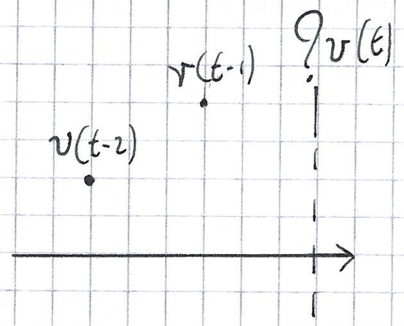
*MIDA 1

1.1: INTRODUZIONE

PREDICTION PROBLEM



OVVERO, DATI $v(t-1)$, $v(t-2)$, ..., $v(t-m)$ INDOVINARE IL VALORE DI $v(t)$

CHIAMEREMO $\hat{v}(t|t-m)$ UN m-STEP BACK PREDICTOR.

UNA REGOLA SEMPLICE PER COSTRUIRE UN PREDICTOR É DATA DAL **LINEAR PREDICTOR**:

$$\hat{v}(t|t-m) = a_1 v(t-1) + \dots + a_m v(t-m) + \varepsilon$$

IL QUALE, PERÒ, PRESENTA DEI PROBLEMI:

1) QUANTI SAMPLES PASSATI DOVREI USARE?
2) COME STIMARE I PARAMETRI $a_1 \dots a_m$?

IN QUELLA FORMULA ABBIAMO ANCHE INTRODOTTO ε, OVVERO IL **PREDICTION ERROR** $\varepsilon(t) = v(t) - \hat{v}(t|t-1, \dots, t-m)$

POSSIAMO ESPRIMERE LA BONTÀ DI UN PREDITTORE IN TERMINI DI PROPRIETÀ CHE DOVREBBE POSSEDERE ε:

1) ZERO-AVERAGE: ALTRIMENTI POSSIAMO AGGIUNGERE UN $\bar{\varepsilon}$ COSTANTE E MIGLIORARE IL PREDICTOR
2) L'IDEALE É CHE ε SIA FULLY UNPREDICTABLE (SE NON LO FOSSE VORREBBE DIRE CHE IL PREDICTOR PUÒ ESSERE MIGLIORATO)

FORMALMENTE QUINDI VOGLIAMO CHE ε SIA UN **WHITE NOISE**:

$\varepsilon(t) \sim WN(0, \lambda^2)$ (ZERO-AVERAGE, COVARIANZA $= \lambda^2$).

SUPPONIAMO CHE ε SIA QUINDI UN WN. ABBIAMO

$$\varepsilon(t) = v(t) - \hat{v}(t|t-1)$$

$$\hookrightarrow v(t) = \hat{v}(t|t-1) + \varepsilon(t) = a_1 v(t-1) + \dots + a_m v(t-m) + \varepsilon(t)$$

INTRODUCIAMO L'OPERATORE Z:

Z: $\quad Z v(t) = v(t+1)$: 1-STEP FORWARD OPERATOR

Z^{-1}: $\quad Z^{-1} v(t) = v(t-1)$: 1-STEP BACWARD OPERATOR

POSSIAMO QUINDI RISCRIVERE:

$$\hat{v}(t|t-n) = a_1 v(t-1) + \dots + a_n v(t-n) = a_1 z^{-1} v(t) + \dots + a_n z^{-n} v(t)$$

DA CUI

$$\varepsilon(t) = v(t)\left(1 - a_1 z^{-1} - \dots - a_n z^{-n}\right)$$

$$\rightarrow v(t) = \overline{W}(z)\,\varepsilon(t) \quad \text{CON} \quad \overline{W}(z) = \frac{1}{1 - a_1 z^{-1} - \dots - a_n z^{-n}}$$

DEFINIAMO $\overline{W}(z)$ FUNZIONE DI TRASFERIMENTO DA $\varepsilon(t)$ A $v(t)$:

$$\varepsilon(t) \longrightarrow \boxed{W(z)} \longrightarrow v(t)$$

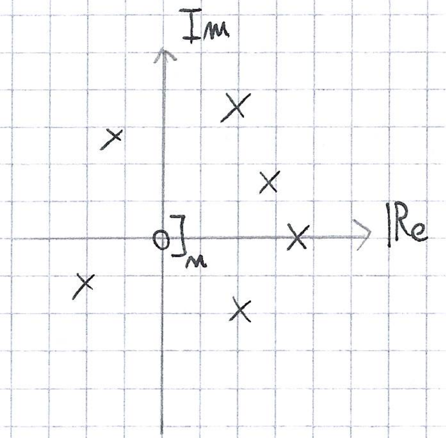
MOLTIPLICANDO SIA A NUM. CHE A DEN. PER z^n ABBIAMO:

$$W(z) = \frac{z^n}{z^n - a_1 z^{n-1} - \dots - a_n}$$

DA CUI:

O: n ZERI NELL'ORIGINE ($z^n = 0$)

X: n POLI ($z^n - a_1 z^{n-1} - \dots - a_n = 0$)



COSTRUIRE UN PREDITTORE È COME DESCRIVERE IL SEGNALE COME OUTPUT DI UN MODELLO DINAMICO AVENTE COME INPUT UN WHITE NOISE.

DOBBIAMO CAPIRE MEGLIO COSA SIA UN WHITE NOISE...

RANDOM CONCEPTS

RANDOM VARIABLE: VARIABILE REALE IL CUI VALORE DIPENDE DALL'ESITO DI UN ESPERIMENTO (FENOMENO) RANDOM

1) MEDIA (AVG. VALUE, EXPECTED VALUE) $= E[v] = m = \bar{V} = V_m$

DOTATA DELLE SEGUENTI PROPRIETÀ:

a) $E[v_1 + v_2] = E[v_1] + E[v_2]$

b) $E[\alpha_1 v_1 + \alpha_2 v_2] = \alpha_1 E[v_1] + \alpha_2 E[v_2]$

2) VARIANZA $= Var[v] = E[(v - E[v])^2] = \sigma^2, \lambda^2, \mu^2$

3) DEVIAZIONE STANDARD $= \sqrt{Var[v]} = \lambda$

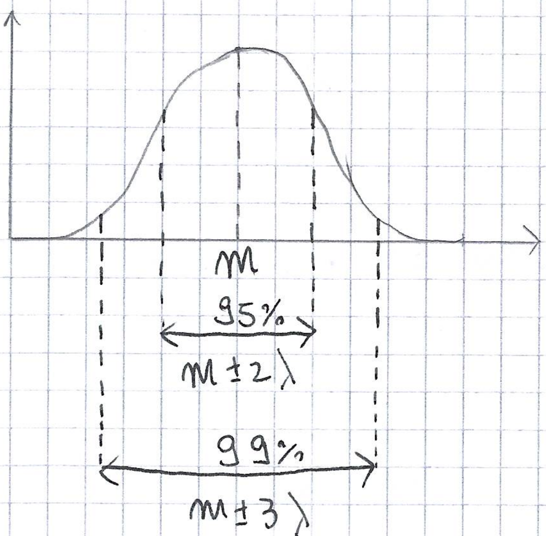
ES

$v \sim G(2, 36)$

$m = 2$

$\lambda^2 = 36$

$\lambda = 6$



4) CROSSED - VARIANCE $\lambda_{12} = E[(v_1 - m_1)(v_2 - m_2)]$

RANDOM VECTORS: ARRAY DI RANDOM VARIABLES $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

1) MEDIA $E[v] = \begin{bmatrix} E[v_1] \\ E[v_2] \end{bmatrix}$

2) VARIANZA $Var[v] = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix}$ CON:
- $\lambda_{11} = \lambda_{22}$
- $\Delta v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \begin{bmatrix} E[v_1] \\ E[v_2] \end{bmatrix}$

$= E[\Delta v \cdot \Delta v^T]$

EVIDENZIAMO CHE $Var[v]$ È <u>POSITIVE SEMI-DEF</u> $\rightarrow \det(Var[v]) \geq 0$

3) COVARIANZA $= \rho = \dfrac{\lambda_{12}}{\sqrt{\lambda_{11}}\sqrt{\lambda_{22}}}$ CON, PER VIA DI ↑, $|\rho| \leq 1$

$\rho = 0 \rightarrow v_1$ E v_2 SONO NON CORRELATE $(v_1 \perp v_2)$

$\rho = \pm 1 \rightarrow v_1$ E v_2 SONO MAX. CORRELATE $(v_2 = \alpha v_1)$

RANDOM PROCESS (STOCHASTIC): SEQUENZA DI RANDOM VARIABLES $v(t)$

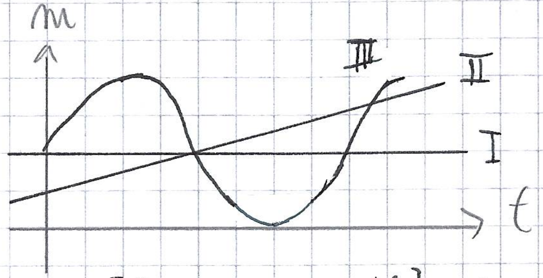
$v(t, s)$ ⟶ SPESSO "ABBREVIAMO" IN $v(t)$

- v → RAND VAR
- t → TEMPO DISCRETO
- s → OUTCOME DELL'ESPERIMENTO

1) MEDIA $= m(t) = E[v|t] = E_s[v(t,s)]$

PUO' ESSERE:

I) COSTANTE
II) PERIODICA
III) LINEARE
IV) VARIABILE



2) VARIANZA $= Var[v(t)] = E\left[(v(t) - m(t))^2\right] = \lambda^2(t)$ (DI SOLITO NON COSTANTE)

3) COVARIANZA $= \gamma(t_1, t_2) = E\left[(v(t_1) - m(t_1))(v(t_2) - m(t_2))\right]$

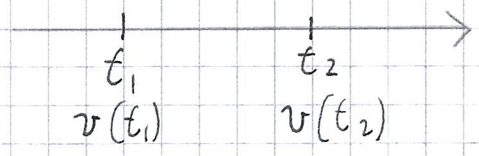
STATIONARY PROCESS: RANDOM PROCESS CON

1) $m(t) = m = \text{CONST}$
2) $\lambda^2(t) = \lambda^2 = \text{CONST}$
3) $\gamma(t_1, t_2)$ DIPENDE SOLO DA $\tau = t_2 - t_1$ $\left(\gamma(0) = \tau = 0 \rightarrow t_1 = t_2\right)$

DA CUI DERIVANO LE PROPRIETA':

1) $\gamma(t_1, t_2) = \gamma(t_2, t_1) \longrightarrow \gamma(\tau) = \gamma(-\tau)$ (FUNZIONE PARI)
2) $\gamma(0)$ = VARIANZA DEL PROCESSO $= \lambda^2 \geq 0$
3) $|\gamma(\tau)| \leq \gamma(0)$

DIM



$$|\rho| = \left| \frac{\lambda_{12}}{\sqrt{\lambda_{11}} \cdot \sqrt{\lambda_{22}}} \right| \leq 1$$

$\lambda_{12} = E[(v_1 - m_1)(v_2 - m_2)] = \gamma(t_1, t_2) = \gamma(\tau)$

$\sqrt{\lambda_{11}}, \sqrt{\lambda_{22}} \rightarrow \sqrt{\lambda^2}$

$$= |\rho| = \left| \frac{\gamma(\tau)}{\lambda^2} \right| \leq 1 \rightarrow |\gamma(\tau)| \boxed{\leq} |\lambda^2| = \lambda^2 = \gamma(0)$$

CVD

WHITE NOISE: PROCESSO STAZIONARIO CON $\eta(t) \sim WN(m, \lambda^2)$

1) $E[v(t)] = m \quad \forall t$
2) $VAR[v(t)] = \lambda^2 \quad \forall t$
3) $\gamma(t_1, t_2) = 0 \quad \forall t_1 \neq t_2 \quad \left(= \gamma(\tau) = \begin{cases} \lambda^2 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases} \right)$

OVVERO, NON C'È CORRELAZIONE TRA DUE RANDOM VAR ESTRATTE CASUALMENTE DAL PROCESSO AD OGNI t_1, t_2 CON $t_1 \neq t_2$

ES

DEFINIAMO UN NUOVO PROCESSO

$\eta(t) \sim WN(0, \lambda^2)$

$v(t) = \alpha \eta(t) + \beta \qquad \alpha, \beta \in \mathbb{R}$

$$E[v(t)] = \alpha \underbrace{E[\eta(t)]}_{0} + \beta = \beta, \quad \forall t$$

$$VAR[v(t)] = E[(v(t) - E[v(t)])^2] = E[(\alpha\eta(t) + \cancel{\beta} - \cancel{\beta})^2] = \alpha^2 E[\eta(t)^2] = \alpha^2\lambda^2, \quad \forall t$$

$$\gamma(t_1, t_2) = E[(v(t_1) - E[v(t_1)])(v(t_2) - E[v(t_2)])] =$$
$$= E[(\alpha\eta(t_1) + \cancel{\beta} - \cancel{\beta})(\alpha\eta(t_2) + \cancel{\beta} - \cancel{\beta})] = \alpha^2 \underbrace{E[\eta(t_1)\eta(t_2)]}_{\text{COV OF } \eta}$$

$$= \begin{cases} 0 & t_1 \neq t_2 \\ \alpha^2\lambda^2 & t_1 = t_2 \end{cases}$$

$\rightarrow$ IL NUOVO PROCESSO È UN WN CON β COME MEDIA E VARIANZA $\alpha^2\lambda^2$

$v(t) \sim WN(\beta, \alpha^2\lambda^2)$

1.2: MA / AR

MOVING AVERAGE PROCESSES

OVVERO, UN MODELLO IN CUI L'OUTPUT DIPENDE LINEARMENTE DAI VALORI CORRENTI E PASSATI DI UN RANDOM PROCESS.

DEFINIAMO UN NUOVO PROCESSO:

$$v(t) = C_0\, \eta(t) + C_1\, \eta(t-1), \quad C_0, C_1 \in \mathbb{R}, \quad \eta(t) \sim WN(0, \lambda^2)$$

TROVIAMONE LE CARATTERISTICHE:

$$E[v(t)] = E[C_0\eta(t)] + E[C_1\eta(t-1)] = C_0 \cdot 0 + C_1 \cdot 0 = 0$$

$$\begin{aligned} VAR[v(t)] &= E[(v(t) - E[v(t)])^2] = E[v(t)^2] = E[(C_0\eta(t) + C_1\eta(t-1))^2] \\ &= E[C_0^2\eta(t)^2] + E[C_1^2\eta(t-1)^2] + E[2C_0C_1\,\eta(t)\eta(t-1)] \\ &= C_0^2\lambda^2 + C_1^2\lambda^2 + 2C_0C_1 \cdot 0 = (C_0^2 + C_1^2)\lambda^2 \end{aligned}$$

$$\begin{aligned} \gamma(t, t-1) &= E[(v(t) - E[v(t)])(v(t-1) - E[v(t-1)])] \\ &= E[(C_0\eta(t) + C_1\eta(t-1) - 0)(C_0\eta(t-1) + C_1\eta(t-2) - 0)] \\ &= C_0^2 \underbrace{\cancel{E[\eta(t)\eta(t-1)]}}_{0} + C_1^2 \underbrace{\cancel{E[\eta(t-1)\eta(t-2)]}}_{0} + C_1C_0\,E[\eta(t-1)\eta(t-1)] \\ &\quad + C_0C_1 \underbrace{\cancel{E[\eta(t)\eta(t-1)]}}_{0} \\ &= \boxed{C_1C_0\lambda^2} \end{aligned}$$

INOLTRE:

$\gamma(t-1, t) =$ UGUALE

$\gamma(t, t-2) = 0$

$\vdots$

$\gamma(t, t-m) = 0$

CONCLUDIAMO CHE $v(t)$ E' UN PROCESSO STAZIONARIO CON

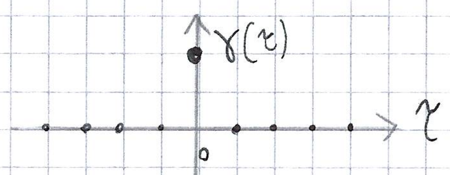
$m = 0, \ \forall t$

$VAR = (C_0^2 + C_1^2)\lambda^2, \ \forall t$

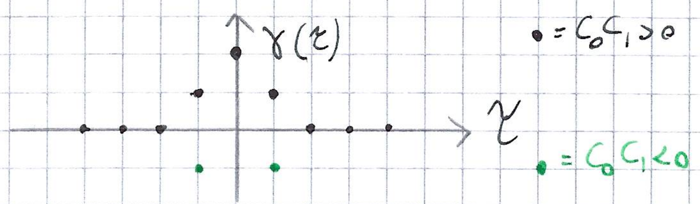
$$\gamma(t_1, t_2) = \begin{cases} 0 & \Delta t > 1 \\ C_0C_1\lambda^2 & t_2 = \pm t_1 \end{cases}$$

NON E' QUINDI UN WN PER VIA DELLA DIVERSA COVARIANZA:

$\boxed{WN}$



$\boxed{v(t)}$



SI TRATTA INFATTI DI UN PROCESSO MA(1).

ES DI PROCESSO MA(1)

$$\Delta_{\text{PREZZO PETROLIO}} = \varepsilon_t + 0.5\,\varepsilon_{t-1}$$

SIA ΔPREZZO $\begin{cases} >0 \rightarrow \text{PREZZO AUMENTA} \\ <0 \rightarrow \text{PREZZO DIMINUISCE} \end{cases}$ E SIA ε_t L'OCCORRENZA DI UN CICLONE IN UNA ZONA PETROLIFERA: SE AVVIENE A t IL PREZZO AUMENTA, E L'AUMENTO RIMANE PARZIALMENTE ANCHE A $t+1$.

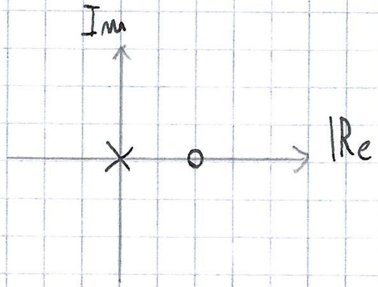
POSSIAMO INOLTRE OSSERVARE:

$$v(t) = c_0\,\eta(t) + c_1\,\eta(t-1) = c_0\,\eta(t) + c_1\,z^{-1}\eta(t)$$
$$= \left(c_0 + z^{-1}c_1\right)\eta(t)$$

$\eta(t) \longrightarrow \boxed{W(z)} \longrightarrow v(t)$

$$W(z) = c_0 + c_1 z^{-1} = c_0 + c_1\frac{1}{z} = \frac{c_0 z + c_1}{z}$$

- UNO ZERO : $c_0 z + c_1 = 0$
- UN POLO : $z = 0$

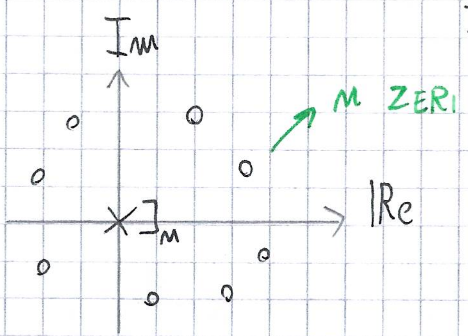


* ESTENSIONE A MA(n)

$$v(t) = c_0\,\eta(t) + c_1\,\eta(t-1) + \dots + c_n\,\eta(t-n)$$
$$v(t) = \left(c_0 + c_1 z^{-1} + \dots + c_n z^{-n}\right)\eta(t)$$

$$W(z) = \frac{c_0 z^n + c_1 z^{n-1} + \dots + c_n}{z^n}$$

- n ZERI
- n POLI IN 0



LA CLASSE DEI PROCESSI MA È UNA CLASSE DI PROCESSI STAZIONARI CON:

$$E[v(t)] = 0 \,, \quad \forall t$$
$$\mathrm{Var}(v(t)) = \left(c_0^2 + \dots + c_n^2\right)\lambda^2$$
$$\gamma(1) = \left(c_0c_1 + c_1c_2 + \dots + c_{n-1}c_n\right)\lambda^2$$
$$\gamma(2) = \left(c_0c_2 + c_1c_3 + \dots + c_{n-2}c_n\right)\lambda^2$$
$$\vdots$$
$$\gamma(k) = 0 \,, \quad k > n$$

FORMULA VELOCE COVARIANZA $\gamma(\tau)$ DI PROCESSI MA(N)

* ESTENSIONE A MA(∞)

SE $n \to \infty$, LA CONDIZIONE NECESSARIA PER AVERE VAR FINITA È CHE

$c_0^2 + c_1^2 + \dots\dots$ SIA FINITO.

AUTOREGRESSIVE PROCESSES

DEFINIAMO

OVVERO CHE DIPENDONO LINEARMENTE DAGLI OUTPUT PASSATI E DA UN TERMINE STOCASTICO.

$$v(t) = a\,v(t-1) + \eta(t) \qquad (AR(1)) \qquad \eta(t) \sim WN(0, \lambda^2)$$

POSSIAMO ESPANDERE $v(t-1) = a\,v(t-2) + \eta(t-1)$ DA CUI

$$v(t) = a^2 v(t-2) + a\,\eta(t-1) + \eta(t)$$ DI NUOVO...

$$v(t) = a^3 v(t-3) + \eta(t) + a^2 v(t-2) \; ...$$

RISCRIVENDO:

$$v(t) = c_0 \eta(t) + c_1 \eta(t-1) + c_2 \eta(t-2) + \ldots + c_m \eta(t-m) \quad] \; MA(\infty)$$

$c_0 \to 1$, $c_1 \to a$, $c_2 \to a^2$, $c_m \to a^m$

SE $|a| < 1$, $* = 0$: IL PROCESSO AR(1) E' WELL-POSED.

SI HA CHE $v(t)$ E' UN PROCESSO $MA(\infty)$ CON

$c_0 = 1$
$c_1 = a$
$c_2 = a^2$
$\vdots$

AVENDO $|a| < 1$, LA SOMMA $1 + c_0 + c_1 + \ldots$ E' $= \dfrac{1}{1-a^2}$

DEFINIAMO:

$$AR(1): \quad v(t) = a\,v(t-1) + \eta(t) \qquad \eta(t) \sim WN(0, \lambda^2), \quad |a| < 1$$

$$Var_{AR}(v(t)) = \frac{1}{1-a^2}\lambda^2 = \gamma(0)$$

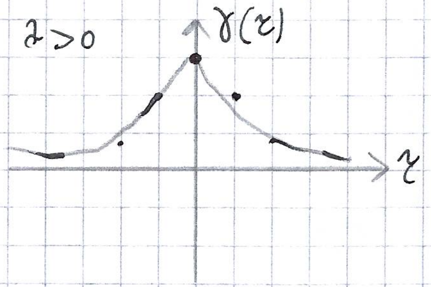
$$\gamma(\tau) = a^\tau \gamma(0)$$

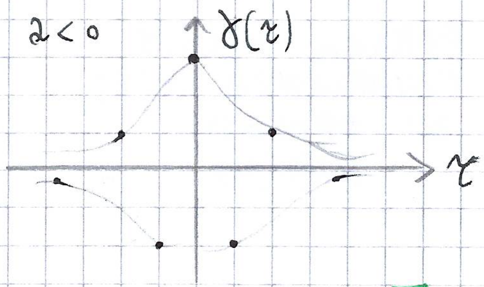
$$\gamma(1) = (c_0 c_1 + c_1 c_2 + \ldots +)\lambda^2 = (1 \cdot a + a \cdot a^2 + \ldots)\lambda^2 = a \frac{1}{1-a^2}\lambda^2$$

$$\gamma(2) = a^2 \frac{1}{1-a^2}\lambda^2$$

$\vdots$

$$\gamma(\tau) = a^{|\tau|} \frac{1}{1-a^2}\lambda^2$$

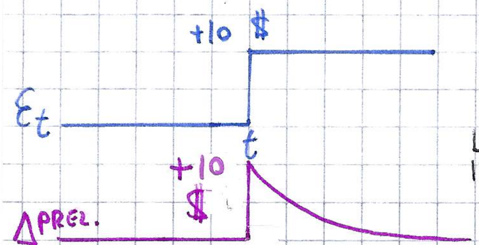




ES DI PROCESSO AR(1)

$$\Delta \text{PREZZO PETROLIO}_t = 0{,}5\,\Delta \text{PREZZO PETROLIO}_{t-1} + \varepsilon_t$$

SIA ε_t L'EFFETTO SUL PREZZO DI UN EVENTO. COSA ACCADE ALL'AVVENIRE DI ε?



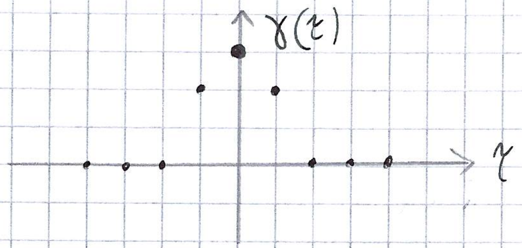
L'EFFETTO SUL PREZZO IMPIEGA TEMPO A SCEMARE.

MA(1) : EFFETTO SI MANIFESTA PER DUE ISTANTI DI TEMPO

AR(1) : EFFETTO POTENZIALMENTE INFINITO (CALA SE $|a|<1$)

* AR(1) VS MA(1)

$v(t) = c_0 \eta(t) + c_1 \eta(t-1)$



$v(t) = a v(t-1) + \eta(t) \quad (|a|<1)$

* STABILITÀ

DA $v(t) = a_1 v(t-1) + a_2 v(t-2) + \dots + a_m v(t-m) + \eta(t)$

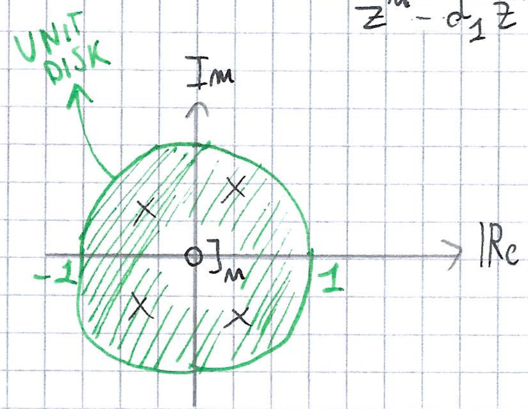
PASSIAMO A

$$v(t) = a_1 z^{-1} v(t) + a_2 z^{-2} v(t) + \dots + a_m z^{-m} v(t) + \eta(t)$$

$$\rightarrow v(t) = W(z)\, \eta(t)$$

$$W(z) = \frac{z^m}{z^m - a_1 z^{m-1} - \dots - a_m}$$

:
- m ZERI IN ORIGINE
- m POLI



GLI m POLI DIPENDONO DA $a_1, \dots, a_m$.

SE TUTTI I POLI SONO ALL'INTERNO DELL'UNIT DISK, ALLORA $v(\cdot)$ È STAZIONARIO.

ARMA PROCESSES

$$v(t) = \underbrace{a_1 v(t-1) + \dots + a_{m_a} v(t-m_a)}_{AR} + \underbrace{c_0 \eta(t) + \dots + c_{m_c} \eta(t-m_c)}_{MA}$$

CON m_a = ORDER DELLA PARTE AR

m_c = ORDER DELLA PARTE MA

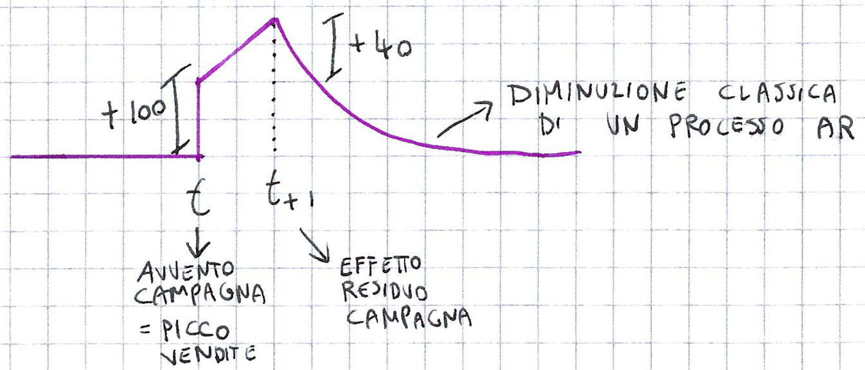
$$W(z) = \frac{c_0 + \dots + c_{m_c} z^{-m_c}}{1 - a_1 z^{-1} - \dots - a_{m_a} z^{-m_a}} = \frac{C(z)}{A(z)}$$

UN ARMA PROCESS È STAZIONARIO SE TUTTI I POLI SONO NELL'UNIT DISK.

ES ARMA (1,1) PROCESS

$$X_t = \underbrace{\rho X_{t-1}}_{AR} + \underbrace{\varepsilon_t + \varepsilon_{t-1}}_{MA}$$

CON
- X_t = VENDITE PRODOTTO
- ε_t = AVVENTO CAMPAGNA PUBBLICITARIA



1.3: ANALISI IN FREQUENZA

SPETTRO DI UN PROCESSO

OVVERO UNA RAPPRESENTAZIONE DI UN PROCESSO STOCASTICO NEL DOMINIO DELLE FREQUENZE.

$$v(t,s) \xrightarrow[\text{CARATTERIZ. DA COVARIANZA}]{} \gamma(\tau) \xrightarrow[\text{(TRASF DI FOURIER)}]{\mathcal{F}} \Gamma(\omega) = \text{SPETTRO}$$

$$\Gamma(\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{+\infty} \gamma(\tau) e^{-j\omega\tau}$$
→ DEFINISCE COME LE FREQUENZE IN INPUT SONO DISTRIBUITE SULL'OUTPUT, PERMETTENDOCI DI "CLASSIFICARE" L'IMPORTANZA DEI DIVERSI INPUT

PROPRIETÀ:

$\Gamma(\omega) =$

$\gamma(-2)e^{j2\omega}$	$\tau=-2$
$\gamma(-1)e^{j\omega}$	$\tau=-1$
$\gamma(0)$	$\tau=0$
$\gamma(1)e^{-j\omega}$	$\tau=1$
$\gamma(2)e^{-j2\omega}$	$\tau=2$

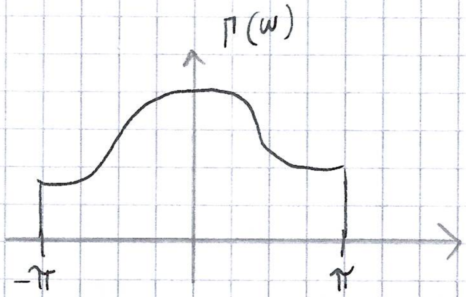
$$\gamma(-1)=\gamma(1) \rightarrow \gamma(1)\left[e^{j\omega}+e^{-j\omega}\right] = 2\gamma(1)\cos\omega$$

$$\rightarrow \gamma(-2)=\gamma(2) \rightarrow \gamma(2)\left[e^{j2\omega}+e^{-j2\omega}\right] = 2\gamma(2)\cos 2\omega$$

$$\Longrightarrow \Gamma(\omega) = \gamma(0) + 2\gamma(1)\cos(\omega) + 2\gamma(2)\cos(2\omega) + \dots$$

È UNA FUNZIONE

1) REALE
2) PARI
3) PERIODICA (2π)
4) $\Gamma(\omega) \geq 0$



$(\omega = 2\pi f)$

$\omega = \pi \rightarrow f = \frac{1}{2}$

$\rightarrow f_{MAX} = 0,5$

DA SPETTRO A $\gamma(\tau)$

$$\gamma(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma(\omega) e^{j\omega\tau} d\omega$$

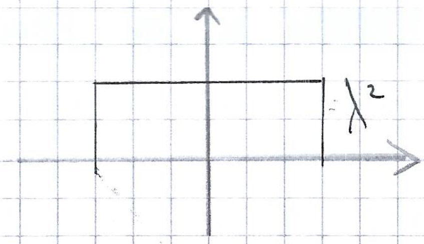
IN PARTICOLARE $\gamma(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Gamma(\omega)\, d\omega$

DA $\gamma(\omega)$ A SPETTRO

$$\Gamma_y(\omega) = \sum_{\tau=-\infty}^{+\infty} \gamma(\tau) e^{-j\omega\tau}$$

ES

$$WN \longrightarrow \gamma(\tau) = \begin{cases} \lambda^2, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}$$



$$\downarrow$$

$$\Gamma(\omega) = \gamma(0) + \cancel{2\gamma(1)\cos\omega + \ldots}$$

SPETTRO DI UN WN SEMPRE COSTANTE

$$MA(1) \longrightarrow v(t) = c_0\,\eta(t) + c_1\,\eta(t-1)$$

$$\downarrow$$

$$\gamma(\tau) = \begin{cases} \tau = 0 & (c_0^2 + c_1^2)\lambda^2 \\ \tau = \pm 1 & c_0 c_1 \lambda^2 \\ \tau = \pm k & 0 \end{cases}$$

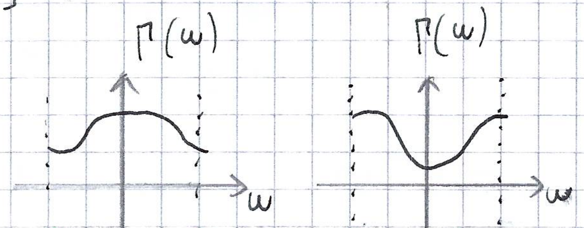
$$\downarrow$$

$$\Gamma(\omega) = \left[(c_0^2 + c_1^2) + 2 c_0 c_1 \cos\omega\right]\lambda^2$$

$$\Gamma(0) = (c_0 + c_1)^2 \lambda^2$$

$$\Gamma(\pi) = (c_0 - c_1)^2 \lambda^2$$

IN BASE AI VALORI DI c_0 E c_1



"DOMINANO BASSE FREQ" "DOMINANO ALTE FREQ"

IN GENERALE, PER COMPUTARE $\Gamma(\omega)$:

1) DATO IL PROCESSO, COMPUTA $\gamma(\tau)$
2) COMPUTA $\Gamma(\omega)$ COME $\mathcal{F}$-TRANSFORM DI $\gamma(\tau)$

... E SE NON POSSIAMO COMPUTARE $\gamma(\tau)$, COME NEL CASO DI:

ARMA(4, 10)

$$\eta(t) \longrightarrow \boxed{W(z)} \longrightarrow v(t)$$

$\longrightarrow$ USIAMO LA MAGIC FORMULA: $\Gamma(\omega) = |W(e^{j\omega})|^2 \lambda^2$

$$= W(e^{j\omega}) \cdot W(e^{-j\omega}) \cdot \lambda^2$$

SPETTRO COMPLESSO

$$\Phi(z) = \sum_{\tau=-\infty}^{+\infty} \gamma(\tau) z^{-\tau} \longrightarrow \Gamma(\omega) = \Phi(z)\Big|_{z=e^{j\omega}}$$

MAGIC FORMULA: $= W(z) W(z^{-1}) \lambda^2$

ES

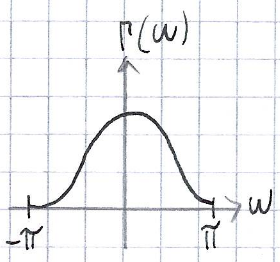
MA(1) $\quad v(t) = \eta(t) + \eta(t-1)$, $\quad c_0 = 1 \quad c_1 = 1$

COMPUTIAMO LO SPETTRO IN 3 MODI:

1) PER DEFINIZIONE: $\gamma(\tau) = \begin{cases} 2\lambda^2 & \tau = 0 \\ \lambda^2 & \tau = \pm 1 \\ 0 & \tau = \pm k \end{cases}$

$$\Phi(z) = \gamma(0) + \gamma(1) z^{-1} + \gamma(-1) z^{1} = (2 + z + z^{-1})\lambda^2$$

$$\Longrightarrow \Gamma(\omega) = (2 + e^{j\omega} + e^{-j\omega})\lambda^2 = (2 + 2\cos\omega)\lambda^2$$



2) MAGIC FORMULA

$$\Phi(z) = W(z) W(z^{-1}) \lambda^2$$

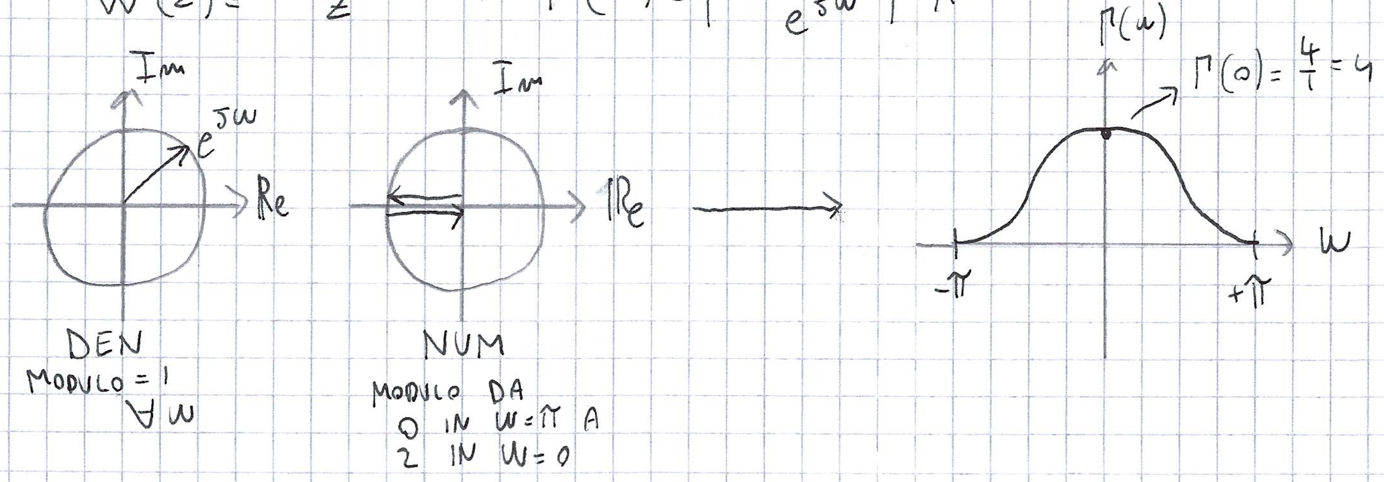
$v(t) = \eta(t) + \eta(t-1) \rightarrow (1 + z^{-1})\eta(t)$

$\longrightarrow W(z) = \frac{z+1}{z}$

$$\hookrightarrow = (1+z^{-1})(1+z)\lambda^2 = (1 + z + z^{-1} + 1) = (2 + z + z^{-1})\lambda^2 \rightarrow \text{COME } (1)$$

3) STUDIO GRAFICO (METODO DEI VETTORI)

$$W(z) = \frac{z+1}{z} \longrightarrow \Gamma(\omega) = \left| \frac{e^{j\omega} + 1}{e^{j\omega}} \right| \lambda^2$$



MULTIPLICITY OF ARMA MODELS FOR STATIONARY STOCH. PROCESSES

ABBIAMO PIÙ MODELLI PER LO STESSO PROCESSO:

$$\underset{WN(0,\lambda^2)}{\xrightarrow{\eta(t)}} \boxed{W(z)} \xrightarrow{v(t)} \qquad \underline{VS} \qquad \underset{WN(0,\tilde{\lambda}^2)}{\xrightarrow{\tilde{\eta}(t)}} \boxed{\tilde{W}(z)} \longrightarrow \tilde{v}(t)$$

$$\tilde{W}(z) = z^{-1} W(z)$$

$$\Phi(z) = \tilde{W}(z)\tilde{W}(z^{-1})\tilde{\lambda}^2 = \cancel{z^{-1}} W(z) \cancel{z} W(z^{-1}) \tilde{\lambda}^2 \overset{\lambda = \tilde{\lambda}}{=} W(z) W(z^{-1}) \lambda^2 = \Phi(z)$$

$\longrightarrow$ DA MODELLI DIFFERENTI ABBIAMO OTTENUTO LO STESSO SPETTRO!
UTILIZZEREMO UNA FORMA CANONICA...

ABBIAMO 4 CAUSE DI MOLTEPLICITÀ:

1) $\tilde{W}(z) = \rho W(z) \qquad \tilde{\lambda}^2 = \frac{1}{\rho^2}\lambda^2$

$$\tilde{\Phi}(z) = \tilde{W}(z)\tilde{W}(z^{-1})\tilde{\lambda}^2 = \cancel{\rho} W(z) \cancel{\rho} W(z^{-1}) \cdot \frac{1}{\cancel{\rho^2}}\lambda^2 = W(z)W(z^{-1})\lambda^2 = \Phi(z)$$

2) $\tilde{W}(z) = \frac{1}{z^n} W(z) \quad n \geq 1 \qquad \tilde{\lambda}^2 = \lambda^2$

$$\tilde{\Phi}(z) = \tilde{W}(z)\tilde{W}(z^{-1})\tilde{\lambda}^2 = \frac{1}{\cancel{z^n}} W(z) \cancel{z^n} W(z^{-1})\lambda^2 = W(z)W(z^{-1})\lambda^2 = \Phi(z)$$

3) $W(z) = \frac{z+\alpha}{z+\beta} \qquad \tilde{W}(z) = \frac{z+\frac{1}{\alpha}}{z+\beta} \qquad \tilde{\lambda}^2 = \alpha^2\lambda^2$

$$\Phi(z) = \frac{z+\alpha}{z+\beta} \cdot \frac{\frac{1}{z}+\alpha}{\frac{1}{z}+\beta} \cdot \lambda^2 = \frac{1+\alpha z+\frac{\alpha}{z}+\alpha^2}{1+\beta z+\beta/z+\beta^2}\lambda^2$$

$$\tilde{\Phi}(z) = \frac{z+1/\alpha}{z+\beta} \cdot \frac{1/z+1/\alpha}{1/z+\alpha} \cdot \alpha^2\lambda^2 = \frac{\alpha^2+\alpha z+\frac{\alpha}{z}+1}{\beta^2+\beta z+\beta/z+1}\lambda^2$$

$$\Phi(z)$$

4) $W(z) = \frac{z+\alpha}{z+\beta} \qquad \tilde{W}(z) = \frac{z+\alpha}{z+1/\beta} \qquad \tilde{\lambda}^2 = 1/\beta^2 \, \lambda^2$

$$\Phi(z) = \frac{z+\alpha}{z+\beta} \cdot \frac{1/z+\alpha}{1/z+\beta} \cdot \lambda^2 = \frac{1+\alpha z+\alpha/z+\alpha^2}{1+\beta z+\beta/z+\beta^2}\lambda^2$$

$$\tilde{\Phi}(z) = \frac{z+\alpha}{z+1/\beta} \cdot \frac{1/z+\alpha}{1/z+1/\beta} \cdot \frac{1}{\beta^2}\lambda^2 = \frac{1+\alpha z+\alpha/z+\alpha^2}{1+\beta z+\beta/z+\beta^2}\lambda^2$$

$$\Phi(z)$$

CANONICAL SPECTRAL FACTORIZATION

DATO $\Phi(z) = W(z)\, W(z^{-1})\, \lambda^2$, TROVA $W(z)$ (SPECTRAL FACTORIZATION PROBLEM)

$\longrightarrow$ PER VIA DELLA MOLTEPLICITÀ DI RAPPRESENTAZIONE IL PROBLEMA HA INFINITE SOLUZIONI!

VOGLIAMO QUINDI SCEGLIERE UNA FORMA CANONICA CHE SODDISFI:

1) NUM & DEN SONO MONIC POLYNOMIALS (TERMINE DI GRADO MAX HA COEFF = 1)
2) NUM & DEN HANNO LO STESSO GRADO
3) NUM & DEN SONO COPRIMI
4) NUM & DEN SONO STABILI (TUTTI ZERI E POLI IN UNIT DISK)

ES

$$W(z) = \frac{z+3}{z+0.5} \longrightarrow \tilde{W}(z) = \frac{\cancel{z+3}}{z+0.5}\, \rho\, \frac{z+\frac{1}{3}}{\cancel{z+3}} = \rho\, \frac{z+\frac{1}{3}}{z+0.5} = \text{CANON. FORM}$$

NO STABILE ZERO NON IN UNIT DISK!

ES 2

$$v(t) = 2\eta(t) + 4\eta(t-1) \quad (MA(1),\ \eta(t) \sim WN(0,1))$$

- FDT?

$$v(t) = (2 + 4z^{-1})\,\eta(t) \longrightarrow \frac{v(t)}{\eta(t)} = 2 + 4z^{-1} = \frac{2z+4}{z}$$

- CANONICA? NO: - NUM NON MONIC
  - ZERO = −2 FUORI DA UNIT DISK

- TROVA FORMA CANONICA

$$\tilde{v}(t) = \tilde{\eta}(t) + 2\tilde{\eta}(t-1) \quad \text{CON} \quad \tilde{\eta}(t) = 2\eta(t) \sim WN(0,4)$$

$$\tilde{W}(z) = 1 + 2z^{-1} = \frac{z+2}{z}$$
$\longrightarrow$ ORA NUM È MONICO, MA HO ANCORA IL PROBLEMA DELLO ZERO

$$\hat{W}(z) = \tilde{W}(z) \cdot \frac{z+1/2}{z+2} = \frac{z+1/2}{z} \longrightarrow \text{CANONICA! CON } \hat{\eta}(t) \sim WN(0,16)$$

FORMA CANONICA: $\hat{v}(t) = \hat{\eta}(t) + 0.5\,\hat{\eta}(t)$, $\hat{\eta}(t) \sim WN(0,16)$

IN GENERALE, PER DERIVARE LA FORMA CANONICA:

SIA $W^\circ(z)$ UNA FDT IN FORMA CANONICA

1) $v(t) = W^\circ(z) z^K \eta(t) \longrightarrow$ N E D NON HANNO STESSO GRADO

$\xrightarrow[\text{IN WN}]{\text{MUOVI } z^K} = W^\circ(z)\eta(t+K) \qquad \eta \sim WN(m, \lambda^2)$

$= W^\circ(z)\tilde{\eta}(t) \qquad \tilde{\eta}(t) = \eta(t+K) \sim WN(m, \lambda^2)$

2) $v(t) = K W^\circ(z)\eta(t)$, $K \neq 1 \longrightarrow$ N NON SARA' MONICO

$\xrightarrow[\text{IN WN}]{\text{MUOVI K}} = W^\circ(z)\tilde{\eta}(t)$, $\tilde{\eta}(t) = K\eta(t) \sim WN(Km, K^2\lambda^2)$

3) ZERO = POLO $\longrightarrow$ POSSO SEMPLIFICARE SE ENTRAMBI SONO STABILI

4) $v(t) = W^\circ(z)(1+Kz^{-1})\eta(t) \longrightarrow$ HO UNO ZERO DI TROPPO IN $z=-K$, $|K|>1$

$\longrightarrow = W^\circ(z) \cdot \frac{K}{K} \cdot \frac{(1+\frac{1}{K}z^{-1})}{(1+\frac{1}{K}z^{-1})}(1+Kz^{-1})\eta(t)$

$= \boxed{\frac{1}{K}\frac{(1+Kz^{-1})}{(1+\frac{1}{K}z^{-1})}}(1+\frac{1}{K}z^{-1})W^\circ(z)\tilde{\eta}(t)$

ALL PASS FILTER (STABILE SE $K>1$) $\qquad \tilde{\eta}(t) \sim WN(Km, K^2\lambda^2)$

AVENDO GLI ALL PASS FILTER GAIN = 1
SPETTRO = Γ IN INPUT RESCALED

POSSO SOSTITUIRLO CON "1":

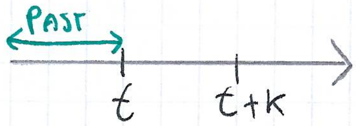
$\longrightarrow v(t) = W^\circ(z)(1+\frac{1}{K}z^{-1})\tilde{\eta}(t) \qquad \tilde{\eta}(t) \sim WN(Km, K^2\lambda^2)$

1.4 - PREDICTION PROBLEM

OVVERO, DATO

$$v(t) = \eta \longrightarrow \boxed{\tilde{W}(t)} \longrightarrow v$$

E



ESSERE IN GRADO DI STIMARE $v(t+k)$ DAI VALORI PASSATI DI v (K = PREDICTION HORIZON)

DISTINGUIAMO DUE PROBLEMI DIVERSI:

1) FAKE PROBLEM: SI HANNO I VALORI PASSATI DI $\eta(\cdot)$

2) REAL PROBLEM: SI HANNO I VALORI PASSATI DI $v(\cdot)$

1) FAKE PROBLEM

ASSUMIAMO DI AVERE IL PASSATO DI $\eta(\cdot)$ MA NEL MONDO REALE QUESTO RARAMENTE É POSSIBILE.

$$\hat{W}(z) = W_0 + W_1 z^{-1} + W_2 z^{-2} + \dots$$

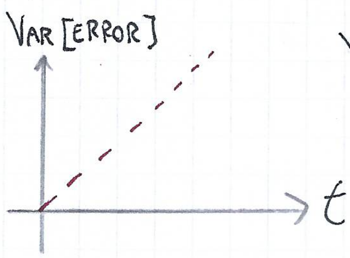
$$v(t+k) = \hat{W}(z)\eta(t+k) = \overbrace{W_0\eta(t+k) + W_1\eta(t+k-1) + \dots + W_{k-1}\eta(t+1)}^{\alpha(t)}$$
$$\underbrace{+ W_k\eta(t) + W_{1+k}\eta(t-1) + \dots}_{\beta(t)} = \alpha(t) + \beta(t)$$

OVVIAMENTE $\beta(\cdot)$ PUO' ESSERE COMPUTATO: E $\alpha(\cdot)$?

$\alpha(\cdot)$ NON PUO' ESSERE COMPUTATO ESSENDO I VALORI FUTURI DI $\eta(\cdot)$ FULLY UNPREDICTABLE.

DENOMINIAMO $\hat{v}(t+k|t) = \beta(t)$ UN OPTIMAL FAKE PREDICTOR.

$\alpha(\cdot)$ SARA' IL PREDICTION ERROR IN QUANTO $\alpha(t) = v(t+k) - \hat{v}(t+k|t)$.



$$Var[ERROR] = \left(\hat{W}_0^2 + W_1^2 + \dots +\right)\lambda^2 \rightarrow$$ AUMENTA ALL' AUMENTARE DEL PREDICTION HORIZON!

NON MOLTO PERFORMANTE...

COME OTTENERLO IN PRATICA?

$$\hat{v}(t+k|t) = \hat{W}_k\eta(t) + \dots = [\underbrace{\hat{W}_k + \hat{W}_{k-1}z^{-1} + \dots}_{\hat{W}_k(z)}]\,\eta(t)$$

$$\eta(t) \longrightarrow \underset{\text{PROCESS}}{\boxed{\hat{W}(z)}} \xrightarrow{v(t)} \qquad \text{VS} \qquad \eta(t) \longrightarrow \underset{\text{PREDICTOR}}{\boxed{\hat{W}_k(z)}} \longrightarrow \hat{v}(t+k|t)$$

COME TROVARE $\hat{W}_K(z)$ DA $\hat{W}(z)$?

$$\hat{W}(z) = \hat{W}_0 + \dots + \hat{W}_{K-1} z^{-K+1} + \hat{W}_K z^{-K} + \hat{W}_{K+1} z^{-K-1} =$$
$$= \hat{W}_0 + \dots + \hat{W}_{K-1} z^{-K+1} + z^{-K} \underbrace{\left[\hat{W}_K + \hat{W}_{K+1} z^{-1} + \dots\right]}_{\hat{W}_K(z)}$$

ES (AR(1))

$$v(t) = a v(t-1) + \eta(t) = a v(t) z^{-1} + \eta(t) \Rightarrow W(z) = \frac{z}{z-a}$$ ($|a|<1$ x STABILITA')

$W(z)$ E' IN FORMA CANONICA $\longrightarrow \hat{W}(z) = W(z) = \frac{z}{z-a}$ $z \to C(z)$, $z-a \to A(z)$

1-step predictor:

I) LONG DIVISION

$$\begin{array}{r|l} z & z-a \\ -z+a & 1 \\ \hline a & \end{array} \longrightarrow \hat{W}(z) = 1 + \frac{a}{z-a}$$

II) OPTIMAL PREDICTOR TRANSFER FUNCTION:

$$\hat{W}(z) = \underset{W_0}{1} + z^{-1} \frac{za}{z-a}$$
$$\downarrow$$
$$\hat{W}_1(z) = \frac{za}{z-a}$$

III) OPTIMAL PREDICTOR:

$$v(t+1|t) = \frac{za}{z-a}\eta(t) = \frac{a}{1-az^{-1}}\eta(t)$$

REGOLA PER r-STEPS PREDICTOR: $\hat{W}_r(z) = \frac{a^r z}{z-a}$

VARIANZA

1-step: $$\hat{v}(t+1|t) = \frac{a}{1-az^{-1}}\eta(t)$$
$$\hat{v}(t+1|t) - a\hat{v}(t|t-1) = a\eta(t)$$
$$\text{VAR}[\text{ERROR}] = \text{VAR}[\hat{v}(t+1|t) - v(t+1)] = \text{VAR}[W_0 \eta(t+1)] = W_0^2 \lambda^2 = \lambda^2$$

2-steps: $\text{VAR}[\text{ERROR}] = (1+a^2)\lambda^2$

3-steps: $\text{VAR}[\text{ERROR}] = (1+a^2+a^4)\lambda^2$

....

2) REAL PROBLEM

OTTENIAMO UN PREDICTOR DAI VALORI PASSATI

$$v(t) \longrightarrow \underbrace{\underset{\text{WHITENING FILTER}}{\boxed{\check{W}(z)}} \xrightarrow{\eta(t)} \underset{\text{OPTIMAL FAKE PREDICTOR}}{\boxed{\hat{W}_r(z)}}}_{\text{OPTIMAL PREDICTOR FROM DATA } v(t)} \longrightarrow v(t+r|t)$$

$$W_r(z) = \check{W}(z) \cdot \hat{W}_r(z) \qquad \text{CON} \qquad \hat{W}(z) = \frac{C(z)}{A(z)} \qquad \check{W}(z) = \frac{A(z)}{C(z)}$$

OTTENUTO CON LONG DIVISION

$$\hat{W}_r(z) = \frac{\dots}{A(z)}$$

$$= \frac{A(z)}{C(z)} \cdot \frac{\dots}{A(z)} = \frac{\dots}{C(z)}$$

L'OPTIMAL TRUE PREDICTOR È DATO DALL'OPTIMAL FAKE PREDICTOR RIMPIAZZANDO $A(z)$ CON $C(z)$ DELLA FORMA CANONICA

ES: AR(1) $\quad v(t) = a\,v(t-1) + \eta(t)$

1_step: $\quad \hat{W}_1 = \dfrac{az}{z-a} \qquad \check{W}(z) = \dfrac{z-a}{z}$

$$\rightarrow W_1(z) = \frac{az}{z} = a \longrightarrow \hat{v}(t+1|t) = a\,v(t)$$

PROCESS: $\quad v(t+1) = a\,v(t) + \underbrace{WN[AR(1)]}_{\text{UNPREDICTABLE}}$

PREDICTOR: $\quad \hat{v}(t+1) = a\,v(t)$

2_steps:

$$\rightarrow W_2(z) = \frac{a^2 z}{z} = a^2$$

$$v(t+2) = a^2 v(t) + \underbrace{a\eta(t+1) + \eta(t+2)}_{\text{UNPREDICTABLE}}$$

$$\hat{v}(t+2) = a^2 v(t)$$

K-steps:

$$\hat{v}(t+k|t) = a^k v(t)$$

ES: ARMA(1,1) $v(t) = a v(t-1) + \underbrace{c_0}_{=1} \eta(t) + c_1 \eta(t-1)$

IPOTIZZANDO FORMA CANONICA

SI HA: $A(z) = 1 - a z^{-1}$

$C(z) = 1 + c z^{-1}$ $\qquad A(z) v(t) = C(z) \eta(t)$

$$\rightarrow W(z) = \frac{C(z)}{A(z)} = \frac{1 + c z^{-1}}{1 - a z^{-1}} = \frac{z + c}{z - a}$$

APPLICHIAMO LONG DIVISION:

$$\begin{array}{l|l} C(z) & A(z) \\ \cline{2-2} -A(z) & 1 \\ \hline C(z)-A(z) & \end{array} \implies W(z) = 1 + z^{-1} \cdot \underbrace{\frac{C(z)-A(z)}{A(z)} \cdot z}_{\text{OPT. 1-Step AHEAD FAKE PREDICTOR}}$$

$$\implies \hat{W}_1(t) = \frac{C(z) - A(z)}{A(z)} \cdot z \rightarrow W_1(t) = \frac{C(z) - A(z)}{C(z)} \cdot z = \text{OPTIMAL PREDICTOR FROM DATA}$$

TORNANDO NEL DOMINIO DEL TEMPO: $\hat{v}(t+1|t) = -c\hat{v}(t|t-1) + (a+c) v(t)$

$$\mathrm{Var}[v(t+1) - \hat{v}(t+1|t)] = \lambda^2$$

IN GENERALE, DATA LA FORMA CANONICA $\hat{W}(z) = 1 + z^{-1} \frac{C(z)-A(z)}{A(z)} z$

CON $A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} \dots$ E $C(z) = 1 + c_1 z^{-1} + c_2 z^{-2} \dots$

OTTENIAMO L'EQUAZIONE GENERALE:

$$\hat{v}(t+1|t) = -c_1 \hat{v}(t|t-1) - c_2 \hat{v}(t-1|t-2) - \dots + (a_1 + c_1) v(t) + (a_2 + c_2) v(t-1) \dots$$

SCORCIATOIA

SE AGGIUNGIAMO E SOTTRAIAMO $C(z) v(t)$ DALLA DEF. DELLA FDT:

$$A(z) v(t) + C(z) v(t) - C(z) v(t) = C(z) \eta(t)$$

POSSIAMO DIVIDERE PER $C(z)$ E OTTENERE IL PREDICTOR, IGNORANDO IL WN:

$$v(t) = \frac{C(z) - A(z)}{C(z)} v(t) + \eta(t) \implies v(t|t-1) = \frac{C(z)-A(z)}{C(z)} v(t)$$

PREDICTION CON SEGNALI ESOGENI

IPOTIZZIAMO CHE IL NOSTRO PROCESSO DIPENDA ANCHE DA UN'ALTRA VARIABILE $u(t)$, LA QUALE È DETERMINISTICA (NON RANDOM):

ES : ARX

$$v(t) = a v(t-1) + \underbrace{u + \eta(t)}_{\text{POSSIAMO UNIRLI IN UN NUOVO WN} \sim (u, \lambda^2)}, \quad \eta(t) \sim WN(0, \lambda^2)$$

1) MEAN VALUE DI $v(t)$

$$m = E[v(t)] = am + u + 0 \quad \rightarrow m = \frac{u}{1-a}$$

2) DEFINISCO UN NUOVO PROCESSO

$$\tilde{v}(t) = v(t) - m \quad \longrightarrow \text{DEBIASED} \quad \rightarrow v(t) = \tilde{v}(t) + m$$

3)
$$v(t) = a v(t-1) + u + \eta$$
$$\tilde{v}(t) + m = a(\tilde{v}(t-1) + m) + u + \eta$$
$$\tilde{v}(t) = \cancel{-m} + a\tilde{v}(t-1) \cancel{+am} \cancel{+u} + \eta \qquad *: am = \frac{au}{1-a}$$

$$\rightarrow \frac{au}{1-a} + u - \frac{u}{1-a} = \frac{\cancel{au} \cancel{+u} \cancel{-au} \cancel{-u}}{1-a} = 0$$

$$\tilde{v}(t) = a\tilde{v}(t-1) + \eta$$

OTTENIAMO IL PREDICTOR RIMUOVENDO WN:

$$\hat{\tilde{v}}(t|t-1) = a\tilde{v}(t-1)$$

SOSTITUENDO, TROVIAMO IL PREDICTOR PER IL PROCESSO ORIGINARIO:

$$\hat{v}(t|t-1) = \hat{\tilde{v}}(t|t-1) + m = a\tilde{v}(t-1) + m = a v(t-1) - am + m$$
$$= a v(t-1) \cancel{-am} + \cancel{am} + u = a v(t-1) + u$$

IN GENERALE, PER UNO 1-step PREDICTOR PER:

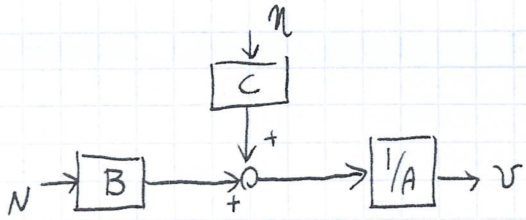
$$A(z) v(t) = C(z)\eta(t) + B(z) u(t)$$

USIAMO:

$$C(z)\hat{v}(t|t-1) = (C(z) - A(z)) v(t) + B(z) u(t-1)$$

PROCESSI ARMAX

$$A(z)\,y(t) = B(z)\,\eta(t) + \underline{C(z)\,N(t-1)}$$



PREDICTOR:

$$A(z)\,y(t) \underbrace{+\,C(z)\,y(t) - C(z)\,y(t)}_{\text{SCORCIATOIA}} = B(z)\,N(t-1) + C(z)\,\eta(t)$$

$$C(z)\,y(t) = [C(z) - A(z)]\,y(t) + B(z)\,N(t-1) + C(z)\,\eta(t)$$

$$y(t) = \frac{C(z)-A(z)}{C(z)}\,y(t) + \frac{B(z)}{C(z)}\,N(t-1) + \cancel{\eta(t)}_{\text{IGNORO}}$$

1.5: IDENTIFICATION PROBLEM

ABBIAMO UN SISTEMA CON INPUT u E OUTPUT y:

INOLTRE, $u(\cdot)$ E $y(\cdot)$ SONO MISURABILI.

POSSIAMO RICAVARE UN ARMA MODEL PER IL SISTEMA?

TENENDO CONTO CHE:

- I DATI $u(t)$ E $y(t)$ SONO NUMERI
- $y(t)$ DI ARMA È UN PROCESSO STOCASTICO

COME CONFRONTIAMO $y(t)$ DEL SYS CON $y(t)$ DELL'ARMA MODEL?

STUDIEREMO I PEM METHODS: PREDICTION ERROR MINIMIZATIONS M.

DATO UN MODELLO, POSSIAMO COMPUTARNE L'OUTPUT E CONFRONTARLO CON QUELLO REALE:

$$\varepsilon(t) = y(t) - \hat{y}(t \mid t-1)$$

IL NOSTRO OBIETTIVO È FARE IN MODO CHE:

* $\varepsilon(\cdot) \sim WN$
* $\varepsilon(\cdot)$ SIA PICCOLO

STEPS USUALI

1. DATA COLLECTION: SIA $u(1), u(2), \dots, u(N)$ CHE $y(1), y(2), \dots, y(N)$
2. SCELTA DEL MODEL FAMILY: OVVERO $\{M(\vartheta) \mid \vartheta \in \Theta\}$ DOVE ϑ È UN VECT. DI PARAM.
   DI SOLITO USIAMO AR/ARMA PER LE TIME SERIES E ARX/ARMAX PER I SISTEMI
3. SCELTA DEL CRITERIO DI OTTIMIZZAZIONE: UNA VOLTA OTTENUTO $M(\vartheta)$ E IL RELATIVO PREDICTION ERROR SI SCEGLIE UN CRITERIO DI OTTIMIZZAZIONE COME MSE.
4. OTTIMIZZAZIONE: SI OTTENGONO I PARAMETRI DEL MODELLO MIGLIORI
5. VERIFICA

LEAST SQUARE METHOD

SI TRATTA DEL PR. ERROR METHOD PIU' SEMPLICE, ADATTO AD ARX E AR.

$$m(\vartheta) = y(t) = a_1 y(t-1) + \dots + a_{m_a} y(t-m_a) + b_1 u(t-1) + \dots + b_{m_b} u(t-m_b)$$

$$= \underbrace{\vartheta^T}_{\text{PARAMETER VECTOR}} \cdot \underbrace{\varphi(t)}_{\text{OBSERVATION VECTOR}}$$

IL PROBLEMA E': TROVARE IL "MIGLIOR" ϑ, OVVERO

$$\min J(\vartheta) = \frac{1}{N} \sum_t \varepsilon_\vartheta(t)^2$$

DOVE $\varepsilon_\vartheta(t)$ E' IL PREDICTION ERROR

PER TROVARE IL MINIMO DERIVIAMO:

$$\frac{dJ}{dt} = 0 \longrightarrow J(\vartheta) = \frac{1}{N} \sum \left(y(t) - \vartheta^T \varphi(t)\right)^2$$
$$= \frac{1}{N} \sum \left(y(t) - \varphi(t)^T \vartheta\right)^2$$

$$\frac{dJ}{dt} = -\frac{1}{N} \sum 2\left(y(t) - \vartheta^T \varphi(t)\right) \varphi(t)' = -\frac{2}{N} \left\{ \sum y(t)\varphi(t)' - \sum \vartheta^T \varphi(t)\varphi(t)' \right\}$$

LA IMPONIAMO $= 0$:

$$\sum_{t=1}^{N} y(t)\varphi(t)' = \sum_{t=1}^{N} \vartheta^T \varphi(t)\varphi(t)'$$

SCAMBIAMO I LATI

$$\sum_{t=1}^{N} \varphi(t)\varphi(t)'\vartheta = \sum_{t=1}^{N} y(t)\varphi(t)'$$

EQUAZIONE NORMALE

$$Ax = b$$

POSSIAMO OTTENERE LA STIMA DEI PARAMETRI:

$$\hat{\vartheta} = \left[\sum_{t=1}^{N} \varphi(t)\varphi(t)'\right]^{-1} \sum_{t=1}^{N} y(t)\varphi(t)'$$

VERIFICHIAMO SIA UN MINIMO CONTROLLANDO CHE $\frac{d^2 J(\vartheta)}{d\vartheta^2}$ SIA POSITIVA

$$\rightarrow \frac{d^2 J(\vartheta)}{d\vartheta^2} = \frac{2}{N}\left(\sum_{t=1}^{N} \varphi(t)\varphi(t)'\right)$$

(VERO SULLA FIDUCIA)

DOBBIAMO INOLTRE VERIFICARE CHE LA STIMA SIA UNICA [IDENTIFIABILITY PROBLEM]

CONSIDERIAMO LA MATRICE $R(N)$:

$$R(N) = \frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi(t)'$$

SE LA MATRICE É POSITIVA SEMI-DEFINITA É ANCHE INVERTIBILE E LE EQ NORMALI HANNO UN'UNICA SOLUZIONE.

IN UN PROC. ARX(1,1) ABBIAMO

$$\varphi(t)\varphi(t)' = \begin{bmatrix} y(t-1) \\ u(t-1) \end{bmatrix} \begin{bmatrix} y(t-1) & u(t-1) \end{bmatrix} = \begin{bmatrix} y(t-1)^2 & y(t-1)u(t-1) \\ u(t-1)y(t-1) & u(t-1)^2 \end{bmatrix}$$

QUINDI POSSIAMO RISCRIVERE $R(N)$:

$$R(N) = \begin{bmatrix} \frac{1}{N}\sum y(t-1)^2 & \frac{1}{N}\sum y(t-1)u(t-1) \\ \frac{1}{N}\sum u(t-1)y(t-1) & \frac{1}{N}u(t-1)^2 \end{bmatrix}$$

〰 = SAMPLE VARIANCE DI y

〰 = SAMPLE VARIANCE DI u

PER $N \to \infty$, ABBIAMO

$$R(N) \to \bar{R} = \begin{bmatrix} \text{VAR}(y) & \cdots \\ \cdots & \text{VAR}(u) \end{bmatrix} = \begin{bmatrix} \gamma_{yy}(0) & \cdots \\ \cdots & \gamma_{uu}(0) \end{bmatrix}$$

$\bar{R}_{yy}$ $\bar{R}_{yu}$ $\bar{R}_{uy}$ $\bar{R}_{uu}$

SE GENERALIZZIAMO IL PROBLEMA AD UN MODELLO ARX GENERICO CON

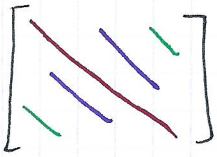
$$\varphi(t) = \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-m_a) \\ u(t-1) \\ \vdots \\ u(t-m_b) \end{bmatrix}$$

ABBIAMO

$$\bar{R}_{uu} = \begin{bmatrix} \gamma_{uu}(0) & \gamma_{uu}(1) & \gamma_{uu}(2) & \ddots \\ \gamma_{uu}(1) & \gamma_{uu}(0) & \ddots & \ddots \\ \gamma_{uu}(2) & \ddots & \ddots & \ddots \\ & \ddots & \ddots & \ddots \end{bmatrix}$$

($\bar{R}_{yy}$ ANALOGA)

É UNA TOEPLITZ MATRIX



$$\Longrightarrow \bar{R} = \begin{bmatrix} \bar{R}_{yy} & \cdots \\ \cdots & \bar{R}_{uu} \end{bmatrix}$$

CONDIZIONE NECESSARIA PER $\bar{R}$ INVERTIBILE É CHE LO SIA $\bar{R}_{uu}$ LA QUALE <u>DIPENDE SOLO DALL'INPUT</u>.

SE $\bar{R}_{uu}$ É INVERTIBILE DICIAMO CHE $u(\cdot)$ É PERSISTENTLY EXCITANT

ES

$$u(\cdot) \sim WN(0, \lambda^2) \longrightarrow \bar{R}_{uu} = \begin{bmatrix} \gamma_{uu}(0) & \gamma_{uu}(1) & \gamma_{uu}(2) & \cdots \\ \gamma_{uu}(1) & \gamma_{uu}(0) & \gamma_{uu}(1) & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} = \begin{bmatrix} \lambda^2 & 0 & 0 & \cdots \\ 0 & \lambda^2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} = \lambda^2 I$$

$\lambda^2 I$ É SEMPRE INVERTIBILE, QUINDI IL SEGNALE É P.E PER OGNI POSSIBILE ORDER DEL MODELLO.

ES 2

ABBIAMO UN SISTEMA S DESCRITTO DALLA TRUE FDT

$$G^{\circ}(z) = \frac{z}{(z+0.5)(z+0.8)}$$

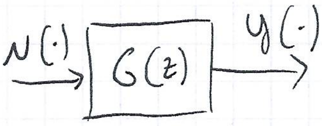
NON ABBIAMO IL MECCANISMO DI GENERAZIONE DEI DATI, QUINDI CONSIDERIAMO UNA FAMIGLIA DI MODELLI:

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) + b_1 u(t-1) + b_2 u(t-2) + \eta(t)$$

CON IL VETT DI PARAMETRI: $\vartheta = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \end{bmatrix}$

SE u E' COSTANTE NON POSSIAMO IDENTIFICARE POLI E ZERI IN QUANTO ANCHE L'OUTPUT LO SAREBBE.
INFATTI $\bar{R}_{uu}$ NON E' INVERTIBILE.

GUADAGNO DI UN DYNAMIC SYSTEM



SE $u(\cdot) = \text{CONST} = \bar{u}$
$\Rightarrow y(\cdot) = \text{CONST} = \bar{y} \longrightarrow \mu = \frac{\bar{y}}{\bar{u}} = \text{GAIN}$

POSSIAMO IN GENERALE TROVARE IL GAIN DA $G(z)$:

$$y(t) = G(z)u(t) = \frac{N(z)}{D(z)} u(t) \Rightarrow y(t) D(z) = N(z) u(t)$$

ESSENDO:

$$D(z) = d_0 z^m + d_1 z^{m-1} + \dots$$

$$\rightarrow D(z) y(t) = d_0 y(t+m) + d_1 y(t+m-1) + \dots$$

SE $y(\cdot) = \text{CONST} = \bar{y}$:

$$D(z) y(t) = d_0 \bar{y} + d_1 \bar{y} + \dots = (d_0 + d_1 + \dots)\bar{y} = D(z)\big|_{z=1} \bar{y}$$

$$N(z) u(t) = N(z)\big|_{z=1} \bar{u}$$

DA CUI $\mu = \frac{\bar{y}}{\bar{u}} = \frac{N(z)}{D(z)}\Big|_{z=1}$ = GAIN E' FDT VALUTATA IN $z=1$

ES

ABBIAMO DATI GENERATI DA $u(t) \to \boxed{G(z)} \to y(t)$

CON S: $G^\circ(z) = \dfrac{z}{(z+0.5)(z+0.8)} = \dfrac{z}{z^2+1.3z+0.4}$

$\mu^\circ = \dfrac{1}{1+1.3+0.4} = \dfrac{1}{2{,}7} = \text{GAIN}$

$\left.\begin{array}{l}\text{ZERI: } z=0 \\ \text{POLI: } z=-0.5,\ z=-0.8\end{array}\right\}$ SYS STABILE

USIAMO $\boxed{\text{ARX:}}$

$\mathcal{M}: \ y(t) = a_1 y(t-1) + a_2 y(t-2) + b_1 u(t-1) + \eta(t)$

$\hookrightarrow G(z) = \dfrac{b_1 z^{-1}}{1-a_1 z^{-1} - a_2 z^{-2}} = \dfrac{b_1 z}{z^2 - a_1 z - a_2}$

UGUAGLIANDO SI HA

$G^\circ(z) = G(z)$ IF $\left.\begin{array}{l} b_1 = b_1^\circ = 1 \\ a_1 = a_1^\circ = -1.3 \\ a_2 = a_2^\circ = -0.4 \end{array}\right\}$

OVVERO... UN POSSIBILE ASSEGNAMENTO!

COME OTTENIAMO QUESTI PARAMETRI DAI DATI (OVVERO LE MISURAZIONI)?

CASO A

$u(t) = \bar{u}$ E $y(t) = \bar{y}$, COSTANTI. IL VETTORE OSSERVAZIONE È:

$$\varphi(t) = \begin{vmatrix} \bar{y} \\ \bar{y} \\ \bar{u} \end{vmatrix}$$

LA MATRICE $\bar{R} = \sum \varphi(t)\varphi(t)'$ È SINGOLARE (NON INVERTIBILE).

QUINDI ABBIAMO SOLO QUELLA INFORMAZIONE SUL GAIN PER TROVARE I PARAMETRI...
MA ABBIAMO INFINITI POSSIBILI ASSEGNAMENTI PER OTTENERE QUEL GAIN!

CASO B

$u(t) \sim WN(0, \lambda^2)$ QUINDI $y(t)$ È UN PROCESSO STOCASTICO.

LO STIMATORE DEI PARAMETRI CON N SAMPLES È: $\hat{\theta}_N = \left(\sum_{t=1}^{N} \varphi(t)\varphi(t)'\right) \sum_{t=1}^{N} \varphi(t) y(t)$

PER $N \to \infty$ ABBIAMO $\hat{\theta}_N \to \theta^\circ$

COSA SUCCEDE SE NON CONOSCIAMO IL MODELLO E NE SCEGLIAMO UNO "OVERSIZED" COME:

$y(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) + b_1 u(t-1) + b_2 u(t-2) + \eta(t)$

HO SBAGLIATO MODELLO...

$G = \dfrac{z(b_1 z + b_2)}{(\quad)(\quad)(\quad)}$ PER AVERE DEGREE$(D) = 2$ DOBBIAMO SEMPLIFICARE $\longrightarrow \infty$ POSSIBILI SEMPLIFICAZIONI!

MAXIMUM LIKELIHOOD METHOD

DA LS → ML

LS → (AR, ARX) NORMAL EQ

ML → (ARMA, ARMAX) NO NORMAL EQ

ARMAX : SYSTEM (INPUT → OUTPUT)

ARMA : TIME SERIES

CONSIDERIAMO ARMAX (PER ARMA E' LO STESSO):

$$\text{ARMAX}: \quad M: A(z)y(t) = B(z)u(t-1) + C(z)\eta(t)$$

CON:

$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} - \dots$$
$$C(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots$$
$$B(z) = b_1 + b_2 z^{-1} + \dots$$

QUINDI IL VETTORE PARAMETRI SARA':

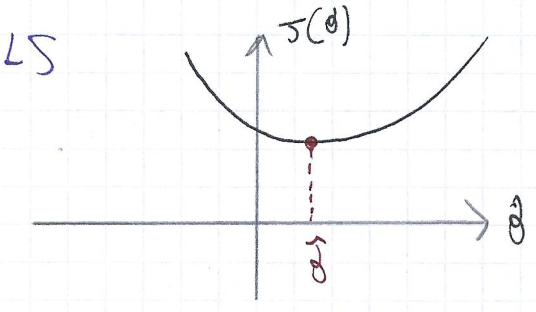
$$\vartheta' = |a_1 \ a_2 \dots a_{M_a} \ b_1 \ b_2 \dots b_{M_b} \ c_1 \ c_2 \dots c_{M_c}|$$

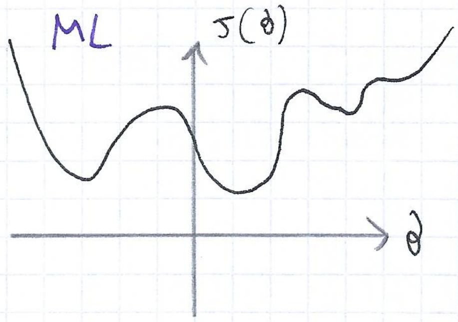
COME NEI CASI PRECEDENTI DOBBIAMO TROVARE UN BUON $\hat{\vartheta}$ DALLE RILEVAZIONI (AKA $y(1), y(2) \dots y(N), u(1), u(2), \dots u(N)$).

IL PERFORMANCE INDEX BASATO SUL PREDICTION ERROR PUO' ANCORA ESSERE IL MEAN SQUARE ERROR:

$$J = \frac{1}{N} \sum_{t=1}^{N} \varepsilon_\vartheta(t)^2$$

LA DIFFERENZA IMPORTANTE CON LS E' CHE $J(\cdot)$ E' NON-CONVEX INVECE CHE QUADRATICA:

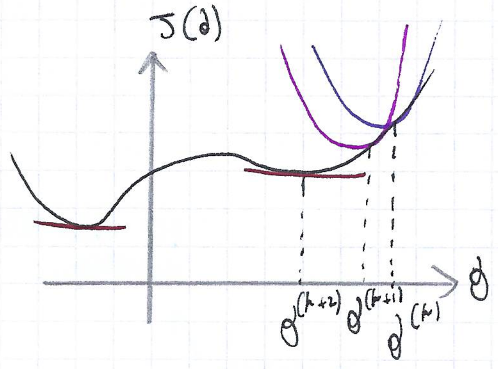




E' MOLTO PIU' DIFFICILE TROVARE IL MIN. ASSOLUTO

METODO DI NEWTON (ITERATIVO)

SUPPONENDO ϑ SCALARE SI APPROSSIMA J AD UNA f QUADRATICA $V(\vartheta)$.



$V(\vartheta)$ "PROXY" DI $J(\vartheta)$

IL MIN DI $V(\vartheta)$ E' $\hat{\vartheta}^{(k)}$

$$\hat{\vartheta}^{(k)} \xrightarrow[k \to +\infty]{} \hat{\vartheta}$$

POTREBBE ESSERE UN MINIMO GLOBALE...

MA ANCHE NO!

→

CONSIDERIAMO UN'APPROSSIMAZIONE QUADRATICA:

$$V(\vartheta) = J(\vartheta)\Big|_{\vartheta=\vartheta^{(k)}} + \frac{dJ(\vartheta)}{d\vartheta}\Big|_{\vartheta=\vartheta^{(k)}}(\vartheta-\vartheta^{(k)}) + \frac{1}{2}(\vartheta-\vartheta^{(k)})' \frac{d^2J(\vartheta)}{d\vartheta^2}\Big|_{\vartheta=\vartheta^{(k)}}(\vartheta-\vartheta^{(k)})$$

PER LA FORMULA DI NEWTON ABBIAMO:

(1) $$\vartheta^{(k+1)} = \vartheta^{(k)} - \left(\frac{d^2J(\vartheta)}{d\vartheta^2}\Big|_{\vartheta=\vartheta^{(k)}}\right)^{-1} \frac{dJ}{d\vartheta}\Big|'_{\vartheta=\vartheta^{(k)}}$$

DOBBIAMO COMPUTARE IL GRADIENTE E L'HESSIANO

$$J(\vartheta) = \frac{1}{N}\sum_t \varepsilon(t)^2 \longrightarrow \frac{dJ(\vartheta)}{d\vartheta} = \frac{2}{N}\sum_t \varepsilon(t)\frac{d\varepsilon(t)}{d\vartheta}$$

$$\frac{d^2J(\vartheta)}{d\vartheta^2} = \frac{2}{N}\sum \frac{d\varepsilon(t)}{d\vartheta}\frac{d\varepsilon(t)}{d\vartheta} + \cancel{\frac{1}{N}\sum \varepsilon(t)\frac{d^2\varepsilon(t)}{d\vartheta^2}}$$

DI SOLITO OMESSO PER SEMPLICITÀ

POSSIAMO DEFINIRE $\psi(t) = -\frac{d\varepsilon(t)}{d\vartheta}$

SOSTITUIAMO IN (1) E OTTENIAMO LA FORMULA DI GAUSS-NEWTON:

$$\vartheta^{(k+1)} = \vartheta^{(k)} + \left(\sum_{t=1}^N \psi(t)\psi(t)'\right)^{-1}\sum_{t=1}^N \psi(t)\varepsilon(t)$$

GIRANDOLA "RICORDA" LA NORMAL EQUATION:

$$\sum_{t=1}^N \psi(t)\psi(t)'\underbrace{\left(\vartheta^{(k+1)} - \vartheta^{(k)}\right)}_{\vartheta} = \sum_{t=1}^N \psi(t)\underbrace{\varepsilon(t)}_{y(t)}$$

CI RIMANE DA COMPUTARE $\varepsilon(\cdot)$ E $\psi(\cdot)$ DAI DATI:

$$M: \; Ay(t) \pm C(y(t)) = Bu(t-1) + C\eta(t)$$

DA CUI

$$Cy(t) = [C-A]y(t) + Bu(t-1) + C\eta(t)$$

$$y(t) = \underbrace{\frac{C-A}{C}y(t)}_{\text{PASSATO DI } y} + \underbrace{\frac{B}{C}u(t-1)}_{\text{PASSATO DI } u} + \underbrace{\eta(t)}_{\text{IGNORIAMO X PREDICTOR}}$$

C E A SONO MONICI:

$$C = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots$$
$$A = 1 - a_1 z^{-1} - a_2 z^{-2} - \dots$$
$$\rightarrow C-A = (a_1+c_1)z^{-1} + (a_2+c_2)z^{-2}\dots$$

MOLTIPLICHIAMO PER C DA ENTRAMBI I LATI

$$C\hat{y}(t) = Cy(t) - Ay(t) + Bu(t-1) \longrightarrow C(\overbrace{y(t)-\hat{y}(t)}^{\varepsilon(t)}) = Ay(t) - Bu(t-1)$$

DA CUI, FINALMENTE: $C^{(k)}\varepsilon(t)^{(k)} = A^{(k)}y(t) - B^{(k)}u(t-1)$: EQUAZIONE DEL PREDICTION ERROR PER L'ITERAZIONE k

RIASSUMENDO:

COMPUTAZIONE ITERATIVA DI $\hat{\vartheta}$: ALLA ITERAZIONE k ABBIAMO LA STIMA $\hat{\vartheta}^{(k)}$ DEI PARAMETRI.

1. DA $\vartheta^{(k)}$ TROVA $A^{(k)}(z)$, $B^{(k)}(z)$ E $C^{(k)}(z)$
2. FILTRA I DATI CON QUEI POLINOMI PER OTTENERE $\varepsilon(t)^{(k)}$ E $\psi(t)^{(k)}$
3. USA LA FORMULA DI G/N PER COMPUTARE $\hat{\vartheta}^{(k+1)}$
4. ITERA SINO ALLA CONVERGENZA

PERFORMANCE DEI PREDICTION ERROR IDENTIFICATION METHODS

SE COSTRUIAMO IL P.E COME AL SOLITO: $\varepsilon_\vartheta(t) = y(t) - \hat{y}_\vartheta(t)$

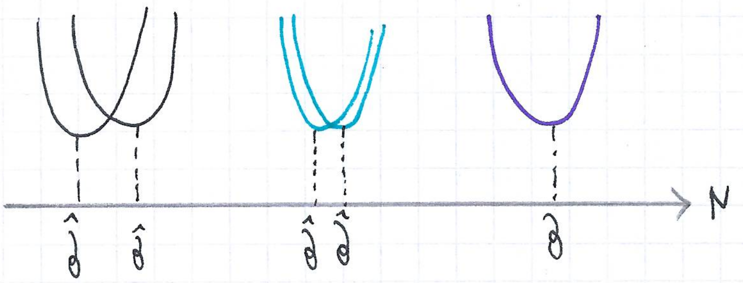
ABBIAMO CHE SIA $y(t)$ CHE $\hat{y}_\vartheta(t)$ SONO SEQUENZE DI PUNTI E QUINDI IL PERF. INDEX DIPENDE DAGLI SPECIFICI PUNTI:

$$J_N(\vartheta) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_\vartheta(t)$$

SE CONSIDERIAMO IL PR. ERROR UN PROCESSO STAZIONARIO ABBIAMO CHE PER $N \to \infty$:

$$J_N(\vartheta) \to \bar{J}(\vartheta) = E\left[\varepsilon_\vartheta(t)^2\right]$$

ASSUMIAMO ϑ SCALARE PER SEMPLICITA' E RAPPRESENTIAMO LA MINIMIZATION FUNC DI $J_N(\vartheta)$ AL VARIARE DI N:



PIU' È ALTO N, PIU' I MINIMI SONO VICINI PER ESPERIMENTI DIFFERENTI (FINO AD ESSERE UNICO PER $N \to \infty$).

PER N FINITO, $\hat{\vartheta}_N$ DIPENDE DALL'OUTCOME DELL'ESPERIMENTO E QUINDI PUO' ESSERE CONSIDERATO UNA VARIABILE RANDOM E QUINDI, PER $N \to \infty$, ESSA TENDE A $\bar{\vartheta}$ DOVE $\bar{\vartheta}$ E' IL RISULTATO DEL MINIMIZATION PROCESS DEL PERFORMANCE INDEX ASINTOTICO.

$$\bar{\vartheta} = \min E\left[\varepsilon_\vartheta(t)^2\right]$$

ES

SIA IL DATA GENERATOR: $S:\ y(t) = a^\circ y(t-1) + \eta(t)$ $\qquad \eta \sim WN(0, \lambda^2)$

SIA $|a^\circ| < 1$ COSI' CHE $y(\cdot)$ SIA UN PROC STAZIONARIO.

CONSIDERIAMO IL MODELLO $M(\vartheta): \hat{y}_\vartheta(t) = a\,y(t-1)$

IL VETTORE DI PARAMETRI QUINDI SI RIDUCE A $\vartheta = [a]$.

PROVIAMO A STIMARLO CON IL METODO LS. $\longrightarrow$

IL VETTORE DI OSSERVAZIONI HA SIZE 1×1

$$\varphi(t) = [y(t-1)]$$

SCRIVIAMO LA EQ NORMALE

$$\sum_{t=1}^{N} y(t-1)^2 \vartheta = \sum_{t=1}^{N} y(t-1)y(t)$$

LA STIMA DEI PARAMETRI E':

$$\hat{\vartheta} = \hat{a} = \frac{\frac{1}{N}\sum_{t=1}^{N} y(t-1)y(t)}{\frac{1}{N}\sum_{t=1}^{N} y(t-1)^2} = \frac{\hat{\gamma}_{yy}(1)}{\hat{\gamma}_{yy}(0)}$$

CON IL SOLITO PERFORMANCE INDEX

$$\bar{J}(\vartheta) = E[\varepsilon_\vartheta(t)^2]$$

SE $\varepsilon_\vartheta(\cdot)$ E' STAZIONARIO, ALLORA $E[\varepsilon_\vartheta(t)^2]$ NON DIPENDE DA t.

IN QUEL CASO IL PREDICT ERROR SARA':

$$\varepsilon_\vartheta(t) = y(t) - \hat{y}(t) = a^\circ y(t-1) + \eta(t) - a y(t-1) = (a^\circ - a) y(t-1) + \eta(t)$$

$$\varepsilon_\vartheta^2(t) = (a^\circ - a)^2 y(t-1) + \eta(t)^2 + 2(a^\circ - a) y(t-1)\eta(t)$$

$$E[\varepsilon_\vartheta(t)^2] = (a^\circ - a) E[y(t-1)^2] + \lambda^2 + 2(a^\circ - a) E[\cancel{y(t-1)\eta(t)}]$$ 0

QUINDI

$$\bar{J}(\vartheta) = (a^\circ - a)\gamma_{yy}(0) + \lambda^2$$

PER $N \to \infty$, $\hat{a}_N \to a^\circ$ FINCHE', PER $a = a^\circ$ ABBIAMO $J(\vartheta) = \lambda^2$.

TESTARE LA VALIDITA' DEL MODELLO

OVVERO VOGLIAMO VERIFICARE CHE IL PRED. ERROR SIA UN WN.

CASO A

$S: AR(1): \quad y(t) = a^\circ y(t-1) + \eta(t)$

$M: AR(1): \quad \hat{y}(t) = a y(t-1)$

CON IL PR. ERROR ID METHOD LA STIMA $\hat{a}_N$ TENDE AD a°.

$$\varepsilon(t) = y(t) - \hat{y}(t) = a^\circ \cancel{y(t-1)} + \eta(t) - \cancel{a^\circ y(t-1)} = \eta(t)$$

CASO B

$S: MA(1): \quad y(t) = \eta(t) + c^\circ \eta(t-1)$

$M: AR(1): \quad \hat{y}(t) = a y(t-1)$

IN QUESTO CASO $\hat{a}_N \to \bar{a} = \frac{c^\circ}{1+c^{\circ 2}}$

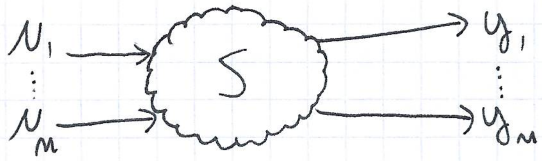
$$\to \varepsilon(t) = y(t) - \hat{y}(t) = \eta(t) + c^\circ \eta(t-1) - \frac{c^\circ}{1+c_0^2} y(t-1)$$

QUINDI DIPENDE ANCHE DA $\eta(t-1), \eta(t-2)$, E NON E' QUINDI UN WN.

RIEPILOGO

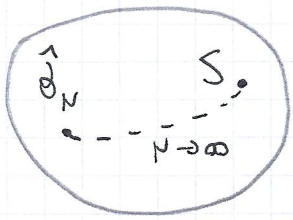
FACENDO L'IDENTIFICAZIONE ABBIAMO

* UN DATA GENERATION MECHANISM S



* UNA FAMIGLIA DI MODELLI IN PREDICTIVE FORM $\hat{\mathcal{M}}(\vartheta)$

4 SITUAZIONI:

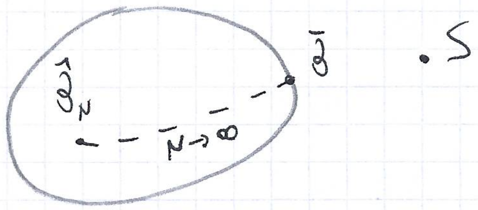
1)


$\hat{\vartheta}_N \to \vartheta^o$

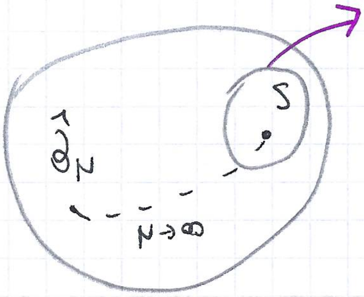
$S = \mathcal{M}(\vartheta^o)$

$S \in \hat{\mathcal{M}}(\vartheta)$

↳ S E' NELLA FAMIGLIA DI MODELLI

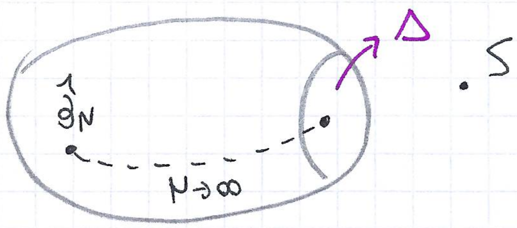
2)


IL MODELLO $\mathcal{M}(\bar{\vartheta})$ E' IL BEST PROXY DI S NELLA FAMIGLIA DI MODELLI CONSIDERATA

3)


Δ SET DI 'OPTIMAL' MODELS

POTREMMO VOLER PASSARE AD UNA FAMIGLIA PIU' SEMPLICE SE VOGLIAMO UN MODELLO UNICO

4)


C'E' UNA MOLTEPLICITA' DI MODELLI Δ CHE NON INCLUDONO S.

VELOCITÀ DI CONVERGENZA

QUANDO INTENDIAMO "$N \to \infty$", A CHE VELOCITÀ LO FACCIAMO?

CONSIDERIAMO IL CASO 1)

$$S = \mathcal{M}(\vartheta^\circ) \to \hat{\vartheta}_N \to \vartheta^\circ \text{ PER } N \to \infty$$

$$\mathrm{VAR}[\hat{\vartheta}_N - \vartheta^\circ] = \underbrace{\frac{1}{N} \lambda^2 \bar{R}^{-2}}_{\text{PRENDERE PER VERO}} \longrightarrow \mathrm{VAR} \to 0 \text{ A } 1/N$$

$$\text{DEV STD} \to 0 \text{ A } 1/\sqrt{N}$$

1.6: MODEL COMPLEXITY SELECTION

ARX(1,1)
ARX(2,2)
ARX(3,3)

$$\hat{y}(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3)$$
$$b_1 u(t-1) + b_2 u(t-2) + b_3 u(t-3)$$

$$\vartheta = \begin{vmatrix} a_1 \\ b_1 \end{vmatrix} \qquad \vartheta = \begin{vmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{vmatrix}$$

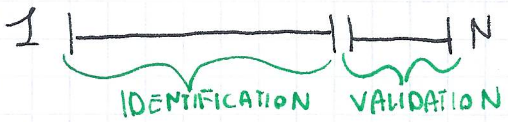
SE a_2 E $b_2 \approx 0$

IN GENERALE ARX(n, n) E' SEMPRE MEGLIO DI ARX($n-1$, $n-1$)... MA DOBBIAMO COMUNQUE SCEGLIERE UN n A CUI FERMARCI!

COMPUTARE IL PERF. INDEX PER VALORI CRESCENTI DI N E' NAIVE IN QUANTO DIMINUIRA' SEMPRE MENO, MA POTREBBE NON ESSERE ADEGUATO PER DATI NUOVI MAI VISTI.

CROSS VALIDATION

TECNICA MIGLIORE CHE PREVEDE DI DIVIDERE I DATA POINTS IN DUE SETS, UNO PER L'IDENTIFICAZIONE E UNO PER LA VALIDAZIONE



TUTTAVIA IN QUESTO MODO "SPRECHIAMO" DEI DATI CHE NON POTREMO USARE NELLA FASE ID...

FPE: FINAL PREDICTION ERROR

LO SCOPO DEL CRITERIO E' VALUTARE $\bar{J}(\vartheta) = E\left[(y(t) - \hat{y}_\vartheta(t))^2\right]$

OVVERO IL PR. ERROR DEL MODELLO $M(\vartheta)$ PER TUTTE LE POSSIBILI SEQ. DI DATA.

STIMIAMO ϑ DAI DATA: $\frac{1}{N}\sum \varepsilon_\vartheta(t)^2 \xrightarrow{min} \hat{\vartheta}_N$

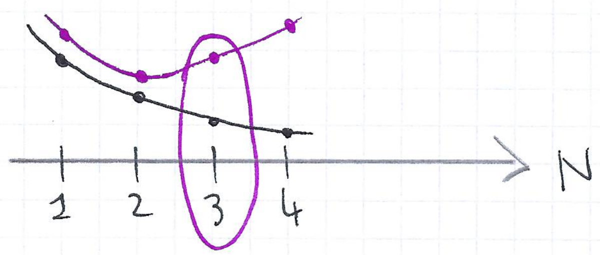
$\bar{J}(\hat{\vartheta}_N)$ = AVG. PRED. ERROR DEL MODELLO $M(\hat{\vartheta}_N)$

RANDOM

$$FPE = E\left[\bar{J}(\hat{\vartheta}_N)\right] = \frac{N+V}{N-V} J(\vartheta)$$

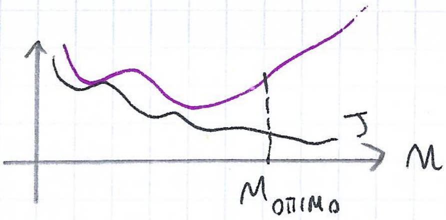
DOVE N = # DI DATA
V = # DI PARAMETRI

J = SAMPLE VARIANCE DEL PRED ERROR PER $M(\hat{\vartheta}_N)$



AIC

$$= 2\frac{M}{N} + \ln J(\vartheta)$$



MDL

$$= (\ln N)\frac{M}{N} + \ln J(\vartheta)$$ PIU' "PARSIMONIOSO"

1.7: DURBIN - LEVINSON ALGORITHM

ABBIAMO LA FAMIGLIA DI MODELLI AR(K), K=1,...,100

1) QUAL E' IL MIGLIOR K?
2) COME STIMARE TALI SET DI MODELLI?

INVERTIRE 100 MATRICI K × K E' COSTOSO...
CON L'ALGORITMO POSSIAMO COMPUTARE LA SOLUZIONE DI AR(K+1) PARTENDO DALLA SOL DI AR(K).

ES AR(1) ⟶ AR(2)

IL MODELLO E' $\hat{y}(t) = \tilde{a}_1 y(t-1) + \tilde{a}_2 y(t-2)$

POSSIAMO COMPUTARE LA COV. $\gamma(\tau)$ DEL MODELLO AR(1) CON:

$$\gamma(0) = a_1\gamma(1) + \lambda^2$$
$$\gamma(1) = a_1\gamma(0)$$
$$\gamma(\tau) = a_1\gamma(\tau-1) \quad \forall \tau > 1$$

[EQ DI YULE-WALKER]

DATO $\gamma(\cdot)$, QUINDI

$$\rightarrow a_1 = \frac{\gamma(1)}{\gamma(0)}, \quad \lambda^2 = \gamma(0) - a_1\gamma(1)$$

DA CUI:

$$\begin{cases} \tilde{a}_1 = a_1 - \tilde{a}_2\frac{\gamma(1)}{\gamma(0)} \\ \tilde{\lambda}^2 = \lambda^2(1-\tilde{a}_2^2) \\ \tilde{a}_2 = \frac{1}{\lambda^2}(\gamma(2) - a_1\gamma(1)) \end{cases}$$

(NOTA $\tilde{\lambda}^2 \leq \lambda^2$)

GENERALE:

$$\begin{cases} a_{m+1}^{(m+1)} = \frac{1}{\lambda^{2(m)}}(\dots) \\ a^{(m+1)} = a_i^{(m)} - a_{(m+1)}^{(m+1)} a_{m+1-i}^{(m)} \\ \lambda^{2(m+1)} = \lambda^{2(m)}\left[1 - \left\{a_{m+1}^{(m+1)}\right\}^2\right] \end{cases}$$

PARCOV

CONSIDERIAMO I DUE MODELLI

$$AR(k-1): \quad y(t) = a_1^{(k-1)} y(t-1) + \dots + a_{k-1}^{(k-1)} y(t-k+1) + \eta(t)$$

$$AR(k): \quad y(t) = a_1^{(k)} y(t-1) + \dots + a_k^{(k)} y(t-k) + \eta(t)$$

IL PARAMETRO $a_k^{(k)}$ E' CHIAMATO PARTIAL COVARIANCE:

$$PARCOV(\tau) = a_\tau^{(\tau)}$$

SE IL MODELLO AR "TRUE" HA ORDINE n, ALLORA

$$PARCOV(\tau) = 0 \quad \forall \tau > n$$

PUO' QUINDI ESSERE USATO PER TROVARE UN MODELLO AR ADEGUATO.

E' PARALLELO AL CASO AM PER CUI:

$$\gamma(\tau) = 0 \quad \forall \tau > n$$

INOLTRE SE DOBBIAMO SCEGLIERE TRA AR E MA:

i) SE COV. FUNC VA A 0 PRIMA DI PARCOV $\longrightarrow$ GO FOR MA

ii) PARCOV VA A 0 PER PRIMA $\longrightarrow$ GO FOR AR

1.8: RECURSIVE LEAST SQUARES

INVECE CHE USARE TUTTI I DATA IN UN COLPO, AGGIORNIAMO MAN MANO LA STIMA.

LA FORMULA PER LS NORMALE È

$$\hat{\vartheta}_t = \left(\sum_{i=1}^{t} \varphi(i) \varphi(i)' \right)^{-1} \sum_{i=1}^{t} \varphi(i) y(i)$$

DEFINIAMO IL PRIMO TERMINE SOMMA COME:

$$S(t) = \sum_{i=1}^{t} \varphi(i) \varphi(i)'$$

$$J = \sum_{i=1}^{t} \mu^{t-i} \varepsilon(i)^2 = \sum_{i=1}^{t} \mu^{t-i} \left(y(t) - \varphi(t)' \vartheta \right)^2$$