ARMA Models for Stationary Time Series

INSR 260, Spring 2009

Bob Stine

Overview

- Stationarity
- Correlation functions
  - Autocorrelation function (TAC and SAC, ACF)
  - Partial autocorrelation function (TPAC and SPAC, PACF)

  (True → TAC, TPAC; Sample → SAC, SPAC)
- Models
  - Autoregression
  - Moving average
  - Combination
- Examples
  - GDP, paper towel sales
  - Simulation of processes

Stationarity

- Stationarity is a key assumption in time series.
  - Stationarity = properties of the time series do not depend on when time starts.
    - Second-order stationary is most common form assumed in practice
  - Constant mean and variance

    $E(Y_t) = \mu \qquad Var(Y_t) = \sigma_y^2$
  - Autocorrelation only depends only on separation

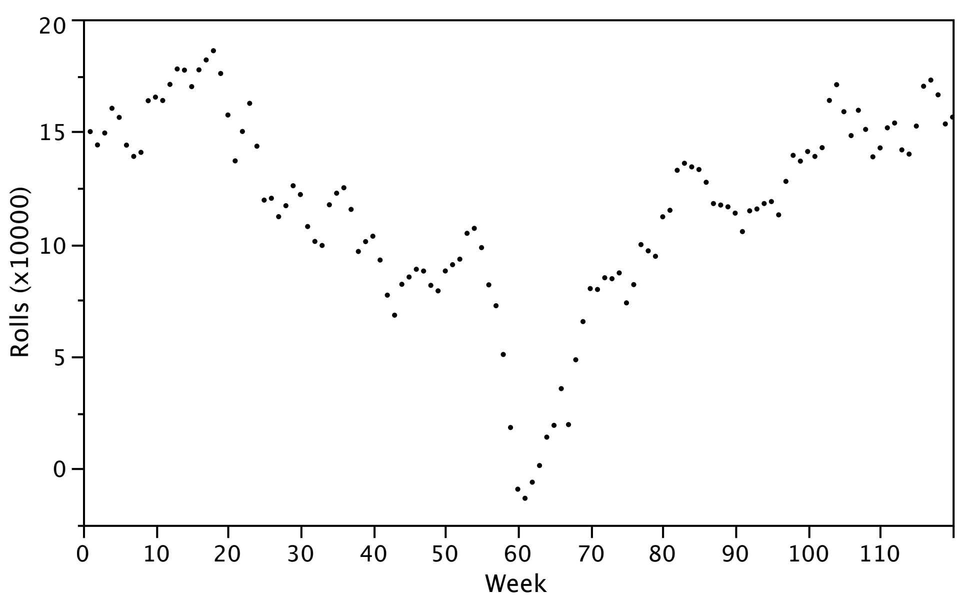
    $Corr(Y_t, Y_{t-k}) = \rho_k$

- Why does stationarity matter?
  - Without stationarity, we've got nothing to average.

- Two approaches to obtaining stationary data
  - Differencing — Model changes $Y_t - Y_{t-1}$
  - Regression

Differencing

- Raw data has possible trends, but differencing removes most of this appearance
  - Visual procedure
    - Heuristic
    - Further differencing is also possible and sometimes needed
  - Later modeling allows more precise analysis



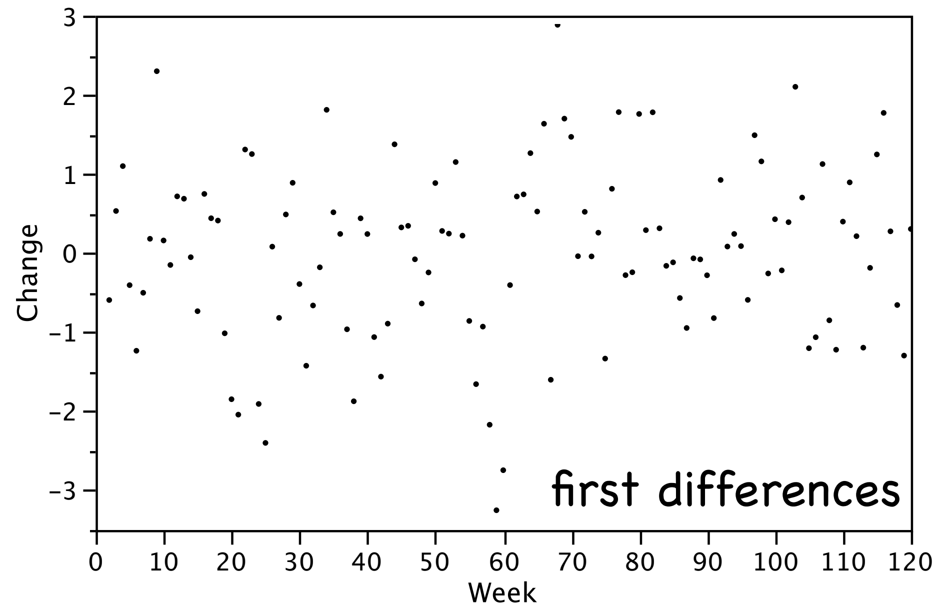
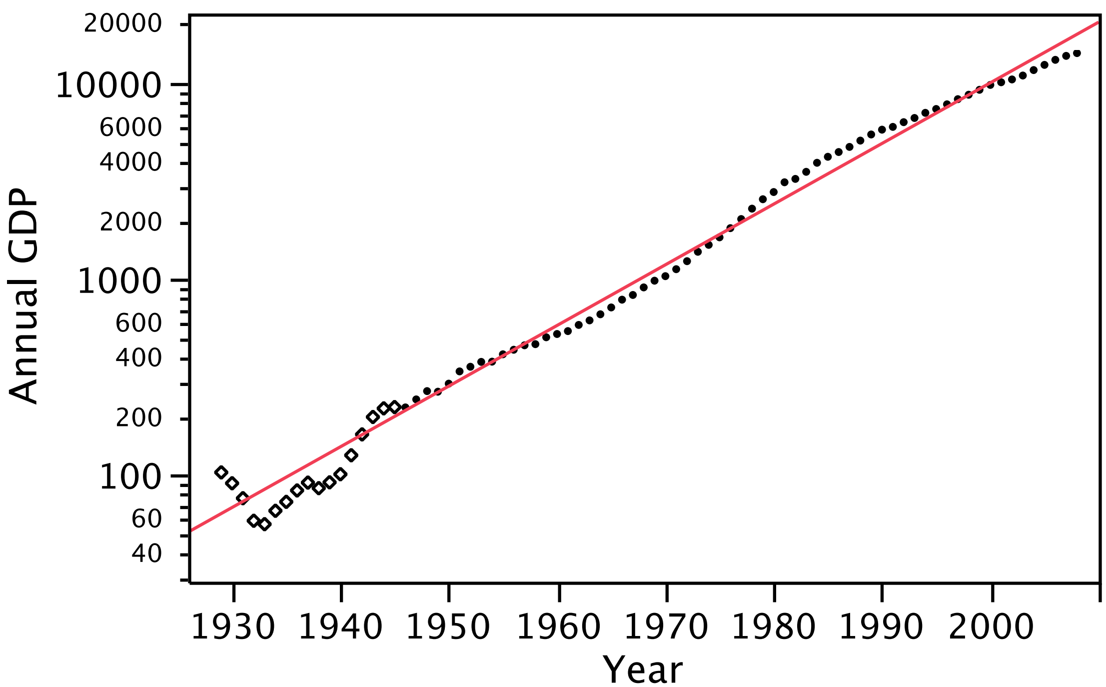


Table 9-1 Paper towels

GDP Example

- Predict annual US GDP (real dollars)
  - Data span 1929-2008
  - Use data after WWII (1946-2008)
- Remarkably linear on log scale
  - $R^2 \approx 99\%$
  - About 7% annual rate of growth (includes inflation)



Parameter Estimates

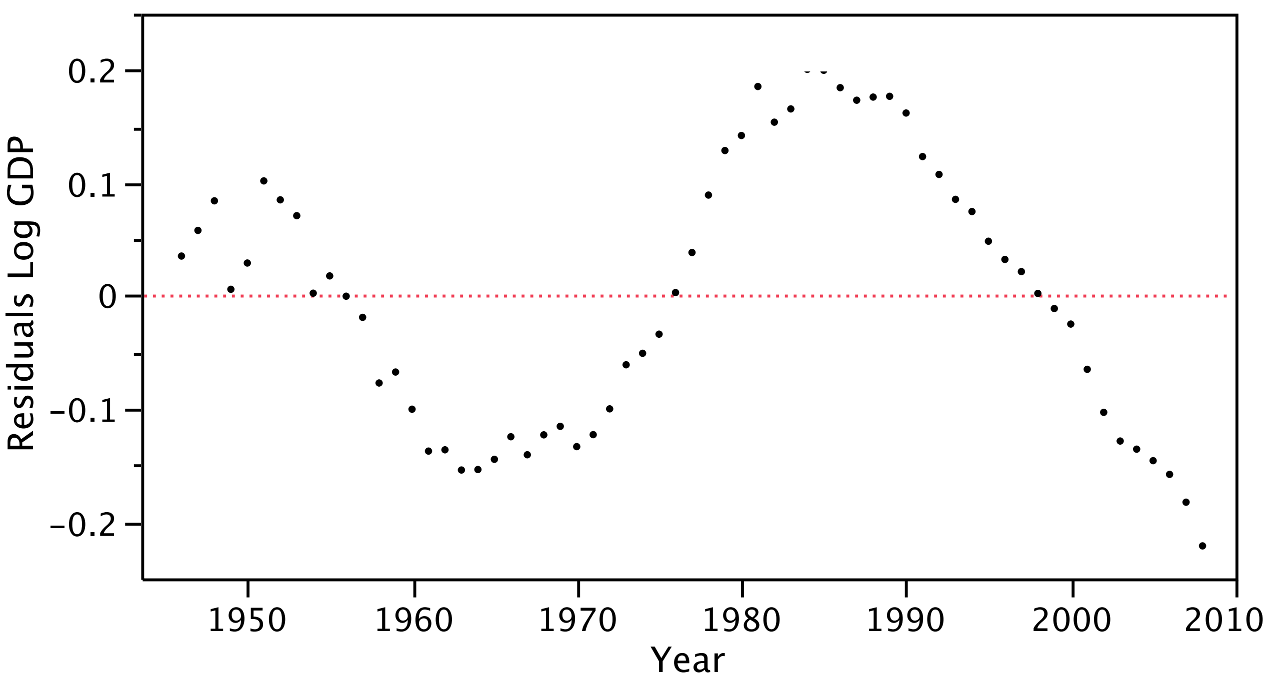
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-133.3041	1.629637	-81.80	<.0001*
Year	0.0712604	0.000824	86.45	<.0001*

Concerns?

GDP Residuals

- Pattern in residuals from linear regression
  - Residuals suggest stationary time series
  - Meandering pattern implies dependence
  - Oscillates around fit of model on log scale
    - Regression over-predicts GDP since 2000
  - Autocorrelation between adjacent residuals

$$corr(e_t, e_{t-1}) \approx 0.94$$



Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation
0.0596695	63	0.9411

Look back:
Pattern becomes evident in scatterplot of GDP on time once you have seen it in the residual plot.

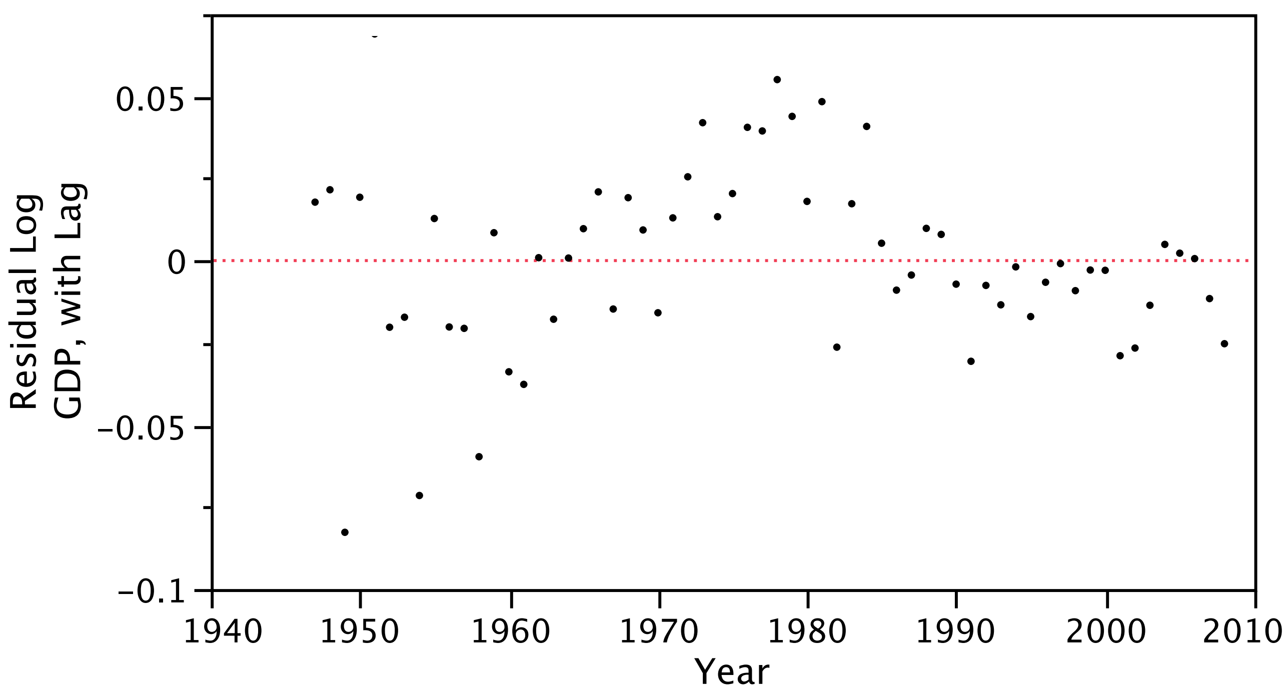
Improved Model

- Follow Durbin-Watson approach
  - Add lag of residuals back into equation
  - Refit regression with lagged variable

Term	Estimate	Std Error	t Ratio
Intercept	-132.7244	0.403924	-328.6
Year	0.0709652	0.000204	347.43
Lag Log Residuals GDP	1.0013018	0.031908	31.38

- Multiple regression with lagged variable
  - Very high R^2 (0.9995)
  - But DW remains significant
  - Does more dependence remain in the residuals?

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
1.5222692	62	0.2290	0.0143*



Closer Look At DW Stat

- Durbin-Watson statistic tests for a specific type of dependence among residuals:
  first-order autocorrelation

- Model: First-order autoregression   AR(1)
  - Equation
    $(y_t-\mu) = \varphi(y_{t-1}-\mu) + a_t$ $\mu = E(y_t)$
  - Errors a_t have mean zero, equal variance, independent
    $E(a_t) = 0$, $Var(a_t) = \sigma^2$, $Corr(a_t,a_s)=0,\ s\neq t$

- Questions
  - How do we identify whether the dependence among the errors in our model is AR(1) or some other type?
  - What other types are there?

Recognizing AR(1) Dependence

- A first-order autoregression is a geometrically weighted sum of past error terms.
  - Assume $|\varphi| < 1$
  - Back-substitute, assuming $\mu = 0$ (e.g. residuals)

$$
\begin{aligned}
y_t &= \varphi y_{t-1} + a_t \\
&= \varphi(\varphi y_{t-2} + a_{t-1}) + a_t \\
&= \varphi(\varphi(\varphi y_{t-3} + a_{t-2}) + a_{t-1}) + a_t \\
&= a_t + \varphi a_{t-1} + \varphi^2 a_{t-2} + \varphi^3 y_{t-3} \\
&= a_t + \varphi a_{t-1} + \varphi^2 a_{t-2} + \ldots + \varphi^m a_{t-m} + \ldots
\end{aligned}
$$

- Autocorrelation decays
  - First two

$$\mathrm{Corr}(y_t, y_{t-1}) = \mathrm{Corr}(\varphi y_{t-1} + a_t,\ y_{t-1}) = \varphi$$

$$\mathrm{Corr}(y_t, y_{t-2}) = \mathrm{Corr}(\varphi^2 y_{t-2} + \varphi a_{t-1} + a_t,\ y_{t-1}) = \varphi^2$$

  - In general

$$\mathrm{Corr}(y_t, y_{t-k}) = \varphi^k,\ k = 1,2,3\ldots$$

Autocorrelation Function

- Identification procedure
  - Autocorrelations decay geometrically if process is AR(1)
  - Plot the autocorrelations; decide if decay looks geometric.
  - Notation

    TAC(k) = ρ_k = Correlation(Y_t, Y_{t-k})   [a.k.a., ACF]

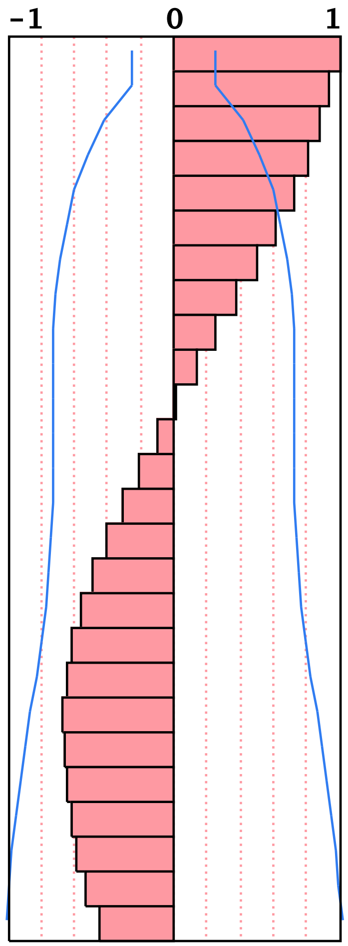
- Complications
  - We observe sample correlations (SAC r_k), not true values
  - Outliers, nonlinearity can make a mess of things
  - How close to geometric is close enough?

- Calculation
  - JMP plots the autocorrelation function for you without having to form all of the lags (beware outliers!)
  - Analyze > Modeling > Time Series

GDP Residual ACF

- SAC starts out positive, then drifts negative
  - Not a geometric decay toward 0
  - Values start out positive, then gradually become negative

r_k



Lag	AutoCorr	Ljung-Box Q	p-Value
0	1.0000	.	.
1	0.9411	58.4964	<.0001*
2	0.8756	109.963	<.0001*
3	0.8053	154.225	<.0001*
4	0.7242	190.629	<.0001*
5	0.6223	217.974	<.0001*
6	0.4995	235.898	<.0001*
7	0.3751	246.184	<.0001*
8	0.2532	250.958	<.0001*
9	0.1378	252.398	<.0001*
10	0.0163	252.418	<.0001*
11	-0.0971	253.160	<.0001*
12	-0.2068	256.593	<.0001*
13	-0.3109	264.511	<.0001*
14	-0.4071	278.360	<.0001*
15	-0.4868	298.580	<.0001*
16	-0.5555	325.470	<.0001*
17	-0.6106	358.660	<.0001*
18	-0.6491	397.006	<.0001*
19	-0.6662	438.309	<.0001*
20	-0.6631	480.180	<.0001*
21	-0.6458	520.846	<.0001*
22	-0.6207	559.322	<.0001*
23	-0.5836	594.190	<.0001*
24	-0.5248	623.104	<.0001*
25	-0.4501	644.932	<.0001*

H_0:
No cumulative autocorrelation

Conclude: Not AR(1) process.

Models of Dependence

- "Box-Jenkins" family of models (ARIMA models)
  - Named for authors of influential book
- Autoregressions AR(p), p = 1,2,...
  - Regression models that describe the current value of the time series Y_t as a weighted sum of p past values of the time series.
- Moving averages MA(q), q = 1,2,...
  - Represent the time series as a weighted sum of q unobserved, uncorrelated error terms.
- Autoregressive, moving-average (ARMA) combine the two types of models

Distinguishing AR(p) Models

- AR(p) model adds lags of the time series
  - $Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} + a_t$ $(\mu=0)$
  - Stationarity constrains the coefficients
    - Analogous to keeping $|\varphi|<1$ in AR(1) model

- Complication:
  All AR(p) models have geometric decay in TAC
  - How do we distinguish an AR(2) from an AR(4)?
  - You cannot, at least not based on estimates.

- Distinguishing among AR(p) models requires another plot
  - Need to determine reasonable choice for p
  - Partial autocorrelation function (TPAC or SPAC, PACF)

Partial Autocorrelation

- Partial correlation
  - Correlation between two random variables conditional on a collection of other variables

    $$Corr(X,Y|Z_1,Z_2,\ldots,Z_k)$$

  - Example: multiple regression of Y on X and $Z_1,Z_2,\ldots,Z_k$.
  - Leverage plot is a plot that shows the partial correlation between Y and X having adjusted for $Z_1,Z_2,\ldots,Z_k$.

- Role in autoregression
  - Suppose Yt is an AR(2) process

    $$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t$$

  - Consider partial correlation conditional on Y_{t-1} and Y_{t-2}.

    $$Corr(Y_t,Y_{t-k}|Y_{t-1},Y_{t-2})$$

  - What's this going to be if $k > 2$?
    - Construction of leverage plots provides a hint.

Graphing PACF

- Construct the sequence

  $Corr(Y_t, Y_{t-1})$ $\quad$ $Y_t = \varphi_1 Y_{t-1} + a_t$

  $Corr(Y_t, Y_{t-2} | Y_{t-1})$ $\quad$ $Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t$

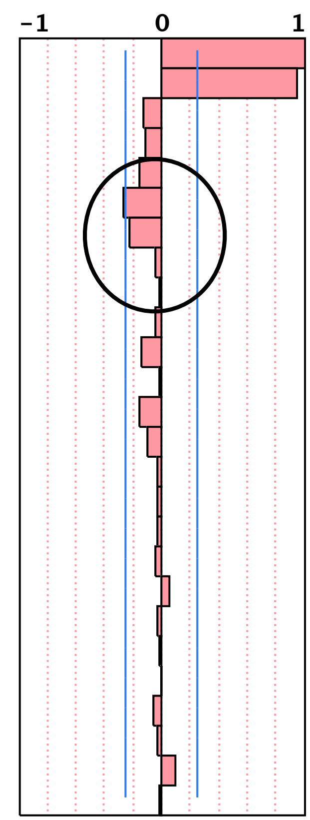
  $Corr(Y_t, Y_{t-3} | Y_{t-1}, Y_{t-2})$ $\quad$ $Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + a_t$

  ...

  $Corr\ (Y_t, Y_{t-k}\ |\ Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k-1})$

- Graph the sequence of estimates versus lag k
  - Denote estimates r_{kk}
  - Dependence at separation of about 5-6 years indicates need for a more complex AR model

Lag	Partial
0	1.0000
1	0.9474
2	-0.1306
3	-0.1058
4	-0.1493
5	-0.2592
6	-0.2174
7	-0.0389
8	-0.0174
9	-0.0460
10	-0.1361
11	-0.0200
12	-0.1602
13	-0.0962
14	-0.0300
15	-0.0290
16	-0.0270
17	-0.0397
18	0.0593
19	-0.0233
20	-0.0126
21	-0.0041
22	-0.0617
23	-0.0267
24	0.0917
25	-0.0150

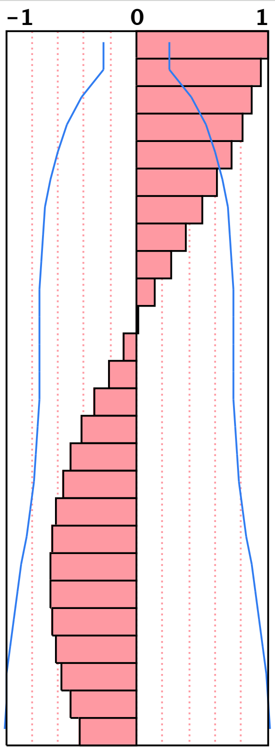


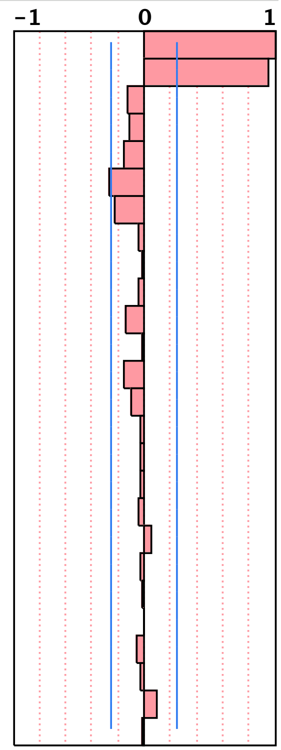
JMP Software

- Time series platform shows SAC, SPAC side-by-side
- Autoregression
  - Geometric decay in the SAC
  - Sharp "cut-off" in the SPAC

Time Series Basic Diagnostics

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.9474	58.3803	<.0001*	1	0.9474
2	0.8841	110.071	<.0001*	2	-0.1306
3	0.8125	154.473	<.0001*	3	-0.1058
4	0.7287	190.798	<.0001*	4	-0.1493
5	0.6221	217.736	<.0001*	5	-0.2592
6	0.4986	235.351	<.0001*	6	-0.2174
7	0.3770	245.605	<.0001*	7	-0.0389
8	0.2588	250.525	<.0001*	8	-0.0174
9	0.1400	251.992	<.0001*	9	-0.0460
10	0.0147	252.009	<.0001*	10	-0.1361
11	-0.1002	252.790	<.0001*	11	-0.0200
12	-0.2136	256.411	<.0001*	12	-0.1602
13	-0.3199	264.698	<.0001*	13	-0.0962
14	-0.4148	278.920	<.0001*	14	-0.0300
15	-0.4965	299.734	<.0001*	15	-0.0290
16	-0.5636	327.134	<.0001*	16	-0.0270
17	-0.6178	360.787	<.0001*	17	-0.0397
18	-0.6482	398.676	<.0001*	18	0.0593
19	-0.6621	439.134	<.0001*	19	-0.0233
20	-0.6589	480.153	<.0001*	20	-0.0126
21	-0.6412	519.938	<.0001*	21	-0.0041
22	-0.6113	557.005	<.0001*	22	-0.0617
23	-0.5674	589.765	<.0001*	23	-0.0267
24	-0.5044	616.334	<.0001*	24	0.0917
25	-0.4345	636.585	<.0001*	25	-0.0150



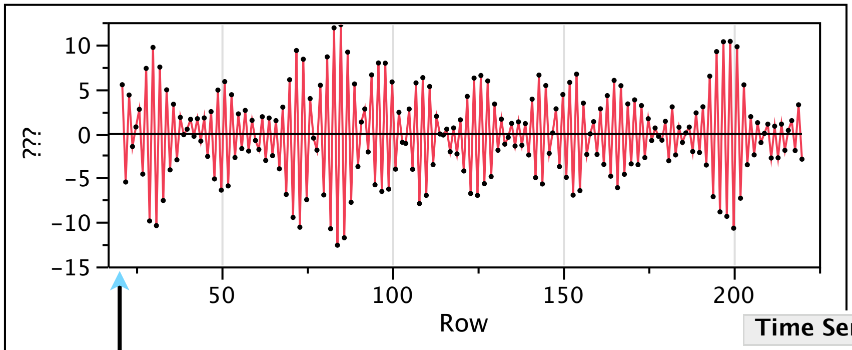


r_k r_{kk}

Fact that we observe estimates rather than population values complicates task.

Simulated Example

- Generate data from an AR process... Which is it?



What happens if observe fewer data than these 200?

Exclude the first 20 rows for "burn in", start the recursion

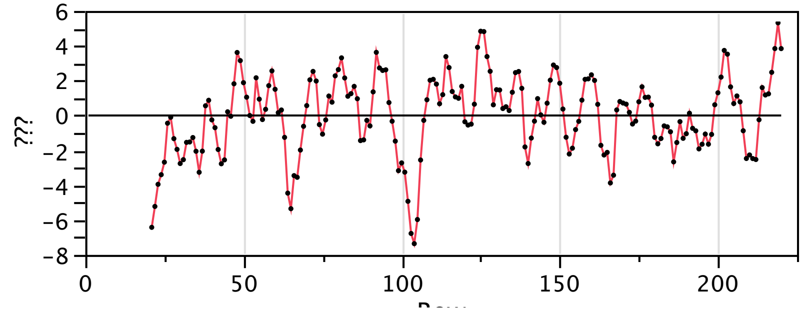
Time Series Basic Diagnostics

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	-0.9499	183.180	<.0001*	1	-0.9499
2	0.8730	338.677	<.0001*	2	-0.3001
3	-0.7362	449.830	<.0001*	3	0.6391
4	0.5888	521.291	<.0001*	4	0.0165
5	-0.4187	557.603	<.0001*	5	-0.0249
6	0.2502	570.636	<.0001*	6	-0.1301
7	-0.0836	572.098	<.0001*	7	0.0389
8	-0.0745	573.268	<.0001*	8	-0.1296
9	0.2110	582.684	<.0001*	9	-0.0560
10	-0.3319	606.105	<.0001*	10	-0.0640
11	0.4200	643.810	<.0001*	11	-0.0033
12	-0.4894	695.274	<.0001*	12	-0.0666
13	0.5257	754.976	<.0001*	13	0.0334
14	-0.5371	817.639	<.0001*	14	0.0956
15	0.5234	877.471	<.0001*	15	0.0878
16	-0.4833	928.748	<.0001*	16	0.0508
17	0.4249	968.596	<.0001*	17	-0.0414
18	-0.3485	995.556	<.0001*	18	-0.0660
19	0.2609	1010.75	<.0001*	19	-0.0801
20	-0.1637	1016.76	<.0001*	20	0.0274
21	0.0615	1017.61	<.0001*	21	-0.0610
22	0.0407	1017.99	<.0001*	22	0.0188
23	-0.1434	1022.68	<.0001*	23	-0.0859
24	0.2276	1034.57	<.0001*	24	-0.1442
25	-0.3106	1056.84	<.0001*	25	-0.1572

arma_sim.jmp

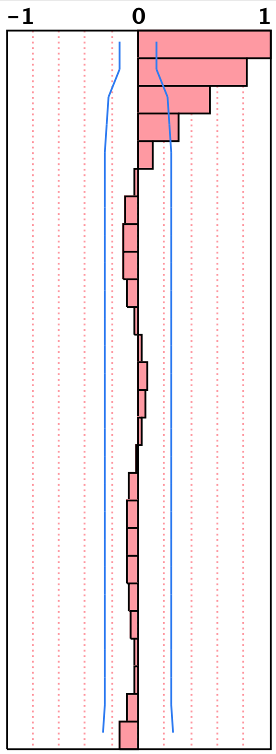
Another Example

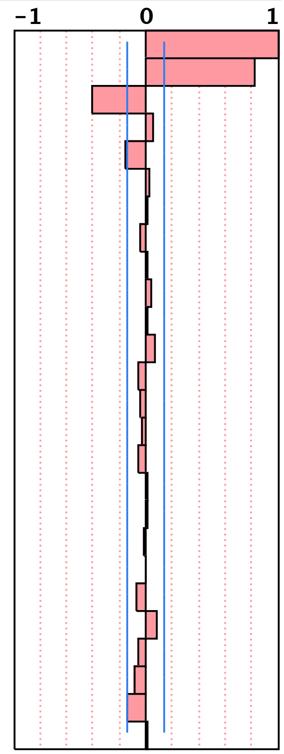
- What about this time series?



Time Series Basic Diagnostics

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.8221	137.220	<.0001*	1	0.8221
2	0.5440	197.599	<.0001*	2	-0.4071
3	0.3090	217.184	<.0001*	3	0.0533
4	0.1098	219.670	<.0001*	4	-0.1602
5	-0.0303	219.859	<.0001*	5	0.0284
6	-0.0948	221.733	<.0001*	6	0.0161
7	-0.1183	224.662	<.0001*	7	-0.0411
8	-0.1151	227.449	<.0001*	8	0.0118
9	-0.0803	228.813	<.0001*	9	0.0388
10	-0.0312	229.020	<.0001*	10	0.0169
11	0.0282	229.190	<.0001*	11	0.0694
12	0.0656	230.115	<.0001*	12	-0.0583
13	0.0621	230.947	<.0001*	13	-0.0374
14	0.0313	231.161	<.0001*	14	-0.0278
15	-0.0179	231.231	<.0001*	15	-0.0510
16	-0.0646	232.147	<.0001*	16	0.0088
17	-0.0854	233.756	<.0001*	17	0.0160
18	-0.0876	235.461	<.0001*	18	-0.0185
19	-0.0792	236.861	<.0001*	19	0.0046
20	-0.0753	238.132	<.0001*	20	-0.0632
21	-0.0507	238.712	<.0001*	21	0.0873
22	-0.0247	238.851	<.0001*	22	-0.0595
23	-0.0327	239.096	<.0001*	23	-0.0885
24	-0.0888	240.904	<.0001*	24	-0.1408
25	-0.1433	245.642	<.0001*	25	0.0180





Moving Average Models

- Moving Average models MA(q)
  - Weighted average of error terms
  - MA(q) $Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots \theta_q a_{t-q}$
- Dual behavior compared to AR models
  - TAC ρ_k
    Cuts off since errors are uncorrelated once the observations are more than q terms apart.
  - TPAC ρ_{kk}
    Geometric decay
  - Estimates may not look much like either
- Relationship to AR models
  - AR models are an infinite weighted average of past a_t
  - MA models only depend on a finite number of a_t.

ARMA Models

- Mixture of autoregressive and moving average components
  - ARMA(p,q)

$$(Y_t - \mu) = \varphi_1 (Y_{t-1} - \mu) + \ldots + \varphi_p (Y_{t-p} - \mu) + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

- Correlation functions
  - TAC
    - Geometric decay
  - TPAC
    - Geometric decay
- Identification procedure is iterative
  - Guess from substantive knowledge, SAC, and SPAC
  - Estimate the model to see how well it does

Summary

- AR, MA and ARMA models for <u>stationary</u> series
  - Particularly useful models for residual processes
- Correlation functions

Fig 9.5–9.6, page 412

	AR(p)	MA(q)	ARMA
TAC	geometric	cuts off	geometric
TPAC	cuts off	geometric	geometric

- Challenges that remain
  - Estimation
  - Diagnostics
  - Prediction

What to use to predict GDP?