Estimating ARMA Models

INSR 260, Spring 2009
Bob Stine

Overview

- Review
- Estimating correlations
  - SAC and SPAC, standard errors
  - Recognizing patterns
- Fitting models
  - Model selection criteria
- Residual diagnostics
  - Tests for any remaining autocorrelation

Review

- Autoregressive, moving average models
  - Autoregression AR(p)
    - $Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} + a_t$
    - Current value is a weighted sum of p past values.
  - Moving Average MA(q)
    - $Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots \theta_q a_{t-q}$
    - Current value is a weighted sum of current and prior error terms.
  - ARMA models combine the two ARMA(p,q)
    - Example: $Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t - \theta_1 a_{t-1}$
- Graphical identification procedure
  - Inspect estimated autocorrelations r_k and partial autocorrelations r_{kk}.

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

Estimating Autocorrelations

- Estimates vary, complicating identification of model
  - Sample autocorrelations r_k subject to sampling variation
  - Estimate of variation uses estimates at lower lags

$$n\,\mathrm{Var}(r_k) \approx 1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_{k-1}^2$$
$$\approx 1 + 2r_1^2 + 2r_2^2 + \dots + 2r_{k-1}^2$$

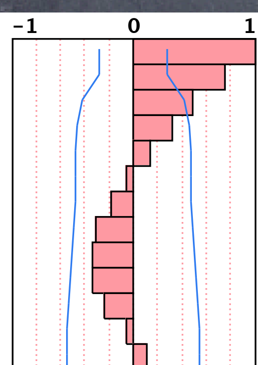
  - This expression determines blue bounds in JMP plots

- Simulated example: AR(1) with $\varphi = 0.8$
  - True autocorrelations drop geometrically

$$\rho_k = \varphi^k = 0.8^k$$

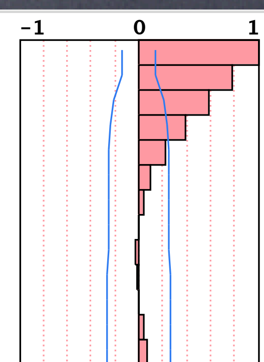
  - Compare ρ_k to estimate r_k for different sample lengths n

Lag	AutoCorr
0	1.0000
1	0.7558
2	0.4964
3	0.3196
4	0.1332
5	-0.0525
6	-0.1754
7	-0.3122
8	-0.3422
9	-0.3296
10	-0.2425
11	-0.0549
12	0.1056



n=
← 50
200 →

Lag	AutoCorr
0	1.0000
1	0.7804
2	0.5862
3	0.3978
4	0.2267
5	0.0973
6	0.0361
7	-0.0061
8	-0.0308
9	-0.0135
10	0.0010
11	0.0476
12	0.0636



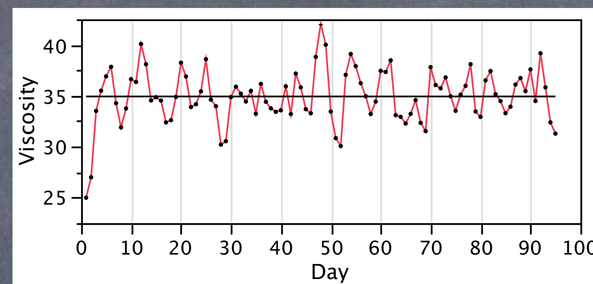
arma_sim.jmp

Example

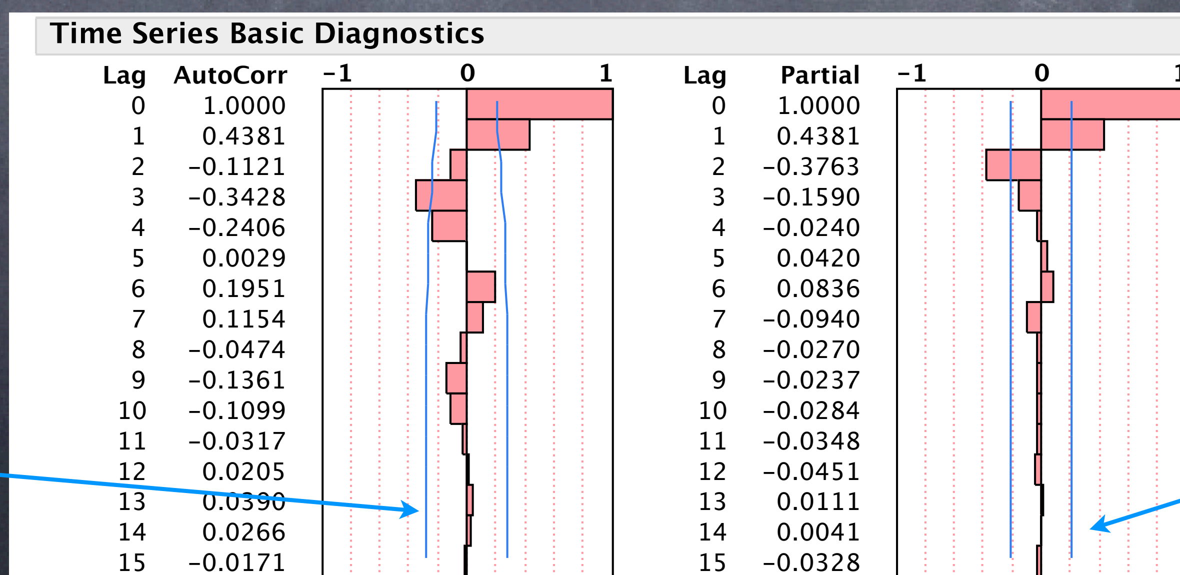
- Chemical viscosity

  Table 9-4, p 433

  - Need to control production process for compound
  - Sample of n=95 consecutive daily readings
  - Visual inspection indicates stationary process



- Estimated correlation functions r_k and r_{kk}



Time Series Basic Diagnostics

Lag	AutoCorr	Lag	Partial
0	1.0000	0	1.0000
1	0.4381	1	0.4381
2	-0.1121	2	-0.3763
3	-0.3428	3	-0.1590
4	-0.2406	4	-0.0240
5	0.0029	5	0.0420
6	0.1951	6	0.0836
7	0.1154	7	-0.0940
8	-0.0474	8	-0.0270
9	-0.1361	9	-0.0237
10	-0.1099	10	-0.0284
11	-0.0317	11	-0.0348
12	0.0205	12	-0.0451
13	0.0390	13	0.0111
14	0.0266	14	0.0041
15	-0.0171	15	-0.0328

± 2/ $SD(r_k)$

± 2/ sqrt(n)

Estimation

- Computer software makes this fast now
  - AR models are easy
    Regress Y_t on lags Y_{t-1}, Y_{t-2},..., Y_{t-p}.
  - MA models are hard
    Don't observe the explanatory variables a_{t-1}, a_{t-2}...
  - Estimating the model is not a problem anymore.
- Task becomes deciding which model to fit
- Two approaches
  - Model selection
    Try many models, use selection criterion to decide best.
  - Model diagnostics
    Inspect residuals for remaining correlation

Estimating Viscosity Model

- Initial visual analysis indicates
  - Apparently stationary process
  - Correlation functions suggest AR(2) model
- Fit model
  - Estimates are significant
  - Beware collinearity

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant Estimate
AR1	1	0.68209	0.0979920	6.96	<.0001*	26.2522703
AR2	2	−0.43330	0.1036814	−4.18	<.0001*	
Intercept	0	34.94641	0.2934785	119.08	<.0001*	

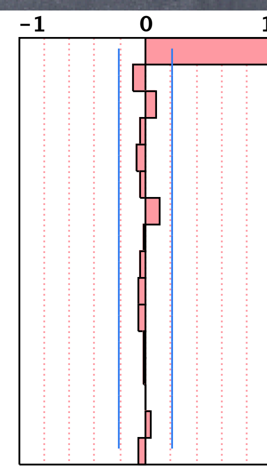
- Residual autocorrelations
  - Software shows test of cumulative residual autocorrelation
  - Expression is approximately

$$Q^* \approx n \sum r_k(\hat{a})^2$$

(See page 459.)

- Conclude: ready to forecast

Lag	AutoCorr	Ljung-Box Q	p-Value
0	1.0000	.	.
1	−0.0911	0.8141	0.3669
2	0.0791	1.4343	0.4882
3	−0.0449	1.6365	0.6511
4	−0.0732	2.1792	0.7028
5	−0.0363	2.3143	0.8042
6	0.1064	3.4870	0.7457
7	−0.0073	3.4926	0.8360
8	−0.0370	3.6376	0.8883
9	−0.0494	3.8996	0.9179
10	−0.0585	4.2702	0.9343
11	−0.0073	4.2761	0.9611
12	−0.0173	4.3094	0.9772
13	−0.0064	4.3139	0.9874
14	0.0486	4.5823	0.9908
15	−0.0589	4.9816	0.9923



Model Selection Approach

- Automatic procedure
  - Useful when initial analysis is vague
    - rather than clear choice suggested for the viscosity data
  - Fit several, judging best using a selection criterion
- Selection criteria
  - Objective
    Find model that will give best predictions out-of-sample or when applied to a new data series of same form.
  - "Penalize" model for adding more predictors
    - Unlike R^2, selection criteria don't automatically improve as model becomes larger. Only improve when added variable demonstrates benefit to predictions.
  - Choices include
    - AIC = Akaike information criterion
    - BIC = SBC = Bayes information criterion
    - RIC = Risk inflation criterion (stronger penalty)

JMP shows these two.

Example:Viscosity

- Fit several models to viscosity data
  - JMP accumulates results in summary table

Model Comparison

Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	AIC Rank	SBC Rank
AR(1)	93	5.5045142	433.92316	439.03091	0.192	429.92316	6	6
AR(2)	92	4.6917561	420.07428	427.73591	0.294	414.07428	5	1
AR(3)	91	4.5222374	417.78483	428.00034	0.314	409.78483	1	2
AR(4)	90	4.5677261	419.70276	432.47214	0.314	409.70276	3	4
ARI(2, 1)	91	6.6359967	447.71640	455.34628	-0.07	441.7164	7	7
ARMA(2, 1)	91	4.5517973	418.37704	428.59255	0.310	410.37704	2	3
ARMA(3, 1)	90	4.5691998	419.72810	432.49749	0.314	409.7281	4	5

resid SS + 2k   resid SS + k log n   ≈residual SS

- Results show...
  - AIC penalty for adding parameters = 2 (# of estimates)

    AIC prefers the AR(3) model
  - SBC penalizes more (log n)(# of estimates)

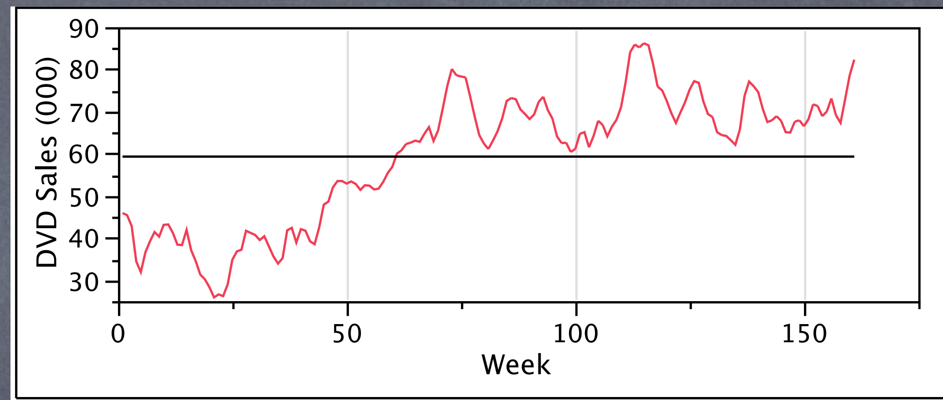
    SBC prefers simpler models, here AR(2)

$\log_e 95 \approx 4.55$

- Theory
  - AIC tends to "overfit", but predicts well.
  - Predictions likely to be similar with either choice.

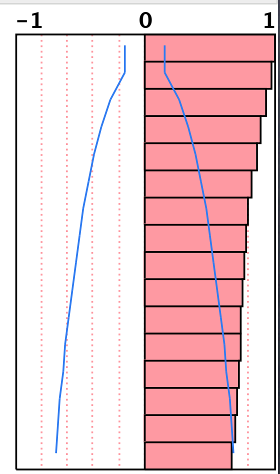
Example: DVD Case Study

- Objective
  - Predict weekly DVD sales (data are in thousands)
  - Observe n=161 past weeks of observations
- Initial visual analysis
  - Plot of data shows "transient" level shift.
  - Did something change around week 25?
  - Sample autocorrelations decay very slowly, as typical of a non-stationary process.



Time Series Basic Diagnostics

Lag	AutoCorr
0	1.0000
1	0.9738
2	0.9348
3	0.8986
4	0.8624
5	0.8247
6	0.7939
7	0.7770
8	0.7652
9	0.7535
10	0.7450
11	0.7392
12	0.7305
13	0.7147
14	0.6948
15	0.6756



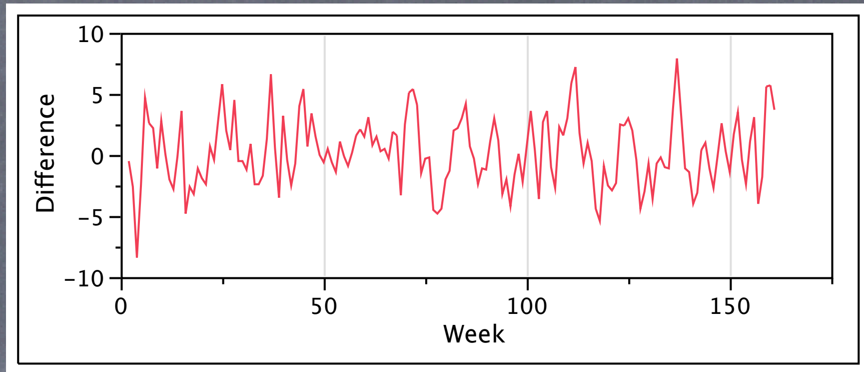
- Conclude

Difference data, try again.

Table 10-2

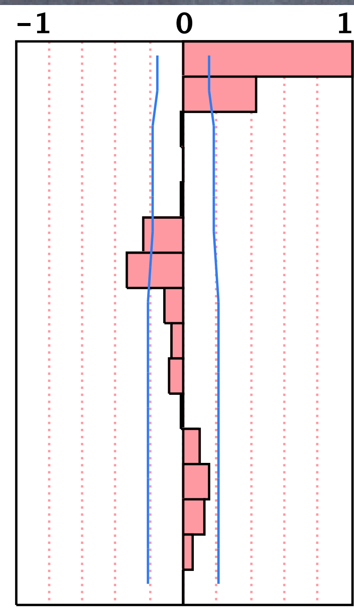
Case Study: Differences

- Visual inspection
  - Sequence plot appears to fluctuate around stable level.
  - Autocorrelations decay at faster rate.

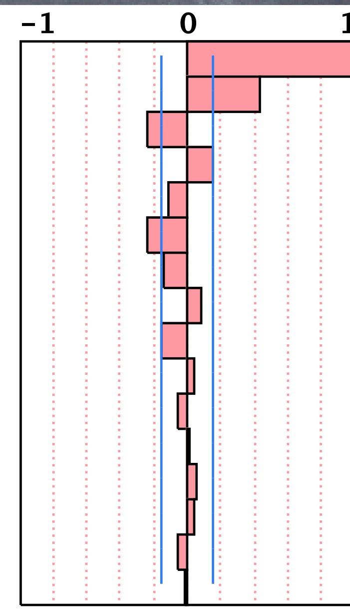


- Autocorrelations of differences r_k and r_{kk} suggest a mix of autoregressive and moving average components

Lag	AutoCorr
0	1.0000
1	0.4349
2	-0.0082
3	0.0020
4	-0.0172
5	-0.2387
6	-0.3355
7	-0.1125
8	-0.0661
9	-0.0801
10	-0.0186
11	0.0918
12	0.1495
13	0.1200
14	0.0617
15	0.0036



Lag	Partial
0	1.0000
1	0.4349
2	-0.2434
3	0.1471
4	-0.1139
5	-0.2396
6	-0.1463
7	0.0850
8	-0.1579
9	0.0356
10	-0.0565
11	0.0194
12	0.0609
13	0.0410
14	-0.0581
15	-0.0104



Case Study: Estimation

- Try several models
  - Increased ARMA(k,k) until seemed little to gain
  - Then removed last AR term that was not significant
  - I(1) model is a reminder of where that large R^2 statistic comes from: differencing.

Model Comparison

Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	AIC Rank	SBC Rank
ARIMA(1, 1, 1) No Constrain	157	6.01	744.2705	753.4960	0.975	738.2705	12	3
ARIMA(2, 1, 2) No Constrain	155	5.79	742.5693	757.9452	0.976	732.5693	10	5
I(1)	159	7.97	787.0895	790.1647	0.966	785.0895	13	13
ARIMA(3, 1, 3) No Constrain	153	5.82	743.7660	765.2922	0.976	729.7660	11	8
ARIMA(4, 1, 4) No Constrain	151	5.56	738.3926	766.0691	0.977	720.3926	9	9
ARIMA(5, 1, 5) No Constrain	149	5.21	732.7144	766.5413	0.979	710.7144	7	10
ARIMA(6, 1, 6) No Constrain	147	5.17	733.6808	773.6581	0.979	707.6808	8	12
ARIMA(5, 1, 6) No Constrain	148	5.14	731.8426	768.7447	0.979	707.8426	6	11
ARIMA(4, 1, 6) No Constrain	149	5.10	729.8626	763.6895	0.979	707.8626	4	7
ARIMA(3, 1, 6) No Constrain	150	5.15	730.3289	761.0807	0.979	710.3289	5	6
ARIMA(2, 1, 6) No Constrain	151	5.12	728.5672	756.2437	0.979	710.5672	3	4
ARIMA(1, 1, 6) No Constrain	152	5.09	726.7272	751.3285	0.979	710.7272	2	2
IMA(1, 6) No Constrain	153	5.06	724.8162	746.3424	0.979	710.8162	1	1

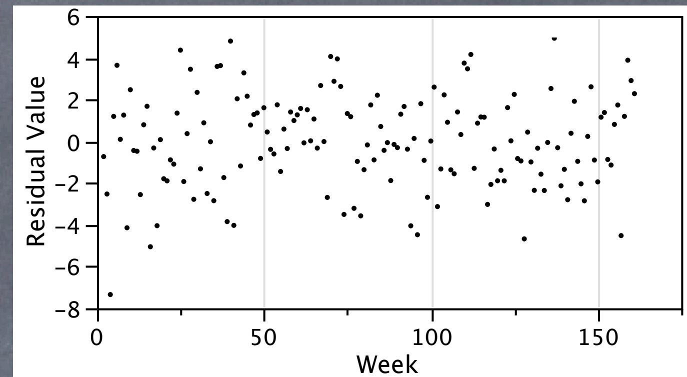
- Choice
  - Model contains all lower order terms
  - Collinearity

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant Estimate
MA1	1	-0.6168536	0.0768915	-8.02	<.0001*	0.23371712
MA2	2	0.0410673	0.0857010	0.48	0.6325	
MA3	3	-0.0121524	0.0846485	-0.14	0.8860	
MA4	4	0.0539297	0.0871212	0.62	0.5368	
MA5	5	0.1817214	0.1070533	1.70	0.0916	
MA6	6	0.4577184	0.0793308	5.77	<.0001*	
Intercept	0	0.2337171	0.1592325	1.47	0.1442	

Case Study: Diagnostics

- Integrated MA(6) model
- Residuals show no evident pattern



- Residual autocorrelations find no remaining dependence over time.

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.0114	0.0213	0.8840	1	0.0114
2	0.0062	0.0275	0.9863	2	0.0060
3	-0.0306	0.1823	0.9804	3	-0.0308
4	0.0493	0.5859	0.9646	4	0.0500
5	0.0097	0.6015	0.9879	5	0.0089
6	-0.0145	0.6368	0.9958	6	-0.0163
7	-0.0678	1.4164	0.9851	7	-0.0647
8	0.0162	1.4613	0.9933	8	0.0163
9	-0.0574	2.0268	0.9910	9	-0.0592
10	-0.0277	2.1598	0.9950	10	-0.0294
11	0.0503	2.5997	0.9950	11	0.0602
12	0.0728	3.5290	0.9905	12	0.0689
13	0.0581	4.1236	0.9898	13	0.0583
14	0.0259	4.2429	0.9938	14	0.0281
15	-0.0035	4.2451	0.9968	15	-0.0048

- Ready to forecast...

Summary

- Estimating correlations
  - SAC and SPAC, standard errors
  - Recognizing patterns
  - Most important use in checking the fit of a model
    Does this model leave dependence in residuals?
- Identification methods
  - Visual inspection works in some cases
  - Model selection criteria helpful when vague
- Residual diagnostics
  - Test for any remaining autocorrelation
- Next step: seasonal models, forecasting