respectively.

Example: ARMA(1,1)

$$X_t - .5X_{t-1} = Z_t + .4Z_{t-1}$$

- Stationary solution? AR polynomial, no roots on unit circle

  $\phi(B) = 1 - 0.5B = 0 \quad \therefore B = 2 \quad \Rightarrow$ stationary

- Causal? $|z| > 1$ ✓

- Invertible? $\theta(B) = 1 + 0.4B = 0 \quad \therefore B = -\frac{10}{4} = -2.5$

  $|B| > 1 \quad \therefore$ invertible

Example: ARMA(2,1)

Consider

$$X_t - .75X_{t-1} + .5625X_{t-2} = Z_t + 1.25Z_{t-1}.$$

Is it causal/invertible/has stationary solution?

Rather complicate to find solution. In R, you can use

constant x x^2

```
> ch = polyroot(c(1, -.75, .5625))
> ch
[1] 0.666667+1.154701i 0.666667-1.154701i
> Mod(ch)
[1] 1.333333 1.333333
```

Therefore, it is a stationary and causal process but not invertible.

Theoretical ACVF of ARMA(p, q)

$\phi(B) X_t = \theta(B) Z_t$

$X_t = \phi(B)^{-1} \theta(B) Z_t = \psi(B) Z_t$

Calculation of theoretical ACVF uses two major tools

- Form linear (causal) process representation

$$\gamma(h) = \text{Cov}(X_{t+h}, X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, \quad \gamma(-h) = \gamma(h)$$

  Provides general formula for any linear causal process, but actual calculation is tedious. Useful for pure MA models.

- Difference equations. Useful when AR part is included. But, still appeals to numerical computation.

Linear process representation

Useful for pure MA(q) process.

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

Thus, we have

$$\psi_0 = 1, \quad \psi_1 = \theta_1, \ldots, \psi_q = \theta_q.$$ $\psi_{q+1} = 0 \cdots\cdots$

Therefore, plug-into formula gives

$$\gamma(h) = \mathrm{Cov}(X_{t+h}, X_t)$$

$$= \mathrm{Cov}(Z_{t+h} + \theta_1 Z_{t+h-1} + \ldots + \theta_q Z_{t+h-q}, Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q})$$

$$= \begin{cases} \sigma^2(1 + \theta_1^2 + \ldots + \theta_q^2), & h = 0 \\ \sigma^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 1 \le h \le q \\ 0, & q < h \end{cases}$$

Linear process representation: ARMA(1,1)

$\psi(B) = \phi(B)^{-1}\theta(B)$

For $|\phi| < 1$ and $Z_t \sim WN(0, \sigma^2)$

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1} \iff (1 - \phi B)X_t = (1 + \theta B)Z_t$$

Need to find ψ_j to plug-into formula. Rather than calculate $\phi(B)^{-1}$, a bit better way is to calculate ψ_j from identity

$$\phi(B)^{-1}\theta(B) = \psi(B) \Rightarrow \theta(B) = \psi(B)\phi(B)$$

$$1 + \theta B = (1 - \phi B)(1 + \psi_1 B + \psi_2 B^2 + \dots)$$

0^{th} order : $1=1$

1^{st} order :
$$\theta = -\phi + \psi_1 \Rightarrow \psi_1 = \phi + \theta$$

2^{nd} order :
$$0 = -\phi\psi_1 + \psi_2 \Rightarrow \psi_2 = \phi(\phi + \theta)$$

⋮

$$\dots \Rightarrow \psi_j = \phi^{j-1}(\phi + \theta)$$

$\ast$ $\sigma^2\left(\sum_{j=0}^{\infty} \psi_j \psi_{j+h}\right)$

$$\gamma(h) = \begin{cases} \sigma^2\left(1 + \sum_{j=1}^{\infty} \phi^{2j-2}(\phi+\theta)^2\right), & h = 0 \\ \sigma^2\left(\phi^{h-1}(\phi+\theta) + \sum_{j=1}^{\infty} \phi^{2j+h-2}(\phi+\theta)^2\right), & h \geq 1 \end{cases}$$

Difference equations

The key idea is to multiply X_{t-k} on ARMA equation and take expectation. System of linear equations of $\gamma(h)$

$$E\left[X_{t-k}\left(X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p}\right)\right] = E\left[\left(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}\right) X_{t-k}\right]$$

$$\gamma(k) - \phi_1 \gamma(k-1) - \ldots - \phi_p \gamma(k-p) = \mathrm{Cov}(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}, X_{t-k})$$

Since $X_{t-k} = \sum_{j=0}^{\infty} \psi_j Z_{t-k-j}$ we can calculate RHS.

$= Z_{t-k} + \psi_1 Z_{t-k-1} + \psi_2 Z_{t-k-2} + \cdots$

Example ARMA(1,1) revisited:

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}$$

$$\gamma(k) - \phi\gamma(k-1) = \mathrm{Cov}(Z_t + \theta Z_{t-1}, X_{t-k})$$

$$= \mathrm{Cov}\left(Z_t + \theta Z_{t-1}, Z_{t-k} + \psi_1 Z_{t-k-1} + \psi_2 Z_{t-k-2} + \ldots\right)$$

Difference equations

$$
\begin{aligned}
k = 0: &\quad \gamma(0) - \phi\gamma(1) = \sigma^2(1 + \theta\psi_1) = \sigma^2(1 + \theta(\theta + \phi)) && (1)\\
k = 1: &\quad \gamma(1) - \phi\gamma(0) = \sigma^2\theta && (2)\\
k = 2: &\quad \gamma(2) - \phi\gamma(1) = 0 && (3)\\
k = h: &\quad \gamma(h) = \phi\gamma(h-1) && (4)
\end{aligned}
$$

From (1) and (2), (initial conditions)

$$\gamma(0) = \sigma^2 \frac{1 + \theta^2 + 2\theta\phi}{(1 - \phi^2)}$$

$$\gamma(1) = \sigma^2 \frac{(\theta + \phi)(1 + \theta\phi)}{1 - \phi^2}$$

and iteratively calculate for $h \geq 2$,

$$\gamma(h) = \phi\gamma(h-1).$$

Numerical Example

$\gamma(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$

Find the theoretical ACF/PACF of ARMA(1,1)

- $X_t = .7X_{t-1} + Z_t + .5Z_{t-1}$

$X_t - 0.7X_{t-1} = Z_t + 0.5Z_{t-1}$

$(1-0.7B)X_t = (1+0.5B)Z_t$

$X_t = (1-0.7B)^{-1}(1+0.5B)Z_t$

$= (1 + 0.7B + 0.7^2B^2 + \cdots)(1+0.5B)Z_t$

$= \underset{\psi_0}{1} + \underset{\psi_1}{1.2}B + \underset{\psi_2}{(0.49+0.35)}B^2 + \cdots$

or

$\psi(B) = \phi(B)^{-1}\theta(B)$

$\psi(B)\phi(B) = \theta(B)$

$(1 + \psi_1 B + \psi_2 B^2 + \cdots)(1-0.7B) = 1+0.5B$

- $X_t = .7X_{t-1} - .1X_{t-2} + Z_t.$ AR(2)

$E(X_{t-k} \cdot X_t) = E(X_{t-k}(0.7X_{t-1} - 0.1X_{t-2} + Z_t))$

$\gamma(k) = 0.7\,\gamma(k-1) - 0.1\,\gamma(k-2) + \mathrm{cov}(X_{t-k}, Z_t)$

i) $k=0 \Rightarrow \gamma(0) = 0.7\,\gamma(1) - 0.1\,\gamma(2) + \mathrm{cov}(X_t, Z_t)$

ii) $k=1 \Rightarrow \gamma(1) = 0.7\,\gamma(0) - 0.1\,\gamma(1) + \mathrm{cov}(X_{t-1}, Z_t)$

$\vdots$

Partial Autocorrelation Function

Recall that ACF is given by $\rho(h) = \mathrm{Corr}(X_t, X_{t+h})$.

Definition (PACF)

PACF (partial autocorrelation function) of a stationary TS is given by

~ adjusted autocorrelation function

$$\alpha(0) = \mathrm{Corr}(X_1, X_1) = 1$$

$$\alpha(1) = \mathrm{Corr}(X_2, X_1) = \rho(1)$$

$$\alpha(k) = \mathrm{Corr}(X_{k+1} - P_k^* X_{k+1}, X_1 - P_k^* X_1), \quad k \geq 2,$$

subtracting the effect of $X_2, \ldots, X_k$ (intermediate values)

where

$$P_k^* X_{k+1} = \textit{ BLP based on } \{1, X_2, \ldots, X_k\}$$

$$P_k^* X_1 = \textit{ BLP based on } \{1, X_2, \ldots, X_k\}$$

Conditional correlation of X_1 and X_{k+1} given intermediate values $X_2, \ldots, X_k$.

PACF

Alternatively, consider the following regression

$$
\begin{aligned}
X_{k+1} &= \phi_{11} X_k + \epsilon_{k+1} \\
X_{k+1} &= \phi_{21} X_k + \phi_{22} X_{k-1} + \epsilon_{k+1} \\
&\vdots \\
X_{k+1} &= \phi_{k1} X_k + \phi_{k2} X_{k-1} + \ldots + \phi_{kk} X_1 + \epsilon_{k+1}
\end{aligned}
$$

PACF $\alpha(k)$

Then, the BLP of X_{k+1} based on $\{X_k, \ldots, X_1\}$ is obtained by

$$
\widehat{X}_{k+1} = \underset{\phi}{\operatorname{argmin}} \, \mathrm{E} \left(X_{k+1} - \phi_{k1} X_k - \phi_{k2} X_{k-1} - \ldots - \phi_{kk} X_1 \right)^2
$$

The coefficient ϕ_{kk} measures correlation between X_{k+1} and X_1 when $X_2, \ldots, X_k$ is fixed.

$$
\boxed{\alpha(k) = \phi_{kk}, \quad k \geq 1}
$$

PACF: Examples

- AR(p)

$$X_t - \phi_1 X_{t-1} - \dots \phi_p X_{t-p} = Z_t$$

BLP based of X_{k+1} based on $\{X_k, \dots, X_1\}$ is given by

$$\widehat{X}_{k+1} = \phi_1 X_k + \dots \phi_p X_{k+1-p} + 0 X_{k-p} + \dots + 0 X_1.$$

Thus, $\alpha(0) := 1$,

$$\alpha(p) = \phi_p, \quad \alpha(k) = 0, \quad k > p.$$

Pure AR(p) has PACF stops at lag p. Other coefficients $\alpha(1), \dots \alpha(p-1)$ comes from the matrix equation.

- MA(1). It can be shown that

$$\alpha(k) = -(-\theta)^k / (1 + \theta^2 + \dots + \theta^{2k})$$

MA(q) has decreasing (tails-off) PACF

PACF: Examples

- For WN$(0, \sigma^2)$ process

$$X_t = Z_t,$$

we deduce that

$$\alpha(k) = 0, \quad k \geq 1.$$

- Therefore, when working on SPACF, we reject test for

$$\mathrm{H}_0 : \widehat{\alpha}(k) = 0$$

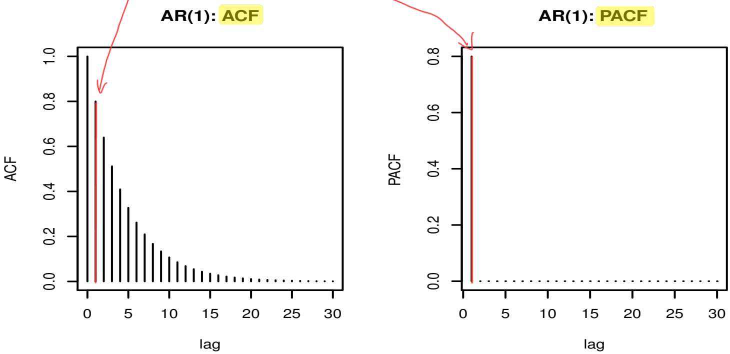
if

$$|\widehat{\alpha}(k)| > z_{\alpha/2} \frac{1}{\sqrt{n}}$$

ACF and PACF for ARMA

AR(1): $X_t = .8X_{t-1} + Z_t.$

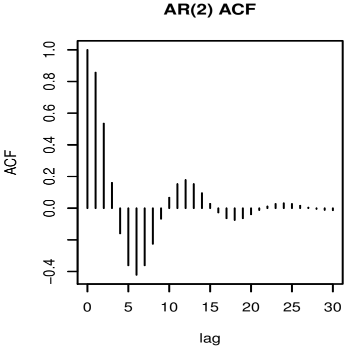
$\phi_1 = \rho(1) = \alpha(1) = 0.8$

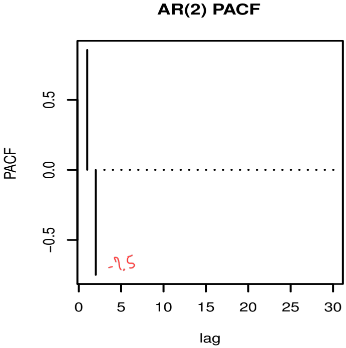


ACF decays fast, but PACF cuts off after lag 1

ACF and PACF for ARMA

AR(2): $X_t = 1.5X_{t-1} - .75X_{t-2} + Z_t$.

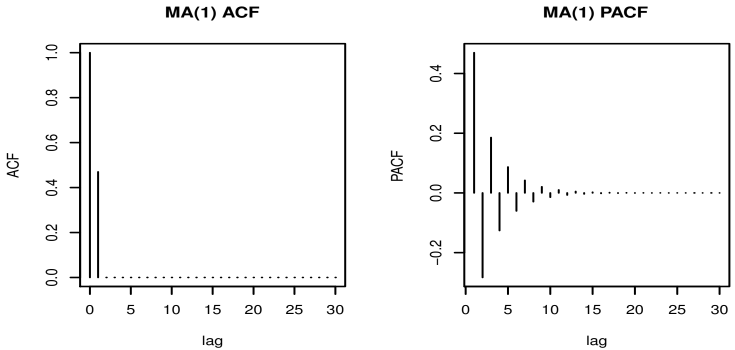




ACF decays fast (though with some sinusoidal decays), but PACF cuts off after lag 2. Also observe that $\alpha(2) = -.75$.

ACF and PACF for ARMA

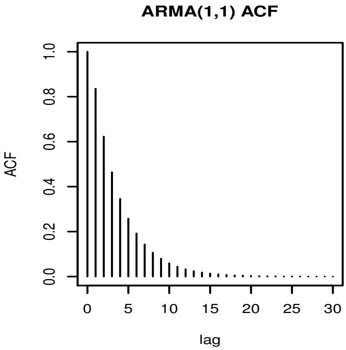
MA(1): $X_t = Z_t + .7Z_{t-1}$.

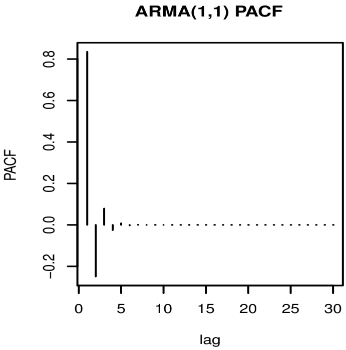


PACF decays fast with alternating sign, but ACF cuts off after lag 1.

ACF and PACF for ARMA

ARMA(1,1): $X_t - .7453X_{t-1} = Z_t + .32Z_{t-1}.$





Both ACF/PACF tails off quickly.

Identifying stationary ARMA(p, q) process

Keep this fact in mind whey you fit ARMA(p, q) model:

Process	ACF	PACF
AR(p)	Tails off exponentially (alternating/sine waves)	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off exponentially (alternating/ sine waves)
ARMA(p, q)	Combination of the above	