7. Stochastische Prozesse und Zeitreihenmodelle

➔ Regelmäßigkeiten in der Entwicklung einer Zeitreihe, um auf zukünftige Entwicklung zu schließen

➔ Verwendung zu Prognosezwecken

Univariate Zeitreihenanalyse

➔ für Prognose werden nur endogene Variablen verwendet, d.h. nur Beobachtungswerte derselben statistischen Variablen grundlegendes Konzept ist der stochastische Prozess:

Definition:

Eine zeitlich geordnete Folge von Zufallsvariablen

$$\{Y\} = Y_1, Y_2, Y_3, \ldots, Y_t, \ldots$$

heißt stochastischer Prozess (Sequenz von endlich vielen oder (abzählbar) unendlich vielen Zufallsvariablen bzw. Wertemenge), $\{Y \mid \text{für alle ganzen Zahlen } t\}$.

Zeitreihenmodelle basieren auf der Annahme, dass die zu prognostizierende Zeitreihe durch einen stochastischen Prozess generiert wird. Bei endlich vielen

ZVs kann der stochastische Prozess durch gemeinsame Verteilungsfunktion vollständig angegeben werden.

Realisation der Länge T des stochastischen Prozess: $y_1, y_2, y_3, \ldots, y_T$, wobei jeder Beobachtungswert von einer anderen ZV generiert ist, d.h. jeder beobachtete Wert einer Zeitreihe entstammt zufällig einer Wahrscheinlichkeitsverteilung. Ein stochastischer Prozess stellt einen Zufallsprozess $\{Y\}$ dar, welcher Werte zwischen $-\infty$ und $+\infty$ annehmen kann. Der beobachtete Wert y_T zum Zeitpunkt T beschreibt eine Realisation dieser stochastischen Prozesse.

7.1 Kennzahlen stochastischer Prozesse

I.d.R. genügt die Angabe der ersten und zweiten Momente der ZV, womit der Prozess hinreichend charakterisiert ist. Erwartungswerte, Varianzen und Kovarianzen sind als eine Funktion von t und j zu sehen.

Definition:

1. $\mu(t) := E(Y_t)$ Mittelwertfunktion
2. $\sigma^2(t) := V(Y_t)$, Varianzfunktion
3. $\gamma_j(t) =: Cov(Y_t, Y_{t-j})$, Autokovarianzfunktion, mit $\sigma^2(t) =: \gamma_0(t)$
4. $\rho_j(t) =: \dfrac{\gamma_j(t)}{\sigma(t) \cdot \sigma(t-j)}$ Autokorrelationsfunktion

des stochastischen Prozesses $\{Y_t\}$

7.2 Stationäre stochastische Prozesse

Ann.: zeitliche Stabilität des stochastischen Prozesses (bestimmte mögliche Prozesse werden vorgegeben)

Sind die ersten beiden Momente eines stochastischen Prozesses unabhängig vom Zeitindex, so ist der Prozess stationär.

Definition: Schwache Stationarität

Ein stochastischer Prozess $\{Y_t\}$ heißt schwach stationär, wenn für alle t und j gilt:

1. $\mu(t) = \mu$ oder $E(Y_t) = \mu = const., \; für \; alle \; t$ mittelwertstationär
2. $\sigma^2(t) = \sigma^2$ oder $\sigma_t^2 = \sigma^2 = const., \; für \; alle \; t$ varianzstationär
3. $\gamma_j(t) = \gamma_j$ oder $Cov(Y_t, Y_{t-j}) = \sigma_{ij} = \sigma_j = const., \; für \; alle \; t$ kovarianzstat.

Kovarianzen hängen lediglich von Entfernung j ab, aber nicht vom Zeitindex t

Bei stationären Prozessen gilt für jedes j:

$$\rho_j = \frac{\gamma_j}{\gamma_0}, \text{ mit } -1 \leq \rho_j \leq +1$$

<u>Definition:</u> Strenge Stationarität

Prozess, bei dem gemeinsame Verteilungsfunktion jedes endlichen zusammenhängenden Teiles von ZVs $Y_1, Y_2, Y_3, \ldots, Y_m$ gleich der Verteilungsfunktion des um k Zeitpunkte verschobenen Teiles $Y_{1+k}, Y_{2+k}, Y_{3+k}, \ldots, Y_{m+k}$ ist, heißt streng stationärer stochastischer Prozess, d.h. gemeinsame Verteilung ist zeitinvariant. $Y(t+1), \ldots, Y(t+m)$ hat die gleiche gemeinsame Verteilung wie $Y(t+1+k), \ldots, Y(t+m+k)$, wobei k eine willkürliche positive ganze Zahl ist.

Alle Y_t sind identisch verteilt, aber nicht notwendigerweise unabhängig.

Strenge Stationarität ist eine stärkere Eigenschaft. Jeder streng stationäre stochastische Prozess ist auch schwach stationär.

<u>Aber:</u>

Eigenschaft der Stationarität ist nicht hinreichend, um konsistente Schätzer zu erhalten.

<u>Vorauss.:</u> unabhängige Stichproben <u>nicht</u> gegeben (Abhängigkeitsstrukturen)

Definition eines ergodischen Prozesses:

Ein stationärer stochastischer Prozess $\{Y_t\}$ mit Mittelwertfunktion μ und Kovarianzfunktionen γ_j heißt

1. mittelwertergodisch, wenn $\lim\limits_{T\to\infty} E\left(\frac{1}{T}\sum_{t=1}^{T} Y_t\right) = \mu$

2. kovarianzergodisch, wenn für alle j

$$\lim_{T\to\infty} E\left(\frac{1}{T}\sum_{t=1}^{T}(Y_t-\mu)(Y_{t-j}-\mu)\right) = \gamma_j$$

Autokovarianz sollte nicht zu groß sein bzw. sollte mit zunehmender Entfernung j schnell genug kleiner werden.

Schätzen der Kennzahlen

empirische Momente eines stationären Prozesses können konsistent geschätzt werden:

Erwartungswert: $\hat{\mu} = \bar{y}$

Autokovarianzen: $\hat{\gamma}_j = \frac{1}{T-j-1} \sum_{t=j+1}^{T} (y_t - \bar{y})(y_{t-j} - \bar{y})$

Autokorrelationsfunktionen: $\hat{\rho}_j = r_j = \frac{\sum_{t=j+1}^{T} (y_t - \bar{y})(y_{t-j} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$, für $j = 0, 1, 2, \ldots$

(empirische Autokorrelationsfunktion)

EViews

File Edit Objects View Procs Quick Options Window Help

Correlogram of LNDAXTRSERIES

Date: 10/26/04 Time: 19:31
Sample: 1964:12 2003:12
Included observations: 469

	AC	PAC	Q-Stat	Prob
1	0.995	0.995	466.89	0.000
2	0.989	0.000	929.72	0.000
3	0.984	-0.015	1388.4	0.000
4	0.978	0.008	1843.0	0.000
5	0.973	-0.022	2293.4	0.000
6	0.967	0.001	2739.5	0.000
7	0.961	-0.014	3181.4	0.000
8	0.956	0.034	3619.4	0.000
9	0.951	0.002	4053.6	0.000
10	0.946	0.081	4484.7	0.000
11	0.941	-0.065	4912.2	0.000
12	0.936	-0.050	5335.6	0.000
13	0.930	-0.043	5754.6	0.000
14	0.924	-0.027	6168.9	0.000
15	0.918	0.035	6579.0	0.000
16	0.913	0.060	6985.3	0.000
17	0.906	-0.111	7386.9	0.000
18	0.900	-0.005	7783.8	0.000
19	0.893	-0.105	8174.9	0.000
20	0.885	-0.051	8560.2	0.000
21	0.877	0.031	8939.8	0.000
22	0.870	-0.034	9313.6	0.000
23	0.862	0.004	9681.6	0.000
24	0.854	0.002	10044.	0.000
25	0.846	-0.036	10400.	0.000
26	0.838	-0.007	10750.	0.000
27	0.831	0.061	11095.	0.000
28	0.824	0.006	11435.	0.000
29	0.817	-0.039	11770.	0.000
30	0.809	-0.001	12099.	0.000
31	0.801	-0.020	12423.	0.000
32	0.793	-0.003	12741.	0.000
33	0.786	0.006	13053.	0.000
34	0.778	0.000	13361.	0.000
35	0.770	0.019	13663.	0.000
36	0.762	-0.042	13959.	0.000

Correlogram of DLNDAXTR

Date: 10/26/04 Time: 19:32
Sample: 1964:12 2003:12
Included observations: 468

	AC	PAC	Q-Stat	Prob
1	0.040	0.040	0.7573	0.384
2	0.026	0.025	1.0847	0.581
3	0.038	0.036	1.7557	0.625
4	-0.008	-0.011	1.7843	0.775
5	-0.006	-0.007	1.8037	0.876
6	0.055	0.054	3.2238	0.780
7	-0.007	-0.010	3.2446	0.861
8	-0.003	-0.004	3.2480	0.918
9	-0.054	-0.057	4.6255	0.866
10	0.087	0.094	8.2614	0.603
11	0.024	0.021	8.5445	0.664
12	0.054	0.049	9.9323	0.622
13	-0.044	-0.058	10.877	0.621
14	-0.042	-0.041	11.722	0.629
15	-0.075	-0.067	14.451	0.492
16	0.012	0.016	14.520	0.560
17	-0.023	-0.020	14.785	0.611
18	0.058	0.058	16.430	0.563
19	-0.052	-0.045	17.755	0.539
20	-0.100	-0.102	22.678	0.305
21	0.022	0.036	22.907	0.349
22	0.024	0.015	23.189	0.391
23	-0.014	-0.003	23.281	0.444
24	0.053	0.041	24.657	0.425
25	-0.009	0.012	24.695	0.480
26	-0.070	-0.065	27.161	0.401
27	0.005	0.016	27.175	0.454
28	-0.071	-0.100	29.730	0.376
29	-0.042	-0.035	30.601	0.384
30	-0.028	-0.017	30.999	0.415
31	-0.002	0.023	31.002	0.466
32	0.036	0.049	31.651	0.484
33	0.008	0.002	31.684	0.533
34	-0.012	-0.037	31.758	0.578
35	0.067	0.055	34.051	0.514
36	0.035	0.047	34.692	0.531

Path = c:\windows DB = none WF = untitled

Signifikanztest:

(Autokorelationskoeffizienten signifikant von Null verschieden?)

$$H_0 : \rho_j = 0$$

ZV R_j ist unter H_0 annähernd normalverteilt mit der Varianz $1/T$. Man verwendet die Faustregel:

$$|r_j| > \frac{2}{\sqrt{T}} \quad \Rightarrow \quad H_0 \; verwerfen$$

≙ zweiseitiger Test auf 5%-Signifikanzniveau ($\alpha = 0.05$).

White-Noise- und Normalprozess

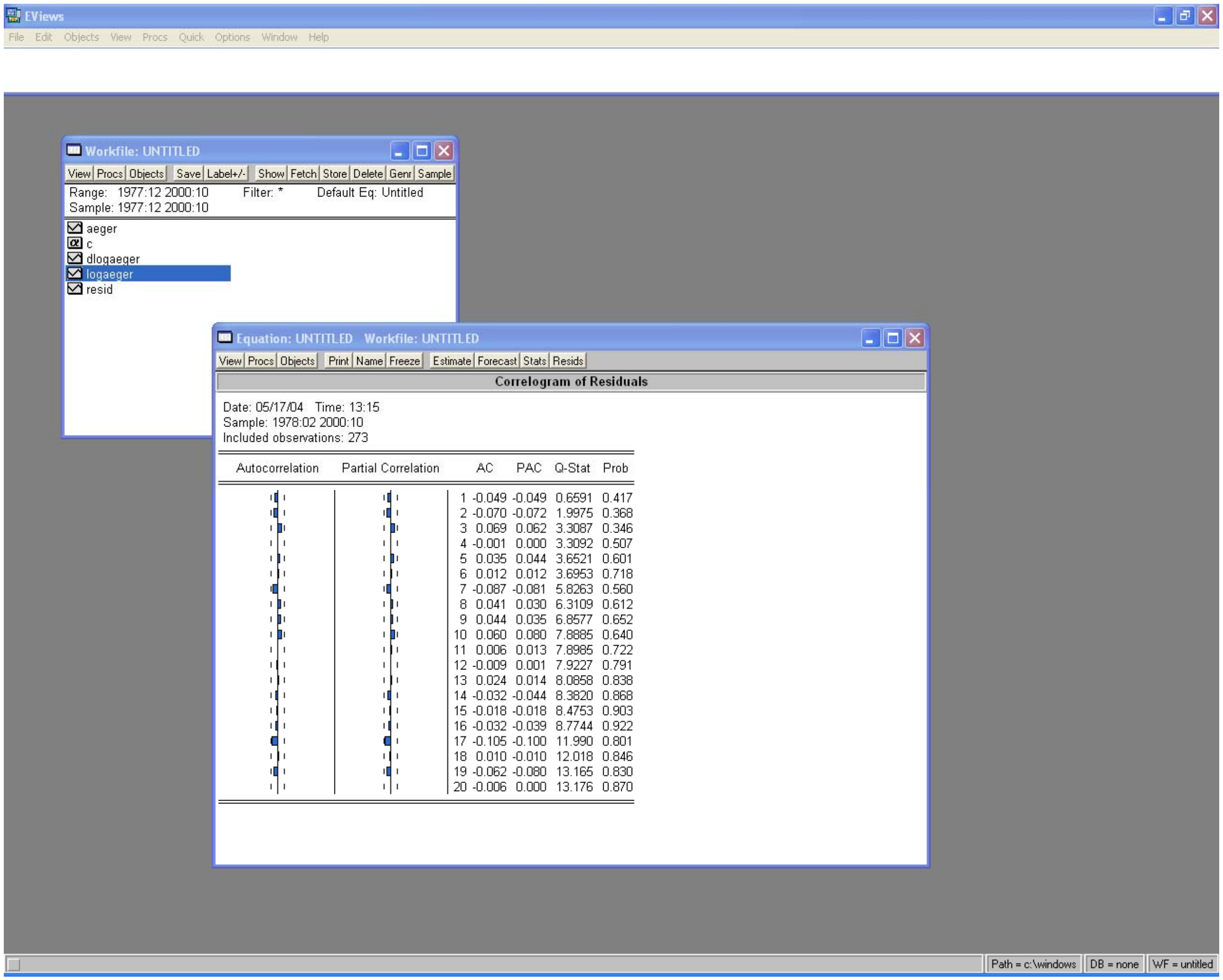
= einfacher stochastischer Prozess

Definition:

Ein stochastischer Prozess $\{\varepsilon_t\} = \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_t, \ldots$ heißt weißes Rauschen, White-Noise-Prozess, oder reiner Zufallsprozess, wenn

1. $E(\varepsilon_t) = 0$
2. $V(\varepsilon_t) = \sigma_\varepsilon^2 = const.$ für jedes t
3. $Cov(\varepsilon_t, \varepsilon_{t-j}) = 0$ für $j > 0$

White-Noise-Prozesse sind stationär. Sind ε_t eines weißen Rauschens normalverteilt, so spricht man von Normalprozess oder Gaussian White Noise.



	AC	PAC	Q-Stat	Prob
1	-0.049	-0.049	0.6591	0.417
2	-0.070	-0.072	1.9975	0.368
3	0.069	0.062	3.3087	0.346
4	-0.001	0.000	3.3092	0.507
5	0.035	0.044	3.6521	0.601
6	0.012	0.012	3.6953	0.718
7	-0.087	-0.081	5.8263	0.560
8	0.041	0.030	6.3109	0.612
9	0.044	0.035	6.8577	0.652
10	0.060	0.080	7.8885	0.640
11	0.006	0.013	7.8985	0.722
12	-0.009	0.001	7.9227	0.791
13	0.024	0.014	8.0858	0.838
14	-0.032	-0.044	8.3820	0.868
15	-0.018	-0.018	8.4753	0.903
16	-0.032	-0.039	8.7744	0.922
17	-0.105	-0.100	11.990	0.801
18	0.010	-0.010	12.018	0.846
19	-0.062	-0.080	13.165	0.830
20	-0.006	0.000	13.176	0.870

7.3 Moving-Average-Prozesse

Autoregressive Moving Average Modelle = ARMA-Modelle

➔ für Kurzfristprognose von guter Qualität

Drei grundlegende Modelltypen

- Autoregressiver Prozess *p*-ter Ordnung, AR(p)
- Moving-Average Prozess *q*-ter Ordnung, MA(q)
- ARMA-Prozess (*p,q*)-ter Ordnung, ARMA(p,q)

außerdem: ARIMA(*p,d,q*)-Modell (*d* = Anzahl der Differenzenbildung)

MA(1)-Prozess

Über Bildung gleitender Durchschnitte über eine White-Noise-Prozess erhält man einen neuen stochastischen Prozess

Definition:

Ein stochastischer Prozess $\{Y_t\}$ mit

$$Y_t := \alpha_0 + \varepsilon_t + \alpha_1 \varepsilon_{t-1}$$

wobei α_0 und α_1 Parameter sind und die ε_t einen White-Noise-Prozess mit $V(\varepsilon_t) = \sigma_\varepsilon^2$ bilden, heißt Moving-Average Prozess erster Ordung oder MA(1)-Prozess.

Eigenschaften:

1. $\mu(t) = E(\alpha_0 + \varepsilon_t + \alpha_1\varepsilon_{t-1}) = \alpha_0$

2. $\gamma_0(t) = V(\alpha_0 + \varepsilon_t + \alpha_1\varepsilon_{t-1}) = V(\varepsilon_t) + V(\alpha_1\varepsilon_{t-1}) = (1+\alpha_1^2)\sigma_\varepsilon^2$

3.
$$\begin{aligned}
\gamma_1(t) &= E\left[Y_t - \alpha_0)(Y_{t-1} - \alpha_0)\right] \\
&= E\left[(\varepsilon_t + \alpha_1\varepsilon_{t-1})(\varepsilon_{t-1} + \alpha_1\varepsilon_{t-2})\right] \\
&= E(\varepsilon_t\varepsilon_{t-1}) + \alpha_1 E(\varepsilon_{t-1}\varepsilon_{t-1}) + \alpha_1 E(\varepsilon_t\varepsilon_{t-2}) + \alpha_1^2 E(\varepsilon_{t-1}\varepsilon_{t-2}) \\
&= 0 + \alpha_1 E(\varepsilon_{t-1}\varepsilon_{t-1}) + 0 + 0 \\
&= \alpha_1\sigma_\varepsilon^2
\end{aligned}$$

Bei MA(1)-Prozess gibt es nur Abhängigkeit der ZV von ihrer jeweils unmittelbaren Vorgängerin. Sämtliche Kovarianzen höherer Ordnung als 1 verschwinden:

$$\gamma_j = \begin{cases} (1+\alpha_1)\sigma_\varepsilon^2 & \text{für } j=0 \\ \alpha_1\sigma_\varepsilon^2 & \text{für } j=1 \\ 0 & \text{für } j>1 \end{cases}$$

Kovarianzfunktion ist unabhängig von t,

MA(q)-Prozess ist immer stationär

Korrelationsfunktion eines MA(1)-Prozesses:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \begin{cases} 1 & \text{für } j = 0 \\ \dfrac{\alpha_1}{1+\alpha_1^2} & \text{für } j = 1 \\ 0 & \text{für } j > 1 \end{cases}$$

d.h. nimmt für Lags größer 1 den Wert Null an

<u>MA(q)-Prozesse</u>

gleitende Durchschnitte mit mehr als zwei White-Noise-Variablen

$$Y_t := \alpha_0 + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \ldots + \alpha_q \varepsilon_{t-q}$$

Erwartungswert ist konstant und α_0

Merke:

1. Jeder MA(q)-Prozess ist schwach stationär
2. Seine Kovarianzfunktion bricht ab für Lags $j > q$, so dass $\gamma_j = 0$
3. Jeder MA(q)-Prozess ist mittelwertergodisch
4. MA(q)-Prozess ist kovarianzergodisch, wenn das ihn erzeugende weiße Rauschen ein Normalprozess ist.

7.4 Autoregressive Prozesse

autoregressiver Prozess erster Ordnung oder AR(1)-Prozess

= einfachster Prozess von AR-Prozessen bei dem aktueller Wert vom vorherigen Wert bestimmt wird:

= stochastischer Prozess $\{Y_t\}$ mit der Prozessgleichung

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

wobei β_0 und β_1 feste Parameter sind und $\{\varepsilon_t\}$ ein White-Noise-Prozess

Abhängigkeit von allen vorangegangenen Werten:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

$$Y_t = \beta_0 + \beta_1(\beta_0 + \beta_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \beta_0(1+\beta_1) + \beta_1^2 Y_{t-2} + (\varepsilon_t + \beta_1\varepsilon_{t-1})$$

$$Y_t = \beta_0(1+\beta_1) + \beta_1^2(\beta_0 + \beta_1 Y_{t-3} + \varepsilon_{t-2}) + (\varepsilon_t + \beta_1\varepsilon_{t-1})$$

$$= \beta_0(1+\beta_1+\beta_1^2) + (\beta_1^3 Y_{t-3} + (\varepsilon_t + \beta_1\varepsilon_{t-1} + \beta_1^2\varepsilon_{t-2})$$

$$\vdots$$

$$Y_t = \beta_0(1+\beta_1+\ldots+\beta_1^{t-2}) + \beta_1^{t-1}(\beta_0 + \beta_1 y_0 + \varepsilon_1) + (\varepsilon_t + \beta_1\varepsilon_{t-1} + \ldots + \beta_1^{t-2}\varepsilon_2)$$

$$= \beta_0(1+\beta_1+\ldots+\beta_1^{t-1}) + \beta_1^t y_0 + (\varepsilon_t + \beta_1\varepsilon_{t-1} + \ldots + \beta_1^{t-1}\varepsilon_1)$$

Einfluss des Startwertes y_0 nimmt mit wachsende zeitlichen Abstand ab, wenn $|\beta_1| < 1$ ist.

Satz:

Jeder AR(1)-Prozess, dessen Parameter $|\beta_1| < 1$, ist mittelwertstationär, mittelwertergodisch, kovarianzstationär. Ist sein weißes Rauschen ein Normalprozess, dann ist er auch kovarianzergodisch.

Eigenschaften:

1. Erwartungswert: $\mu = \dfrac{\beta_0}{1-\beta_1}$

2. $Var(Y_t) = \dfrac{\sigma_u^2}{1-\alpha_1^2}$

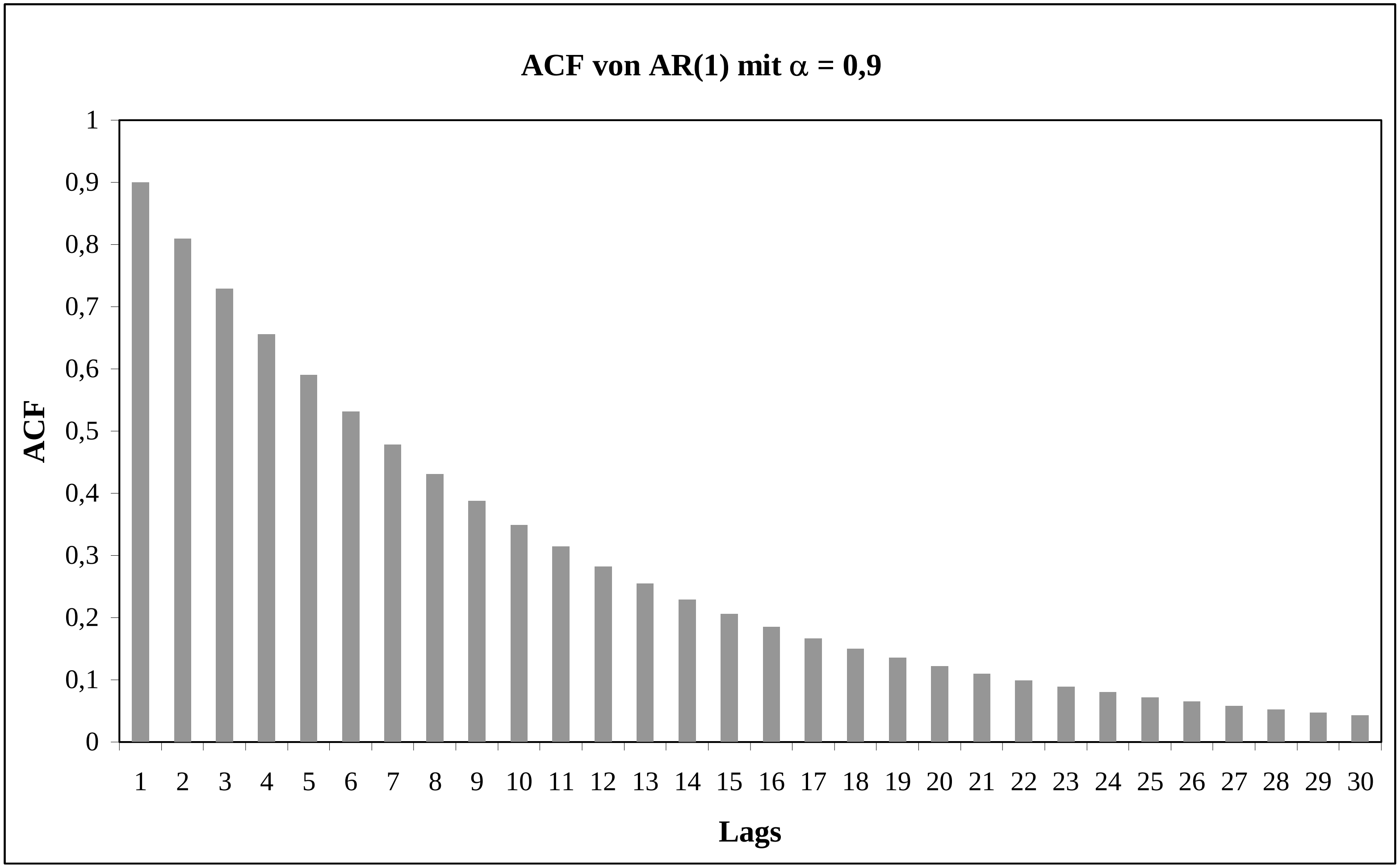
3. Kovarianzfunktion: $\gamma_j(t) = E\left[(Y_t - \mu)(Y_{t-j} - \mu)\right] = \dfrac{\beta_1^j}{1-\beta_1^2}\sigma_\varepsilon^2 = \gamma_j$

Weitere wichtige Eigenschaft des AR(1)-Prozess:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \beta_1^j$$

d. h. Autokorrelationsfunktion nimmt mit wachsen-dem Lag j ab, wenn β_1 betragsmäßig kleiner eins ist.

- ($0 < \beta_1 < 1$) exponentiell (geometrisch) abnehmendes Verhalten
- ($-1 < \beta_1 < 0$) oszillierendes Verhalten
- „die-out“-Eigenschaft der Autokorrelationsfunktion


ACF von AR(1) mit α = 0,9
ACF
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Lags

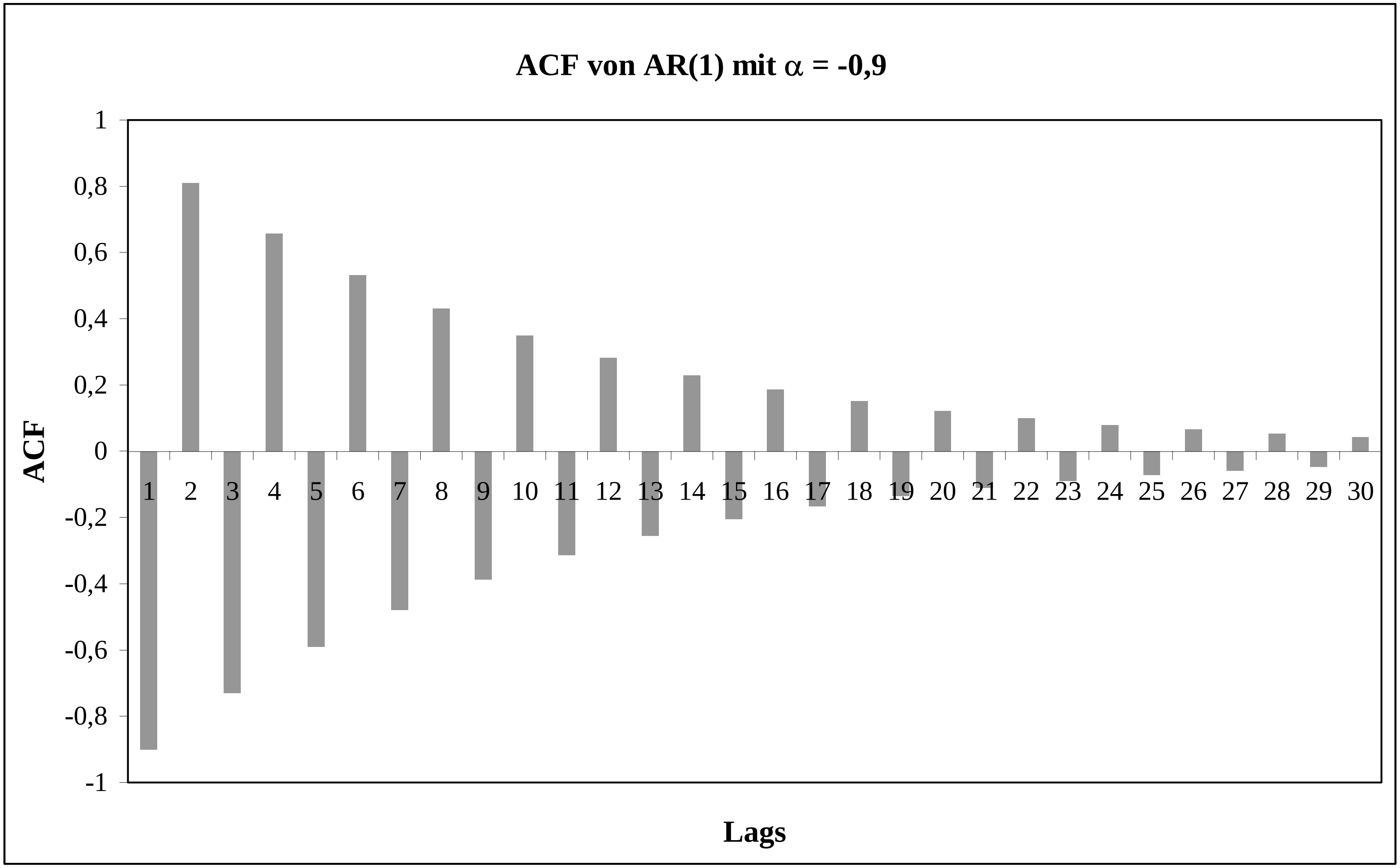
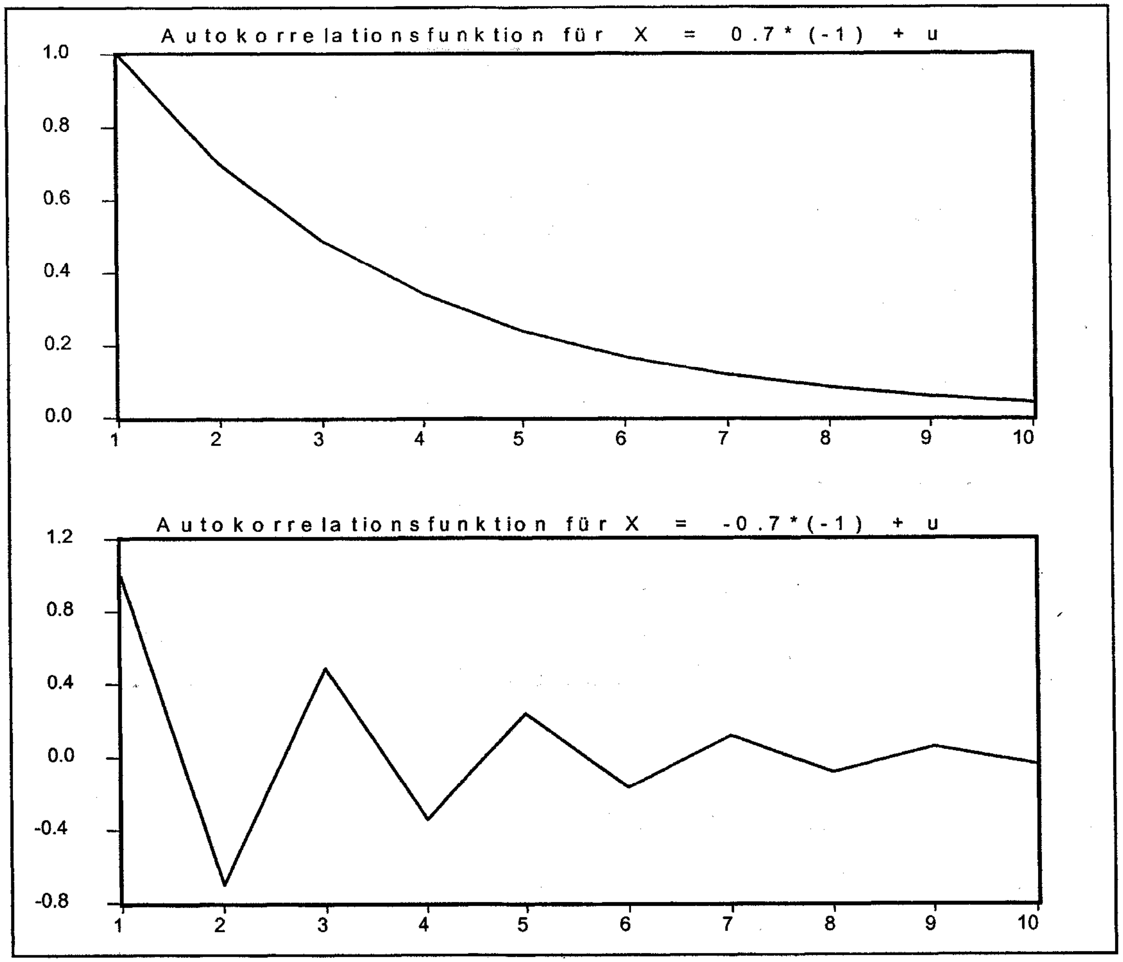
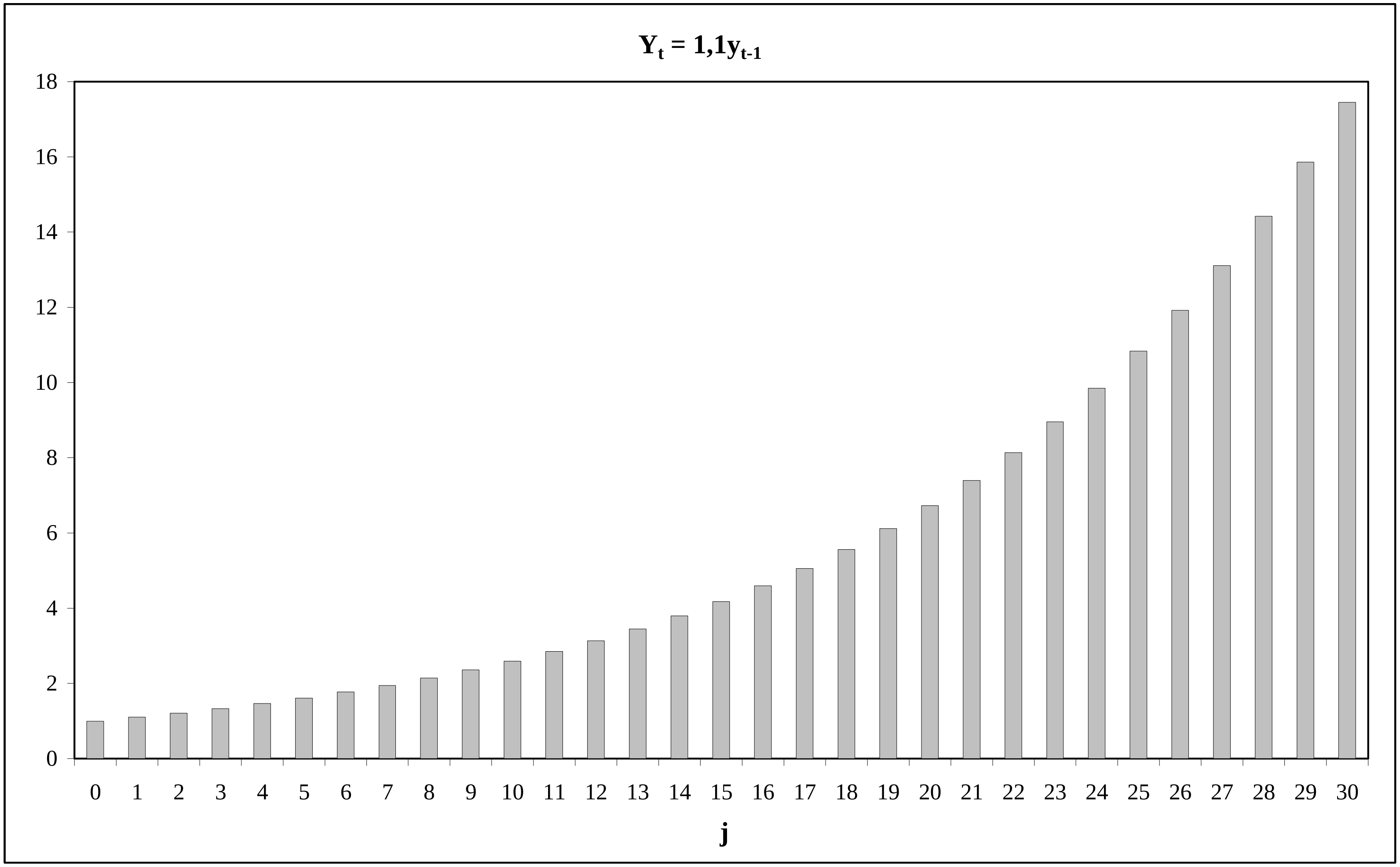
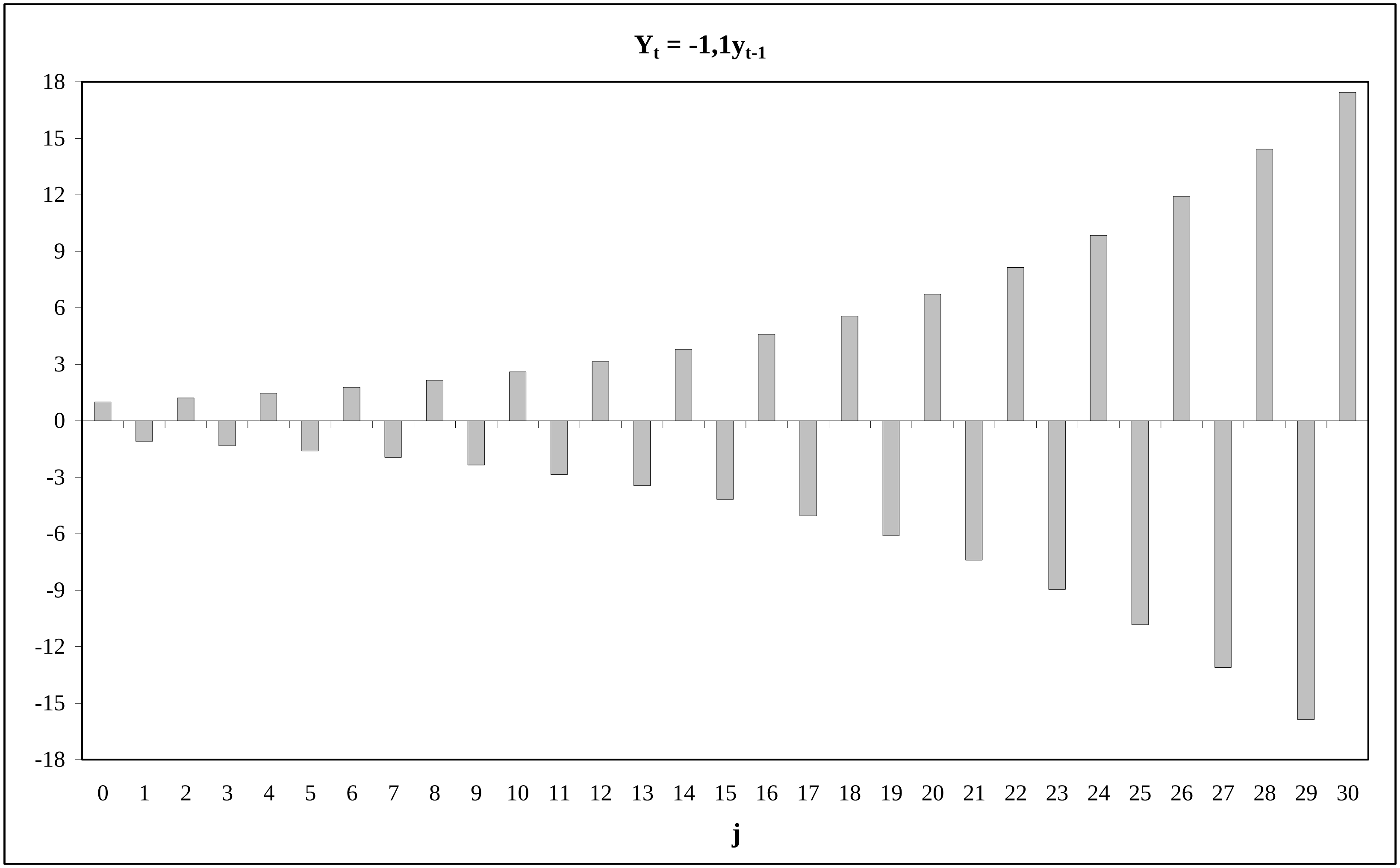

ACF von AR(1) mit α = -0,9
1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
-1
ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Lags

Abbildung 3: Autokorrelationsfunktionen für einen theoretischen AR(1)-Prozess.





$Y_t = 1{,}1y_{t-1}$
18
16
14
12
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
j


Yt = -1,1yt-1
18
15
12
9
6
3
0
-3
-6
-9
-12
-15
-18
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
j

AR(p)-Prozesse

Definition:

Ein stochastischer Prozess $\{Y_t\}$ mit

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots \beta_p Y_{t-p} + \varepsilon_t$$

wobei die β_i konstante Parameter sind und $\{\varepsilon_t\}$ ein White-Noise-Prozess ist,

heißt autoregressiver Prozess der Ordnung p oder AR(p)-Prozess.

$$\beta(L)Y_t = \varepsilon_t \text{ (komplexe Notation)}$$

$$\text{mit } \beta(L) = (1 - \beta_1 L - \beta_2 L^2 - \ldots - \beta_p L^p)$$

Folgende Bedingungen müssen für Stationarität erfüllt sein:

- $\beta_1 + \beta_2 < 1$,
- $\beta_2 - \beta_1 < 1$ und
- $|\beta_2| < 1$

Partielle Autokorrelationsfunktion:

= Instrument zur Bestimmung der *p*-ten Ordnung eines AR(p)-Prozesses.

$$Y_t = \phi_{0,1} + \phi_{1,1} Y_{t-1} + u_{1t}$$

$$Y_t = \phi_{0,2} + \phi_{1,2} Y_{t-1} + \phi_{2,2} Y_{t-2} + u_{2t}$$

$$Y_t = \phi_{0,3} + \phi_{1,3} Y_{t-1} + \phi_{2,3} Y_{t-2} + \phi_{3,3} Y_{t-3} + u_{3t}$$

…

wobei $\phi_{0,j}$, $\phi_{i,j}$, und $\{u_{jt}\}$ Konstante, Koeffizienten von $Y_{t\text{-}j}$ und der Störvariable eines AR(j)-Modells sind.

= multiple lineare Regression (Schätzung durch OLS-Methode)

Der geschätzt Koeffizient $\hat{\phi}_{2,2}$ in der zweiten Gleichung ist die Stichproben-PACF für Lag 2 von Y_t. Der geschätzte Koeffizient $\hat{\phi}_{3,3}$ in der dritten Gleichung ist die Stichproben-PACF für Lag 3 von Y_t und so weiter.

Nach der Definition erfasst der Partielle Autokorrelationskoeffizient $\hat{\phi}_{2,2}$ 2. Ordnung den zusätzlichen Beitrag von $Y_{t\text{-}2}$ zu Y_t über das AR(1)-Modell $Y_t = \phi_0 + \phi_1 Y_{t-1} + u_{1t}$ hinaus. Die PACF für den Lag p zeigt den zusätzlichen Beitrag von Y_{t-p} zu Y_t über das AR(p)-Modell hinaus, die Stichproben-PACF p-ter Ordnung sollte nicht Null sein, aber $\hat{\phi}_{j,j}$ sollte für alle $j > p$ nahe Null sein.

Stichproben-PACF eines AR(p)-Prozesses hat folgende Eigenschaften:

- $\hat{\phi}_{p,p}$ konvergiert gegen ϕ_p wenn Stichprobenumfang T gegen unendlich geht,
- $\hat{\phi}_{j,j}$ konvergiert gegen Null für alle $j > p$

= Cut-off-Eigenschaft, d.h. für einen AR(p)-Prozess, bricht Stichproben-PACF bei Lag p ab.