TSA03: Univariate Stationary Time Series Models

Jakey BLUE

Properties of a Stationary Stochastic Process

Autocorrelation or Spectrum

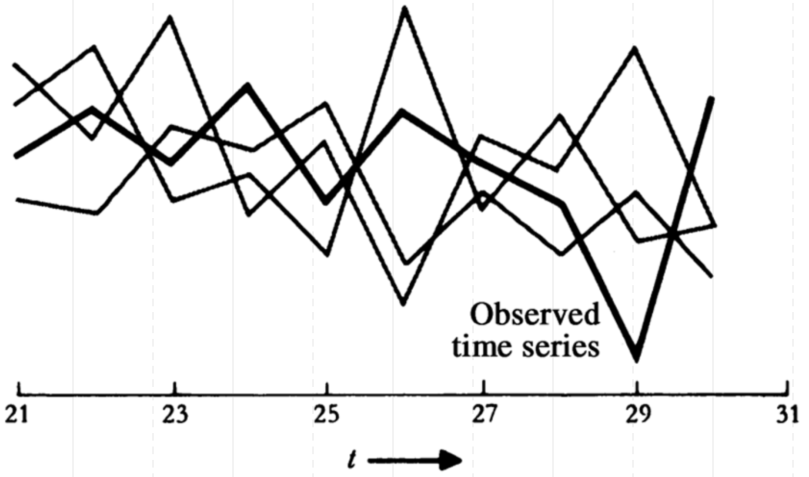
1

What is a Time Series?

- If a series of observations is collected at the timestamps $\{t_1, t_2, \dots, t_n\}$
  - $\{y(t_1), y(t_2), \dots, y(t_n)\}$ is a time series
- $t_2 - t_1 = t_9 - t_8$?
- Is a time series Continuous or Discrete?
  - Can we really capture a Continuous Time Series?
  - Discrete Time Series is usually collected via
    - sampling a *continuous time series*
      - e.g., gas flow
    - accumulating a variable over a period of time
      - e.g., rainfall, material batch

Stochastic Process

- Can a time series be "deterministic"?
- A stochastic process contains a statistical phenomenon that evolves in time according to probabilistic laws.
  - e.g., a Markovian series, random walk series
- A time series is the realization (or outcome) of a stochastic process.



Stationarity of a Stochastic Process

$$\mu = \mathrm{E}[y_t] = \int_{-\infty}^{\infty} yp(y)dy$$

$$\sigma_y^2 = \mathrm{V}[y_t] = \mathrm{E}\left[(y_t - \mu)^2\right] = \int_{-\infty}^{\infty} (y-\mu)^2 p(y)dy$$

- Strictly stationary: if the joint probability distribution associated with m observations $\{y_1, y_2, \dots, y_m\}$, made at any set of times $\{t_1, t_2, \dots, t_m\}$, is the same as that associated with m observations $\{y_{1+k}, y_{2+k}, \dots, y_{m+k}\}$, made at times $\{t_{1+k}, t_{2+k}, \dots, t_{m+k}\}$.

Autocovariance/Autocorrelation Functions

$$\gamma_k = \mathrm{COV}[y_t, y_{t+k}] = \mathrm{E}[(y_t - \mu)(y_{t+k} - \mu)]$$

$$\rho_k = \frac{\mathrm{E}[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{\mathrm{V}[y_t]\mathrm{V}[y_{t+k}]}} = \frac{\mathrm{E}[(y_t - \mu)(y_{t+k} - \mu)]}{\sigma_y^2}$$

- $\rho_k = ?$
- $\rho_0 = ?$
- $\boldsymbol{\Gamma}_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix} =?\, f(\rho)$

Positive Definiteness of the Autocovariance MX

- The determinant of $\mathbf{\Gamma}_n$ or $\mathbf{P}_n$ is greater than 0, why?
  - Think about a linear combination of the random variables

$$\{y_t, y_{t-1}, y_{t-2}, \dots, y_{t-n+1}\} \rightarrow L_t = l_1 y_t + l_2 y_{t-1} + \cdots + l_n y_{t-n+1}$$

$$\because \mathrm{COV}\left[y_i, y_j\right] = \gamma_{|i-j|}, \mathrm{V}[L_t] = \sum_{i=1}^{n}\sum_{j=1}^{n} l_i l_j \gamma_{|i-j|}$$

  - $\mathrm{V}[L_t]$ must be larger than 0 for any non–zero $\{l_1, l_2, \dots, l_n\}$ such that
    - $\sum_{i=1}^{n}\sum_{j=1}^{n} l_i l_j \gamma_{|i-j|} > 0$
    - $\mathbf{\Gamma}_n$ or $\mathbf{P}_n$ is p.d.

Conditions to Positive Definiteness
(Check the Principal Minors of the Matrices)

- $n = 2, \mathbf{P}_2 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}$

- $n = 3, \mathbf{P}_3 = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}$

Is L_t Stationary? if y_t is not stationary.

$$\mathrm{COV}[L_t, L_{t-k}] = \sum_{i=1}^{n}\sum_{j=1}^{n} l_i l_j \,\mathrm{COV}\left[y_{t+1-i}, y_{t+1-k-j}\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} l_i l_j \gamma_{|k+j-i|}$$

- to extend, we check if the first-order difference of y_t: $\nabla y_t = y_t - y_{t-1}$, or higher differences $\nabla^d y_t$ are all stationary.
  - This simply reveals the idea of handling a non-stationary time series.

Gaussian Process ➔ Weak Stationarity

- If the probabilistic law a stochastic process follows is a multivariate normal distribution, the process is called "Gaussian Process".
  - Fixed μ and $\boldsymbol{\Gamma}_n$ will be sufficient to ensure the stationarity.
- Compared to "strict stationarity", weak stationarity only ensures fixed μ and $\boldsymbol{\Gamma}_n$, that is to say, "second-order stationary".
- The most fundamental example of a stationary process:
  - A sequence of i.i.d. random variables ➔ strict or weak? μ and $\boldsymbol{\Gamma}_n$ =?

Autocovariance or Autocorrelation

- Unit-invariant property
  - autocovariance or autocorrelation?
- Symmetric property
  - $\gamma_{-k} = \gamma_k$
  - $\rho_{-k} = \rho_k$
  - We will only look at the positive side.

Estimating Autocovariance/Autocorrelation

- We were only talking about the theoretical parts of a stochastic process.
- What if we have collected the real data: $\{y_1, y_2, \dots, y_n\}$
  - sample mean (accuracy):

$$\bar{y} = \sum_{t=1}^{n} \frac{y_t}{n} \Rightarrow \mathrm{E}[\bar{y}] = \mu.$$

  - Precision of $\bar{y}$:

$$\mathrm{V}[\bar{y}] = \frac{1}{n^2} \sum_{t=1}^{n} \sum_{s=1}^{n} \gamma_{t-s} = \frac{\gamma_0}{n} \left[1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k \right].$$

When n is big enough, $n\mathrm{V}[\bar{y}] \to \gamma_0 \left[1 + 2 \sum_{k=1}^{\infty} \rho_k \right]$.

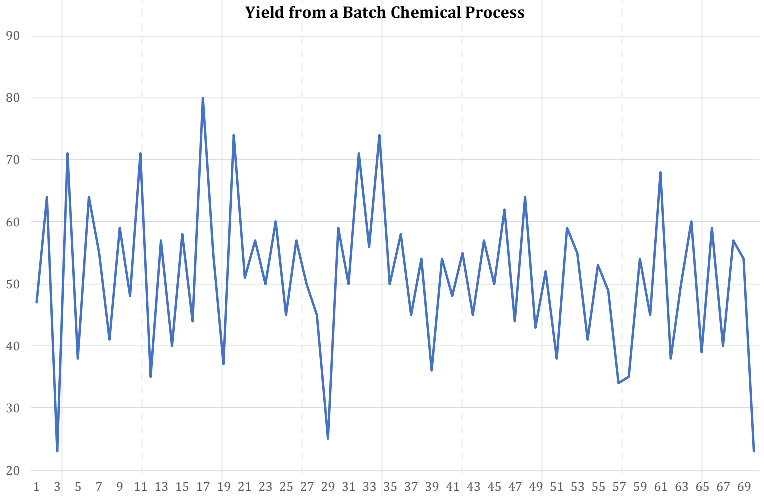
The Most Satisfactory Estimate of ρ_k

not too far from the intuition

$$\hat{\rho}_k = r_k = \frac{c_k}{c_0}$$

$$c_k = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y}), k = 1, 2, \dots, K$$

- r_k (or $\hat{\rho}_k$) is usually called the sample autocovariance (or autocorrelation) function.



k	r_k	k	r_k	k	r_k
1	−0.39	6	−0.05	11	0.11
2	0.30	7	0.04	12	−0.07
3	−0.17	8	−0.04	13	0.15
4	0.07	9	0.00	14	0.04
5	−0.10	10	0.01	15	−0.01

Standard Errors of Autocorrelation Estimates

Bartlett, M. S. "On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series." *Supplement to the Journal of the Royal Statistical Society*, vol. 8, no. 1, 1946, pp. 27–41. JSTOR, www.jstor.org/stable/2983611.

- For a Gaussian Process

$$\mathrm{V}[r_k] \approx \frac{1}{n}\sum_{v=-\infty}^{\infty}\left(\rho_v^2 + \rho_{v+k}\rho_{v-k} - 4\rho_k\rho_v\rho_{v-k} + 2\rho_v^2\rho_k^2\right)$$

  - If $\rho_k = \phi^{|k|}\ (-1 < \phi < 1)$

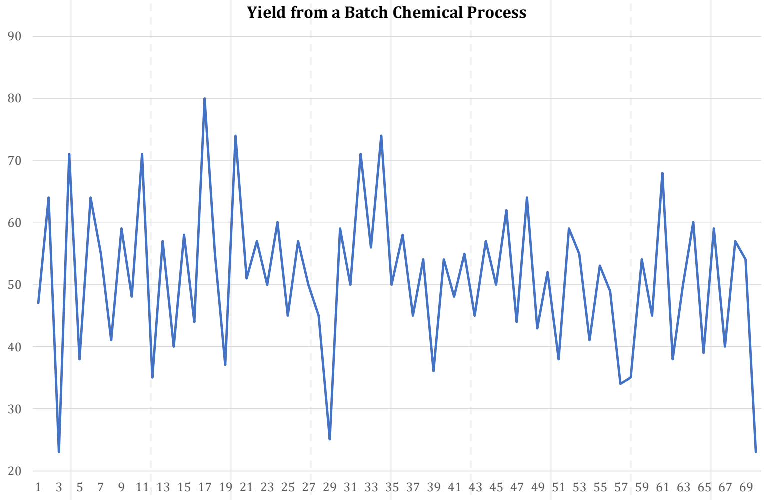
$$\mathrm{V}[r_k] \approx \frac{1}{n}\left[\frac{(1+\phi^2)(1-\phi^2)}{1-\phi^2} - 2k\phi^{2k}\right], \mathrm{V}[r_1] \approx \frac{1}{n}[1-\phi^2]$$

- Sometimes autocorrelations may die-out after certain lag, say q.

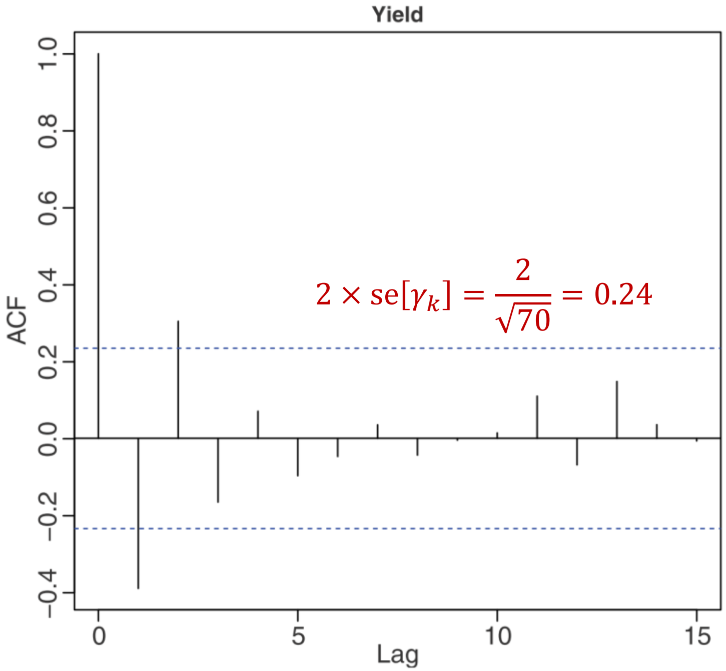
$$\mathrm{V}[r_k] \approx \frac{1}{n}\left(1 + 2\sum_{v=1}^{q}\rho_v^2\right), k > q.$$

- The simplest form (Bartlett 1946): $\mathrm{V}[r_k] = \frac{1}{n}$

Why Do We Need the Standard Errors?



k	r_k	k	r_k	k	r_k
1	−0.39	6	−0.05	11	0.11
2	0.30	7	0.04	12	−0.07
3	−0.17	8	−0.04	13	0.15
4	0.07	9	0.00	14	0.04
5	−0.10	10	0.01	15	−0.01



$$2 \times \mathrm{se}[\gamma_k] = \frac{2}{\sqrt{70}} = 0.24$$

Stationary Models May Have Spectral Properties

Periodogram of a Time Series

- A time series can be viewed as the composition of sine and cosine waves with different frequencies.
  - Assuming the number of observations is odd: $n = 2q + 1$

$$y_t = \alpha_0 + \sum_{i=1}^{q} [\alpha_i \cos(2\pi f_i t) + \beta_i \sin(2\pi f_i t)] + e_t,$$

where $f_i = \frac{i}{n}$ is the i^{th} harmonic of the basic frequency $\frac{1}{n}$ associated with the i^{th} sine wave component.

  - The least square estimates of the α_0 and (α_i, β_i) are

$$a_0 = \bar{y}, a_i = \frac{2}{n} \sum_{t=1}^{n} y_t \cos(2\pi f_i t)\,, b_i = \frac{2}{n} \sum_{t=1}^{n} y_t \sin(2\pi f_i t)\,.$$

Intensity of the Basic Frequency f_i

$$I(f_i) = \frac{n}{2}\left(a_i^2 + b_i^2\right), i = 1, 2, \dots, q,$$

is called the intensity at frequency f_i.
(recall the number of observations is odd: $2q + 1$)

- When n is even, set $n = 2q$
  - $a_q = \frac{1}{n}\sum_{t=1}^{n}(-1)^t y_t \, , b_q = 0$
  - $I\left(f_q\right) = I(0.5) = n a_q^2$

Analysis of Variance (on the Intensities)

- When n is odd, we can have $\frac{n-1}{2}$ pairs of degrees of freedom, i.e., (a_i, b_i).

$$\sum_{t=1}^{n}(y_t - \bar{y})^2 = \sum_{i=1}^{q} I(f_i)$$

- When n is even, there become $\frac{n-2}{2}$ pairs of degrees of freedom and a single one associated with the coefficient a_q.

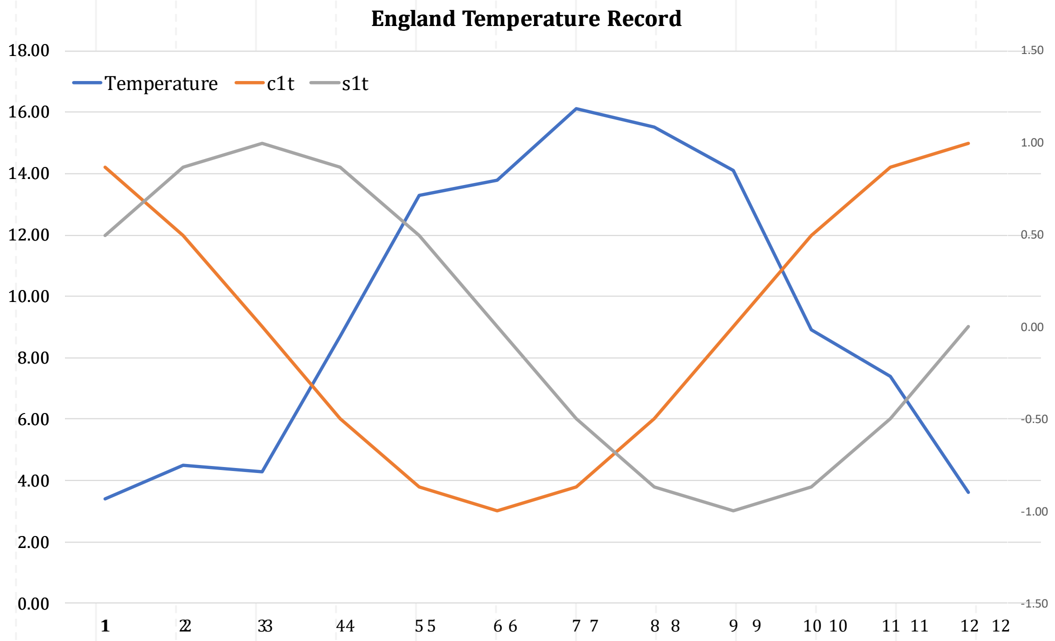
- Compared to "if the process is truly random" ➔ $y_t = \alpha_0 + e_t$
  - $E[I(f_i)] = 2\sigma^2$ and $I(f_i)$ is distributed as $\sigma^2\chi^2(2)$.

Monthly Temperatures for Central England in 1964

Month	Temperature	c1t	s1t
1	3.40	0.87	0.50
2	4.50	0.50	0.87
3	4.30	0.00	1.00
4	8.70	-0.50	0.87
5	13.30	-0.87	0.50
6	13.80	-1.00	0.00
7	16.10	-0.87	-0.50
8	15.50	-0.50	-0.87
9	14.10	0.00	-1.00
10	8.90	0.50	-0.87
11	7.40	0.87	-0.50
12	3.60	1.00	0.00

$$y_t = a_1 \cos(2\pi f_1 t) + b_1 \sin(2\pi f_1 t)$$

$$a_1 = -5.30,\ b_1 = -3.82$$



Analysis of the Variance of the Periodogram

i	a_i	b_i
1	−5.30	−3.82
2	0.05	0.17
3	0.10	0.50
4	0.52	−0.52
5	0.09	−0.58
6	−0.30	

	Frequency				
i	f_i	Period	Periodogram $I(f_i)$	Degrees of Freedom	Mean Square
1	1/12	12	254.96	2	127.48
2	1/6	6	0.19	2	0.10
3	1/4	4	1.56	2	0.78
4	1/3	3	3.22	2	1.61
5	5/12	12/5	2.09	2	1.05
6	1/2	2	1.08	1	1.08
			263.10	11	23.92

From Periodogram to Spectrum

- We can relax the frequency setting, from $f = \frac{i}{n}$ to $0 < f < 0.5$
  - Given a fixed f,

$$I(f) = 2\left[c_0 + 2\sum_{k=1}^{n-1} c_k \cos(2\pi f k)\right],$$

which is the Fourier cosine transform of the estimate of the **autocovariance**.

  - Furthermore, there are issues related to spectrum, spectral density function.

Key remark: “Spectral Density” and “Autocorrelation” are describing a time series equivalently but in different perspectives.

2

Linear Stationary Models

A time series is generated by a linear aggregation of random shocks.

“

Without loss of generality,
From the rest of this slides, we let y_t
denoted the ***centered*** *series, i.e.,*

$$y_t = y_t - \mathrm{E}[y_t].$$

Definitions in Premise

- An infinite series is absolutely summable if the sum of the absolute values of its summands converges, i.e.,

$$\sum_{j=0}^{\infty} |\theta_j| < \infty.$$

- Thus, the partial sum of the series $\sum_{j=0}^{\infty} \theta_j a_{t-j}$ can converge in mean square to the random variable y_t, i.e.,

$$\lim_{n\to\infty} \mathrm{E}\left[\left(\sum_{j=0}^{n} \theta_j a_{t-j} - y_t\right)^2\right] = 0$$

Considering a General Linear Process

$$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}\, y.$$

a_t is the ***white noise*** (shock, innovation, error, etc.), where

$\mathrm{E}[a_t] = 0, \mathrm{V}[a_t] = \sigma_a^2$.

$$\mathrm{E}[a_t a_{t-k}] = \begin{cases} \sigma_a^2 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

Is y a stationary stochastic process? Any assumptions?

Alternative Form of the General Linear Process

$$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \cdots + a_t = \sum_{j=1}^{\infty} \pi_j y_{t-j} + a_t.$$

- The alternative form is to express the process in terms of the past y_t and an added shock a_t.

- Introducing 2 important operators:
  - Backward shift operator B: $By_t = y_{t-1}$, and hence $B^j y_t =$?
  - Forward shift operator F: $Fy_t = y_{t+1}$, and $F^j y_t =$?

- The general linear process becomes

$$y_t = \left(1 + \sum_{j=1}^{\infty} \psi_j B^j\right) a_t \text{ or } y_t = \left(\sum_{j=1}^{\infty} \pi_j B^j\right) y_t + a_t.$$

Simple Stochastic Process: $y_t = a_t - \theta a_{t-1}$

$$y_t = (1 - \theta B)a_t = \theta(B)a_t$$

Is it stationary?

$$\Rightarrow (1 - \theta B)^{-1} y_t = a_t$$

HOW?

$$\Rightarrow \left(1 + \theta B + \theta^2 B^2 + \cdots\right) y_t = a_t.$$

$$\Rightarrow y_t = -\theta y_{t-1} - \theta^2 y_{t-2} - \theta^3 y_{t-3} - \cdots + a_t$$

Let $\pi_j = -\theta^j$

$$y_t = \sum_{j=1}^{\infty} \pi_j B^j y_t + a_t \Rightarrow \left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) y_t = a_t = \pi(B) y_t$$

Invertibility to Stationarity

- To ensure $y_t = (1 - \theta B)a_t$ an invertible process, one needs

$$|\theta| < 1$$

$$\sum_j^{\infty} |\theta|^j = \sum_j^{\infty} |\pi_j| < \infty.$$

Absolutely Summable

- That is to say, the inverted series is stationary.

$$\pi(B) = (1 - \theta B)^{-1} = \sum_{j=0}^{\infty} \theta^j B^j \text{ converges for all } |\theta| < 1$$

Stationary Condition for $\tilde{y}_t = \pi_1 \tilde{y}_{t-1} + a_t$

- On what condition do you think this time series is stationary?

$$y_t = \pi_1 y_{t-1} + a_t$$

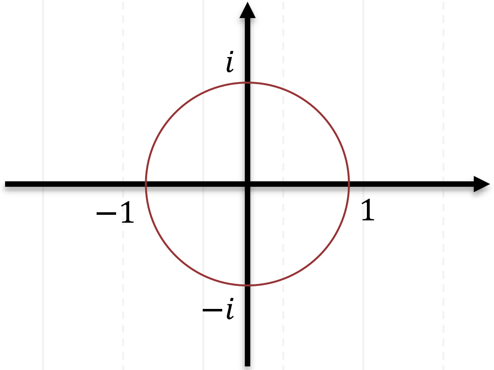
  - Intuition: $|\pi_1| < 1$
  - Characteristic function: $\pi(B) = 1 - \pi_1 B$
  - By assigning $\pi(B) = 0$, one gets

$$B = \frac{1}{\pi_1}.$$

  - Combining the intuition:

$$|B| > 1$$

  - As B can be a complex number, $|B| > 1$ indicates the region outside the unit circle of the complex plane.



Viewpoint from the Autocovariances

- For the General Linear Process: $y_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}$

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

- $\gamma_0 = \sigma_y^2 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2$

- It implies that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ will ensure the **stationarity**, because the autocovariances converge.

*Autocovariance Generating Function (AGF)

- Definition

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k$$

- For white noise, $\gamma(B) = \sigma_a^2$
- For $y_t = \psi(B)a_t, \gamma(B) = \sigma_a^2\psi(B)\psi\left(B^{-1}\right) = \sigma_a^2\psi(B)\psi(F)$ HOW?
- For $y_t = a_t - \theta a_{t-1} = (1-\theta B)a_t$,

$$\gamma(B) = \sigma_a^2(1+\theta B)\left(1+\theta B^{-1}\right) = \sigma_a^2\left[-\theta B^{-1} + \left(1+\theta^2\right) - \theta B\right]$$

- Compared with $\gamma_0 = \left(1+\theta^2\right)\sigma_a^2; \gamma_1 = -\theta\sigma_a^2; \gamma_k = 0$ for $k \geq 2$. Something Similar?

Derivation of AGF

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

$$\gamma(B) = \sigma_a^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+k} B^k = \sigma_a^2 \sum_{j=0}^{\infty} \sum_{k=-j}^{\infty} \psi_j \psi_{j+k} B^k$$

$$= \sigma_a^2 \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h B^{h-j} = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j B^j \sum_{h=0}^{\infty} \psi_h B^h = \sigma_a^2 \psi(B)\psi(F)$$

Let $h = j + k$

Summary: Stationarity and Invertibility

- For a linear process $y_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}$ to be stationary

$$\sum_{j=0}^{\infty} |\psi_j| < \infty.$$

- For a linear process $y_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}$ to be invertible

$$\sum_{j=0}^{\infty} |\pi_j| < \infty,$$

  - where$\pi_j = \psi^j$ and $\pi(B) = \psi^{-1}(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$.

Preview of Time Series Models

AutoRegressive

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + a_t$$

$$(1 - \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p) y_t = a_t$$

$$\phi(B) y_t = a_t$$

$$y_t = \frac{1}{\phi(B)} a_t$$

MovingAverage

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

$$(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t = y_t$$

$$\theta(B) a_t = y_t$$

ARMA

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

$$\phi(B) y_t = \theta(B) a_t$$

$$y_t = \frac{\theta(B)}{\phi(B)} a_t = \frac{1 - \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p}{1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q} a_t$$

Moving Average Processes

3

Invertible Conditions for MA Processes

- An MA(q) process is expressed as

$$\begin{aligned} y_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \\ &= \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right) a_t \\ &= \theta(B) a_t, \end{aligned}$$

  - where $a_t \sim^{iid} N\left(0, \sigma_a^2\right)$ for $t = 1, 2, \dots, \infty$.

- As seen, a MA(1) process, $y_t = (1 - \theta_1 B) a_t$, is invertible if

$$|\theta_1| < 1,$$

  - such that $\pi(B) = (1 - \theta_1 B)^{-1} = \sum_{j=0}^{\infty} \theta_1^j B^j$ will converge on/within an unit circle.
  - Equivalently, $B = \theta_1^{-1}$ of $(1 - \theta_1 B) = 0$ lies outside the unit circle.

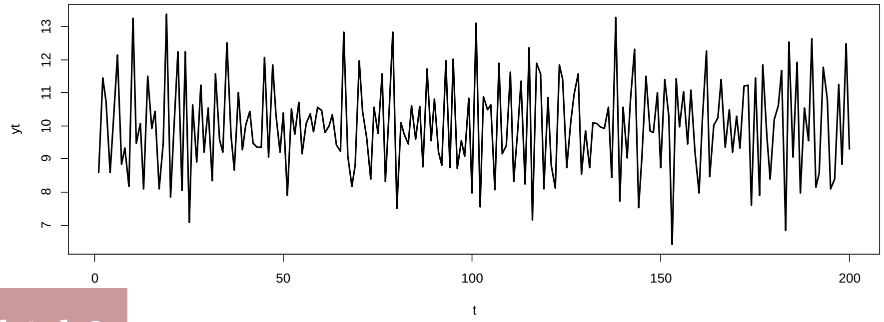
The Simplest MA Process: MA(1)

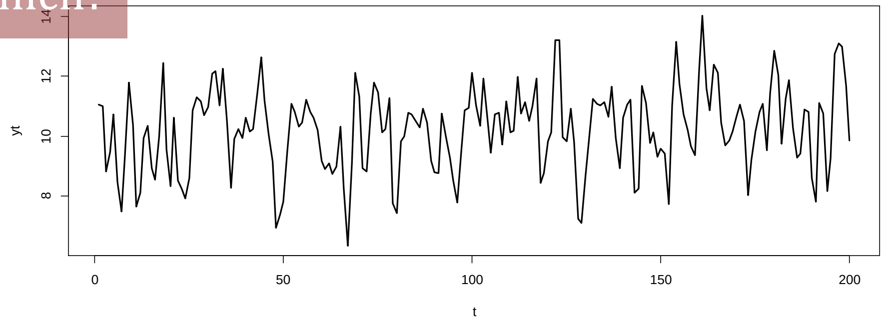
- MA(1) model: $y_t = (1 - \theta_1 B)a_t$
  - $\mathrm{E}[y_t] = ?$
  - $\gamma_0 = \mathrm{V}[y_t] = ?$
  - $\gamma_1 = \mathrm{COV}[y_t, y_{t-1}] = ?$
  - $\gamma_k = \mathrm{COV}[y_t, y_{t-k}] = ?, \forall k \geq 2$
- Therefore, the ACF (AutoCorrelation Function) of MA(1) is: $\rho_k = \begin{cases} ? & , k = 0 \\ ? & , k = 1 \\ ? & , k \geq 2 \end{cases}$

Insight of the ACF of MA(1)

- $$\rho_k = \begin{cases} 1 & , k = 0 \\ -\dfrac{\theta_1}{1+\theta_1^2} & , k = 1 \\ 0 & , k \geq 2 \end{cases}$$
  - $\theta_1 < 0 \Rightarrow \rho_1 > 0$
  - $\theta_1 > 0 \Rightarrow \rho_1 < 0$

Which is which?





What's the behavior of the ACF?

MA(2) Process

- MA(2): $y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = \left(a_t - \theta_1 B - \theta_2 B^2\right) a_t$
  - Mean: $\mathrm{E}[y_t] = ?$
  - Variance: $\gamma_0 = \mathrm{V}[y_t] = ?$
  - ACF: $\rho_k = \begin{cases} ? & , k = 0 \\ ? & , k = 1 \\ ? & , k = 2 \\ ? & , k \geq 3 \end{cases}$

What's the behavior of the ACF?

MA(2) Process

- MA(2): $y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = (a_t - \theta_1 B - \theta_2 B^2) a_t$
  - Mean: $\mathrm{E}[y_t]$ =?
  - Variance: $\gamma_0 = \mathrm{V}[y_t]$ =?
  - ACF: $\rho_k = \begin{cases} 1 & , k = 0 \\ \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & , k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & , k \geq 2 \\ 0 & , k \geq 3 \end{cases}$

What's the behavior of the ACF?

General MA(q) Process

- MA(q): $y_t = \left(1 - -\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right) a_t$
  - Mean: $\mathrm{E}[y_t] = 0$
  - Variance: $\gamma_0 = \mathrm{V}[y_t] = \sigma_a^2 \left(\sum_{i=0}^{q} \theta_i^2\right), \theta_0 = 1$
  - Autocovariance:

$$\gamma_k = \mathrm{E}[y_t y_{t-k}] - \mathrm{E}[y_t]\mathrm{E}[y_{t-k}]$$
$$= \mathrm{E}\left\{\left(a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}\right)\left(a_{t-k} - \theta_1 a_{t-k-1} - \cdots - \theta_q a_{t-k-q}\right)\right\}$$
$$= \sigma_a^2 \left(-\theta_k + \sum_{j=1}^{q-k} \theta_j \theta_{j+k}\right), 1 \le k \le q$$

HOW?

What's the behavior of the ACF?

AutoRegressive Processes

4

Stationary Conditions for AR Processes

- An AR(p) process is expressed as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + a_t \text{ or}$$

$$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) y_t = a_t.$$

- AR(1): $(1 - \phi_1 B) y_t = a_t$

  $|\phi_1| < 1 \Rightarrow |B| > 1,$

  i.e., the root of the characteristic equation $1 - \phi_1 B = 0$ lies outside the unit circle.

Stationary Condition for AR Processes

- As for an AR(p) process: $\phi(B)y_t = a_t$, let's guess the roots of the characteristic equation $\phi(B) = 0$

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B),$$

  - where $G_1^{-1}, \dots, G_p^{-1}$ are the roots to $\phi(B) = 0$.

- We can also rearrange the process

$$y_t = \phi^{-1}(B) a_t = \sum_{i=1}^{p} \frac{K_i}{1 - G_i B} a_t \,.$$

  - If we want $\psi(B) = \phi^{-1}(B)$ to be convergent, i.e., if we want $\psi_j = \sum_{i=1}^{p} K_i G_i^j$ to be absolutely summable such that the AR(p) is stationary, we need

$$|G_i| < 1, i = 1, 2, \dots, p$$

  - equivalently, the roots to $\phi(B) = 0$ must lie outside the unit circle.

Invertibility of AR Processes

- Since $\psi(B)\phi(B) = 1$,

$$\psi_j = \phi_1\psi_{j-1} + \phi_2\psi_{j-2} + \cdots + \phi_p\psi_{j-p}, \text{ for } j > 0.$$

  - with $\psi_0 = 1$, and $\psi_j = 0$ for $j < 0$, we can solve

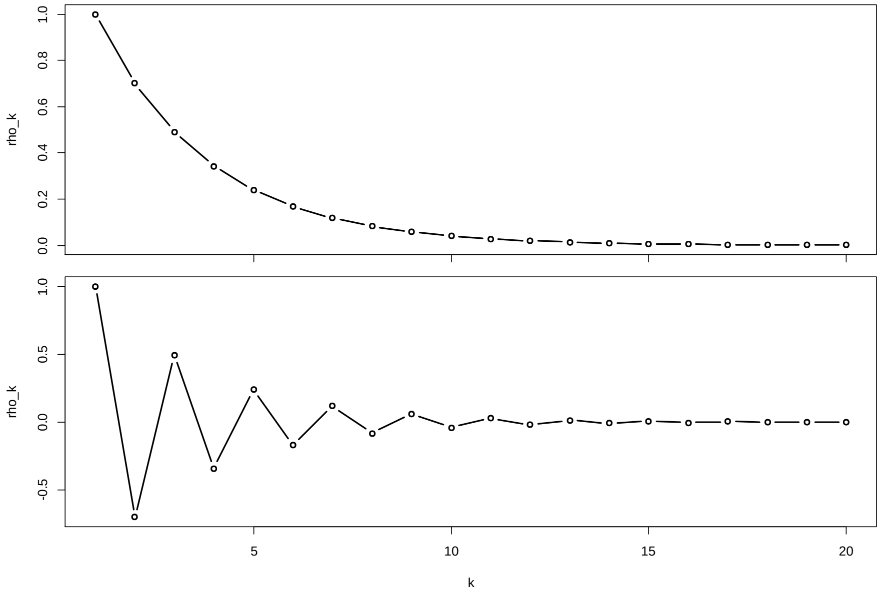
$$\psi_j = \sum_{i=1}^{p} K_i G_i^j \, .$$

First-Order AR Process: AR(1)

- AR(1) model: $y_t = \phi_1 y_{t-1} + a_t$,
  - consequently, $y_{t-1} = \phi_1 y_{t-2} + a_{t-1}$, i.e., y_{t-1} is independent of a_t
  - $\mathrm{E}[y_t] = ?$
  - $\mathrm{V}[y_t] = ?$
  - $\gamma_k = \mathrm{E}[y_t y_{t-k}] - \mathrm{E}[y_t]\mathrm{E}[y_{t-k}] = \mathrm{E}[y_t y_{t-k}] = ?$

The ACF, ρ_k, Behavior of AR(1)

- $\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k, \forall k = 0, 1, 2, \ldots$
  - $0 < \phi_1 < 1$
  - $-1 < \phi_1 < 0$



AR(2) Process

- AR(2) model: $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t$,
  - $\mathrm{E}[y_t] = 0$
  - $\mathrm{V}[y_t] = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma_a^2 = \gamma_0$

    $\gamma_0 = \frac{\sigma_a^2}{1-\phi_1\rho_1-\phi_2\rho_2}$

    HOW?

    where $\rho_1 = \frac{\phi_1}{1-\phi_2}$, $\rho_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2$

    $\mathrm{V}[y_t] = \frac{\sigma_a^2}{1-\phi_1\frac{\phi_1}{1-\phi_2}-\phi_2\left(\frac{\phi_1^2}{1-\phi_2}+\phi_2\right)} = \frac{(1-\phi_2)\sigma_a^2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)} > 0$

Scenarios for $\mathrm{V}[y_t] = \frac{(1-\phi_2)\sigma_a^2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)} > 0$

- $(1-\phi_2)(1+\phi_2) > 0$ and $1-\phi_1-\phi_2 > 0$ and $1+\phi_1-\phi_2 > 0$

$$1-\phi_2^2 > 0 \Rightarrow |\phi_2| < 1 \text{ and } \begin{cases} \phi_1+\phi_2 < 1 \\ \phi_2-\phi_1 < 1 \end{cases}$$

- $(1-\phi_2)(1+\phi_2) > 0$ and $1-\phi_1-\phi_2 < 0$ and $1+\phi_1-\phi_2 < 0$
  - no common solution set

- $(1-\phi_2)(1+\phi_2) < 0$ and $1-\phi_1-\phi_2 > 0$ and $1+\phi_1-\phi_2 < 0$

$$1-\phi_2^2 < 0 \Rightarrow |\phi_2| > 1 \text{ and } \begin{cases} \phi_1+\phi_2 < 1 \\ \phi_2-\phi_1 > 1 \end{cases}$$

  - However, $\rho_1 = \frac{\phi_1}{1-\phi_2}$ and $|\rho_1| < 1 \Rightarrow$ it conflicts against the conditions above.

Stationary Conditions for AR(2) Process

- For the AR(2) to be stationary, one needs $\begin{cases} |\phi_2| < 1 \\ \phi_1 + \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases}$

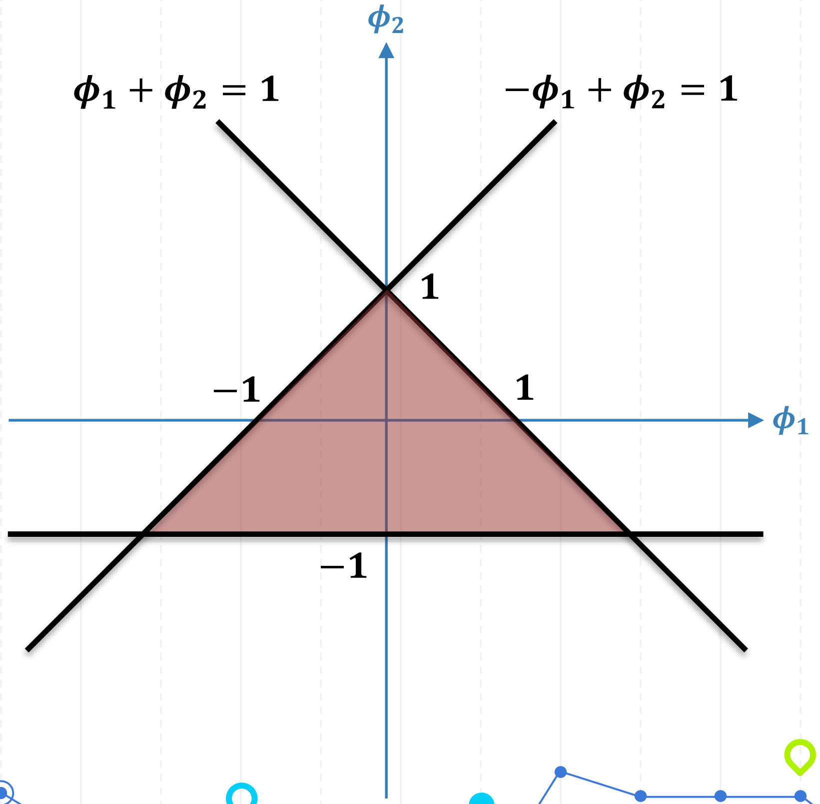
- This is equivalent to the conditions that make the solution to the "characteristic equation" of AR(2), i.e., the roots to $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$, outside the unit circle.

$$\Rightarrow |B_1| = \left| \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \right| > 1 \text{ and } |B_2| = \left| \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \right| > 1$$

The Region for AR(2) Stationarity

$$\begin{cases} |\phi_2| < 1 \\ \phi_1 + \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases}$$

- $|B_1| = \left| \dfrac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \right| > 1$
- $|B_2| = \left| \dfrac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \right| > 1$



ACF of AR(2)

$$\gamma_k = \mathrm{E}[y_t y_{t-k}] - \mathrm{E}[y_t]\mathrm{E}[y_{t-k}]$$
$$= \mathrm{E}[y_t y_{t-k}]$$
$$= \phi_1 \mathrm{E}[y_{t-1} y_{t-k}] + \phi_2 \mathrm{E}[y_{t-2} y_{t-k}] + \mathrm{E}[a_t y_{t-k}]$$
$$= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}$$

<u>dividing γ_0 on the both sides</u>

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \forall k$$

Yule-Walker Equations

$$\begin{cases} \rho_1 = \phi_1 \rho_0 + \phi_2 \rho_{-1} \\ \rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 \end{cases} \Rightarrow \begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 \\ \rho_2 = \phi_1 \rho_1 + \phi_2 \end{cases} \Rightarrow \begin{cases} \rho_1 = \dfrac{\phi_1}{1 - \phi_2} \\ \rho_2 = \dfrac{\phi_1^2}{1 - \phi_2} + \phi_2 \end{cases}$$

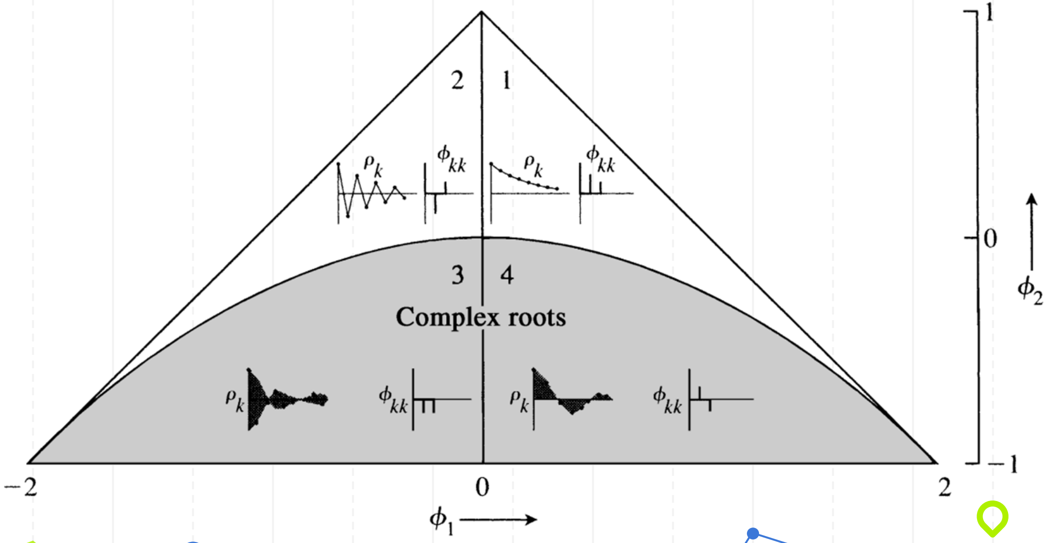
More Interpretations of AR(2) Stationarity

Stralkowski, C.M., Wu, S. M., DeVor, R. E. "Characterization of grinding wheel profiles by autoregressive-moving average models." *International Journal of Machine Tool Design and Research*, vol. 9, no. 2, 1969, pp. 145-163. DOI: 0020-7357(69)90013-4

- When roots of $\phi(B) = 0$ are real, ACFs are damped exponentials (1)(2).
- If $\phi_1^2 + 4\phi_2 < 0$, roots of $\phi(B) = 0$ are complex ➔ (3)(4), the ACFs are damped sine waves:

$$\rho_k = \frac{D^k \sin(2\pi f_0 k + F)}{\sin F},$$

  - $D = \sqrt{-\phi_2}$: damping factor;
  - f_0: frequency;
  - F: phase.



Extension to AR(p)

- AR(p) model: $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + a_t$
  - $\mathrm{E}[y_t] = 0$
  - $\mathrm{V}[y_t] = \phi_1 \mathrm{E}[y_{t-1} y_t] + \phi_2 \mathrm{E}[y_{t-2} y_t] + \cdots + \phi_p \mathrm{E}\left[y_{t-p} y_t\right] + \mathrm{E}[a_t y_t] = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma_a^2 = \gamma_0$
    $\left(1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \cdots - \phi_p \rho_p\right)\gamma_0 = \sigma_a^2$
    $\gamma_0 = \frac{\sigma_a^2}{\left(1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \cdots - \phi_p \rho_p\right)}$
  - $\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \Rightarrow \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$

Yule-Walker Equations of AR(p)

$$\begin{cases} \rho_1 = \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\ \rho_2 = \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\ \quad\vdots \\ \rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p \end{cases}$$

- If the operator B is applied

$$\rho_k = \phi_1 B\rho_k + \phi_2 B^2\rho_k + \cdots + \phi_p B^p\rho_k$$

$$\phi(B)\rho_k = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)\rho_k = 0, \forall k = 1, 2, \ldots$$

How to get $\{\rho_1, \rho_2, \ldots, \rho_p, \rho_{p+1}, \rho_{p+2}, \ldots\}$?

The Beauty of Yule-Walker Equations

$$\begin{cases} \rho_1 = \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\ \rho_2 = \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\ \vdots \\ \rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p \end{cases} \Rightarrow \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}$$

$$\boldsymbol{\rho} = \mathbf{P}\boldsymbol{\phi} \Rightarrow \boldsymbol{\phi} = \mathbf{P}^{-1}\boldsymbol{\rho}$$

- When playing with sample data, calculate the sample autocorrelations, and the estimates of ϕ_i can be solved as the regression coefficients.

Give it a Try: Yule-Walker Equations

- AR(1)
  - $\phi_1 = [1]^{-1}\rho_1 = \rho_1$

- AR(2)
  - $\begin{bmatrix}\phi_1 \\ \phi_2\end{bmatrix} = \begin{bmatrix}1 & \rho_1 \\ \rho_1 & 1\end{bmatrix}^{-1}\begin{bmatrix}\rho_1 \\ \rho_2\end{bmatrix} = \frac{1}{1-\rho_1^2}\begin{bmatrix}1 & -\rho_1 \\ -\rho_1 & 1\end{bmatrix}\begin{bmatrix}\rho_1 \\ \rho_2\end{bmatrix}$
  - $\begin{cases}\phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}\end{cases}$

As long as we can calculate $\hat{\rho}_i$...

Stationary Conditions for AR(p) Processes

- In a similar fashion, one can derive the stationary conditions of AR(p)

$$\left(1-\phi_1 B+\phi_2 B^2+\cdots+\phi_p B^p\right) y_t=a_t \Rightarrow \phi(B) y_t=a_t$$

$$p \text{ roots of } \phi(B)=0 \text{ should satisfy } |B_i|>1$$

- Recall the procedures: find the ranges of $\{\phi_1, \phi_2, \ldots, \phi_p\}$ for $\sigma_y^2>0$ and $|\rho_i|<1$. Since B_i are functions of $\{\phi_1, \phi_2, \ldots, \phi_p\}$, we get the aforementioned conditions.

How Do We Know the Value p in AR(p)?

- By numerating the value p from 1, 2, ..., p, we have p AR processes.
  - AR(1): $y_t = \phi_{11} y_{t-1} + a_t$;
  - AR(2): $y_t = \phi_{21} y_{t-1} + \phi_{22} y_{t-2} + a_t$;
  - AR(3): $y_t = \phi_{31} y_{t-1} + \phi_{32} y_{t-2} + \phi_{33} y_{t-p} + a_t$;

    ……
  - AR(p): $y_t = \phi_{p1} y_{t-1} + \phi_{p2} y_{t-2} + \cdots + \phi_{pp} y_{t-p} + a_t$.
- If the real $p = 2$, $\phi_{33}, \phi_{44}, \dots, \phi_{pp}$ shall be ZEROs.
  - How do we calculate $\{\phi_{11}, \phi_{22}, \phi_{33}, \dots, \phi_{pp}\}$?

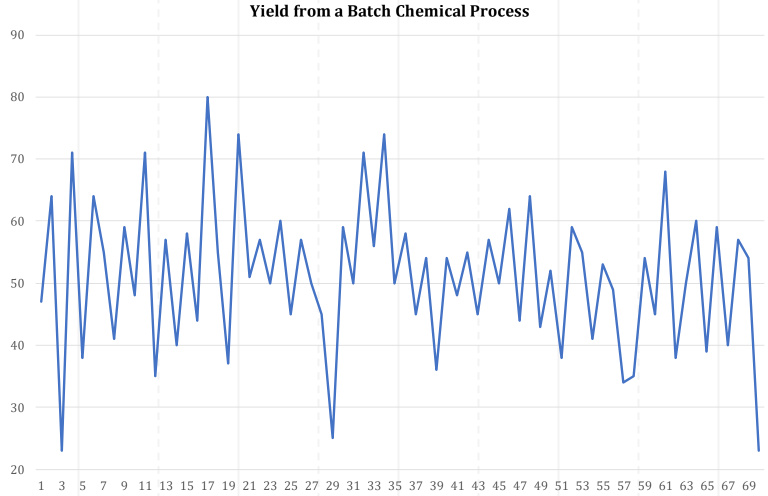
Partial Autocorrelation Function (PACF)

- $\{\phi_{11}, \phi_{22}, \phi_{33}, \dots, \phi_{pp}\}$ are defined as the PACF of y_t.
  - Can be solved by p sets of Yule-Walker equations.

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{p1} \\ \phi_{p2} \\ \vdots \\ \phi_{pp} \end{bmatrix}.$$

$$\phi_{11} = \rho_1; \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}; \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}.$$

Example for Applying PACF



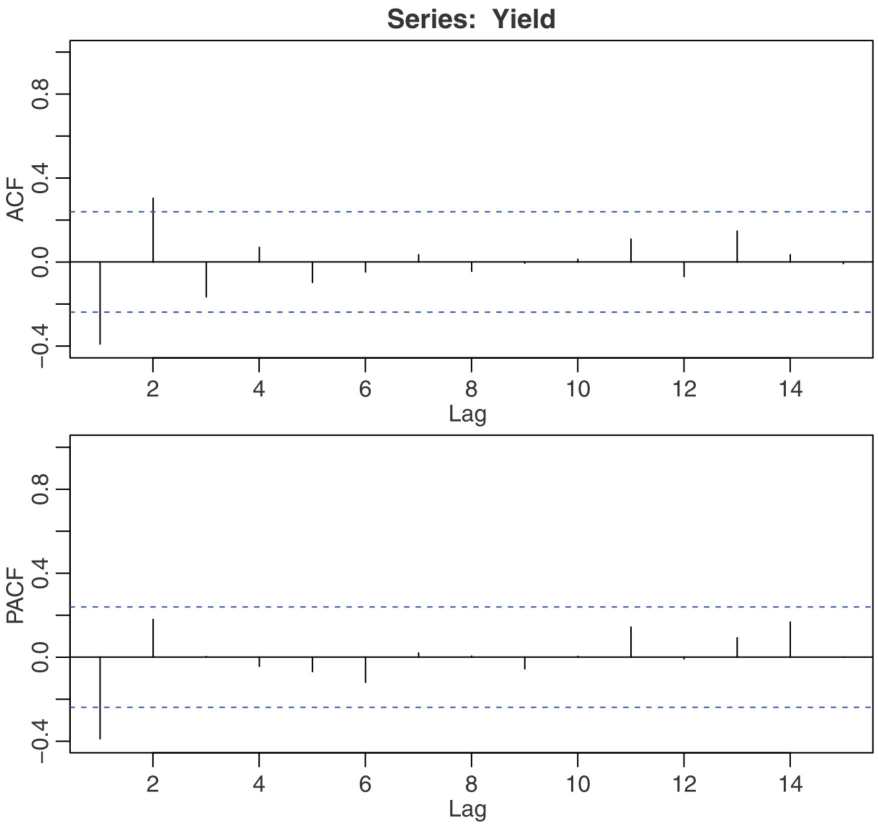
k	r_k	k	r_k	k	r_k
1	−0.39	6	−0.05	11	0.11
2	0.30	7	0.04	12	−0.07
3	−0.17	8	−0.04	13	0.15
4	0.07	9	0.00	14	0.04
5	−0.10	10	0.01	15	−0.01

k	$\hat{\phi}_{kk}$	k	$\hat{\phi}_{kk}$	k	$\hat{\phi}_{kk}$
1	−0.39	6	−0.12	11	0.14
2	0.18	7	0.02	12	−0.01
3	0.00	8	0.00	13	0.09
4	−0.04	9	− 0.06	14	0.17
5	−0.07	10	0.00	15	0.00

ACF vs. PACF: Yield Data

- What we need to determine the p?
  - $\mathrm{E}\left[\hat{\phi}_{kk}\right] = 0$
  - $\mathrm{V}\left[\hat{\phi}_{kk}\right] = \frac{1}{n}, k \geq p + 1$

Quenouille, M. H. "Approximate Tests of Correlation in Time-Series." *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 11, no. 1, 1949, pp. 68–84., www.jstor.org/stable/2983696.



Mix AR with MA
ARMA

(AutoRegressive Moving Average)

5

ARMA(p, q) Process

- An ARMA(p, q) process is expressed as:

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$
$$\Rightarrow \phi(B) y_t = \theta(B) a_t,$$

  - where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p;$$
$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q.$$

  - $\mathrm{E}[y_t] = 0$
  - $\mathrm{VAR}[y_t] = ?$
  - $\gamma_k = ?$

ACF of ARMA(p, q) is not trivial.

$$\gamma_k = \mathrm{E}[y_t y_{t-k}]$$
$$= \mathrm{E}\left[\phi_1 y_{t-1} y_{t-k} + \cdots + \phi_p y_{t-p} y_{t-k} + a_t y_{t-k} - \theta_1 a_{t-1} y_{t-k} - \cdots - \theta_q a_{t-q} y_{t-k}\right]$$
$$= \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{ay}(k) - \theta_1 \gamma_{ay}(k-1) - \cdots - \theta_q \gamma_{ay}(k-q),$$

- where $\gamma_{ay}(j) = \mathrm{cov}\left[a_t, y_{t-j}\right] = \mathrm{E}\left[a_t y_{t-j}\right]$.
  - $j > 0 \Rightarrow \gamma_{ay}(j) = 0$
  - $k > q \Rightarrow j > 0 \Rightarrow \gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p}$
  - $k \le q \Rightarrow \gamma_k$ is function of $\{\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q\}$
    - $k = q \Rightarrow \gamma_q = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{ay}(q) - \theta_1 \gamma_{ay}(q-1) - \cdots - \theta_q \gamma_{ay}(0)$
    - $k = q - 1 \Rightarrow \gamma_{q-1} = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{ay}(q-1) - \theta_1 \gamma_{ay}(q-2) - \cdots - \theta_{q-1} \gamma_{ay}(0) - \theta_q \gamma_{ay}(-1)$
    - $k = q - 2 \ldots$

ARMA(1, 1) Process

- ARMA(1, 1) model: $(1-\phi B)y_t = (1-\theta B)a_t$
  - Stationary Conditions:

$$y_t = (1-\phi B)^{-1}(1-\theta B)a_t = \left(1+\phi B+\phi^2 B^2+\cdots\right)(1-\theta B)a_t$$
$$= \left\{1+(\phi-\theta)B+\phi(\phi-\theta)B^2+\phi^2(\phi-\theta)B^3+\cdots\right\}a_t$$

  - Let $\psi_i = \phi^{i-1}(\phi-\theta)$

$$y_t = \left\{1+\psi_1 B+\psi_2 B^2+\psi_3 B^3+\cdots\right\}a_t$$

  - $\sum_{i=1}^{\infty}\psi_i^2 = (\phi-\theta)\sum_{i=1}^{\infty}\left(\phi_i^2\right)^{i-1} < \infty \Rightarrow |\phi| < 1$

ACF of ARMA(1, 1)

- $\gamma_k = \phi\gamma_{k-1} + \gamma_{ay}(k) - \theta\gamma_{ay}(k-1)$

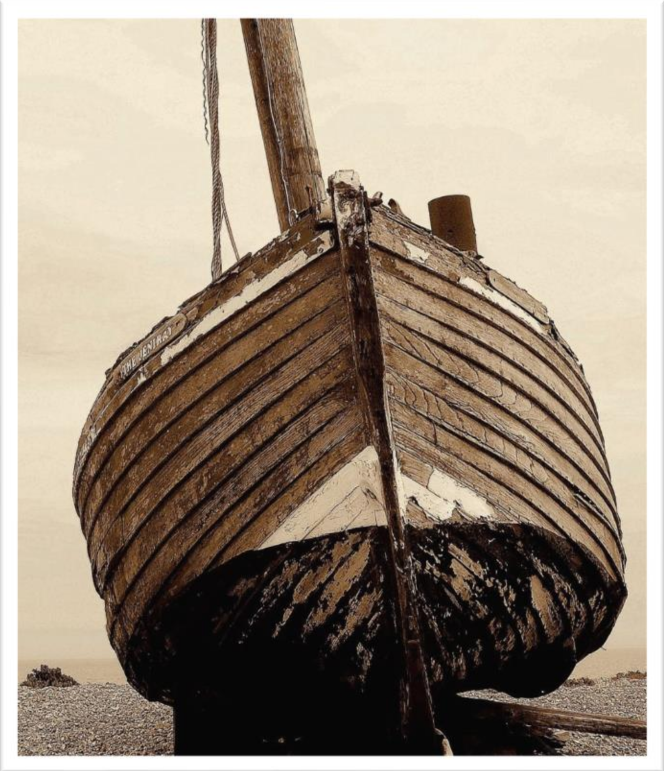
$$\gamma_0 = \phi\gamma_{-1} + \gamma_{ay}(0) - \theta\gamma_{ay}(-1) = \phi\gamma_1 + \sigma_a^2 - \theta(\phi-\theta)\sigma_a^2$$
$$\gamma_1 = \phi\gamma_0 + \gamma_{ay}(1) - \theta\gamma_{ay}(0) = \phi\gamma_0 + 0 - \theta\sigma_a^2$$
$$\gamma_2 = \phi\gamma_1 + \gamma_{ay}(2) - \theta\gamma_{ay}(1) = \phi\gamma_1 + 0 - 0$$
$$\vdots$$
$$\gamma_k = \phi\gamma_{k-1}$$

$$\Rightarrow \gamma_0 = \left(\frac{1-2\theta\phi+\theta^2}{1-\phi^2}\right)\sigma_a^2 \Rightarrow \gamma_1 = \left(\frac{(\phi-\theta)(1-\theta\phi)}{1-\phi^2}\right)\sigma_a^2$$

$$\Rightarrow \rho_1 = \frac{(\phi-\theta)(1-\theta\phi)}{1-2\theta\phi+\theta^2}; \rho_2 = \phi\rho_1; \ldots; \rho_k = \phi^{k-1}\rho_1 = \phi^{k-1}\frac{(\phi-\theta)(1-\theta\phi)}{1-2\theta\phi+\theta^2}$$

What's the behavior of this ACF?

After such a long journey…



we have successfully visited these islands.

- Stationarity and Invertibility
- MA Processes, always stationary
- AR Processes, always invertible
- ARMA, mixed feelings…