NOT ALL MA(q) ARE INVERTIBLE :

Invertible means, the MA(q) can be written as AR(∞)
useful for forecasting, as a_t is a function of past z

example: $Z_t = a_t - \theta_1 a_{t-1}$

$$= a_t - \theta_1 (Z_{t-1} - \theta_1 a_{t-2})$$

$$= -\theta_1 Z_{t-1} + a_t + \theta_1^2 a_{t-2}$$

substitute for a_{t-2}, etc

the AR(∞) is finite if the powers of θ_1 converge
to 0, i.e. when $|\theta_1| < 1$

in general :

$$Z_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

$$= (1 - \theta_1 B - \dots - \theta_1 B^q) a_t$$

the moving average polynomial has to have the
roots outside the unit circle.

Wold theorem:

Implies that any stationary AR(p) can be written as MA(∞) (Moving Average)

example:

$$Z_t = \phi_1 Z_{t-1} + a_t$$

$$Z_t = \phi_1 (\phi_1 Z_{t-2} + a_{t-1}) + a_t$$

$$\vdots$$

$$Z_t = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots$$

$$Z_t = \sum_{i=0}^{\infty} \phi_1^i a_{t-i} \quad \text{MA}(\infty)$$

Z_t is finite as long as ϕ_1^i are summable, i.e the powers of ϕ_1 converge to 0. This is ensured by stationarity.

The Wold theorem says that any stationary process can be arbitrarily well approximated by a linear model, which is MA(∞).

MA(∞) involves a polynomial of infinite order in B, which by definition can always be written as a ratio of two finite polynomials

$$Z_t = \frac{B(B)}{A(B)} a_t \Rightarrow A(B) Z_t = B(B) a_t$$

ARMA(p,q)

$$(1-\phi_1 B - \dots - \phi_p B^p)\, z_t = (1-\theta_1 B - \dots - \theta_q B^q)\, a_t$$

or, with a constant:

$$z_t = \delta + \underbrace{\phi_1 z_{t-1} + \dots + \phi_p z_{t-p}}_{\text{require stationarity}} + a_t \underbrace{- \theta_1 a_{t-1} - \dots - \theta_q z_{t-q}}_{\text{require invertibility}}$$

- $E(z_t) = \mu = \dfrac{\delta}{1-\phi_1 - \dots - \phi_p}$ like in AR(p)

- **ex: ARMA(1,1)**

$$z_t = \delta + \phi_1 z_{t-1} + a_t + \theta_1 a_{t-1}$$

- **$\gamma_0 = var(z_t)$** $= E(\dot{z}_t^2) = E[\phi_1 \dot{z}_{t-1} + a_t - \theta_1 a_{t-1}]^2$

$$= \phi_1^2 \gamma_0 - 2\phi_1\theta_1 E(\dot{z}_{t-1} a_{t-1}) + \sigma_a^2 + \theta_1^2 \sigma_a^2$$

$$E(\dot{z}_{t-1} \cdot a_{t-1}) = E[(\phi_1 \dot{z}_{t-2} + a_{t-1} - \theta_1 a_{t-2})\, a_{t-1}]$$

$$= E(a_{t-1}^2) = \sigma_a^2$$

$$\gamma_0 = \left(\frac{1+\theta_1^2 - 2\phi_1\theta_1}{1-\phi_1}\right)\sigma_a^2$$

autocovariances

$$\gamma_1 = E(\dot{z}_{t-1}\,\dot{z}_t) = E\left[\dot{z}_{t-1}(\phi_1 z_{t-1} + a_t - \theta_1 a_{t-1})\right]$$

$$= \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1-\phi_1^2}\,\sigma_a^2$$

$$= \phi_1\gamma_0 - \theta_1\sigma_e^2$$

$$\gamma_K = \phi_1\,\gamma_{K-1} \quad ; \quad K \geqslant 2$$

autocorrelations:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\theta_1\phi_1}$$

$$\rho_K = \phi\,\rho_{K-1} \quad , \quad K>1$$

MA part contributes to the ACF up to lag q only. After, the ACF behaves like for AR(p):

$$\rightarrow \quad \rho_K = \phi_1\rho_{K-1} + \phi_2\rho_{K-2} + \cdots + \phi_p\rho_{K-p}$$

at $K>q$

BOX JENKINS

method of estimation for stationary ARMA(p,q)

step 1: IDENTIFICATION

plot ACF

plot PACF

AR : PACF cuts off

MA : ACF cuts off

ARMA : none of these

step 2 ESTIMATION

AR : OLS or MLE

MA, ARMA : MLE

STEP 3: DIAGNOSTIC CHECKING

STEP 4: FORECASTING

DIAGNOSTIC CHECKING

1) USE the goodness of fit criteria to assess the fit:

AKAIKE:

- $AIC(p,q) = -2\ln L + 2\cdot(p+q) = \ln\hat{\sigma}_a^2 + \frac{2(p+q)}{n}$

SCHWARTZ (BAYESIAN)

- $BIC(p,q) = \ln(\hat{\sigma}_a^2) + (p+q)\frac{\ln n}{n}$

HANNAN-QUINN

- $\phi(p,q) = \ln(\hat{\sigma}_a^2) + c(p+q)\frac{\ln[\ln(n)]}{n}, \quad c>2$

the criteria "penalize" for using too many coefficients, i.e. for too high orders p,q $\Rightarrow$ a parsimonious specification is preferred

IF THE FIT IS CORRECT, THE RESIDUALS FROM THE MODEL ARE WHITE NOISE:

2) USE BOX PIERCE OR LJUNG BOX STATISTICS to test that the residuals jointly have ACF 0 up to lag K:

B-P $\quad Q(\hat{\rho}) = n\sum_{k=1}^{K}\hat{\rho}_k^2(\hat{a})$

L-B $\quad Q(\hat{\rho}) = n(n+2)\sum \hat{\rho}_k^2(\hat{a})/(n-k)$

under Ho:

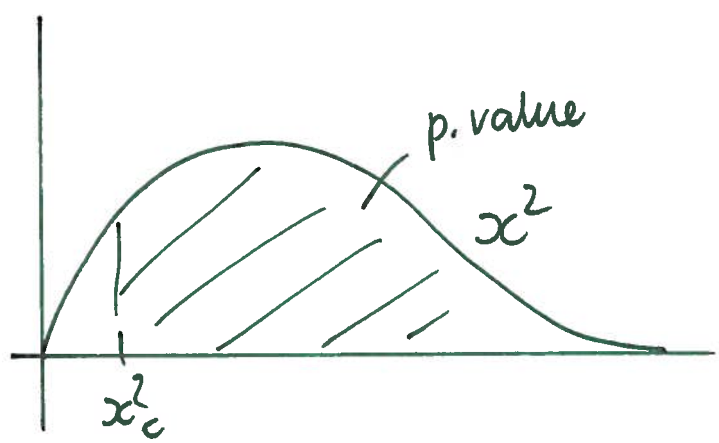
$$\hat\rho_1(\hat a) = \dots = \hat\rho_K(\hat a) = 0$$

L-B, B-P $\overset{a}{\sim}$ $\chi^2(K - l)$ where K is up to which lag the residuals are tested

$l = p+q$ or $p+q+1$ ("intercept")

intuition:

if $\hat\rho$'s are very small, pos or negat then $\hat\rho^2(a)$ are small positive. Reject Ho: "residuals are WN up to lag K" or "$\rho_1(\hat a) = \dots = \rho_K(\hat a)$" if the sum of their squared values is "very" different from 0.



hence "acept" Ho when p-value 0.8, 0.9 etc. close to 1
reject otherwise.

WATCH OUT: some software test at $(1-\alpha)$ and very small p-values indicate WN (S+)

FORECASTING WITH AR(1)

1. h=1

$$Z_{T+1} = \delta + \phi_1 Z_T + a_{T+1}$$

- at time t we know the history and present value of Z_t

$$\underset{t}{E}[Z_t] = E[Z_t | Z_{t-1}, Z_{t-2} \dots Z_0] = Z_t$$

$$\underset{t}{E}[a_t] = a_t$$

- we don't know the future for $h > 0$

$$\underset{t}{E}[a_{t+h}] = E[a_t]$$

↑ UNCONDITIONAL EXPECTATION

- the forecast at horizon h=1 :

$$\hat{Z}_{T+1} = \underset{T}{E}[Z_{T+1}] = \underset{T}{E}[\delta + \phi_1 Z_T + a_{T+1}]$$

$$= \delta + \phi_1 Z_T$$

- forecast error at h=1

$$\hat{a}_{T+1} = Z_{T+1} - \hat{Z}_{T+1} = (\delta + \phi_1 Z_T + a_{T+1}) - (\delta + \phi_1 Z_T)$$
$$= a_{T+1}$$

variance of the forecast error:

$$var(\hat{a}_{T+1}) = var(a_{T+1}) = \sigma_a^2$$

2. $h=2$

$$z_{T+2} = \delta + \phi_1 z_{T+1} + a_{T+2}$$

$$\hat{z}_{T+2} = E_T[z_{T+2}] = \delta + \phi_1 E_T[z_{T+1}] = \delta + \phi_1 \hat{z}_{T+1}$$

$$\hat{a}_{T+2} = z_{T+2} - \hat{z}_{T+2} = (\delta + \phi_1 z_{T+1} + a_{T+2}) - (\delta + \phi_1 \hat{z}_{T+1})$$

$$= a_{T+2} + \phi_1 (z_{T+1} - \hat{z}_{T+1}) = a_{T+2} + \phi_1 \hat{a}_{T+1}$$

$$= a_{T+2} + \phi_1 a_{T+1} \leftarrow \text{MA}(1)$$

$$var(\hat{a}_{T+2}) = var(a_{T+2} + \phi_1 a_{T+1}) = \sigma_a^2 + \phi_1^2 \sigma_a^2 =$$

$$= \sigma_a^2 (1 + \phi_1^2)$$

3. any h:

$$\hat{z}_{T+h} = \delta + \phi_1 E_T[z_{T+h-1}] = \delta + \phi_1 \hat{z}_{T+h-1}$$

Note :

the forecast function converges to the unconditional expectation $E(z_t)$ as $h \to \infty$

the variance of forecast error converges to the unconditional variance of z_t

- forecast error at horizon h follows $\underline{\text{MA}(h-1)}$

CONFIDENCE INTERVAL FOR THE FORECAST:

$$\hat{z}_{t+h} \pm 1.96 \cdot \sqrt{\hat{\text{var}} \text{ of forecast error}}$$

widens up when $h \to \infty$

Forecasting with MA(1):

$$z_t = \mu + a_t - \theta_1 a_{t-1}$$

1. $h=1$

$$z_{T+1} = \mu + a_{T+1} - \theta_1 a_T$$

- forecast: $\underset{T}{E}[z_{T+1}] = \underset{T}{E}[\mu + a_{T+1} - \theta_1 a_T]$
$$= \mu - \theta_1 a_T$$

- error: $\hat{a}_{T+1} = z_{T+1} - \hat{z}_{T+1} = (\mu - \theta_1 a_T + a_{T+1}) - (\mu - \theta_1 a_T)$
$$= a_{T+1}$$

- $\text{var}(\hat{a}_{T+1}) = \sigma_a^2$

2. $h=2$

$$z_{T+2} = \mu + a_{T+2} - \theta_1 a_{T+1}$$

$$\hat{z}_{T+2} = \underset{T}{E}[z_{T+2}] = \mu$$

$$\text{var}(\hat{a}_{T+2}) = \text{var}(a_{T+2} - \theta_1 a_{T+1}) = \sigma_a^2(1+\theta_1^2)$$

FORECASTING WITH ARMA(1,1)

$$Z_t = \delta + \phi_1 Z_{t-1} + a_t - \theta_1 a_{t-1}$$

- $h=1$

$$Z_{T+1} = \delta + \phi_1 Z_T + a_{T+1} - \theta_1 a_T$$

$$\hat{Z}_{T+1} = \underset{T}{E}[Z_{T+1}] = \delta + \phi_1 Z_T - \theta_1 a_T$$

- $\hat{a}_{T+1} = Z_{T+1} - \hat{Z}_{T+1} = a_{T+1}$
- $\text{var}(\hat{a}_{T+1}) = \text{var}(a_{T+1}) = \sigma_a^2$

- $h=2$

$$\hat{Z}_{T+2} = \underset{T}{E}[\delta + \phi_1 Z_{T+1} + a_{T+2} - \theta_1 a_{T+1}]$$

$$= \delta + \phi_1 \underset{T}{E} Z_{T+1} = \delta + \phi_1 \hat{Z}_{T+1}$$

- $\hat{a}_{T+2} = (\phi_1 - \theta_1) a_{T+1} + a_{T+2}$
- $\text{var}(\hat{a}_{T+2}) = \sigma_a^2 [(\phi_1 - \theta_1)^2 + 1]$

- $h=3$

$$\hat{Z}_{T+3} = \delta + \phi_1 \hat{Z}_{T+2}$$

$$\text{var}(\hat{a}_{T+3}) = \sigma_a^2 [1 + (\phi_1 - \theta_1)^2 + (\phi_1^2 - \phi_1\theta_1)]^2$$