

ELABORAZIONE DI SEGNALI BIOLOGICI

MODELLI ARMA

Per elaborare segnali di interesse biologico, spesso si ricorre all'impiego di modelli ARMA, cioè modelli a tempo discreto, lineari, tempo invarianti, descritti da un'equazione alle differenze. Si tratta di "modelli di dati" che si limitano a descrivere le relazioni ingresso-uscita, a differenza dei "modelli di sistema" che invece hanno l'obiettivo di descrivere il sistema che produce i dati. Per mettere in luce le differenze tra queste due classi di modelli, possiamo dire in breve che la messa a punto di un modello di sistema richiede in generale la descrizione dei fenomeni complessi che stanno alla base del processo di generazione dei dati. Tale fenomeno sono generalmente interpretati come combinazione di fenomeni elementari, per la cui descrizione fa ricorso a 'leggi' ampiamente collaudate dall'esperienza. È facile riconoscere che questo approccio può portare ad una serie di difficoltà. In primo luogo, non è detto che, nel patrimonio di conoscenze proprie della disciplina interessata, possa davvero essere reperito un insieme di 'leggi elementari' e di 'regole di composizione' capaci di descrivere con l'affidabilità e l'accuratezza richiesta gli aspetti cruciali del fenomeno in esame. Una seconda difficoltà deriva dal carattere fortemente composito di certi fenomeni. Secondo l'approccio classico bisognerà innanzitutto analizzare una ad una le singole 'componenti' del fenomeno per descriverne il comportamento mediante leggi elementari; quindi si dovranno descrivere mediante opportune equazioni le interazioni tra i fenomeni elementari stessi. Il modello cercato sarà costituito dall'insieme di tutte le relazioni così ottenute. Supponendo anche di disporre di leggi accurate ed affidabili per ogni singola componente, è ovvio che potrebbero nascere serie difficoltà nell'uso del modello così ottenuto a causa della sua elevata complessità in rapporto alla finalità per cui è stato costruito e/o alle potenzialità di calcolo disponibile, al suo costo, al tempo necessario e così via.

Viceversa un modello di dati si limita a dare una descrizione dei dati astraendoli dal contesto fisico che li ha generati. Questa identificazione di tipo 'a scatola nera' può sembrare brutale, però si è rilevata molto utile, e talvolta insostituibile, per la risoluzione di problemi pratici di grande interesse. La generica procedura di identificazione va svolta seguendo i passi:

- 1) osservazione dei dati;
- 2) scelta di una classe di modelli
- 3) stima del modello per mezzo di opportuni criteri di adeguatezza del modello ai dati
- 3) validazione del modello, con eventuale ritorno ai passi precedenti in caso di risposta negativa ai test di validazione.

1. EQUAZIONI DEL MODELLO

I modelli ARMA sono modelli a tempo discreto. Con riferimento ai segnali biologici, che per loro natura sono per lo più a tempo continuo, segnali a tempo discreto sono generalmente ottenuti attraverso campionamento con passo di campionamento costante. Quindi una serie temporale x_n con $n=0,1,2,\dots$ può essere ottenuta campionando il segnale continuo $x(t)$ con passo T_s : $x_n = x(nT_s)$.

Assumiamo dapprima che i segnali di ingresso e uscita, indicati rispettivamente con x_n e w_n siano segnali deterministici a tempo discreto. Il modello ARMA è un modello lineare invariante alla traslazione a coefficienti costanti descritto dall'equazione ingresso-uscita:

$$w_n = -\sum_{k=1}^n c_k w_{n-k} + \sum_{k=0}^m d_k x_{n-k}$$

con ordini n e m interi e finiti e coefficienti d_k e c_k reali. La sua funzione di trasferimento alle trasformate zeta è un rapporto di polinomi in zeta:

$$\frac{W(z)}{X(z)} = \frac{\sum_{k=0}^m d_k z^{-k}}{1 + \sum_{k=1}^n c_k z^{-k}}$$

Si parla invece di modello aleatorio quando i segnali di ingresso e uscita, ora indicati con u_n e y_n , sono processi aleatori stazionari. Si assume che l'ingresso x_n sia rumore bianco, stazionario a media nulla. La relazione ingresso-uscita è simile all'equazione per modelli deterministici

$$y_n = -\sum_{k=1}^p a_k y_{n-k} + \sum_{k=0}^q b_k u_{n-k} \quad 1$$

con ordini p e q interi e finiti e coefficienti a_k e b_k reali. La funzione di trasferimento è ancora un rapporto di polinomi in zeta

$$H(z) = \frac{Y(z)}{U(z)} = \frac{\sum_{k=0}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}}$$

Il modello ARMA fornisce una descrizione parametrica del processo aleatorio stazionario y di uscita, nel senso che ad ogni realizzazione del processo di ingresso corrisponde una realizzazione del processo di uscita.

Un caso più generale è quello in cui il segnale di uscita ξ_n contiene sia una parte deterministica che quella aleatoria, cioè $\xi_n = w_n + y_n$. Per ricavare l'equazione alle differenze che descrive ξ_n non è restrittivo porre $p=n$ e $a_i=c_i$, $i=1,2,..p$. Pensando alla rappresentazione delle due equazioni ingresso-uscita alle trasformate zeta, ciò corrisponde a rendere uguali i polinomi a denominatore delle due funzioni di trasferimento attraverso il calcolo del loro minimo comune multiplo. Pertanto la relazione tra i due ingressi (deterministico e aleatorio) e l'uscita (aleatoria) è descritto dall'equazione

$$\xi_n = -\sum_{k=1}^p a_k \xi_{n-k} + \sum_{k=0}^q b_k u_{n-k} + \sum_{k=0}^n d_k x_{n-k} .$$

Si parla allora di modello ARMAX, cioè modello ARMA (aleatorio) con ingresso esogeno (deterministico).

Nel seguito, tratteremo l'identificazione di modelli ARMA per processi aleatori, cioè di modelli descritti dall'eq.1. In particolare parleremo di modelli AR con riferimento al caso in cui l'uscita in un certo istante dipenda dall'uscita negli istanti precedenti e dall'ingresso nell'istante attuale

$$y_n = -\sum_{k=1}^p a_k y_{n-k} + u_n$$

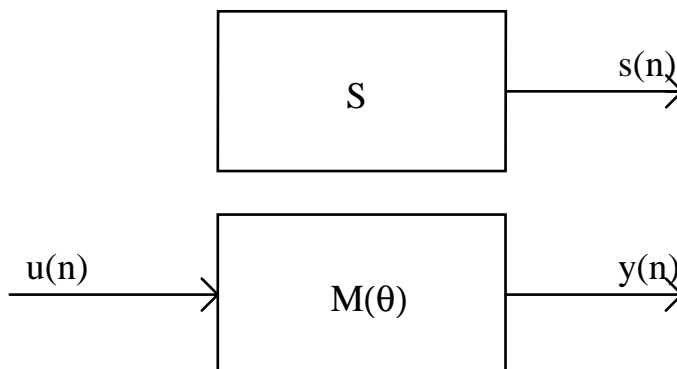
e di modelli MA nel caso in cui non ci sia legame diretto tra l'uscita all'istante n e l'uscita agli istanti precedenti:

$$y_n = \sum_{k=0}^q b_k u_{n-k}$$

Esemplificheremo più avanti l'impiego di modelli ARMAX con riferimento ad una applicazione specifica.

2. IDENTIFICAZIONE: PRINCIPI DI BASE

In termini generali, il problema che ci poniamo è quello di determinare un modello matematico, $M(\theta)$, per descrivere una sequenza di dati sperimentali misurata da un sistema fisico S . θ è un vettore che contiene i parametri del modello.

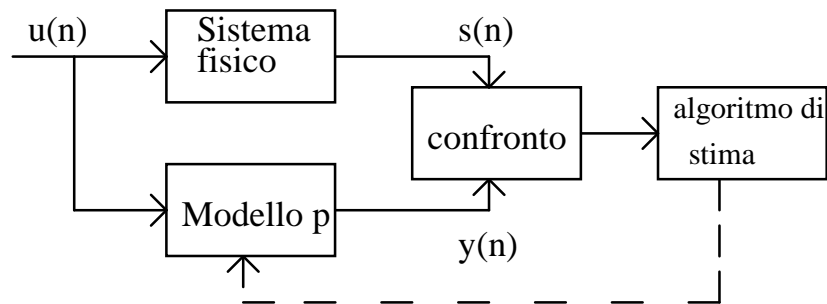


Lavoriamo in un contesto aleatorio, cioè facciamo l'ipotesi che la sequenza di dati $s(n)$ (in seguito indicata come s_n) sia una realizzazione di un processo aleatorio. Il modello che descrive tale processo sia l'uscita di un sistema lineare invariante alla traslazione che ha in ingresso un rumore stazionario, a valor medio nullo, bianco, con una certa varianza. L'obiettivo è di inferire, dalla sequenza di misura s_n la descrizione del processo S che l'ha generata. Con riferimento all'equazione 1, si devono allora determinare i valori numerici degli ordini p e q , dei parametri a_i e b_i , e della varianza del rumore di ingresso. E' chiaro come sia necessaria, oltre all'ipotesi di stazionarietà, anche quella di ergodicità del processo di uscita, ipotesi che garantisce di poter stimare le proprietà dell'intero processo a partire da una singola realizzazione.

Per quanto riguarda il processo di identificazione, un modello è migliore di un altro se interpreta meglio i dati sperimentali, pertanto è necessario operare un confronto tra i dati sperimentali e il valore delle grandezze in uscita del modello. La differenza tra queste due

grandezze può rappresentare un buon indice dell'errore commesso nell'identificazione. Si sceglierà quindi il modello i cui parametri rendono minima tale quantità.

Se la sequenza di ingresso fosse nota, si potrebbe adottare per l'identificazione il seguente schema:



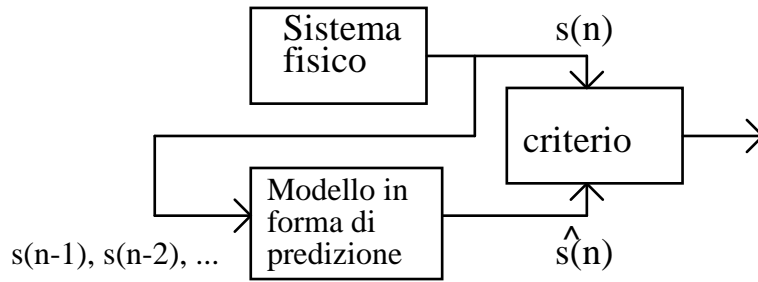
cioè applicare al modello lo stesso ingresso che, applicato al sistema fisico, ha generato i dati misurati, e quindi valutare quei valori dei parametri che fanno sì che l'uscita del modello e quella del sistema fisico siano "simili". Bisogna quindi fissare un criterio di somiglianza, che nel caso di stima ai minimi quadrati, se indichiamo con $e_n = y_n - s_n$ l'errore di modello o residuo consiste nel minimizzare rispetto ai parametri la seguente funzione costo

$$J(\theta) = \sum_{t=1}^N (s_n - y_n)^2$$

dove N è il numero di campioni osservati. Nel caso che stiamo esaminando, l'ingresso non è in generale una entità fisica misurabile, quindi non è nota la particolare realizzazione di ingresso che ha generato la particolare realizzazione misurata. Quindi, mentre l'uscita del sistema fisico è una sequenza di dati (una realizzazione), l'uscita del modello è il processo aleatorio generato come trasformazione lineare del processo aleatorio di ingresso. Ci sono allora due approcci alternativi:

- **Metodo A:** per effettuare un confronto tra uscita del modello (processo aleatorio) e uscita del sistema (sequenza di dati), si confrontano le statistiche del modello e del sistema, cioè la funzione di autocorrelazione del modello (sequenza deterministica, funzione dei parametri del modello) con la funzione di autocorrelazione del sistema (stimata dai dati). Si calcolano i parametri del modello in modo da far sì che la funzione di autocorrelazione del modello "assomigli" alla funzione di autocorrelazione del sistema
- **Metodo B:** si tratta di ricavare dal modello di partenza $M(\theta)$ il modello in forma di predizione $M_{pred}(\theta)$ che: a) dipende dagli stessi parametri del modello M; b) consente di stimare il valore del segnale all'istante n a partire dai suoi valori negli istanti precedenti. In altre parole, se l'unica informazione disponibile è rappresentata dalla sequenza di uscita del sistema, un modo per risolvere il problema è tentare di scoprire il legame tra un dato della sequenza e i suoi valori precedenti. Una volta che la dinamica nascosta sia stata trovata e descritta da un opportuno modello, una ragionevole predizione del valore successivo della serie potrà essere ricavata applicando il modello stesso agli ultimi dati osservati. A questo punto è possibile confrontare i valori veri del segnale con quelli stimati dal predittore, e calcolare i parametri del modello di predizione in modo tale da soddisfare un opportuno criterio di minimo.

Di seguito esamineremo in dettaglio l'identificazione di modelli analizzando in sequenza modelli AR, MA e ARMA.



3. MODELLI AR

I modelli AR sono molto usati per la loro semplicità e perché sono disponibili molti algoritmi per la loro identificazione. L'equazione alle differenze per tali modelli è:

$$y_n = -\sum_{k=1}^p a_k y_{n-k} + b_0 u_n \quad 2$$

dove u_n è rumore bianco, a valor medio nullo e varianza σ_u^2 , cioè:

$$E[u_n] = 0, \quad R_u(k) = E[u_n u_{n+k}] = \sigma_u^2 \delta_m, \quad \delta_m = \begin{cases} 1 & \text{per } m=0 \\ 0 & \text{altrimenti} \end{cases}$$

Per caratterizzare il processo in uscita ad un modello AR, esaminiamo le proprietà di stazionarietà e calcoliamo media e funzione di autocorrelazione.

1. Si può dimostrare che il processo è stazionario se i parametri a_n soddisfano le condizioni di stabilità del sistema, cioè se le radici dell'equazione caratteristica:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p} = 0 \quad 3$$

sono tutte interne al cerchio unitario del piano zeta. Quindi la stabilità del sistema è collegata alla stazionarietà nel meccanismo di generazione dei dati.

2. Il processo di uscita è a valor medio nullo. Infatti, valutando l'aspettazione dei termini dell'eq. 2 e sfruttando la linearità dell'operatore aspettazione:

$$E[y_n] = -\sum_{k=1}^p a_k E[y_{n-k}] + b_0 E[u_n]$$

Per la stazionarietà del processo, $E[y_n] = E[y_{n-k}]$ per ogni k . Essendo il processo di ingresso a media nulla, possiamo riscrivere come

$$E[y_n] \left(1 + \sum_{k=1}^p a_k \right) = 0$$

Dato che il secondo fattore del prodotto è in generale diverso da zero, il fatto che il prodotto sia nullo comporta che sia nullo il primo fattore, ovvero la media del processo.

3. Calcoliamo la funzione di autocorrelazione di y_n per $k > 0$.

$$\begin{aligned} R_y(k) &= E[y_n y_{n-k}] = E \left[\left(-\sum_{i=1}^p a_i y_{n-i} + b_o u_n \right) y_{n-k} \right] = \\ &= -\sum_{i=1}^p a_i R_y(k-i) + b_o E[u_n y_{n-k}] \end{aligned} \quad 4$$

Per calcolare l'ultimo addendo dell'eq. (4), sfruttiamo il fatto che il sistema è causale, quindi per $k > 0$ tale addendo è nullo perchè l'uscita all'istante $n-k$ non dipende dall'ingresso nell'istante futuro n . Per $k=0$, sostituendo a y_n l'equazione (2), si ottiene $E[u_n^2] = b_o \sigma_u^2$.

Riassumendo

$$E(y_{n-k} u_n) = \begin{cases} 0, & \text{per } k > 0 \\ b_o \sigma_u^2 & \text{per } k = 0 \end{cases} \quad 5$$

Sostituendo l'eq.5 nell'eq.4 si ottiene:

$$R_y(k) = \begin{cases} -\sum_{i=1}^p a_i R_y(k-i) & \text{per } k > 0 \\ -\sum_{i=1}^p a_i R_y(k-i) + b_o^2 \sigma_u^2 & \text{per } k = 0 \\ R_y(-k) & \text{per } k < 0 \end{cases} \quad 6$$

L'insieme di equazioni (6) è noto con il nome di equazione di Yule-Walker. Dall'espressione di $R_y(0)$ è evidente come i termini b_o e σ_u^2 non giocano singolarmente, ma solo come prodotto. Pertanto non è restrittivo considerare, in alternativa al modello originale

$b_o = 1$ e varianza del rumore di ingresso = $b_o^2 \sigma_u^2$

$\sigma_u^2 = 1$ e guadagno del sistema = $b_o \sigma_u$

Infatti i tre modelli portano alla stessa descrizione statistica (media e funzione di autocorrelazione) Nel seguito seguiremo la prima alternativa, quindi considereremo l'equazione alle differenze e la funzione di trasferimento di un modello AR nelle seguenti forme

$$y_n = -\sum_{k=1}^p a_k y_{n-k} + u_n$$

$$H_{AR}(z) = \frac{Y(z)}{U(z)} = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}}$$

Esempio: modello AR di ordine 1 (modello di Markov)

$$y_n = -a_1 y_{n-1} + u_n$$

Equazioni di Yule Walker

$$R_y(k) = \begin{cases} -a_1 R_y(k-1) & \text{per } k > 0 \\ -a_1 R_y(k-1) + \sigma_u^2 & \text{per } k = 0 \\ R_y(-k) & \text{per } k < 0 \end{cases}$$

Ovvero

$$R_y(0) = -a_1 R_y(-1) + \sigma_u^2 = -a_1 R_y(1) + \sigma_u^2$$

$$R_y(1) = -a_1 R_y(0)$$

$$R_y(2) = -a_1 R_y(1)$$

$$R_y(3) = -a_1 R_y(2)$$

ecc.

Le equazioni di Yule Walker forniscono una descrizione ricorsiva dei valori della funzione di autocorrelazione. Per ricavare tale funzione in forma esplicita, consideriamo le prime due equazioni e le risolviamo rispetto a $R_y(0)$ e poi esplicitiamo in funzione di $R_y(0)$ e k le equazioni successive ottenendo

$$R_y(0) = \frac{\sigma_u^2}{1 - a_1^2}$$

$$R_y(k) = R_y(0)(-a_1)^k$$

NB. La condizione di stazionarietà richiede la stabilità del sistema, quindi a_1 in modulo minore di 1. In tali condizioni la funzione di autocorrelazione assume valor massimo nell'origine, come deve essere per P.A. stazionari. Se a_1 è negativo, $R_y(k)$ scende con k con andamento esponenziale decrescente. Se a_1 è positivo, $R_y(k)$ ha segni alterni, e il suo modulo scende con k con andamento esponenziale decrescente.

Esempio: modello AR di ordine 2

$$y_n = -a_1 y_{n-1} - a_2 y_{n-2} + u_n$$

equazioni di Yule Walker

$$R_y(k) = \begin{cases} -a_1 R_y(k-1) - a_2 R_y(k-2) & \text{per } k > 0 \\ -a_1 R_y(k-1) - a_2 R_y(k-2) + \sigma_u^2 & \text{per } k = 0 \\ R_y(-k) & \text{per } k < 0 \end{cases}$$

Si può dimostrare che la soluzione esplicita è del tipo

$$R_y(k) = A_1(\alpha_1)^k + A_2(\alpha_2)^k$$

dove α_1 e α_2 sono le radici del polinomio caratteristico:

$$1 + a_1 z + a_2 z^2 = 0$$

La condizione di stazionarietà richiede la stabilità del sistema, quindi α_1 e α_2 in modulo minori di 1. Pertanto anche in questo caso la funzione di autocorrelazione assume valor massimo nell'origine e poi dipende da k con un inviluppo pari ad una funzione esponenziale decrescente

Identificazione di modelli AR

Assumiamo dapprima che l'ordine p sia noto. Identificare il modello vuol dire associare un valore numerico ai parametri del modello a_1, \dots, a_p e σ_u^2 ; in tutto, $p+1$ incognite. Vedremo in seguito dei criteri per selezionare in modo ottimo l'ordine del modello. Partiamo con il metodo di stima dei parametri del modello AR basato sulla predizione (Metodo B).

Il modello in forma di predizione descrive l'uscita all'istante n come combinazione lineare delle uscite negli istanti precedenti con coefficienti a_k pari a quelli del modello originale. Infatti, essendo il rumore di ingresso bianco, si rinuncia a qualsiasi tentativo di predizione e lo si pone pari al suo valor medio, cioè a zero

$$\tilde{y}_n = -\sum_{k=1}^p a_k y_{n-k} \quad 7$$

dove il simbolo \sim indica valore stimato. Alle trasformate zeta:

$$H_{pred}(z) = \frac{\tilde{Y}(z)}{Y(z)} = -\sum_{k=1}^p a_k z^{-k} = 1 - \frac{1}{H_{AR}(z)} \quad 8$$

Applicando il modello di predizione all'uscita del sistema fisico si ottiene la predizione del segnale all'istante n , noti i suoi valori precedenti:

$$\tilde{s}_n = -\sum_{k=1}^p a_k s_{n-k}$$

Per il campione n -esimo possiamo calcolare l'errore di predizione e_n (noto anche come residuo) come differenza tra il valore vero e il valore predetto:

$$e_n = s_n - \tilde{s}_n = s_n + \sum_{k=1}^p a_k s_{n-k} \quad 9$$

Il metodo dei minimi quadrati determina una stima dei parametri definendo un indice J pari all'errore quadratico medio:

$$J = E[e_n^2] = E\left[\left(s_n + \sum_{k=1}^p a_k s_{n-k}\right)^2\right] \quad 10$$

e poi trovando il minimo di J rispetto ai parametri a_1, \dots, a_p

$$a_k = \min_{a_k} J \quad k = 1, \dots, p$$

Essendo J una forma quadratica, il punto di minimo globale è quello in cui si annullano le derivate parziali

$$\frac{\partial J}{\partial a_k} = 0, \quad k=1, \dots, p \quad 11$$

Sostituendo a J la sua espressione, eq.10, e derivando rispetto a a_k , $k=1, \dots, p$ si ottengono p equazioni lineari

$$-\sum_{i=1}^p a_i R_s(k-i) = R_s(k) \quad 1 \leq k \leq p \quad 12$$

dove R_s è la funzione di autocorrelazione dei dati

$$R_s(i-k) = E[s_{n-k} s_{n-i}] = R_s(k-i) \quad 13$$

Essendo queste equazioni della stessa forma delle equazioni di Yule-Walker (eq.6) possiamo concludere che le stime di a_1, \dots, a_p che minimizzano l'errore di predizione si possono trovare considerando le equazioni di Yule-Walker $k=1, 2, \dots, p$, uguagliando in esse la funzione di autocorrelazione del modello con quella dei dati

$$R_s(k) = R_y(k) \quad k=1, 2, \dots, p \quad 14$$

e infine risolvendo tali equazioni in a_1, \dots, a_p

Per stimare la varianza del rumore di ingresso, riconsideriamo l'uscita del modello y_n e osserviamo che la differenza tra uscita del modello e la sua predizione coincide con il rumore di ingresso:

$$\text{modello} \quad y_n - \tilde{y}_n = u_n$$

mentre per il sistema fisico la differenza tra dati e predizione coincide con l'errore di predizione

$$\text{sistema fisico} \quad s_n - \tilde{s}_n = e_n$$

Da qui possiamo interpretare l'errore di predizione come una stima del rumore di ingresso, e quindi stimare la varianza del rumore di ingresso attraverso il calcolo della varianza dell'errore di predizione, che coincide con il calcolare J in corrispondenza del punto di minimo

$$\tilde{\sigma}_u^2 = E[(s_n - \tilde{s}_n)^2] = J_{\min} \quad 15$$

Sostituendo nell'equazione 15 l'equazione 9 e tenendo conto che i valori di a_k sono le stime ottenute risolvendo le equazioni di Yule Walker si ottiene:

$$\sigma_u^2 = E[(s_n - \tilde{s}_n)^2] = R_s(0) + \sum_{i=1}^p a_i R_s(i)$$

Quest'ultima equazione completa le equazioni (12) e perciò si può concludere che è possibile stimare i parametri a_1, \dots, a_p e la varianza del rumore di ingresso σ_u^2 considerando le equazioni di Yule-Walker per $k=0, 1, \dots, p$ e uguagliando in esse la funzione di autocorrelazione del modello con quella dei dati

$$R_s(i) = R_y(i) \quad k=0, 1, \dots, p \quad 16$$

Si ottengono così delle equazioni che vanno risolte nelle incognite a_1, \dots, a_p e σ_u^2 . Quindi, per un modello AR, minimizzare l'errore di predizione (Metodo B) coincide con uguagliare i primi $p+1$ valori della funzione di autocorrelazione del modello $R_y(i)$ con quelli dei dati $R_s(i)$ (Metodo A).

L'insieme delle equazioni da risolvere può essere espressa in forma matriciale.

$$\begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{p-1} \\ R_1 & R_0 & R_1 & \cdots & R_{p-2} \\ R_2 & R_1 & R_0 & \cdots & R_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{p-1} & R_{p-2} & R_{p-3} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_p \end{bmatrix} \quad 17$$

La matrice del sistema è Hermitiana e di Toeplitz, poiché ogni diagonale è caratterizzata da un unico elemento. Inoltre la matrice è definita positiva. Queste condizioni garantiscono l'invertibilità della matrice, tranne il caso molto particolare in cui i dati siano armoniche pure. Pertanto, è praticamente sempre possibile risolvere il sistema di equazioni 17 e la soluzione è unica. Se al sistema (17) si aggiunge l'equazione per $k=0$ si ottiene una nuova matrice:

$$\begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_p \\ R_1 & R_0 & R_1 & \cdots & R_{p-1} \\ R_2 & R_1 & R_0 & \cdots & R_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_p & R_{p-1} & R_{p-2} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \sigma_u^2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad 18$$

Per determinare i parametri a_i e σ_u^2 è necessario risolvere il sistema (18) considerando come termini noti i $p+1$ valori della funzione di autocorrelazione. Tali coefficienti verranno stimati dalla sequenza di dati $\{s_n\}$, $n=0, 1, 2, \dots, (N-1)$, usando ad esempio lo stimatore:

$$\tilde{R}_s(i) = \frac{1}{N-i} \sum_{n=0}^{N-i-1} s_n s_{n+i} \quad 19$$

Algoritmo di Levinson-Durbin.

L'algoritmo di Durbin Levinson costituisce un'efficiente soluzione per il sistema di eq. (18). L'algoritmo procede ricorsivamente calcolando in successione i parametri di modelli di ordine via via più elevato, quindi gli insiemi di parametri $\{a_1^{(1)} (\sigma_u^2)^{(1)}\}$, $\{a_1^{(2)}, a_2^{(2)}, (\sigma_u^2)^{(2)}\}$, ..., $\{a_1^{(p)}, a_2^{(p)}, \dots, a_p^{(p)}, (\sigma_u^2)^{(p)}\}$, dove l'apice si riferisce all'ordine del modello. L'algoritmo di ricorsione è inizializzato da:

$$a_1^{(1)} = -\frac{R_s(1)}{R_s(0)} \quad 20$$

$$(\sigma_u^2)^{(1)} = (1 - |a_1^{(1)}|^2) R_s(0) \quad 21$$

e le formule di ricorsione permettono di stimare i parametri del modello di ordine i noti i parametri del modello di ordine $i-1$

$$a_i^{(i)} = - \left[R(i) + \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) \right] / (\sigma_u^2)^{(i-1)}$$

$$a_j^{(i)} = a_j^{(i-1)} + a_i^{(i)} a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 \quad 22$$

$$(\sigma_u^2)^{(i)} = (1 - a_i^{(i)2}) (\sigma_u^2)^{(i-1)}$$

L'algoritmo di Levinson Durbin genera i parametri del modello AR dagli ordini più bassi via fino ai più alti. Questa proprietà è utile perché generalmente non si conosce a priori l'ordine del modello e l'applicazione dei criteri di scelta dell'ordine (vd sotto) prevedono di identificare modelli di ordine via via più elevato.

Scelta dell'ordine del modello

Fino a questo momento abbiamo supposto che l'ordine del modello sia noto. In realtà l'ordine non è noto a priori, ma viene anch'esso valutato dai dati. Si potrebbe pensare di costruire modelli AR di ordine crescente finché l'errore di predizione J_{min} non raggiunga un valore

minimo, garantendo così che il modello approssima al meglio i dati sperimentali..Ma dalla terza equazione di Levison Durbin possiamo prevedere che l'errore quadratico medio J_{\min} (che coincide con la stima della varianza del rumore di ingresso) per il modello di ordine p sarà senz'altro inferiore al valore di J_{\min} per il modello di ordine $p-1$, in quanto il coefficiente $a_p^{(p)}$ è il termine noto del polinomio a denominatore della funzione di trasferimento, quindi è il prodotto dei suoi poli che, per la condizione di stazionarietà, sono tutti in modulo minori di 1. In altre parole, come è intuitivo aspettarsi, incrementando l'ordine del modello, e quindi la sua complessità, il modello approssima meglio i dati. Allora un criterio realistico è di arrestarsi all'ordine p quando il miglioramento introdotto da un modello di ordine $p+1$ non è significativo, ad es. è inferiore ad un certo ε prefissato

$$\frac{J_{\min}^{(p+1)}}{J_{\min}^{(p)}} \geq 1 - \varepsilon$$

In alternativa, vengono spesso usati dei criteri di valutazione dell'ordine ottimo del modello che si basano sul criterio di parsimonia, cioè considerano contemporaneamente la capacità di un modello di approssimare i dati (J_{\min}) e la sua complessità (numero di parametri), con l'obiettivo di selezionare l'ordine del modello che è in grado di riprodurre bene i dati con un numero ridotto di parametri.

Akaike ha introdotto due criteri: il primo sceglie l'ordine del modello AR in modo che l'indice FPE (final prediction error) sia minimo:

$$FPE(p) = \frac{N+p-1}{N-p-1} J_{\min} \quad 23$$

dove N è il numero di campioni considerati.

Da notare che il primo fattore cresce all'aumentare di p , mentre il secondo fattore diminuisce. L'ordine più appropriato del modello è dato dal valore p che minimizza l'indice FPE.

Il secondo criterio basato sul principio di massima verosimiglianza determina l'ordine del modello minimizzando una teorica funzione informazione. Assumendo che il processo abbia una statistica Gaussiana, l'indice da minimizzare AIC (Akaike Information Criterion) è

$$AIC = \ln J_{\min} + 2(p+1)/N. \quad 24$$

Anche in questo caso, il primo addendo decresce al crescere di p , mentre il secondo cresce con p , rappresentando la penalità introdotta dall'utilizzo di ordine maggiore a quello ottimo. AIC e FPE danno risultati simili se N molto elevato.

4. MODELLI MA

L'equazione alle differenze di un modello MA è data da

$$y_n = \sum_{k=0}^q b_k u_{n-k} \quad 25$$

dove al solito:

$$E[u_n] = 0, \quad R_u(k) = E[u_n u_{n+k}] = \sigma_u^2 \delta_m, \quad \delta_m = \begin{cases} 1 & \text{per } m=0 \\ 0 & \text{altrimenti} \end{cases}$$

In analogia a quanto visto per i modelli AR, caratterizziamo il processo in uscita ad un modello MA esaminando le proprietà di stazionarietà e calcolando media e funzione di autocorrelazione.

1. Non ci sono problemi legati alla stazionarietà del processo di uscita poiché anche per modelli MA la stazionarietà è legata alla stabilità del sistema, e questa è sempre garantita perchè i poli della funzione di trasferimento stanno tutti nell'origine.
2. E' immediato verificare (linearità dell'aspettazione!) che anche il processo di uscita è a valor medio nullo.
3. Per calcolare l'autocorrelazione del processo MA sfruttiamo al solito la linearità dell'aspettazione, che ci permette di scrivere

$$R_y(k) = E \left[\sum_{i=0}^q b_i u_{n-i} \sum_{j=0}^q b_j u_{n-k-j} \right] = \sum_{i=0}^q \sum_{j=0}^q b_i b_j E[u_{n-i} u_{n-k-j}]$$

L'aspettazione che compare a secondo membro dell'uguaglianza è diverso da zero (pari alla varianza del processo in ingresso) solo per: $n-i = n-k-j$, ovvero per $i=k+j$. Pertanto l'autocorrelazione, per k positivi, ha la seguente espressione

$$R_y(k) = \begin{cases} E\{y_n y_{n-k}\} = \sigma_u^2 \sum_{j=0}^{q-k} b_{k+j} b_j & 0 < k \leq q, \quad b_0 = 1 \\ 0, & \text{per } k > q \end{cases} \quad 26$$

quindi è descritta da equazioni non lineari nei parametri. Per k negativi vale la solita relazione

$$R_y(k) = R_y(-k)$$

Dalle eq.26, si può osservare che non è restrittivo porre $b_0=1$, in quanto è sempre possibile scalare tutti i parametri di b_0 e conglobare il quadrato di b_0 nella varianza del rumore di ingresso. Nel seguito quindi considereremo l'equazione alle differenze di un modello MA e la sua funzione di trasferimento nelle seguenti forme

$$y_n = u_n + \sum_{k=1}^q b_k u_{n-k}$$

$$H_{MA} = \frac{Y(z)}{U(z)} = 1 + \sum_{k=1}^q b_k z^{-k}$$

Esempio: modello MA di ordine 1

$$y_n = u_n + b_1 u_{n-1}$$

Funzione di autocorrelazione:

$$R_y(0) = \sigma_u^2 (1 + b_1^2)$$

$$R_y(1) = \sigma_u^2 b_1$$

$$R_y(2) = R_y(3) = \dots = 0$$

$R_y(0)$ e $R_y(1)$ sono non lineari nei parametri !

Esempio: modello MA di ordine 2

$$y_n = u_n + b_1 u_{n-1} + b_2 u_{n-2}$$

Funzione di autocorrelazione:

$$R_y(0) = \sigma_u^2 (1 + b_1^2 + b_2^2)$$

$$R_y(1) = \sigma_u^2 (b_1 + b_1 b_2)$$

$$R_y(2) = \sigma_u^2 b_2$$

$$R_y(3) = R_y(4) = \dots = 0$$

$R_y(0)$, $R_y(1)$ e $R_y(2)$ sono non lineari nei parametri !

Identificazione di modelli MA

La definizione del predittore per un modello MA (metodo B) non è così immediata come per modelli AR e richiede alcuni passaggi. Iniziamo rappresentando il modello alla trasformata z

$$Y(z) = \left[1 + \sum_{i=1}^q b_i z^{-i} \right] U(z) \quad 27$$

Posto

$$B(z) = 1 + \sum_{i=1}^q b_i z^{-i} \quad 28$$

si può scrivere:

$$U(z) = \frac{Y(z)}{B(z)} + Y(z) - Y(z) \quad 29$$

da cui

$$Y(z) = Y(z) \left(1 - \frac{1}{B(z)} \right) + U(z) \quad 30$$

Ricordando che il termine noto del polinomio a denominatore è 1

$$1 - \frac{1}{B(z)} = \frac{\sum_{i=1}^q b_i z^{-i}}{1 + \sum_{i=1}^q b_i z^{-i}} = c_1 z^{-1} + c_2 z^{-2} + \dots \quad 31$$

Il risultato interessante è che il polinomio ottenuto come divisione tra i polinomi a numeratore e denominatore non contiene il termine noto, quindi nell'eq. 30 il termine

$$Y(z) \left(1 - \frac{1}{B(z)} \right)$$

dipende solo dalla storia passata dell'uscita mentre il termine $U(z)$ fa riferimento a rumore bianco, quindi non prevedibile. Allora possiamo rinunciare ad alcuna previsione sull'ingresso, assimilarlo al suo valor medio cioè zero e definire il predittore ad un passo del modello MA come

$$\tilde{Y}(z) = Y(z) \left(1 - \frac{1}{B(z)} \right) = Y(z) \left(1 - \frac{1}{H_{MA}(z)} \right)$$

Quindi la funzione di trasferimento del modello in forma di predizione è:

$$H_{PRED}(z) = \frac{\tilde{Y}(z)}{Y(z)} = 1 - \frac{1}{H_{MA}(z)}$$

Il modello in forma di predizione dipende dagli stessi parametri b_i del modello originale, e applicato alla sequenza di dati consente di stimare il valore del segnale all'istante k noti i suoi valori agli istanti precedenti. L'indice che da minimizzare è ancora:

$$J = E \left[\left(s_n - \tilde{s}_n \right)^2 \right] \quad 32$$

e la minimizzazione richiede di risolvere delle equazioni del tipo

$$\frac{\partial J}{\partial b_i} = 0, \quad 1 \leq i \leq q \quad 33$$

che però, a differenza di quanto accade con i modelli AR, non sono lineari nei parametri. Pertanto la soluzione richiede un algoritmo di ricerca di minimo della funzione J .

In alternativa (Metodo A) possiamo uguagliare le funzioni di correlazione del modello (R_y) e dei dati (R_s) per $k=0, \dots, q$, ottenendo quindi l'insieme di equazioni non lineari

$$R_s(k) = \begin{cases} E[s_n s_{n-k}] = \sum_{j=0}^{q-k} \sigma_u^2 b_{k+j} b_j & 0 < k \leq q, \quad b_0 = 1 \\ 0, & \text{per } k > q \end{cases} \quad 34$$

e risolvere nei parametri b_j . Il problema è che queste equazioni non sono lineari, quindi bisogna utilizzare un metodo numerico di soluzione.

Metodo dei momenti

Un metodo iterativo che sfrutta la particolare struttura delle equazioni da risolvere è il metodo dei momenti. Si parte dalle eq.26, riscritte esplicitando le sommatorie

$$\begin{aligned} R_s(0) &= \sigma_u^2 [1 + b_1^2 + b_2^2 + \dots] \\ R_s(1) &= \sigma_u^2 [b_1 + b_1 b_2 + b_2 b_3 + \dots b_{q-1} b_q] \\ R_s(2) &= \sigma_u^2 [b_2 + b_1 b_3 + b_2 b_4 + \dots b_{q-2} b_q] \end{aligned}$$

Dalla prima si esprime σ_u^2 , dalla seconda b_1 ecc

$$\begin{aligned} \sigma_u^2 &= \frac{R_s(0)}{1 + \sum_{j=1}^q b_j^2} \\ b_k &= \frac{R_s(k)}{\sigma_u^2} - \sum_{j=1}^{q-k} b_j b_{j+k}, \quad k=1, 2, \dots, q \end{aligned} \quad 35$$

che possono essere risolte iterativamente attraverso:

$$\begin{aligned} b_k^{(m+1)} &= \frac{R_s(k)}{(\sigma_u^2)^{(m)}} - \sum_{j=1}^{q-k} b_j^{(m)} b_{j+k}^{(m)} \quad k=1, 2, \dots, q \\ (\sigma_u^2)^{(m+1)} &= \frac{R_s(0)}{1 + \sum_{j=1}^q (b_j^{(m+1)})^2} \end{aligned} \quad 36$$

dove $(.)^{(m)}$ è il valore alla i -esima iterazione di $(.)$. Le condizioni iniziali sono tutte nulle per b_k e pari a $R_s(0)$ per σ_u^2 . Le equazioni (36) convergono linearmente, e generalmente si assume per il criterio di arresto che la distanza tra il valore di tutti i parametri tra due iterazioni successive sia minore di un valore prefissato. Da notare che mentre nel caso di modelli AR il metodo A e B di stima convergono, perché portano a risolvere le stesse equazioni, così non è nel caso di modelli MA.

Scelta dell'ordine

Mentre il procedimento di stima di modelli MA è più complesso che per modelli AR, più semplice risulta selezionare l'ordine ottimo. Infatti abbiamo visto che la funzione di autocorrelazione di un sistema MA è nulla per $|k| > q$, dove q è l'ordine del modello. Allora si può sfruttare questa proprietà per inferire dai dati sull'ordine ottimo, ricavando dai dati una stima della funzione di autocorrelazione per k crescenti, e controllando per quale valore di k si verifica che la stima assuma valori trascurabili. Si sceglie allora l'ordine q in modo che per $|k| > q$ si verifichi, con riferimento ai valori stimati, che $|R_s(k)/R_s(0)| < \varepsilon$ dove ε è un valore "piccolo" prefissato (es. 1%).

5. MODELLI ARMA

Nel modello ARMA l'uscita è espressa come somma di una parte autoregressiva e di una parte a media mobile

$$y_n = -\sum_{k=1}^p a_k y_{n-k} + \sum_{k=0}^q b_k u_{n-k} \quad 38$$

dove u_n è al solito rumore bianco, a media nulla e varianza σ_u^2 . Ancora una volta si può verificare che non è restrittivo porre $b_0=1$, quindi equazione alle differenze e funzione di trasferimento possono essere riscritte nella forma

$$y_n = -\sum_{k=0}^p a_k y_{n-k} + u_n + \sum_{k=1}^q b_k u_{n-k}$$

$$H_{\text{ARMA}}(z) = \frac{Y(z)}{U(z)} = \frac{1 + \sum_{k=1}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}}$$

In analogia a quanto visto per i modelli AR e MA, caratterizziamo il processo in uscita ad un modello ARMA esaminando le proprietà di stazionarietà e calcolando media e funzione di autocorrelazione.

1. Il processo in uscita è stazionario se il sistema è stabile, cioè se le radici dell'equazione caratteristica stanno entro il cerchio unitario del piano zeta.
2. Con una dimostrazione analoga a quella riportata per modelli AR, si può dimostrare che anche il processo in uscita ad un modello ARMA è a media nulla
3. La funzione di autocorrelazione di un modello ARMA risulta

$$R_y(k) = -\sum_{i=1}^p a_i R_y(k-i) + \sum_{i=0}^q b_i E\{u_{n-i} y_{n-k}\} \quad 39$$

dove l'ultimo addendo

$$\sum_{i=0}^q b_i E[u_{n-i} y_{n-k}] = 0 \quad \text{per } k > q \quad 40$$

mentre per $k < i < q$ si ha

$$E[u_{n-i} y_{n-k}] = \sigma_u^2 h_{i-k}, \quad \text{per } 0 \leq k \leq q \quad 41$$

dove h è la risposta impulsiva del sistema. Pertanto

$$R_y(k) = \begin{cases} -\sum_{i=1}^p a_i R_y(k-i) & \text{per } k > q \\ -\sum_{i=1}^p a_i R_y(k-i) + \sigma_u^2 \sum_{i=k}^q b_i h_{i-k} & \text{per } 0 \leq k \leq q \\ R(-k) & \text{per } k < 0 \end{cases} \quad 42$$

Per $k > q$, $R_y(k)$ è lineare nei parametri a_i (come per il modello AR) mentre per $0 \leq k \leq q$ non è lineare nei parametri a_i e b_i perché h è una funzione non lineare di a_i e b_i .

Esempio: modello ARMA di ordini 1,1

$$y_n = -a_1 y_{n-1} + u_n + b_1 u_{n-1}$$

Per prima cosa calcoliamo la risposta impulsiva h , che risulta essere

$$\begin{aligned} h_0 &= 1 \\ h_1 &= -a_1 + b_1 \\ h_2 &= -a_1(-a_1 + b_1) \\ h_3 &= a_1^2(-a_1 + b_1) \\ &\dots \end{aligned}$$

Funzione di autocorrelazione

$$\begin{aligned} R_y(0) &= -a_1 R_y(1) + \sigma_u^2 (b_0 h_0 + b_1 h_1) = -a_1 R_y(1) + \sigma_u^2 (1 + b_1^2 - a_1 b_1) \\ R_y(1) &= -a_1 R_y(0) + \sigma_u^2 (b_1 h_0) = -a_1 R_y(0) + \sigma_u^2 b_1 \\ R_y(2) &= -a_1 R_y(1) \\ R_y(3) &= -a_1 R_y(2) \end{aligned}$$

$R_y(0)$ e $R_y(1)$ sono nonlineari nei parametri !

Identificazione di modelli ARMA

Anche per la stima dei parametri del modello ARMA si può definire un predittore ad un passo. Seguendo un procedimento simile a quello già visto per modelli MA, a partire dal modello ARMA

$$Y(z) = \frac{B(z)}{A(z)} U(z) \quad 43$$

il modello in forma di predizione diventa:

$$\tilde{Y}(z) = \left(1 - \frac{A(z)}{B(z)}\right) Y(z) = \left(1 - \frac{1}{H_{\text{ARMA}}(z)}\right) Y(z)$$

quindi

$$H_{\text{PRED}}(z) = \frac{\tilde{Y}(z)}{Y(z)} = 1 - \frac{1}{H_{\text{ARMA}}(z)} \quad 44$$

Si tratta di un predittore NON lineare per cui non è possibile trovare il minimo della funzione costo, $J = E[e_n^2]$, per via analitica. Anche il Metodo A di stima porta a dover risolvere equazioni non lineari, e non ci sono algoritmi ad hoc per la loro soluzione. Verranno di seguito presentate due tecniche subottime, che riportano il problema di stima dei parametri di un modello ARMA a quelli, più semplici, di stima dei parametri di modelli AR e MA.

Approssimazione con la serie di un modello AR e di un modello MA

La prima è una tecnica sub-ottima basata sul criterio della minimizzazione dell'errore quadratico e richiede la soluzione di equazioni lineari. Questo metodo stima i parametri della parte AR e della parte MA in modo separato.

Per stimare i parametri a_i si ricorre alle proprietà dell'autocorrelazione del modello ARMA che per $k > q$ sono:

$$R_y(k) = -\sum_{i=1}^p a_i R_y(k-i) \quad 45$$

Utilizzando p di queste equazioni, da $k=q+1$ a $k=q+p$, si ottengono p equazioni lineari che possono essere risolte nelle p incognite a_1, \dots, a_p . Il sistema di eq.45 è detto “equazione di Yule-Walker estese”. Infatti non sono vere e proprie equazioni di Yule-Walker perché se scritte in forma matriciale portano ad una matrice ancora di Toeplitz ma non più simmetrica, quindi non sempre invertibile.

$$-\overbrace{\begin{bmatrix} R_y(q) & R_y(q-1) & \cdots & R_y(q-p+1) \\ R_y(q+1) & R_y(q) & \cdots & R_y(q-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ R_y(q+p-1) & R_y(q+p-2) & \cdots & R_y(q) \end{bmatrix}}^{\bar{R}'} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R_y(q+1) \\ R_y(q+2) \\ \vdots \\ R_y(q+p) \end{bmatrix} \quad 46$$

In particolare per un modello ARMA di ordini (p,q), la matrice R' risulta invertibile, così come lo sono le sue sottomatrici. Le matrici simili ad R' ma di dimensione maggiore di p sono invece non invertibili. Questa proprietà viene utilizzata per determinare l'ordine ottimo della parte AR: se $|R'_s|_i$ indica il determinante della matrice delle autocorrelazioni stimate dai dati di dimensione i, si valuta il valore di i per cui $|R'_s|_i$ è prossimo a zero, quindi inferiore ad un valore "piccolo" ϵ . L'ordine p della parte AR sarà i-1.

Una volta trovati i parametri della parte AR del modello, si passa al calcolo dei parametri del modello MA. A partire dal modello AR appena determinato si costruisce il filtro inverso, attraverso il quale si fanno passare i dati. Si ottiene perciò una sequenza di uscita puramente MA, ai quali si applicano le tecniche viste per la stima dei parametri di un modello MA. Il problema è che per scrivere le equazioni 46 l'ordine della parte MA deve essere noto. Si procede allora iterativamente: si parte con q=1, si calcolano i parametri della parte AR risolvendo il sistema 46, si filtrano i dati con il filtro inverso e sulla sequenza di uscita si testa se l'ordine q=1 è ragionevole, cioè se la stima della funzione di autocorrelazione di questa sequenza è trascurabile a partire da k=2. Se ciò non è vero, si aumenta di 1 l'ordine della parte a media mobile (q=2) e si ricomincia daccapo.

Approssimazione con un modello AR di ordine elevato

Il secondo metodo di identificazione dei modelli ARMA si basa su una approssimazione del modello ARMA con un modello AR di ordine eventualmente infinito, cosa sempre possibile perché $C(z)$ è il polinomio ottenuto dalla divisione dei polinomi $A(z)$ e $B(z)$.

$$\frac{B(z)}{A(z)} = \frac{1}{C(z)} \quad 47$$

$$\text{dove } C(z) = 1 + \sum_{k=1}^{\infty} c_k z^{-k} \quad 48$$

$C(z)$ viene poi approssimato con un sistema $\tilde{C}(z)$ di ordine M finito, sufficientemente elevato ($M > p+q$)

$$\tilde{C}(z) = 1 + \sum_{k=1}^M c_k z^{-k} \quad 49$$

Risulta allora

$$\frac{B(z)}{A(z)} \cong \frac{1}{\tilde{C}(z)} \quad \text{ovvero}$$

$$\tilde{C}(z)B(z) \cong A(z) \quad 50$$

Uguagliare i polinomi a destra e sinistra del segno dell'uguale comporta di uguagliare i coefficienti delle potenze di z di ugual grado, ovvero

$$\text{Grado 1 : } b_1 + c_1 = a_1$$

$$\text{Grado 2 : } b_2 + b_1 c_1 + c_2 = a_2$$

$$\text{Grado 3 : } b_3 + b_2 c_1 + b_1 c_2 + c_3 = a_3$$

.....

$$\text{Grado } p+1 : b_q c_{p+1-q} + b_{q-1} c_{p+2-q} + \dots b_1 c_p + c_{p+1} = 0$$

$$\text{Grado } p+2 : b_q c_{p+2-q} + b_{q-1} c_{p+3-q} + \dots b_1 c_{p+1} + c_{p+2} = 0$$

Le equazioni dal grado $p+1$ al grado $p+q$ sono lineari nelle incognite b_1, \dots, b_q . La loro soluzione fornisce quindi i parametri della parte MA del modello. Sostituendo i valori così trovati nelle equazioni dal grado 1 al grado p si possono determinare i valori dei coefficienti a_1, \dots, a_p della parte AR del modello.

L'approssimazione sarà tanto migliore quanto più elevato è l'ordine M del modello AR equivalente.

6. VALIDAZIONE DI MODELLI ARMA

Una volta messo a punto un modello, è essenziale giudicare oggettivamente la bontà del modello nel rappresentare il meccanismo di generazione dei dati. Oltre ai criteri di parsimonia, già visti in precedenza a proposito della scelta dell'ordine del modello, si possono ricavare informazioni sulla bontà del modello analizzando le caratteristiche dell'errore di predizione.

Errore di predizione

In precedenza abbiamo definito degli stimatori ad un passo per modelli AR, MA ed ARMA e in tutti i casi la relazione alle trasformate zeta tra ingresso e uscita del predittore è:

$$H_{\text{PRED}}(z) = \frac{\hat{Y}(z)}{Y(z)} = 1 - \frac{1}{H(z)}$$

Applicando tale relazione ai dati si ottiene

$$\hat{S}_{(z)} = \left[1 - \frac{1}{H(z)} \right] S(z)$$

La differenza $s_n - \hat{s}_n$ è l'errore di predizione.

$$\text{Posto } e_n = s_n - \hat{s}_n \quad \Rightarrow \quad E(z) = \frac{1}{H(z)} S(z)$$

Si può concludere che per un modello AR, MA o ARMA con generica f.d.t. $H(z)$, l'errore di predizione rappresenta l'uscita del filtro inverso, alimentato in ingresso con la sequenza di dati s_n :

$$E(z) = 1/H(z) S(z) \quad \text{ovvero} \quad S(z) = H(z) E(z)$$

Confrontando tale schema con il modello:

$$Y(z)=H(z)U(z)$$

si deduce che, se il modello è una rappresentazione adeguata del processo s_n allora l'errore di predizione e_n approssima una realizzazione di rumore bianco u_n . Pertanto si valuta la bontà del modello verificando la bianchezza dell'errore di predizione. Un test spesso usato è il test di Anderson.

Test di Anderson

Se una sequenza è una realizzazione di rumore bianco, la sua funzione di correlazione è diversa da zero nell'origine e nulla altrove:

cioè se e_n bianco, $R_e(0) = \sigma_e^2$, $R_e(k) = 0$ per $k \neq 0$.

Di conseguenza, il coefficiente di correlazione definito come

$$\rho_e(k) = \frac{R_e(k)}{R_e(0)}$$

è pari a 1 per $k=0$ e zero altrove.

Avendo a disposizione un numero finito N di campioni dell'errore di predizione, si può solo calcolare una stima di ρ_e :

$$\hat{\rho}_e(k) = \frac{\sum_{n=0}^{N-1-k} e_n e_{n+k}}{\sum_{n=0}^{N-1} e_n^2} \quad K = 0, 1, 2, \dots$$

Poichè $\hat{\rho}_e$ è solo una stima non possiamo pretendere che sia $\hat{\rho}_e(k) = 0$ $k \neq 0$, ma solo che $\hat{\rho}_e(k)$ sia "piccolo" per $k \neq 0$. Si segue il procedimento proprio dei test statistici:

1. Si fissa l'ipotesi statistica H_0 : e_n è rumore bianco.
2. Sotto H_0 , si valuta la descrizione statistica dello stimatore $\hat{\rho}_e$. Si può dimostrare che esso approssima una distribuzione gaussiana con valor medio nullo e varianza pari a $1/N$, cioè

$$\hat{\rho}_e(k) \in \mathcal{N}(0, \frac{1}{N}) \quad \text{per } k \neq 0$$
3. Si fissa un certo livello di significatività, ad esempio $\alpha=5\%$ e si valuta dalla distribuzione gaussiana l'intervallo di valori $\pm\beta$ in cui cade una frazione pari a $(1-\alpha)\%$ dei valori di ρ .
4. Si definisce la regola di decisione: si accetta l'ipotesi H_0 se al più una frazione α dei valori stimati di $\hat{\rho}_e(k)$ cade fuori della fascia $\pm\beta$.

Esempio

Se $N=100 \Rightarrow$ fissiamo $\alpha=5\% \Rightarrow \beta=\pm 0.196$ e quindi conclude che e_n è una realizzazione di rumore bianco se non più di 5 stime di ρ cadono fuori dell'intervallo ± 0.196 .