ARMA Modeling Strategy (Chapter 4)

How can I find order of a time-series model?

i) If we're lucky, the <u>sample ACF</u> may drop off rapidly & we could guess the order

Example: MA(1) $X_t = a_t - \theta_1 a_{t-1}$

$$\gamma_0 = (1+\theta_1^2)\sigma_a^2 \quad \gamma_1 = -\theta_1\sigma_a^2 \quad \gamma_2 = \gamma_3 = \dots = 0$$

$\Rightarrow$ If $\gamma_k \approx 0$ for $k \geq 2$, I can assume it is an MA(1) model

$\Rightarrow$ If $\gamma_k \approx 0$ for $k \geq n$ $\Rightarrow$ I can assume it is an MA$(n-1)$ model

Problems: - Sample auto-correlation is a poor estimator for large lags (I really do not know how auto-correlation behaves for large lags)

- What if γ_k does not die out rapidly enough?

ii) <u>Brute force</u> → methodically check ARMA models.

$(1,0) \rightarrow (2,0) \rightarrow (2,1) \rightarrow (3,0) \rightarrow (3,1) \rightarrow (3,2)$

Stop fitting once adding more parameters does not decrease the RSS significantly (we'll see what "significant" is) pp. 2

Problem: every time I increase model order by 1, I check n models!

To get to ARMA(n, n-1), I need to check $1+2+3+\dots+n$ models!

iii) System based approach

It consists of successively fitting ARMA(n, n-1) models untill the RSS does not decrease significantly. Why the "fancy name"?

a) Intuitively nice! Fewer models to check! Increase model order by 1 => only 1 model to check

b) Plausible from the systems point of view!

When we described G.F., we said that a time-series X_t can be seen as an output of a system whose input was white noise

$$X_t = \sum_{j=0}^{\infty} G_j a_{t-j} = G_t * a_t$$

$$a_t \ \text{NIID} \sim N(0, \sigma_a^2)$$

and the G.F. coefficients were the impulse response coeffs of the system that, when driven by white noise, produced X_t.

Let's observe dynamic behavior of this impulse response & see what kinds of models we obtain

i) No dynamics $\quad G_j = 0, \ j = 1, 2, \ldots, n \qquad X_t = a_t \quad$ (ARMA(0,0))

ii) First order dynamics $\quad G_j = \lambda^j \Rightarrow X_t = (\sum \lambda^j B^j) a_t$

$$= \frac{1}{1 - \lambda B} a_t$$

$\Rightarrow (1 - \lambda B) X_t = a_t \quad \longrightarrow \quad$ ARMA(1,0)

iii) Second order dynamics $\quad G_j = g_1 \lambda_1^j + g_2 \lambda_2^j \quad (g_1 + g_2 = 1)$

$$\Rightarrow X_t = g_1 \frac{1}{1-\lambda_1 B} a_t + g_2 \frac{1}{1-\lambda_2 B} a_t = \frac{(g_1+g_2) - (\lambda_1 g_2 + \lambda_2 g_1) B}{1 - (\lambda_1 + \lambda_2) B + \lambda_1 \lambda_2 B^2} a_t$$

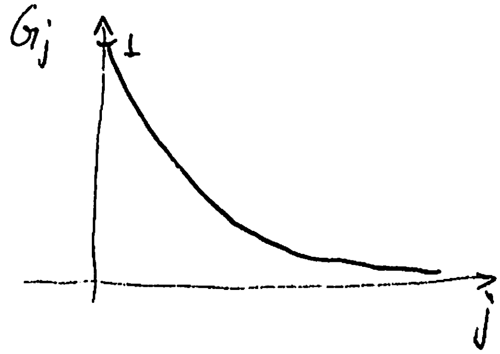
$$\Rightarrow \left(1-\underbrace{(\lambda_1+\lambda_2)}_{\phi_1}B+\underbrace{\lambda_1\lambda_2}_{-\phi_2}B^2\right)X_t=\left(1-\underbrace{(\lambda_1 g_2+\lambda_2 g_1)}_{\theta_1}B\right)a_t$$

we get ARMA(2,1) model.

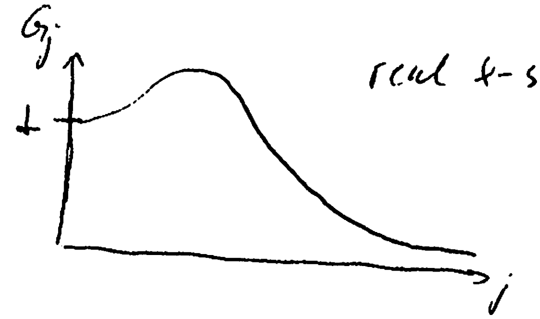
Only in a special case when $\lambda_1 g_2 + \lambda_2 g_1 = 0$, we get an AR(2) model (i.e. ARMA(2,0)).

Hence, it makes no sense to check AR(2) regularly!

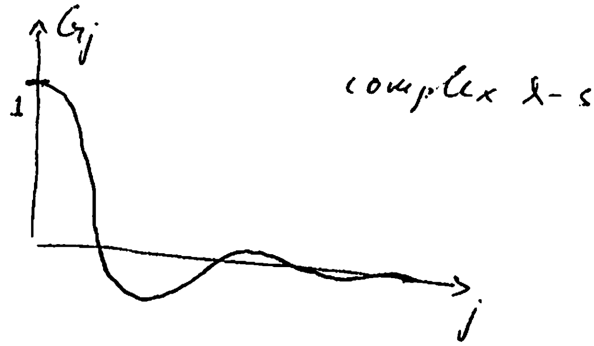
What do we get when we transit from 1st to 2nd order dynamics?



AR(1)



ARMA(2,1)



We get "richer" G-s when we go from (1,0) to (2,1).

We'll see that when we sample a simple mass-spring-damper system, we will get an ARMA(2,1) model

$\Rightarrow$ it's a very important model!

- third order dynamics $G_j = g_1\lambda_1^j + g_2\lambda_2^j + g_3\lambda_3^j$

$$X_t = \frac{\varphi_2(B)}{\varphi_3(B)} a_t$$

$$\varphi_2(B) = 1 - [(g_1\lambda_1 + g_2\lambda_2 + g_3\lambda_3) - (\lambda_1+\lambda_2+\lambda_3)]B \quad \leftarrow \theta_1$$

$$\theta_2 \rightarrow [g_1\lambda_1^2 + g_2\lambda_2^2 + g_3\lambda_3^2 + (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) - (\lambda_1+\lambda_2+\lambda_3)(g_1\lambda_1 + g_2\lambda_2 + g_3\lambda_3)]B^2$$

$$\varphi_3(B) = 1 - (\lambda_1+\lambda_2+\lambda_3)B + (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)B^2 - \lambda_1\lambda_2\lambda_3 B^3$$

Hence, I get an ARMA(3,2) model!

ARMA(3,1) & ARMA(3,0) are very special cases!

Obviously, n^{th} order dynamics $G_j = g_1\lambda_1^j + g_2\lambda_2^j + \dots + g_n\lambda_n^j$ leads to an ARMA(n, n-1) model!

~~Henq~~ Hence, by going from ARMA(n, n-1) to ARMA(n+1, n), I am increasing order of model dynamics by 1!

$\boxed{Q}$ Can I represent any dynamics as a sum of exponentials?

$\boxed{A}$ very important, Any stable discrete function G_j can be approximated to within an arbitrarily selected accuracy by some function

$$G_j = g_1 \lambda_1^j + g_2 \lambda_2^j + \ldots + g_n \lambda_n^j$$

ficiently large n. More formally,

$(\forall \varepsilon > 0)$ $(\exists n \in \mathbb{N},\ g_1, g_2, \ldots, g_n,\ \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R})$ such that

$$\sum_{j=0}^{\infty} [\, G_j - (g_1 \lambda_1^j + g_2 \lambda_2^j + \ldots + g_n \lambda_n^j) \,] < \varepsilon$$

Hence!

Any stationary STABLE time series can be approximated arbitrarily close ~~by~~ by an ARMA(n, n-1) model of a sufficiently high order!

Q What if the time-series is generated by an unstable system?

A Locally, we can approximate G_j as close as we want, but globally no guaranties can be given.

You'll hate predicting from unstable time-series models cause then you are exiting the area where our model approximation is good!

We've seen that from the point of view of dynamics of Green's Fcn, we should be checking models of order $(2,1) \rightarrow (3,2) \rightarrow \ldots \rightarrow (n, n-1) \rightarrow \ldots$

When we introduce some more physics, we can make it even faster

i) Mass spring damper system

$$m\ddot{x} + b\dot{x} + kx = \xrightarrow{\text{sampling}} ARMA(2,1)$$ (Ch. 7. describe this)

ii) Mass spring damper with 2 d.o.f. will have 2 equations of 2nd order, which when combined, gives a 4th order ODE $\xrightarrow{\text{sampling}}$ ARMA(4,3)

iii) Each new d.o.f raises the ARMA model order by 2.

$\Rightarrow$ perhaps we can do $(2,1) \rightarrow (4,3) \rightarrow (6,5) \rightarrow \ldots$

<u>REASON 2</u>: λ-s are coming either as a real root or as a pair of complex roots. If I fit a model of an odd

order. I force one root (at least) to be real. If I have sinusoids only in G_t, this approximation becomes bad (or adds virtually nothing to reduction of RSS). Then, I am likely to get an insignificant RSS reduction & stop model fitting (even though adding one more order would reduce the RSS significantly).

If indeed one root is real & order of the model is odd, one of the roots in the even-ordered ARMA model I forced upon it will be close to zero (meaning that the highest order ϕ coefficient will be close to zero) & that would tip me off that perhaps I can lower the order of the model by 1.

Reason 3 to do $(2,1) \to (4,3) \to (6,5) \to \ldots \to (2n, 2n-1) \to \ldots$

Twice less effort!

You'll see, in such systems, as we predict further & further, prediction errors will grow beyond any bounds

Now, we can also answer better, why we always have ARMA(n, n-1)!

i) n-th order dynamics of the impulse response:

$$X_t = \left(\frac{g_1}{1-\lambda_1 B} + \frac{g_2}{1-\lambda_2 B} + \dots + \frac{g_n}{1-\lambda_n B} \right) a_t = \frac{P_{n-1}(B)}{P_n(B)}$$

$\Rightarrow$ $P_n(B) X_t = P_{n-1}(B) a_t$ $\Rightarrow$ MA order MUST be n-1

ii) Another way: when we have AR portion of order n, that means, the G.F. is described by a difference equation

$$(1-\phi_1 B - \phi_2 B^2 \dots - \phi_n B^n) G_t = 0 \qquad t \geq n$$

If time-series begun at $t=0$, there can be only n initial conditions for initiating G-s!

Since $G_0 = 1$, I can play with only n-1 params!

Modeling Strategy (modified Fig. 4.1)

$n \leftarrow 1$

fit ARMA$(2n+2, 2n+1)$

Did RSS of $2n+2, 2n+1$ drop significantly compared to the RSS of $2n, 2n-1$?

YES: $n \leftarrow n+1$

NO (adequate model is $2n, 2n-1$)

Does RSS of $2n-1, 2n-2$ jump significantly compared to $2n, 2n-1$?

YES (keep $2n, 2n-1$)

Keep dropping highest order MA parameters and checking the jumps in the RSS (using F tests). This leads to an ARMA$(2n, m)$ model

NO (keep $2n-1$, $2n-2$)

keep dropping the highest order MA parameters and checking for changes in the RSS using F-tests. This leads to an ARMA$(2n-1, m)$ model

Confirm with the autocorrelation of residuals. If they are NOT white, keep increasing model orders!

Checks of Adequacy and Modeling Procedure

Let A_0 be the RSS of the unrestricted model (H_1)

Let A_1 be the RSS of the restricted model (H_0)

Then, let us define ratio

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)}$$

where:

s - number of restricted parameters

N - number of samples

r - number of estimated parameters

If both ARMA$(2n+2, 2n+1)$ and ARMA$(2n, 2n-1)$ are adequate, then $F \sim$ Fischer distr $(s, N-r)$

Hence, reduction in RSS is significant if

$$F > F_{s, N-r, \alpha}$$

α - Confidence rate, usually set to 0.95 or 0.99

You can get it from table D (pp. 508-513)

If $F > F_{s, N-r, \alpha}$ => I must continue fitting higher and higher orders

If $F < F_{s, N-r, \alpha}$ the new model did NOT improve RSS significantly and the old model can be considered adequate

<u>Problem</u> The test is correct ONLY if both models are adequate! $\Rightarrow$ you can get a bogus small RSS just because theoretically, the test is off!

That's why you'll do that last check to see if $\hat{\rho}_k$-s are small ($\hat{\rho}_k \leq \frac{2}{\sqrt{N}}$) i.e. if residuals are indeed white!

So, when I have N-samples and am testing

ARMA$(2n+2, 2n+1)$ vs ARMA$(2n, 2n-1)$
A_0 A_1

# of restricted params $s = 4$

# of estimated params for unrestricted model?

$$r = \underbrace{2n+2}_{\phi\text{-s}} + \underbrace{2n+1}_{\theta\text{-s}} + \underbrace{1}_{\mu} = 4n+4$$

	ARMA( )				
Parameter	(2,1)	(3,2)	(4,2)	**(4,3)**	(6,5)
Φ_1	1.867 ± .1058	1.538 ± .9546	1.085 ± .1409	**1.825 ± .3694**	.7341 ± .3470
Φ_2	-.8666. ± .1055	-.7543 ± 1.621	-1.063 ± .1905	**-1.907 ± .4545**	-.3836 ± .4525
Φ_3		.2118 ± .7368	.8748 ± .1616	**1.716 ± .4413**	.9424 ± .4757
Φ_4			.0567 ± .1047	**-.6387 ± .3465**	-.1900 ± .4756
Φ_5					.6316 ± .4472
Φ_6					-.7308 ± .2628
θ_1	.9751 ± .0613	.6269 ± .9630	.1048 ± .1108	**.9025 ± .3322**	-.2031 ± .3674
θ_2		-.1151 ± .8355	-.9242 ± .1086	**-1.018 ± .1406**	-.4908 ± .2842
θ_3				**.7447 ± .2929**	.5012 ± .2886
θ_4					.2911 ± .2836
θ_5					.7624 ± .3326
RSS	5.35E+08	5.11E+08	5.14E+08	**4.89E+08**	4.65E+08

Table for example A
sales in the UH sandwitch shop.

Example of Modeling Sandwitch Sales in Ann Arbor

* First test $(4,3)$ vs $(2,1)$

$N = 157$

$A_1 = RSS_{2,1} = 5.35 \cdot 10^8$

$A_0 = RSS_{4,3} = 4.89 \cdot 10^8$

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)}$$

$s = \#$ of restricted params $= 4$

$$= 3.4795$$

$r = \underset{\uparrow \\ \text{As}}{4} + \underset{\uparrow \\ \text{Cs}}{3} + \underset{\uparrow \\ \mu}{1} \rightarrow \#$ of estimated params in $(4,3)$

$$F^{95\%}_{s,N-r} = F^{95}_{4,149} = 2.37$$

$F > F^{95}_{4,149} \Rightarrow$ reduction is significant $\Rightarrow$ keep fitting

* Second test $(6,5)$ vs $(4,3)$

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)}$$

$N = 157$

$A_1 = RSS_{4,3} = 4.89 \cdot 10^8$

$F = 1.9041$

$A_0 = RSS_{6,5} = 4.65 \cdot 10^8$

$$F^{95\%}_{s,N-r} = F^{95\%}_{4,145} = 2.37$$

$s = 4$

$r = 6 + 5 + 1 = 12$

$F < F^{95\%}_{4,145} \Rightarrow$ no need to keep increasing the order!

Hence, both (4,3) & (6,5) are adequate (potentially)

We need to check $\hat{\rho}_k$-s now for residuals, but let's pretend they are indeed uncorrelated.

Even though $\hat{\phi}_4$ and $\hat{\theta}_3$ do not include 0 inside their confidence intervals, let's check if (3,2) is potentially correct (adequate)

Third test: (3,2) vs (4,3)

$$F = \frac{(A_1 - A_0)/s}{A_0/N-r}$$

$A_1 = RSS_{3,2} = 5.11 \cdot 10^8$

$A_0 = RSS_{4,3} = 4.89 \cdot 10^8$

$F = 3.3314$

$s = 2$

$r = 4+3+1 = 8$

$F^{95\%}_{s,N-r} = F^{95\%}_{2,149} = 3.056$

$N = 157$

$F > F^{95\%}_{2,149}$ => change of RSS is significant! => keep (4,3)

How about trying to drop highest order θ-s?

Test 4: (4,2) vs (4,3) (dropped θ_3)

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)}$$

$A_1 = RSS_{4,2} = 5.14 \cdot 10^8$

$A_0 = RSS_{4,3} = 4.89 \cdot 10^8$

$F = 7.5025$

$s = 1$; $r = 4+3+1 = 8$; $N = 157$

$F^{95\%}_{1,149} = F^{95\%}_{s,N-r} = 3.84$; $F > F^{95\%}_{1,149}$ => KEEP (4,3)!