Statistical properties of linear time series

(Hamilton: Chapters 3.1-3.2)

Isabel Casas

SYDDANSK UNIVERSITET

Summary statistics, stationarity and ergodicity

- Definitions
  - Expectation, variance, covariance and autocorrelation
  - Stationarity
  - Ergodicity

Summary statistics: expectation

Expectations and variances:

- Consider I sequences of the process $\{Y_t\}_{t=-\infty}^{\infty}$
- For example, computer 1 generates a sequence $y^{(1)} = \{\ldots, y_{-1}^{(1)}, y_0^{(1)}, y_1^{(1)}, \ldots\}$, computer 2 generates $y^{(2)}$ .... computer I generates $y^{(I)}$
- Select the t-th observation from each of the I sequences. This gives the sample (denoted I realizations of the random variable $Y_t^{(i)}$)

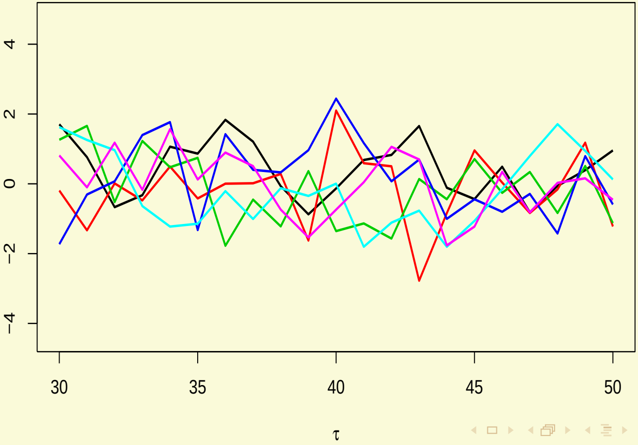
$$y_t^{(1)}, y_t^{(2)}, ..., y_t^{(I)}.$$

- Let $f_{Y_t}(y_t)$ denote the unconditional density of Y_t. The **unconditional expectation** of Y_t is then given as

$$E(Y_t) = \int y_t f_{Y_t}(y_t) dy_t,$$

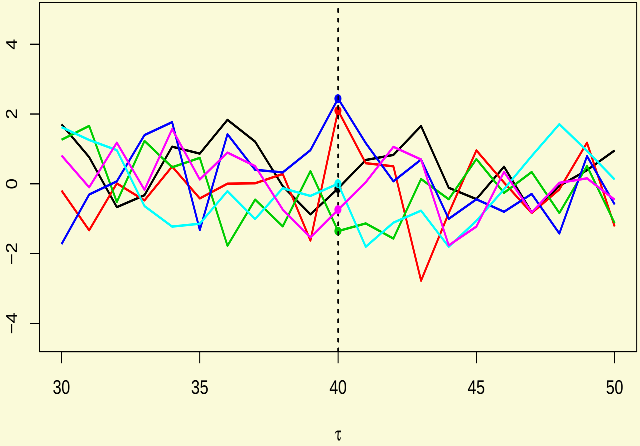
Summary statistics: 6 different processes

AR(1) $\quad y_t = 0.1 + 0.2 y_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$



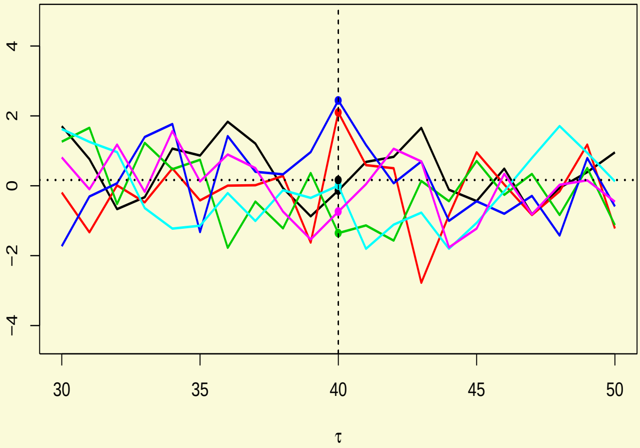
Summary statistics: ensemble average

Value at time $t = 40$ of each process



Summary statistics: ensemble average

The ensemble average at time $t = 40$



Summary statistics: expectation

- The unconditional expectation of Y_t is equal to the probability limit of the **ensemble average**

$$\begin{aligned} E(Y_t) &= \text{plim}\frac{1}{I}\sum_{i=1}^{I} Y_t^{(i)}, \\ &\equiv \mu_t \end{aligned}$$

Summary statistics: expectation

$$I = 1000$$

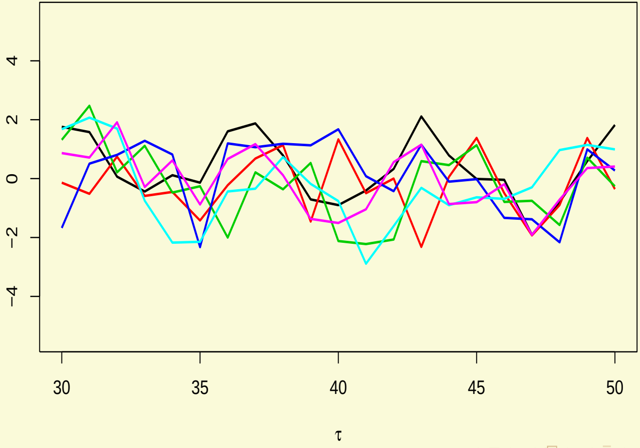
$$y_t = 0.1 + 0.2y_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$$

$$\mu_t = 0.1/(1-0.2) = 0.125$$

t	y.1	y.2	y.3	y.4	y.5	y.6	ensemble_mean
5	1.04	-0.09	0.34	1.21	-1.85	0.98	0.09
10	-0.33	-0.81	0.14	0.25	1.14	0.22	0.12
11	-0.74	0.59	-0.23	-0.02	-0.98	0.76	0.12
12	1.13	0.54	-2.12	-1.05	-1.89	-0.20	0.11
13	2.17	0.35	-0.03	-0.91	-1.92	0.17	0.12
45	-0.44	0.96	0.71	-0.44	-1.06	-1.23	0.16
50	0.96	-1.22	-1.13	-0.59	0.13	-0.46	0.13

Summary statistics: 6 different processes

AR(1) $z_t = \cos(t) + 0.2 z_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$



Summary statistics: expectation

$$I = 1000$$

$$z_t = \cos(t) + 0.2 z_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$$

$$\mu_t = \frac{1}{0.8}\cos(t)$$

t	z.1	z.2	z.3	z.4	z.5	z.6	mean
5	1.22	0.10	0.53	1.40	-1.66	1.17	0.28
10	-1.27	-1.75	-0.80	-0.69	0.20	-0.72	-0.82
11	-0.83	0.50	-0.32	-0.12	-1.08	0.66	0.02
12	1.87	1.29	-1.38	-0.31	-1.14	0.54	0.86
13	2.98	1.16	0.78	-0.10	-1.11	0.97	0.92
45	-0.01	1.38	1.13	-0.02	-0.64	-0.80	0.59
50	1.82	-0.36	-0.26	0.27	0.99	0.41	0.99

Exercise: expectation

Calculate the unconditional expectation of the following first-order moving average model:

$$\text{MA(1)} \qquad y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim N(2,1)$$

$$E(y_t) = E(3) + E(\epsilon_t) + 0.5E(\epsilon_{t-1}) = 3 + 2 + 1 = 6$$

What about the expectation of:

$$y_t = 6t + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim N(0,1)?$$

Summary statistics: variance

- The **unconditional variance** of Y_t

$$\gamma_{0t} \equiv E(Y_t - \mu_t)^2 = \int (y_t - \mu_t)^2 f_{Y_t}(y_t)\, dy_t$$
$$= \text{plim} \frac{1}{I} \sum_{i=1}^{I} (y_t^{(i)} - \mu_t)^2.$$

SYDDANSK UNIVERSITET

Summary statistics: variance

$$I = 1000$$

$$y_t = 0.1 + 0.2y_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$$

t	y.1	y.2	y.3	y.4	y.5	y.6	var
5	1.04	-0.09	0.34	1.21	-1.85	0.98	1.12
10	-0.33	-0.81	0.14	0.25	1.14	0.22	1.00
11	-0.74	0.59	-0.23	-0.02	-0.98	0.76	1.00
12	1.13	0.54	-2.12	-1.05	-1.89	-0.20	1.01
13	2.17	0.35	-0.03	-0.91	-1.92	0.17	1.07
45	-0.44	0.96	0.71	-0.44	-1.06	-1.23	1.06
50	0.96	-1.22	-1.13	-0.59	0.13	-0.46	0.94

Syddansk Universitet

Summary statistics: variance

$$I = 1000$$

$$z_t = \cos(t) + 0.2 z_{t-1} + \epsilon_t \quad \epsilon \sim N(0,1)$$

t	z.1	z.2	z.3	z.4	z.5	z.6	var
5	1.22	0.10	0.53	1.40	-1.66	1.17	1.12
10	-1.27	-1.75	-0.80	-0.69	0.20	-0.72	1.00
11	-0.83	0.50	-0.32	-0.12	-1.08	0.66	1.00
12	1.87	1.29	-1.38	-0.31	-1.14	0.54	1.01
13	2.98	1.16	0.78	-0.10	-1.11	0.97	1.07
45	-0.01	1.38	1.13	-0.02	-0.64	-0.80	1.06
50	1.82	-0.36	-0.26	0.27	0.99	0.41	0.94

SYDDANSK UNIVERSITET

Exercise: variance

Calculate the unconditional variance of the following first-order moving average model (MA(1)):

$$y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim N(2,1)$$

$$\begin{aligned} Var(y_t) =& Var(3) + Var(\epsilon_t) + Var(0.5\epsilon_{t-1}) \\ =& 0 + 1 + 0.5^2(1) = 1.25 \end{aligned}$$

What about the variance of:

$$y_t = 6t + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim N(0,1)?$$

Summary statistics: autocovariance

Autocovariances:

- Consider a particular realisation $\{y_t^{(i)}\}_{-\infty}^{\infty}$ and define the vector $\mathbf{x}_t^{(i)} = (y_t^{(i)}, y_{t-1}^{(i)}, ..., y_{t-j}^{(i)})'$
- We are interested in the distribution of $\mathbf{x}_t^{(i)}$ across realizations of i, i.e., the joint distribution of $Y_t, Y_{t-1}, ..., Y_{t-j}$, which we denote $F_{Y_t, Y_{t-1}, ..., Y_{t-j}}(\mathbf{x}_t^{(i)})$
- The jth autocovariance of Y_t, defined as $\gamma_{jt} \equiv E(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})$, equals (for a given realization of i)

$$\gamma_{jt} = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} (y_t^{(i)} - \mu_t)(y_{t-j}^{(i)} - \mu_{t-j}) f_{Y_t, ..., Y_{t-j}}(\mathbf{x}_t^{(i)}) \partial y_t^{(i)} ... \partial y_{t-j}^{(i)}$$

Syddansk Universitet

Summary statistics: autocovariance

- Again, we can think of the autocovariance γ_{jt} as the probability limit of an **ensemble average**, i.e,

$$\gamma_{jt} = \text{plim} \frac{1}{I} \sum_{i=1}^{I} (y_t^{(i)} - \mu_t)(y_{t-j}^{(i)} - \mu_{t-j}).$$

- Since $\gamma_{jt} = Cov(Y_t, Y_{t-j}) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, it is also called the **autocovariance function**

SYDDANSK UNIVERSITET

Exercise: autocovariance

Calculate the autocovariance of the following first-order moving average model (MA(1)):

$$y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim IIDN(2,1)$$

$$\begin{aligned}
\gamma_j = Cov(y_t, y_{t-j}) =& Cov(3 + \epsilon_t + 0.5\epsilon_{t-1}, 3 + \epsilon_{t-j} + 0.5\epsilon_{t-j-1}) \\
=& Cov(3,3) + Cov(3, \epsilon_{t-j}) + Cov(3, 0.5\epsilon_{t-j-1}) \\
& + Cov(\epsilon_t, 3) + Cov(\epsilon_t, \epsilon_{t-j}) + Cov(\epsilon_t, 0.5\epsilon_{t-j-1}) \\
& + Cov(0.5\epsilon_{t-1}, 3) + Cov(0.5\epsilon_{t-1}, \epsilon_{t-j}) + Cov(0.5\epsilon \\
=& 0 + 0 + 0 \\
& + 0 + Cov(\epsilon_t, \epsilon_{t-j}) + 0.5 Cov(\epsilon_t, \epsilon_{t-j-1}) \\
& + 0 + 0.5 Cov(\epsilon_{t-1}, \epsilon_{t-j}) + 0.5^2 Cov(\epsilon_{t-1}, \epsilon_{t-j-1})
\end{aligned}$$

Exercise: autocovariance

Calculate the **autocovariance** of the following first-order moving average model (MA(1)):

$$y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim IIDN(2,1)$$

If ϵ_t is iid or serially uncorrelated $\Rightarrow Cov(\epsilon_t, \epsilon_{t-j}) = 0$ for $j \neq 0$:

$$\begin{aligned} j = 0: &\quad \gamma_0 = Var(y_t) = Var(\epsilon_t) + 0.5^2\, Var(\epsilon_{t-1}) = 1.25 \\ j = 1: &\quad \gamma_1 = 0.5\, Var(\epsilon_t) = 0.5 \\ j = 2: &\quad \gamma_2 = 0 \end{aligned}$$

Summary statistics: autocorrelation

Autocorrelation or ACF:

- The **autocorrelation function** $\rho_{jt} = corr(Y_t, Y_{t-j}) : \mathbb{R} \times \mathbb{R} \to [-1, 1]$ is defined as

$$\rho_{jt} \equiv \frac{\gamma_{jt}}{\sqrt{\gamma_{0t}\gamma_{0t-j}}}.$$

SYDDANSK UNIVERSITET

Exercise: autocorrelation

Calculate the autocorrelation of the following first-order moving average model (MA(1)):

$$y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim IIDN(2,1)$$

We know that the autocovariance is:

$$\begin{aligned}
j = 0: &\quad \gamma_0 = Var(y_t) = Var(\epsilon_t) + 0.5^2\, Var(\epsilon_{t-1}) = 1.25 \\
j = 1: &\quad \gamma_1 = 0.5\, Var(\epsilon_t) = 0.5 \\
j = 2: &\quad \gamma_2 = 0
\end{aligned}$$

Q:What is the autocorrelation?

SYDDANSK UNIVERSITET

Exercise: autocorrelation in R

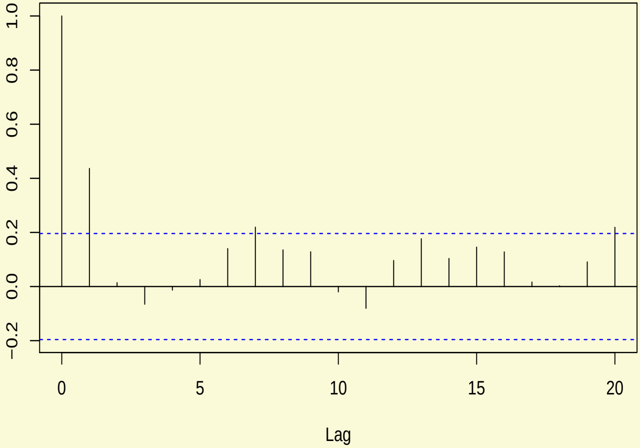
First we generate the process:

$$y_t = 3 + \epsilon_t + 0.5\epsilon_{t-1} \quad \epsilon \sim IIDN(2,1)$$

```
> epsilon <- rnorm(101, mean = 2, sd = 1)
> y.1 <- 3 + epsilon[2:101] + 0.5 * epsilon[1:100]
> y.1.acf <- acf(y.1)
> plot(y.1.acf)
> y.2 <- arima.sim(n = 100, list(ma = 0.5)) + 3
> y.2.acf <- acf(y.2)
```

The acf plot in the next page:

Exercise: autocorrelation in R



Question

Find the correlation of the AR(2) process:

$$y_t = 0.3 + 0.1 y_{t-1} - 0.5 y_{t-2} + \epsilon_t$$

Generate the process and plot its autocorrelation in R.

Weak stationarity

Stationarity:

- Definition (**weak stationarity**): If

$$\begin{aligned} E(Y_t) &= \mu \quad \forall t, \\ E(Y_{t-k}-\mu)(Y_{t-s}-\mu) &= \gamma_{|s-k|} \quad \forall t \end{aligned}$$

  then Y_t is said to be **covariance-stationary** also called weakly stationary)

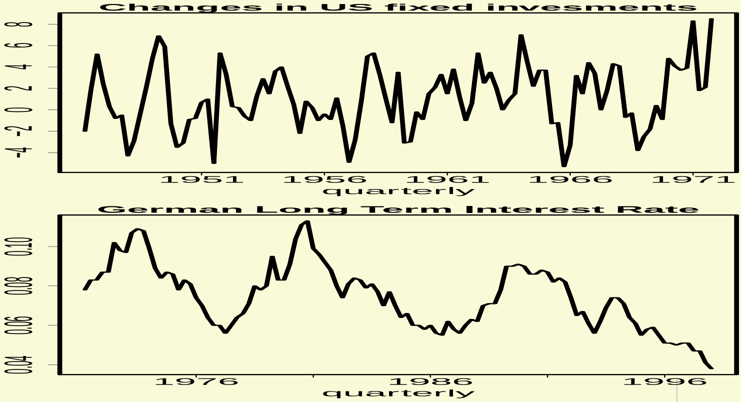
Weakly stationarity

A stochastic process whose mean function is constant and whose covariance function $\gamma_{|t-u|}$ depends of $|t-u|$ only and not separately of t and/or u is termed *weakly stationary* or *wide sense stationary* or *second-order stationary* or *covariance stationary*.

- In other words, its first and second moments are time–invariant.
- A time series generated by a (weakly) stationary process will fluctuate around the mean value, and does not have a trend.
- The variance are also time–invariant
- The covariances do not depend on time but only on the distance between the two observations
- All this implies that the first and second moments are finite, i.e. they exist.

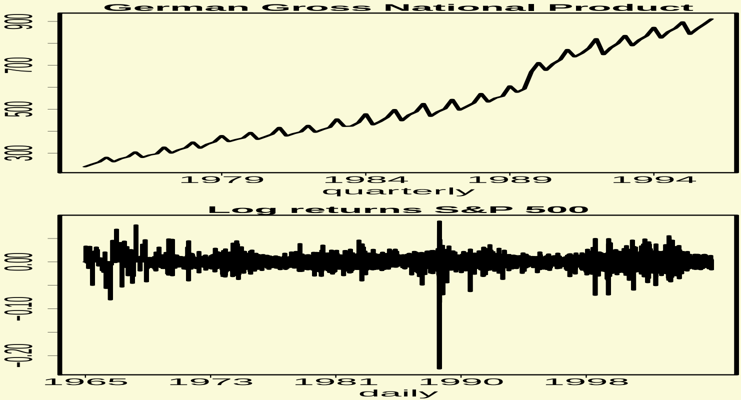
Weakly Stationarity

Do you think the following time series are generated by a weakly stationary process?



Weakly Stationarity

and these?

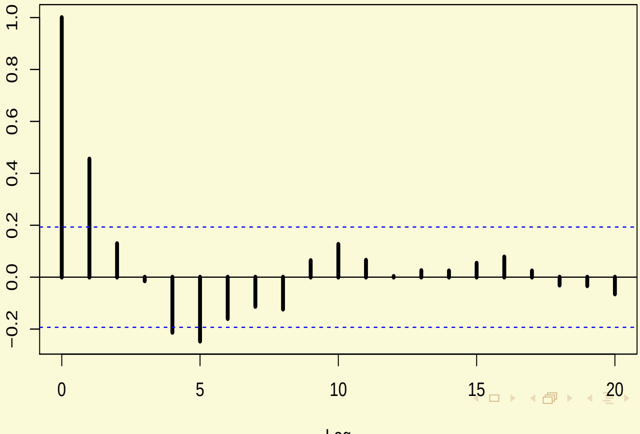


SYDDANSK UNIVERSITET

Weakly Stationarity

The sample autocorrelation of series with stationary DGP dies out quickly.

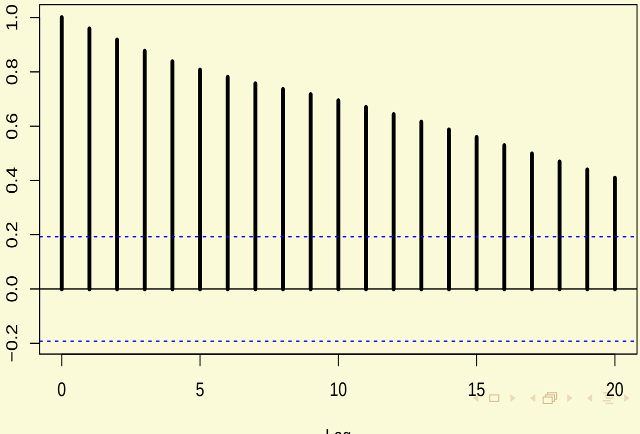
ACF of Quaterly Change in US fixed investments:



Weakly Stationarity

The sample autocorrelation of series with nonstationary DGP does not die out quickly.

ACF of Quaterly US fixed investments:



Weakly Stationarity

- So stationarity is rare in economic time series.
- However, some transformations (like the first difference) can help us to get stationary series.

Strict stationarity

- Definition (**strict stationarity**): The joint distribution of all collections of $(Y_t, Y_{t+1}, ..., Y_{t+k})$ for $k > 0$ do not depend (in any way) on t.
- Example: A covariance stationary Gaussian process is also strictly stationary since it is fully characterized by the mean and variance which by covariance stationarity are independent of t.
- This property is used in proofs but it is difficult to prove in practice.

SYDDANSK UNIVERSITET

Partial autocorrelation function

Partial Autocorrelation Function (conditional correlation) or PACF:

- This function gives the correlation between two random variables that are j periods apart when the in-between linear dependence (between t and $t-j$) is removed.
- Let Y_t and Y_{t-j} be two random variables. The PACF is then given as

$$\alpha_j = \rho_j(Y_t, Y_{t-j} | Y_{t-1}, ..., Y_{t-j+1})$$

- We need to define a linear projection.

Partial autocovariance function

The **projection** of Y_{t+1} onto the space spanned by $Y_t, Y_{t-1}, \ldots, Y_{t-k+1}$ is the best linear predictor of Y_{t+1} given $Y_t, Y_{t-1}, \ldots, Y_{t-k+1}$:

$$Y^*_{t+1|t} - \mu = \sum_{i=0}^{k-1} \alpha_{i+1}(Y_{t-i} - \mu)$$

where α_i minimises $E[(Y_{t+1} - \sum_{i=0}^{k-1} \alpha_{i+1} Y_{t-i})^2]$

The partial covariance between Y_t and Y_{t-j} with $j > 0$ is α_j the correlation between Y_t and Y_{t-j} conditioning out all variables in between.

SYDDANSK UNIVERSITET

Partial autocorrelation function

Find the partial autocovariance between Y_t and Y_{t-2} for the process:

$$Y_t = 0.5\,Y_{t-1} + \epsilon_t \quad \epsilon_T \sim N(0,1)$$

The linear projection of Y_t onto Y_{t-1} is $0.5\,Y_{t-1}$
The linear projection of Y_{t-2} onto Y_{t-1} is $\frac{E(Y_{t-1}E(Y_{t-2}))}{E(Y_{t-1}^2)}$

$$\alpha_2 = Cov(Y_t - Y^*_{t|t-1}, Y_{t-2} - Y^*_{t-2|t-1}) = 0$$

Partial autocorrelation function

Find the partial autocorrelation between Y_t and Y_{t-2}, i.e. γ_2:

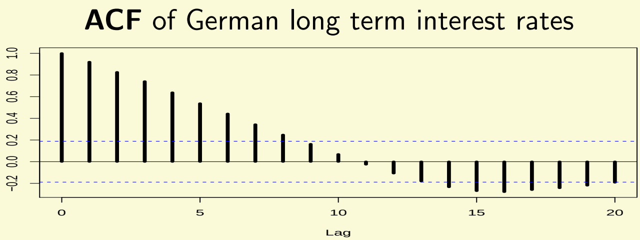
$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 \\ \gamma_1, \gamma_0 & \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$$

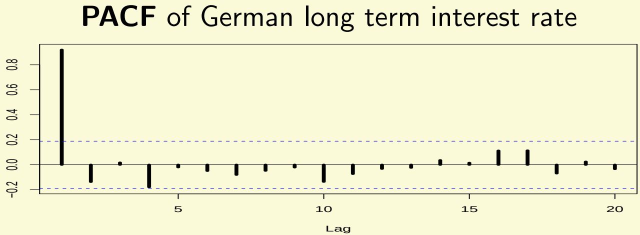
In our example $\gamma_0 = \frac{4}{3}$, $\gamma_1 = \frac{2}{3}$, $\gamma_2 = \frac{1}{3}$

Partial autocorrelation function

```
> Gamma = matrix(c(4/3, 2/3, 2/3, 4/3),
ncol = 2, byrow = T)
> gamma = c(2/3, 1/3)
> solve(Gamma) %*% gamma
     [,1]
[1,]  0.5
[2,]  0.0
```

PACF





Note: Initial point

Ergodicity

- A time series is a single realisation of the generating stochastic process.
- Can the sample mean of this realisation gives us any information about the ensemble mean at each point in time?
- What about the variance?
- The answer to these two questions is yes if the process is ergodic

Ergodicity

- The introduction of the ensemble averages serves to understand the concepts of ergodicity.
- Consider the sequence $\{y_t^{(1)}\}_{t=1}^T$. From this sequence we could compute the time average and time covariance:

$$\begin{aligned}\overline{y} =& \frac{1}{T}\sum_{t=1}^{T} y_t^{(1)}, \\ \overline{\gamma}_j =& \frac{1}{T-j}\sum_{t=j+1}^{T}(y_t^{(1)}-\mu)(y_{t-j}^{(1)}-\mu)\end{aligned}$$

- Does $\overline{y}$ converge to $E(Y_t)$? Does $\overline{\gamma}_j$ converges to γ_j? Is the process ergodic?

Ergodicity

A covariance stationary process is said to be *ergodic for the mean* if

$$\overline{y} \rightarrow^{P} E(Y_t) \quad \text{as } T \rightarrow \infty.$$

Comment:

If

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

for a covariance stationary process Y_t, then it will be ergodic in the mean. This condition (sufficient) is referred to as "absolute summability of the autocovariances". Proven in Chapter 7 of Hamilton.

Ergodicity

A covariance stationary process is said to be *ergodic for the second moment* if

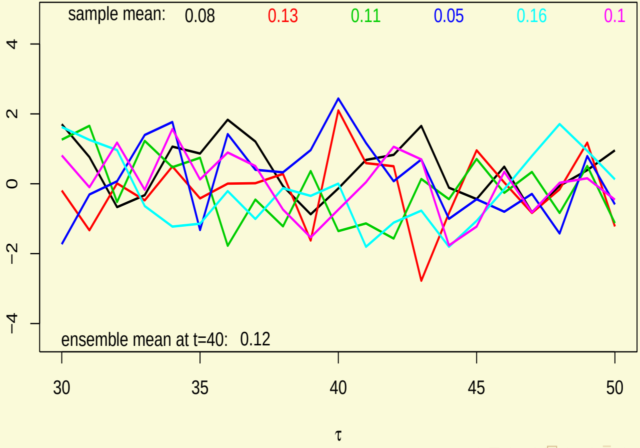
$$\overline{\gamma}_j \to^P \gamma_j, \quad \forall j.$$

Comment:

If Y_t is a stationary Gaussian process, absolute summability of the covariances is sufficient to ensure ergodicity for all the moments. Proven in Chapter 7 of Hamilton.

Example: A stationary process that is ergodic in mean

AR(1) $\quad y_t = 0.1 + 0.2y_{t-1} + \epsilon_t \quad \epsilon_t \sim N(0,1)$



Example: A stationary process that is ergodic in mean

$$\text{AR(1)} \qquad y_t = 0.1 + 0.2y_{t-1} + \epsilon_t \quad \epsilon_t \sim N(0,1)$$

$$\begin{aligned} \mu_t = E(y_t) =& 0.1 + 0.2E(y_{t-1}) + E(\epsilon_t) \\ (1-0.2)\mu_t =& 0.1 \\ \mu_t = \mu =& 0.1/0.8 = 0.125 \end{aligned}$$

$$\begin{aligned} \overline{y} =& \frac{1}{T}\sum_{t=1}^{T} y_t = 0.1 + \frac{0.2}{T}\sum_{t=1}^{T} y_{t-1} + \frac{1}{T}\sum_{t=1}^{T}\epsilon_t \\ \overline{y} \rightarrow \mu & \quad \text{as } T \rightarrow \infty \end{aligned}$$

Ergodicity

If $\{ Y_t \}$ is a stationary iid process, for example a stationary Gaussian process, then it is ergodic.

The constant process is also ergodic.

Processes that are strictly stationary and whose autocorrelation converges sufficiently fast are ergodic in mean

Do we find ergodicity in reality? It is not sure!

Some people uses certain transformations to obtain ergodicity

Example: A stationary process that is not ergodic in mean

$$Y_t^{(i)} = \mu^{(i)} + \epsilon_t \quad \epsilon_t \sim IIDN(0,1) \quad \mu^{(i)} \sim N(0,\lambda^2)$$

$$\begin{aligned} E(Y_t^{(i)}) =& E(\mu^{(i)}) + E(\epsilon_t) = 0 \\ \gamma_{0t} =& E(\mu^{(i)} + \epsilon_t)^2 = \lambda^2 + \sigma^2 \\ \gamma_{jt} =& E(\mu^{(i)} + \epsilon_t)E(\mu^{(i)} + \epsilon_{t-j}) = \lambda^2 \quad \text{for } j \neq 0 \end{aligned}$$

$$\overline{Y}_t^{(i)} = \frac{1}{T}\sum_{t=1}^{T} Y_t^{(i)} = \mu_t^{(i)} + \frac{1}{T}\sum_{t=1}^{T}\epsilon_t \rightarrow_{T\rightarrow\infty} \mu^{(i)}$$

Syddansk Universitet