# 6. ARIMA and SARIMA models

Purpose: modelling of covariance stationary time series where the condition of constant mean is relaxed: $\mu_X(t)$ is allowed to change in time provided that causal ARMA process is obtained after removing $\mu_X(t)$ by repeated differencing. [a piecewise polynomial nature of $\mu_X(t)$ is usually sufficient].

Differencing operators:

$$\left.\begin{aligned} \Delta X_t &= X_t - X_{t-1} = X_t - BX_t = (1-B)X_t \\ \Delta^d X_t &= \Delta(\Delta^{d-1}X_t) = \dots = (1-B)^d X_t \end{aligned}\right\} \text{lag 1 differencing}$$

$$\left.\begin{aligned} \Delta_s X_t &= X_t - X_{t-s} = X_t - B^s X_t = (1-B^s) X_t \\ \Delta_s^D X_t &= \Delta_s(\Delta_s^{D-1} X_t) = \dots = (1-B^s)^D X_t \end{aligned}\right\} \begin{array}{l}\text{lag } s \text{ differencing} \\ s>1\end{array}$$

## 6.1. ARIMA(p,d,q) models

### ! 6.1.1. Definition

If $d \geq 0$ is an integer, then $\{X_t\}$ is said to be an ARIMA(p,d,q) process if $W_t := \Delta^d X_t = (1-B)^d X_t$ is a causal ARMA(p,q) process. [ARIMA = Autoregressive Integrated Moving Average]

We write more explicitly:

$$(6.1a) \quad \{X_t\} \sim ARIMA(p,d,q): \; \phi(B) W_t = \theta(B) Z_t \;, \; \{Z_t\} \sim WN(0,\sigma^2)$$
$$W_t = \Delta^d X_t \;, \; \{W_t\} \sim ARMA(p,q)$$

or equivalently in operator form:

$$(6.1b) \quad \{X_t\} \sim ARIMA(p,d,q): \; \phi(B)(1-B)^d X_t = \theta(B) Z_t \;, \quad \{Z_t\} \sim WN(0,\sigma^2).$$

where $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$ are polynomials of degrees $p$ and $q$ respectively ($\phi_p \neq 0, \theta_q \neq 0$) and $\phi(z) \neq 0$ for $|z| \leq 1$ guaranteeing causality of $W_t$ by 4.30(i)

! **6.1.2. Remark**

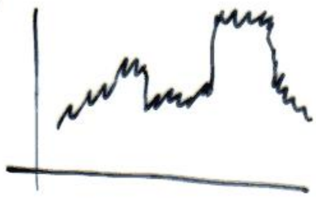
Denoting $\phi^*(z) := \phi(z)(1-z)^d$, we can rewrite (6.1b) as

(6.1c) $\{X_t\} \sim ARMA(p+d, q)$: $\phi^*(B)X_t = \theta(B)Z_t$, $\{Z_t\} \sim WN(0, \sigma^2)$

which is not causal, root $z=1$ of multiplicity $d$ being the only root of $\phi^*(z)$ violating causality (it lies on the unit circle, not outside)

! **6.1.3. Example (Time Series with polynomial trend)**

Assume $\mu_X(t)$ to be (piecewise) polynomial of degree at most $k$. Let $p(t) = \beta_0 + \beta_1 t + \dots + \beta_k t^k$ of degree $k$ ($\beta_k \neq 0$) be one of polynomial pieces. Applying $\Delta$, we get:



$$\Delta p(t) = p(t) - p(t-1) =$$

$$\underbrace{\beta_0 - \beta_0}_{0} + \beta_1 \underbrace{(t-(t-1))}_{1} + \beta_2 \underbrace{[t^2-(t-1)^2]}_{2t-1} + \dots + \beta_k \underbrace{[t^k-(t-1)^k]}_{k\,t^{k-1}+\dots} =:$$

$$=: \beta_0^{(1)} + \beta_1^{(1)} t + \dots + \beta_{k-1}^{(1)} t^{k-1} \text{ is of degree } k-1: \beta_{k-1}^{(1)} = \beta_k \cdot k \neq 0$$

After applying $\Delta$ $k$-times:

$$\Delta^k p(t) = \beta_0^{(k)} = \beta_k \cdot k \cdot (k-1) \dots 1 = \beta_k k! \neq 0$$

$$\Delta^{k+1} p(t) = 0$$

Choice of $d$:

(1) $\mu_X(t) = p(t)$ ... one piece $\Rightarrow$ $d=k$ is sufficient:

$$W_t = \Delta^d (\underbrace{p(t) + V_t}_{X_t}) = \Delta^d p(t) + \Delta^d V_t = \beta_k k! + \Delta^d V_t \quad ; \quad \Delta^d V_t \sim ARMA(p,q): \ \phi(B)\Delta^d V_t = \theta(B) Z_t$$

$$\Downarrow$$

$$\phi(B) W_t = \phi(B)\beta_k k! + \phi(B)\Delta^d V_t = \underbrace{(1-\phi_1-\dots-\phi_p)\beta_k k!}_{=:c} + \theta(B) Z_t$$

where we have used $B^j \beta_k k! = \beta_k k!$ (constant time series does not change with shifting)

Thus $W_t$ has generally nonzero mean $\mu_W = c$ which may be added to the other parameters as the so-called 'nuisance' parameter. Most SW packages allow this (if not, we proceed as with (2))

(2) $\mu_x(t)$ piecewise polynomial $\Rightarrow d = k+1$

Method (1) cannot be applied because after k-th differencing we arrive at a piecewise constant function which is more difficult to be estimated as an additional part of an ARMA model.

6.1.4. Remark

See [BD93, §9.1] for numerical examples illustrated with figures[pages 274 – 279]

## 6.2. SARIMA(p,d,q,P,D,Q,s) models

These models are suitable for seasonal series which are characterized by a strong serial correlation at the seasonal lag (and possibly multiples thereof). Unlike decomposition models which assume 'exact' seasonal behaviour, the SARIMA models are more flexible allowing for randomness in the seasonal pattern from one cycle to the next, incl. the seasonal period itself. [see 'sunspots.dat' data set: the Wolfer sunspot numbers 1770–1869].

! 6.2.1. Definition

If $d \geq 0$, $D \geq 0$ are integers, then $\{X_t\}$ is said to be a seasonal ARIMA process with period $s$ ($s > 1$ integer), or more precisely, SARIMA(p,d,q,P,D,Q,s) process if the differenced process $W_t := \Delta^d \Delta_s^D X_t = (1-B)^d (1-B^s)^D X_t$ is a causal ARMA(p,q) process; more explicitly:

(6.2a) $\{X_t\} \sim SARIMA(p,d,q,P,D,Q,s)$: $\phi(B)\tilde{\phi}(B^s) W_t = \theta(B)\tilde{\theta}(B^s) Z_t$, $\{Z_t\} \sim WN(0,\sigma^2)$

$W_t = \Delta^d \Delta_s^D X_t$, $\{W_t\} \sim ARMA(p,q)$

or equivalently

(6.2b) $\{X_t\} \sim SARIMA(p,d,q,P,D,Q,s)$: $\phi(B)\tilde{\phi}(B^s)(1-B)^d(1-B^s)^D X_t = \theta(B)\tilde{\theta}(B^s) Z_t$, $\{Z_t\} \sim WN(0,\sigma^2)$

where $\Phi(z)=1-\phi_1 z-\dots-\phi_p z^p$, $\tilde{\Phi}(z)=1-\tilde{\Phi}_1 z-\dots-\tilde{\Phi}_P z^P$, $\Theta(z)=1+\theta_1 z+\dots+\theta_q z^q$ and $\tilde{\Theta}(z)=1+\tilde{\Theta}_1 z+\dots+\tilde{\Theta}_Q z^Q$ are polynomials of degrees $p, P, q, Q$ respectively ($\phi_p\neq 0, \tilde{\Phi}_P\neq 0, \theta_q\neq 0, \tilde{\Theta}_Q\neq 0$), $\Phi(z)\neq 0$ and $\tilde{\Phi}(z)\neq 0$ for $|z|\leq 1$.

## 6.2.2. Remark

Denoting $\Phi^*(z) := \Phi(z)\tilde{\Phi}(z^s)(1-z)^d(1-z^s)^D$ [of degree $p+Ps+d+Ds$]

$\Theta^*(z) := \Theta(z)\tilde{\Theta}(z^s)$ [of degree $q+Qs$]

we can rewrite (6.2b) as

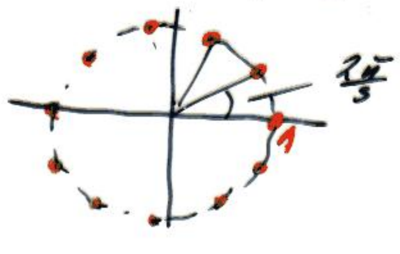
(6.2c) $\{X_t\}\sim ARMA(p+d+(P+D)s,\ q+Qs): \Phi^*(B)X_t=\Theta(B)Z_t$

$\{Z_t\}\sim WN(0,\sigma^2)$

which is not causal, the root $z_0=1$ with multiplicity $d+D$ and $s-1$ roots $z_k=e^{i\frac{2\pi k}{s}}$, $k=1,\dots,s-1$, each with multiplicity $D$ being the only roots of $\Phi^*(z)$ violating the causality.

[they are uniformly spread on the unit circle with angle step $\frac{2\pi}{s}$ not outside]

To verify this, let us check the roots of all component polynomials:




(i) $\Phi(z)\neq 0$ for $|z|\leq 1$ by assumption

(ii) $\tilde{\Phi}(z^s)\neq 0$ for $|z^s|\leq 1$ by assumption;
as $|z|\leq 1 \Rightarrow |z^s|=|z|^s\leq 1$, we have $\tilde{\Phi}(z^s)\neq 0$ for $|z|\leq 1$.

(iii) $(1-z)^d$ has root 1 with multiplicity $d$

(iv) $(1-z^s)^D=[-(z-z_0)\dots(z-z_{s-1})]^D=(-1)^D(z-z_0)^D\dots(z-z_{s-1})^D$
where $z_k$, $k=0,1,\dots,s$ are all roots of $1-z^s$, i.e. solutions of eq. $z^s=1$ which are all complex $s$-th roots of 1: $z_k=e^{i\frac{2\pi k}{s}}$ [clearly $z_k^s=e^{i\frac{2\pi ks}{s}}=1$]. Because of the $D$-th power, each of them has multiplicity $D$, except $z_0=1$ which has total multiplicity $d+D$ in view of (iii).

## 6.2.3. Construction of the SARIMA model

Let $X^{(k)} = \{X_\tau^{(k)}\} := \{X_{k+\tau s} \mid \tau \in \mathbb{Z}\}$ be between-seasonal partial series for each $k = 1, \dots, s$.

For example, if $\{X_t\}$ are monthly data with seasonal period $s = 12$ months, then $X^{(1)}$ are all January values, $X^{(2)}$ all February values, etc.

Then the same ARIMA$(P,D,Q)$ model is applied to each $X_\tau^{(k)}$:

$V_\tau^{(k)} = \Delta^D X_\tau^{(k)} = (1-B)^D X_{k+\tau s}$ for each $k = 1, 2, \dots, s$

As $\Delta X_\tau^{(k)} = X_\tau^{(k)} - X_{\tau-1}^{(k)} = X_{k+\tau s} - X_{k+(\tau-1)s} = X_{\underbrace{k+\tau s}_{t}} - X_{\underbrace{k+\tau s}_{t} - s} = (1-B^s)X_t$

This differencing (for all $k = 1, \dots, s$ simultaneously) is equivalent with lag $s$ differencing $\boxed{V_t = (1-B^s)^D X_t}$ (6.3)

Then $V^{(k)} \sim$ ARMA$(P,Q)$ with the same parameters given by $\tilde{\Phi}$ and $\tilde{\Theta}$:

(6.3a) $\tilde{\Phi}(B) V_\tau^{(k)} = \tilde{\Theta}(B) E_\tau^{(k)}$ ; $\{E_\tau^{(k)}\} = \{E_{k+\tau s}\} \sim WN(0, \tilde{\sigma}^2)$ for all $k = 1, 2, \dots, s$.
decorrelated at lag $s$

Clearly (6.3a) is equivalent with any of the following:

(6.3b) $V_{k+\tau s} - \tilde{\Phi}_1 V_{k+(\tau-1)s} - \dots - \tilde{\Phi}_P V_{k+(\tau-P)s} = E_{k+\tau s} + \tilde{\Theta}_1 E_{k+(\tau-1)s} + \dots + \tilde{\Theta}_Q E_{k+(\tau-Q)s}$

where $k + (\tau - j)s = k + \tau s - js = t - js$

(6.3c) $V_t - \tilde{\Phi}_1 \underbrace{V_{t-s}}_{B^s V_t} - \dots - \tilde{\Phi}_P \underbrace{V_{t-Ps}}_{(B^s)^P V_t} = E_t + \tilde{\Theta}_1 \underbrace{E_{t-s}}_{B^s E_t} + \dots + \tilde{\Theta}_Q \underbrace{E_{t-Qs}}_{(B^s)^Q E_t}$

(6.3d) $\tilde{\Phi}(B^s) V_t = \tilde{\Theta}(B^s) E_t$

$E_t$ are decorrelated at lag $s$ only: $\rho_E(s \cdot h) = 0$ for $h = 1, 2, \dots$

Therefore $E_t$ is modelled as $\{E_t\} \sim$ ARIMA$(p,d,q)$ because we expect relicts of the global trend after the lag $s$ differencing:

(6.4) $\phi(B)(1-B)^d E_t = \theta(B) Z_t$ ; $\{Z_t\} \sim WN(0, \sigma^2)$.

Combining (6.3), (6.3d) and (6.4) we arrive exactly to (6.2b) as follows:

$$\tilde{\Phi}(B^s)V_t \overset{(6.3d)}{=} \tilde{\Theta}(B^s)E_t \Rightarrow \phi(B)(1-B)^d\tilde{\Phi}(B^s)V_t = \phi(B)(1-B)^d\tilde{\Theta}(B^s)E_t \Rightarrow$$

$$\Rightarrow \underline{\phi(B)\tilde{\Phi}(B^s)(1-B)^d\underbrace{(1-B^s)^D X_t}_{V_t \text{ by } (6.3)}} = \tilde{\Theta}(B^s)\phi(B)(1-B)^d E_t \overset{(6.4)}{=}$$

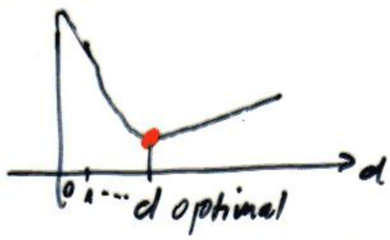
$$= \tilde{\Theta}(B^s)\Theta(B)Z_t = \underline{\Theta(B)\tilde{\Theta}(B^s)Z_t} \quad \text{which is (6.2b).}$$

6.2.4. Remark

See [BD 93, §9.6] for numerical examples illustrated with figures [pages 320–326].

## 6.3 Practical recommendations

ARIMA: choice of $d$: we repeat differencing until the acf and pacf looks like from stationary process.
A good indicator of that is the estimated variance of the differenced series which should be minimized.
Denoting $Y_t = \Delta^d X_t$, $\hat{\mu}_Y = \frac{1}{n}(Y_1 + \dots + Y_n)$
Then $\hat{\sigma}_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \hat{\mu}_Y)^2 \to \min$




SARIMA: Usually $0 \le d, D \le 1$. The right orders $d, D$ are best to find by trial.
The orders $p, P, q, Q$ are usually small as well.
When searching them by trial, start either with $p = P = 0$, or $q = Q = 0$. Combinations $p = 0, q > 0$ with $P > 0, Q = 0$ or vice versa $p > 0, q = 0$ with $P = 0, Q > 0$ are less probable, or give models with high orders. So avoid this in the first stage.

## 6.4. Generalized SARIMA models

We have seen in 6.1 and 6.2. that the roots on unit circle give a special structure to the $\phi^*(z)$ which is characteristic for certain type of nonstationarity.

$z_0 = 1$ ... the only multiple root, $|z| = 1 \Rightarrow$ ARIMA

$z_k = e^{i\frac{2\pi k}{s}}$ ... $s$ roots uniformly distributed on unit circle $\Rightarrow$ SARIMA.

In general the roots violating stationarity may lie anywhere on the unit circle:

$z = -1$, giving term $(1+B)^d$ instead of $(1-B)^d$ in $\phi^*(B)$

or any pair of complex conjugate roots $z_1 = e^{i\theta}$, $z_2 = e^{-i\theta}$

giving term $(1-e^{-i\theta}B)^d \cdot (1-e^{i\theta}B)^d = (1-(e^{i\theta}+e^{-i\theta})B + B^2)^d =$

$= (1 - 2(\cos\theta)B + B^2)^d$.

The above terms play the role of 'generalized' (or weighted) differencing operators and may be combined if more such roots lie on the unit circle:

$$D(B) = (1-B)^{d_1}(1+B)^{d_2} \prod_{j=3}^{J} (1 - 2(\cos\theta_j)B + B^2)^{d_j}$$

See [BD93, §9.1] for such numerical examples [pages 279–283].

The art of modeling is in finding the special structure of $\phi^*(z)$ where most coefficients are usually zeros. $D(z)$ is a multiplicative factor helping us to approach that structure.

With regard to practical modeling the amount of fitted parameters may be significantly reduced using this procedure. Unfortunately many SW packages for time series modeling do not allow fixing of selected parameters to zero values. The PEST program which is part of ITSM package is an exception – see [BD94]. Reduced number of parameters gives both narrower confidence intervals as well as better numerical stability.