Probabilità e Processi Stocastici (455AA)

Lezione 19

Dario Trevisan

2/12/2024

Section 1

Processi a stati continui (Gaussiani)

Presentazione

Affrontiamo lo studio dei processi stocastici a **stati continui** (e tempi discreti) X_t a valori reali.

- Introduciamo i concetti fondamentali: funzione di media, di autocovarianza o di **autocorrelazione**, stazionarietà in senso lato, e processi gaussiani.

Presentazione

Affrontiamo lo studio dei processi stocastici a **stati continui** (e tempi discreti) X_t a valori reali.

- Introduciamo i concetti fondamentali: funzione di media, di autocovarianza o di **autocorrelazione**, stazionarietà in senso lato, e processi gaussiani.
- Tre esempi fondamontali di processi (gaussiani): il rumore bianco, la passeggiata aleatoria e l'equazione lineare con smorzamento.

Presentazione

Affrontiamo lo studio dei processi stocastici a **stati continui** (e tempi discreti) X_t a valori reali.

- Introduciamo i concetti fondamentali: funzione di media, di autocovarianza o di **autocorrelazione**, stazionarietà in senso lato, e processi gaussiani.
- Tre esempi fondamontali di processi (gaussiani): il rumore bianco, la passeggiata aleatoria e l'equazione lineare con smorzamento.
- I modelli ARIMA: definizione e proprietà (stazionarietà, funzione di autocovarianza e stima dei parametri).

Presentazione

Affrontiamo lo studio dei processi stocastici a **stati continui** (e tempi discreti) X_t a valori reali.

- Introduciamo i concetti fondamentali: funzione di media, di autocovarianza o di **autocorrelazione**, stazionarietà in senso lato, e processi gaussiani.
- Tre esempi fondamontali di processi (gaussiani): il rumore bianco, la passeggiata aleatoria e l'equazione lineare con smorzamento.
- I modelli ARIMA: definizione e proprietà (stazionarietà, funzione di autocovarianza e stima dei parametri).
- Infine studiamo il problema generale di stimare la funzione di autocovarianza di un processo stazionario (gaussiano), tramite **autocorrelazione campionaria** o **densità spettrale di potenza**.

Funzione di media e di autocovarianza

Dato un processo stocastico $(X_t)_{t\in\mathcal{T}}$ avente come stati $E = \mathbb{R}$, si definiscono

- la **funzione di media** del processo,

$$t \in \mathcal{T} \mapsto \mathbb{E}\left[X_t\right],$$

Funzione di media e di autocovarianza

Dato un processo stocastico $(X_t)_{t\in\mathcal{T}}$ avente come stati $E = \mathbb{R}$, si definiscono

- la **funzione di media** del processo,

$$t \in \mathcal{T} \mapsto \mathbb{E}\left[X_t\right],$$

- la **funzione di autocovarianza** del processo

$$(s,t) \in \underbrace{\mathcal{T}^2}_{\mathcal{T}\times\mathcal{T}} \mapsto \mathrm{Cov}\left(X_s, X_t\right) = K_{X_s X_t} = C(s,t),$$

Funzione di media e di autocovarianza

Dato un processo stocastico $(X_t)_{t\in\mathcal{T}}$ avente come stati $E = \mathbb{R}$, si definiscono

- la **funzione di media** del processo,

$$t \in \mathcal{T} \mapsto \mathbb{E}\left[X_t\right],$$

- la **funzione di autocovarianza** del processo

$$(s,t) \in \mathcal{T}^2 \mapsto \mathrm{Cov}\left(X_s, X_t\right) = K_{X_s X_t} = C(s,t),$$

- Notiamo che la varianza è

$$C(s,s) = \mathrm{Cov}\left(X_s, X_s\right) = \mathrm{Var}\left(X_s\right)$$

Funzione di autocorrelazione

A volte si considera anche la funzione

$$R(s,t) = \mathbb{E}[X_s X_t],$$

che è legata alle funzioni di autocovarianza e di media tramite la formula alternativa per il calcolo della covarianza:

$$R(s,t) = C(s,t) + \mathbb{E}[X_s]\mathbb{E}[X_t]$$

- Una terza funzione collegata a queste è la **funzione di autocorrelazione** (ACF) data da

$$(s,t) \in \mathcal{T}^2 \mapsto \mathrm{ACF}(s,t) = \rho_{X_s X_t} = \frac{\mathrm{Cov}(X_s, X_t)}{\sqrt{\mathrm{Var}(X_s)\,\mathrm{Var}(X_t)}} \in [-1,1],$$

Funzione di cross-covarianza

Ci limiteremo a processi a valori reali. Ma nel caso vettoriale ossia X_t a valori in $\mathbb{R}^d$, la funzione di autocovarianza si estende alla funzione di **covarianza incrociata** *cross-covariance* in inglese)

- per ogni coppia di componenti $i, j \in \{1, \dots, d\}$, definita come

$$K_{X_i X_j}(s, t) = \mathrm{Cov}\left(X_{i,s}, X_{j,t}\right),$$

Stazionarietà e funzioni di media e covarianza

Se il processo è stazionario, le funzioni di media e covarianza dipendono da *un parametro in meno*.

- Precisamente, se $\mathcal{T} = \{0, 1, 2, \dots, n\}$ oppure $\mathcal{T} = \mathbb{N}$ e il processo $(X_t)_{t \in \mathcal{T}}$ è stazionario, allora

$0 \le t_1 \le t_2 < \dots \le t_n \in \mathcal{T}$

legge di $(X_{t_i})_{i=1 \dots n}$

$\Delta t \ge 0$

$\|$ $\downarrow$

legge di $(X_{t_i + \Delta t})_{i=1 \dots n}$

Stazionarietà e funzioni di media e covarianza

Se il processo è stazionario, le funzioni di media e covarianza dipendono da *un parametro in meno*.

- Precisamente, se $\mathcal{T} = \{0, 1, 2, \ldots, n\}$ oppure $\mathcal{T} = \mathbb{N}$ e il processo $(X_t)_{t \in \mathcal{T}}$ è stazionario, allora
- il valor medio è costante nel tempo,

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] \quad \text{per ogni } t \in \mathcal{T},$$

$$\parallel$$

$$\int_{\mathbb{R}} x \, \underbrace{p(X_t = x)}_{\text{non dipende da } t} dx = \int x \, p(X_0 = x) dx$$

Stazionarietà e funzioni di media e covarianza

Se il processo è stazionario, le funzioni di media e covarianza dipendono da *un parametro in meno*.

- Precisamente, se $\mathcal{T} = \{0, 1, 2, \ldots, n\}$ oppure $\mathcal{T} = \mathbb{N}$ e il processo $(X_t)_{t \in \mathcal{T}}$ è stazionario, allora
- il valor medio è costante nel tempo,

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] \quad \text{per ogni } t \in \mathcal{T},$$

- l'autocovarianza dipende solamente dalla differenza (assoluta) dei due tempi,

$$C(s,t) = C(0, |t-s|), \quad \text{per ogni } s,\, t \in \mathcal{T}.$$

$0 \le s \le t \quad R(s,t) = \int_{\mathbb{R}\times\mathbb{R}} x \cdot y \, p(X_s = x, X_t = y)\, dx\, dy = \int_{\mathbb{R}\times\mathbb{R}} xy \, p(X_0 = x, X_{t-s} = y)\, dx\, dy$

Stazionarietà e funzioni di media e covarianza

Se il processo è stazionario, le funzioni di media e covarianza dipendono da *un parametro in meno*.

- Precisamente, se $\mathcal{T} = \{0, 1, 2, \ldots, n\}$ oppure $\mathcal{T} = \mathbb{N}$ e il processo $(X_t)_{t\in\mathcal{T}}$ è stazionario, allora
- il valor medio è costante nel tempo,

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] \quad \text{per ogni } t \in \mathcal{T},$$

- l'autocovarianza dipende solamente dalla differenza (assoluta) dei due tempi,

$$C(s,t) = C(0, |t-s|), \quad \text{per ogni } s,\, t \in \mathcal{T}.$$

- In particolare, la varianza $C(s,s) = C(0,0)$ è costante.

Dimostrazione

$0 \leq s \leq t$

$$R(s,t) = \int_{\mathbb{R}\times\mathbb{R}} x \cdot y \, p(X_s = x, X_t = y) \, dx \, dy = \int_{\mathbb{R}\times\mathbb{R}} x y \, p(X_0 = x, X_{t-s} = y) \, dx \, dy = R(0, t-s)$$

$$C(s,t) = R(s,t) - \mathbb{E}[X_s]\mathbb{E}[X_t]$$

$$= R(0,t-s) - \mathbb{E}[X_0]\mathbb{E}[X_{t-s}] = C(0,t-s)$$

Stazionarietà in senso lato

Il risultato sopra motiva un indebolimento del concetto di stazionarietà.

- Si dice che $(X_t)_{t\in\mathcal{T}}$ è **stazionario in senso lato**, se

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] \quad \text{per ogni } t,$$

e

$$C(s,t) = C(0, |t-s|), \quad \text{per ogni } s,\, t \in \mathcal{T}.$$

in particolare se stazionario $\Rightarrow$ staz. in senso lato

Stazionarietà in senso lato

Il risultato sopra motiva un indebolimento del concetto di stazionarietà.

- Si dice che $(X_t)_{t \in \mathcal{T}}$ è **stazionario in senso lato**, se

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] \quad \text{per ogni } t,$$

e

$$C(s,t) = C(0, |t-s|), \quad \text{per ogni } s, t \in \mathcal{T}.$$

- Questa nozione è più debole della stazionarietà *in senso stretto*. Ad esempio non dice nulla sui momenti terzo, quarto ecc. delle marginali.

Esempio $X_t = \begin{cases} \text{Bernoulli}(1/2) & t < 10 \\ N\left(\frac{1}{2}, \frac{1}{4}\right) & t \geq 10 \end{cases}$ indipendenti

Processi gaussiani

Diciamo che il processo $(X_t)_{t\in\mathcal{T}}$ è **gaussiano**, se ogni variabile congiunta

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_d}), \quad \text{è gaussiana vettoriale,}$$

per qualsiasi scelta $t_1, t_2, \ldots, t_d \in \mathcal{T}$.

- Se il processo è gaussiano e stazionario in senso lato, allora lo è anche in senso stretto: la legge congiunta del processo dipende solo dalla funzione di media e di covarianza, e in particolare i vettori

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_d}), \qquad (X_{t_1+s}, X_{t_2+s}, \ldots, X_{t_d+s})$$

  hanno gli stessi parametri di media e covarianza (quindi la stessa legge).

$$\mathrm{Cov}(X_{t_1}, X_{t_2}) = C(t_1, t_2) = C(0, |t_2 - t_1|) = \mathrm{Cov}(X_{t_1+s}, X_{t_2+s})$$

Section 2

Esempi

Esempio 1: Rumore bianco gaussiano

? ?

Il più semplice processo a stati continui che consideriamo consiste di variabili aleatorie $(W_t)_{t\in\mathcal{T}}$, tutte con la medesima legge *gaussiana e indipendenti*, con media nulla e varianza σ^2: la densità della marginale è quindi

$$p(W_t = w) = \exp\left(-\frac{w^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}. \qquad w \in \mathbb{R}$$

- $(W_t)_{t\in\mathcal{T}}$ è detto **rumore bianco gaussiano** di **intensità** σ^2.

Il termine *rumore* è motivato da modelli di teoria dell'informazione.

- Un messaggio $(M_t)_{t \in \mathcal{T}}$ è trasmesso tramite un mezzo di comunicazione reale affetto da *distorsione*, perdite ecc.

Il termine *rumore* è motivato da modelli di teoria dell'informazione.

- Un messaggio $(M_t)_{t\in\mathcal{T}}$ è trasmesso tramite un mezzo di comunicazione reale affetto da *distorsione*, perdite ecc.
- Si suppone che il ricevitore osservi il processo

$$(M_t + W_t)_{t\in\mathcal{T}},$$

dove $(W_t)_{t\in\mathcal{T}}$ è un rumore bianco gaussiano di una intensità $\boxed{\sigma^2}$ (è un parametro che si dovrà stimare).

Il termine *rumore* è motivato da modelli di teoria dell'informazione.

- Un messaggio $(M_t)_{t\in\mathcal{T}}$ è trasmesso tramite un mezzo di comunicazione reale affetto da *distorsione*, perdite ecc.
- Si suppone che il ricevitore osservi il processo

$$(M_t + W_t)_{t\in\mathcal{T}},$$

dove $(W_t)_{t\in\mathcal{T}}$ è un rumore bianco gaussiano di una intensità σ^2 (è un parametro che si dovrà stimare).

- Il rumore è sommato, per questo a volte è detto *additivo*.

Funzione di media e di autocovarianza

Sia $(W_t)_{t \in \mathcal{T}}$ un rumore bianco gaussiano di intensità σ^2.

- La funzione di media è identicamente nulla:

$$t \mapsto \mathbb{E}[W_t] = 0.$$

Funzione di media e di autocovarianza

Sia $(W_t)_{t \in \mathcal{T}}$ un rumore bianco gaussiano di intensità σ^2.

- La funzione di media è identicamente nulla:

$$t \mapsto \mathbb{E}[W_t] = 0.$$

- Ricordando che variabili indipendenti non sono correlate, e inoltre $\mathrm{Var}(W_t) = \sigma^2$ si ha

$$C(s,t) = \begin{cases} 0 & \text{se } s \neq t, \\ \sigma^2 & \text{se } s = t. \end{cases} \qquad \begin{array}{l} \text{se } t-s \neq 0 \\ \text{se } t-s = 0 \end{array}$$

$\Rightarrow (W_t)_{t \in \mathcal{T}}$ è stazionario

Funzione di media e di autocovarianza

Sia $(W_t)_{t \in \mathcal{T}}$ un rumore bianco gaussiano di intensità σ^2.

- La funzione di media è identicamente nulla:

$$t \mapsto \mathbb{E}\left[W_t\right] = 0.$$

- Ricordando che variabili indipendenti non sono correlate, e inoltre $\mathrm{Var}\left(W_t\right) = \sigma^2$ si ha

$$C(s,t) = \begin{cases} 0 & \text{se } s \neq t, \\ \sigma^2 & \text{se } s = t. \end{cases}$$

- Notazione $C(s,t) = \sigma^2 \delta_0(t-s)$ usando la delta di Dirac discreta:

$$\delta_0(x) = \begin{cases} 0 & \text{se } x \neq 0, \\ 1 & \text{se } x = 0. \end{cases}$$

Funzione di media e di autocovarianza

Sia $(W_t)_{t\in\mathcal{T}}$ un rumore bianco gaussiano di intensità σ^2.

- La funzione di media è identicamente nulla:
$$t \mapsto \mathbb{E}\left[W_t\right] = 0.$$
- Ricordando che variabili indipendenti non sono correlate, e inoltre $\mathrm{Var}\left(W_t\right) = \sigma^2$ si ha
$$C(s,t) = \begin{cases} 0 & \text{se } s \neq t, \\ \sigma^2 & \text{se } s = t. \end{cases}$$
- Notazione $C(s,t) = \sigma^2 \delta_0(t-s)$ usando la delta di Dirac discreta:
$$\delta_0(x) = \begin{cases} 0 & \text{se } x \neq 0, \\ 1 & \text{se } x = 0. \end{cases}$$
- Il processo è gaussiano e stazionario.

Stima dell'intensità σ^2 dalle osservazioni

Si può stimare σ^2 a partire da n osservazioni $W_{t_i} = w_i$.

- Il problema è lo stesso della stima della varianza di un campione gaussiano (di cui la media è nota e uguale a zero).

Stima dell'intensità σ^2 dalle osservazioni

Si può stimare σ^2 a partire da n osservazioni $W_{t_i} = w_i$.

- Il problema è lo stesso della stima della varianza di un campione gaussiano (di cui la media è nota e uguale a zero).
- In particolare la stima di massima verosimiglianza in questo caso è

$$\sigma^2_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} w_i^2 .$$

- Stime Bayesiane ...

Perché "bianco"?

- Supponiamo che l'insieme dei tempi sia finito $\mathcal{T} = \{0, 1, \ldots, n-1\}$.

.

Perché “bianco”?

- Supponiamo che l’insieme dei tempi sia finito $\mathcal{T} = \{0, 1, \ldots, n-1\}$.
- La trasformata di Fourier di $(W_t)_{t=0}^{n-1}$ è

finita

$$\hat{W}(\xi) = \sum_{t=0}^{n-1} W_t e^{-2\pi i \xi t/n}. \qquad \xi \in \{0, 1, 2, \ldots n-1\}$$

v.v.a. gaussiane

Perché "bianco"?

- Supponiamo che l'insieme dei tempi sia finito $\mathcal{T} = \{0, 1, \ldots, n-1\}$.
- La trasformata di Fourier di $(W_t)_{t=0}^{n-1}$ è

$$\hat{W}(\xi) = \sum_{t=0}^{n-1} W_t e^{-2\pi i \xi t/n}.$$

- Per ciascuna frequenza $\xi \in \{0, \ldots, (n-1)\}$ la variabile aleatoria $\hat{W}(\xi)$ è una combinazione lineare (a coefficienti complessi) di variabili gaussiane indipendenti, quindi è gaussiana.

Perché “bianco”?

- Supponiamo che l’insieme dei tempi sia finito $\mathcal{T} = \{0, 1, \ldots, n-1\}$.
- La trasformata di Fourier di $(W_t)_{t=0}^{n-1}$ è

$$\boxed{\hat{W}(\xi) = \sum_{t=0}^{n-1} W_t e^{-2\pi i \xi t/n}.}$$

- Per ciascuna frequenza $\xi \in \{0, \ldots, (n-1)\}$ la variabile aleatoria $\hat{W}(\xi)$ è una combinazione lineare (a coefficienti complessi) di variabili gaussiane indipendenti, quindi è gaussiana.
- Il valor medio è per linearità

$$\mathbb{E}\left[\hat{W}(\xi)\right] = \sum_{t=0}^{n-1} \mathbb{E}\left[W_t\right] e^{-2\pi i \xi t/n} = 0,$$

$$\mathbb{E}\left[\frac{|\hat{W}(\xi)|^2}{n}\right] = \sigma^2$$

- Il valor medio dell'energia $|\hat{W}(\xi)|^2$ sulla frequenza ξ è costante:

$$\mathbb{E}\left[|\hat{W}(\xi)|^2\right] = \sigma^2 n.$$

$$\mathbb{E}\left[|\hat{W}(\xi)|^2\right] = \mathbb{E}\left[\left|\sum_{t=0}^{n-1} W_t e^{-2\pi i t\xi/n}\right|^2\right]$$

$$= \mathbb{E}\left[\left(\sum_{t=0}^{n-1} W_t e^{-2\pi i t\xi/n}\right)\left(\sum_{s=0}^{n-1} W_s e^{+2\pi i s\xi/n}\right)\right]$$

$$= \sum_{s,t=0}^{n-1} \mathbb{E}\left[W_t W_s \, e^{-2\pi i t\xi/n} e^{+2\pi i s\xi/n}\right] = n\sigma^2$$

- Il valor medio dell'energia $|\hat{W}(\xi)|^2$ sulla frequenza ξ è costante:

$$\boxed{\mathbb{E}\left[|\hat{W}(\xi)|^2\right] = \sigma^2 n.}$$

- Dividendo per il tempo n si trova che la *potenza* media sulla frequenza ξ è costante e pari all'intensità σ^2.

Esempio 2: passeggiata aleatoria gaussiana

Posto $\mathcal{T} = \{0, 1, \ldots, n\}$ o eventualmente $\mathcal{T} = \mathbb{N}$, definiamo

$$S_0 = 0, \quad \boxed{S_t = W_1 + W_2 + \ldots + W_t} = \boxed{\sum_{s=1}^{t} W_s,}$$

dove $(W_s)_s$ è un rumore bianco gaussiano di instensità σ^2.

- Una definizione alternativa *ricorsiva* pone $S_0 = 0$ e per ogni $t \in \mathcal{T}$, $t \geq 1$,

$$\boxed{S_t = S_{t-1} + W_t.}$$

Esempio 2: passeggiata aleatoria gaussiana

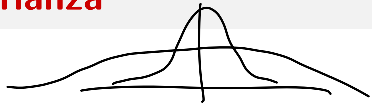
Posto $\mathcal{T} = \{0, 1, \ldots, n\}$ o eventualmente $\mathcal{T} = \mathbb{N}$, definiamo

$$S_0 = 0, \quad S_t = W_1 + W_2 + \ldots + W_t = \sum_{s=1}^{t} W_s,$$

dove $(W_s)_s$ è un rumore bianco gaussiano di instensità σ^2.

- Una definizione alternativa *ricorsiva* pone $S_0 = 0$ e per ogni $t \in \mathcal{T}$, $t \geq 1$,
$$S_t = S_{t-1} + W_t.$$
- Il processo si intepreta come una "passeggiata", in cui ogni nuovo "passo" W_t sposta da S_{t-1} in $S_t = S_{t-1} + W_t$. È detto **passeggiata aleatoria gaussiana**.

Funzione di media e di autocovarianza



La media di ciascuna S_t è nulla:

$$\mathbb{E}\left[S_t\right] = \mathbb{E}\left[\sum_{s=1}^{t} W_s\right] = \sum_{s=1}^{t} \mathbb{E}\left[W_s\right] = 0.$$

- La varianza vale

$$\mathrm{Var}\left(S_t\right) = \mathrm{Var}\left(\sum_{i=1}^{t} W_i\right) \overset{\text{indipendenti}}{=} \sum_{i=1}^{t} \mathrm{Var}\left(W_i\right)$$

$$= \sum_{i=1}^{t} \sigma^2 = t\sigma^2.$$

e non è costante. **La passeggiata aleatoria non è stazionaria.**

Funzione di media e di autocovarianza

La media di ciascuna S_t è nulla:

$$\mathbb{E}[S_t] = \mathbb{E}\left[\sum_{s=1}^{t} W_s\right] = \sum_{s=1}^{t} \mathbb{E}[W_s] = 0.$$

- La varianza vale

$$\begin{aligned}\mathrm{Var}(S_t) &= \mathrm{Var}\left(\sum_{i=1}^{t} W_i\right) = \sum_{i=1}^{t} \mathrm{Var}(W_i) \\ &= \sum_{i=1}^{t} \sigma^2 = t\sigma^2.\end{aligned}$$

  e non è costante. **La passeggiata aleatoria non è stazionaria.**

- Possiamo anche calcolare la funzione di autocovarianza $C(s,t) = \sigma^2 \min\{s,t\}$. $\neq C(0, |t-s|)$

Stima del parametro σ^2

$$S_t = S_{t-1} + W_t \qquad (S_0 = 0)$$

Si può stimare σ^2 partendo da n osservazioni $S_t = s_t$, per $t = 1, 2, \ldots, n$.

- Tramite una *differenza finita* (o derivata discreta) passiamo dalla passeggiata aleatoria al rumore bianco gaussiano:

$$w_t = s_t - s_{t-1}$$

si trovano n osservazioni.

Stima del parametro σ^2

Si può stimare σ^2 partendo da n osservazioni $S_t = s_t$, per $t = 1, 2, \ldots, n$.

- Tramite una *differenza finita* (o derivata discreta) passiamo dalla passeggiata aleatoria al rumore bianco gaussiano:

$$w_t = s_t - s_{t-1}$$

  si trovano n osservazioni.

- La stima di massima verosimiglianza è

$$\boxed{\sigma^2_{\mathrm{MLE}} = \frac{1}{n} \sum_{t=1}^{n} (s_t - s_{t-1})^2.}$$

Esempio 3: equazione lineare con smorzamento

Consideriamo una variante della passeggiata aleatoria: prima di ogni nuovo passo *trasformiamo* lo stato tramite una *dilatazione* di un parametro α. In formule:

- posto $(W_i)_i$ un rumore bianco gaussiano di instensità σ^2,

$$S_t = \alpha S_{t-1} + W_t \qquad \alpha$$

Esempio 3: equazione lineare con smorzamento

Consideriamo una variante della passeggiata aleatoria: prima di ogni nuovo passo *trasformiamo* lo stato tramite una *dilatazione* di un parametro α. In formule:

- posto $(W_i)_i$ un rumore bianco gaussiano di instensità σ^2,
- l'equazione ricorsiva è, per ogni $t \in \mathcal{T}$, $t \geq 1$,

$$X_t = \alpha X_{t-1} + W_t.$$

Esempio 3: equazione lineare con smorzamento

Consideriamo una variante della passeggiata aleatoria: prima di ogni nuovo passo *trasformiamo* lo stato tramite una *dilatazione* di un parametro α. In formule:

- posto $(W_i)_i$ un rumore bianco gaussiano di instensità σ^2,
- l'equazione ricorsiva è, per ogni $t \in \mathcal{T}$, $t \geq 1$,

$$X_t = \alpha X_{t-1} + W_t.$$

$X_0 = ?$

- Se $\alpha = 1$ è la passeggiata aleatoria.

Esempio 3: equazione lineare con smorzamento

Consideriamo una variante della passeggiata aleatoria: prima di ogni nuovo passo *trasformiamo* lo stato tramite una *dilatazione* di un parametro α. In formule:

- posto $(W_i)_i$ un rumore bianco gaussiano di instensità σ^2,
- l'equazione ricorsiva è, per ogni $t \in \mathcal{T}$, $t \geq 1$,

$$X_t = \alpha X_{t-1} + W_t.$$

- Se $\alpha = 1$ è la passeggiata aleatoria.
- Se $|\alpha| < 1$, l'effetto è di riavvicinare X_{t-1} verso l'origine, uno *smorzamento*. Se non ci fosse il rumore, sarebbe esponenziale:

$$\boxed{X_t = \alpha X_{t-1}} = \alpha^2 X_{t-2} = \ldots = \alpha^t X_0.$$

Funzione di media e varianza

Supponiamo che X_0 abbia densità gaussiana di parametri $\mathcal{N}(0, \sigma_0^2)$ e sia indipendente dal rumore bianco gaussiano.

- La funzione di media del processo è costante e nulla:

$$\mathbb{E}[X_t] = \mathbb{E}[\alpha X_{t-1} + W_t] = \alpha \mathbb{E}[X_{t-1}] + \cancel{\mathbb{E}[W_t]} = \alpha \mathbb{E}[X_{t-1}],$$

e quindi, ripetendo t volte,

$$\mathbb{E}[X_t] = \alpha \mathbb{E}[X_{t-1}] = \alpha^2 \mathbb{E}[X_{t-2}] = \ldots = \underline{\alpha^t \mathbb{E}[X_0]} = 0.$$

$t \mapsto \mathbb{E}[X_t] = 0$

Funzione di media e varianza

Supponiamo che X_0 abbia densità gaussiana di parametri $\mathcal{N}(0, \sigma_0^2)$ e sia indipendente dal rumore bianco gaussiano.

- La funzione di media del processo è costante e nulla:

$$\mathbb{E}[X_t] = \mathbb{E}[\alpha X_{t-1} + W_t] = \alpha \mathbb{E}[X_{t-1}] + \mathbb{E}[W_{t-1}] = \alpha \mathbb{E}[X_{t-1}],$$

e quindi, ripetendo t volte,

$$\mathbb{E}[X_t] = \alpha \mathbb{E}[X_{t-1}] = \alpha^2 \mathbb{E}[X_{t-2}] = \ldots = \alpha^t \mathbb{E}[X_0] = 0.$$

- Per la varianza usiamo che $X_{t-1} = f(X_0, W_1, \ldots, W_{t-1})$ è indipendente da W_t,

$$\mathrm{Var}(X_t) = \mathrm{Var}(\alpha X_{t-1} + W_t) = \alpha^2 \mathrm{Var}(X_{t-1}) + \sigma^2.$$

$$\Rightarrow \boxed{\sigma_0^2 = \alpha^2 \sigma_0^2 + \sigma^2}$$

Condizione di stazionarietà (Necessaria)

Sotto quali condizioni la varianza è costante $\mathrm{Var}(X_t) = \sigma^2 = \sigma_0^2$? deve valere

$$\sigma_0^2 = \alpha^2 \sigma_0^2 + \sigma^2,$$

da cui

$$\boxed{\sigma_0^2 = \frac{\sigma^2}{1-\alpha^2}.}$$

- Il termine $1-\alpha^2$ deve essere positivo, e quindi troviamo la condizione

$$|\alpha| < 1.$$

Funzione di autocovarianza

$s \le t$ $\quad$ $\underline{s = t \text{ ok}} \Rightarrow \underline{s < t}$

$$C(s,t) = Cov(X_s, X_t) = Cov(X_s, \alpha X_{t-1} + W_t) =$$

$$= \alpha\, Cov(X_s, X_{t-1}) + \cancel{Cov(X_s, W_t)}$$

$$= \alpha\, C(s, t-1)$$

$$\text{indip} \begin{cases} X_s = f(X_0, W_1, \dots, W_s) \\ W_t \end{cases} \quad \underline{t > s}$$

$$C(s,t) = \alpha\, C(s,t-1) = \alpha^2 C(s,t-2) = \dots = \alpha^k C(s, t-k)$$

?

$$\boxed{C(s,t) = \alpha^{t-s}\, Var(X_s)}$$

$\underline{\text{fino a } t-k=s} \Rightarrow k = t-s$

Se $\sigma_0^2 = \dfrac{\sigma^2}{1-\alpha^2} \Rightarrow Var(X_s) = \sigma_0^2 \Longrightarrow$

$$C(s,t) = \alpha^{t-s} \cdot \sigma_0^2$$

<u>Stazionario</u>

Funzione di autocorrelazione

La funzione di autocorrelazione è $\rho(t) = \alpha^t$. (nel caso stazionario)

$\rho(s, t) = \mathrm{cov}(X_s, X_t)/\sigma_0^2 = \alpha^{t-s}$

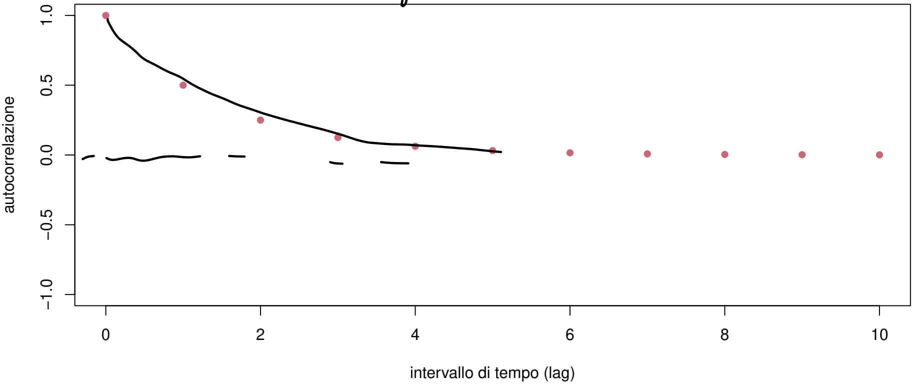


Figure 1: funzione di autocorrelazione dell'equazione lineare con smorzamento per $\alpha = 1/2$

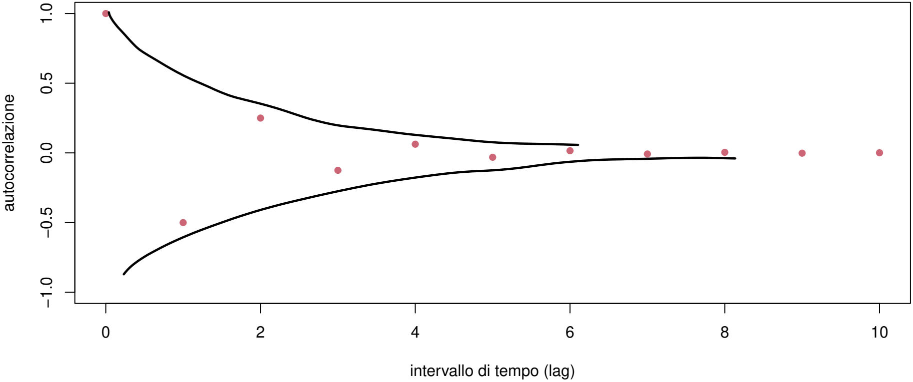


Figure 2: funzione di autocorrelazione dell'equazione lineare con smorzamento per $\alpha = -1/2$

Stima dei parametri

Possiamo stimare i parametri α e σ^2 da $n+1$ osservazioni $X_t = x_t$ per $t = 0, 1, \ldots, n$.

- Ci riconduciamo a n osservazioni di rumore bianco gaussiano

$$w_t = x_t - \alpha x_{t-1}.$$

$$X_t = \alpha X_{t-1} + W_t \Rightarrow W_t = X_t - \alpha X_{t-1}$$

osservate

Stima dei parametri

Possiamo stimare i parametri α e σ^2 da $n+1$ osservazioni $X_t = x_t$ per $t = 0, 1, \ldots, n$.

- Ci riconduciamo a n osservazioni di rumore bianco gaussiano

$$w_t = x_t - \alpha x_{t-1}.$$

- La stima di massima verosimiglianza si ottiene minimizzando

$$(\alpha, \sigma^2) \mapsto \sum_{t=1}^{n} \frac{(x_t - \alpha x_{t-1})^2}{\sigma^2} - n \log(\sigma^2)$$

In particolare, α_{MLE} minimizza la somma dei quadrati dei "residui"

$$\boxed{\alpha \mapsto \sum_{t=1}^{n} (x_t - \alpha x_{t-1})^2,}$$

Stima dei parametri

Possiamo stimare i parametri α e σ^2 da $n+1$ osservazioni $X_t = x_t$ per $t = 0, 1, \ldots, n$.

- Ci riconduciamo a n osservazioni di rumore bianco gaussiano

$$w_t = x_t - \alpha x_{t-1}.$$

- La stima di massima verosimiglianza si ottiene minimizzando

$$(\alpha, \sigma^2) \mapsto \sum_{t=1}^{n} \frac{(x_t - \alpha x_{t-1})^2}{\sigma^2} - n \log(\sigma^2)$$

  In particolare, α_{MLE} minimizza la somma dei quadrati dei "residui"

$$\alpha \mapsto \sum_{t=1}^{n} (x_t - \alpha x_{t-1})^2,$$

- Si trova $\alpha_{\mathrm{MLE}} = \frac{\sum_{t=1}^{n} x_t x_{t-1}}{\sum_{t=1}^{n} x_{t-1}^2}$, $\sigma^2_{\mathrm{MLE}} = \frac{1}{n} \sum_{t=1}^{n} (x_t - \alpha_{\mathrm{MLE}} x_{t-1})^2$.

Section 3

Modelli ARIMA

Introduzione

Introduciamo una famiglia generale di processi, detti ARIMA: **A**uto**R**egressive **I**ntegrated **M**oving **A**verage.

- Studiamo separatamente i tre "ingredienti" di un processo ARIMA:

Introduzione

Introduciamo una famiglia generale di processi, detti ARIMA: **A**uto**R**egressive **I**ntegrated **M**oving **A**verage.

- Studiamo separatamente i tre "ingredienti" di un processo ARIMA:

1. la componente autoregressiva (AR)

Introduzione

Introduciamo una famiglia generale di processi, detti ARIMA: **A**uto**R**egressive **I**ntegrated **M**oving **A**verage.

- Studiamo separatamente i tre "ingredienti" di un processo ARIMA:

1. la componente autoregressiva (AR)
2. quella a media mobile (MA)

Introduzione

Introduciamo una famiglia generale di processi, detti ARIMA: **A**uto**R**egressive **I**ntegrated **M**oving **A**verage.

- Studiamo separatamente i tre "ingredienti" di un processo ARIMA:

1. la componente autoregressiva (AR) $\simeq$ Eq. lineare con smorz.
2. quella a media mobile (MA)
3. il procedimento di integrazione (I) a tempi discreti. $\approx$ passeggiata aleatoria

Introduzione

Introduciamo una famiglia generale di processi, detti ARIMA:
Auto**R**egressive **I**ntegrated **M**oving **A**verage.

- Studiamo separatamente i tre "ingredienti" di un processo ARIMA:

1. la componente autoregressiva (AR)
2. quella a media mobile (MA)
3. il procedimento di integrazione (I) a tempi discreti.

- Supponiamo che che $\mathcal{T} = \{0, 1, \ldots, n\}$ oppure $\mathcal{T} = \mathbb{N}$ o anche $\mathcal{T} = \mathbb{Z}$, e che $(W_t)_{t \in \mathcal{T}}$ sia un rumore bianco gaussiano di intensità σ^2.

Operatore di Lag

Introduciamo l'operatore L che trasforma

$$(X_t)_{t\in\mathcal{T}} \mapsto ((LX)_t)_{t\in\mathcal{T}} = (X_{t-1})_{t\in\mathcal{T}}$$

- Scriviamo semplicemente LX_t invece di $(LX)_t$.

$$\parallel$$
$$X_{t-1}$$

Operatore di Lag

Introduciamo l'operatore L che trasforma

$$(X_t)_{t\in\mathcal{T}} \mapsto ((LX)_t)_{t\in\mathcal{T}} = (X_{t-1})_{t\in\mathcal{T}}$$

- Scriviamo semplicemente LX_t invece di $(LX)_t$.
- L è lineare:

$$L(X+Y)_t = X_{t-1} + Y_{t-1} = LX_t + LY_t, \quad L(cX)_t = cLX_t.$$

Operatore di Lag

Introduciamo l'operatore L che trasforma

$$(X_t)_{t\in\mathcal{T}} \mapsto ((LX)_t)_{t\in\mathcal{T}} = (X_{t-1})_{t\in\mathcal{T}}$$

- Scriviamo semplicemente LX_t invece di $(LX)_t$.
- L è lineare:

$$L(X+Y)_t = X_{t-1} + Y_{t-1} = LX_t + LY_t, \quad L(cX)_t = cLX_t.$$

- Componendo L con se stesso si ottengono ritardi di ordine superiore: $L^2X_t = LLX_t = X_{t-2}$, $L^3X_t = X_{t-3}$, ecc.

Operatore di Lag

$X_t + 30X_{t-1} + 4X_{t-10} = p(L)X_t$

Introduciamo l'operatore L che trasforma

$$(X_t)_{t\in\mathcal{T}} \mapsto ((LX)_t)_{t\in\mathcal{T}} = (X_{t-1})_{t\in\mathcal{T}}$$

- Scriviamo semplicemente LX_t invece di $(LX)_t$.
- L è lineare:

$$L(X+Y)_t = X_{t-1} + Y_{t-1} = LX_t + LY_t, \quad L(cX)_t = cLX_t.$$

- Componendo L con se stesso si ottengono ritardi di ordine superiore: $L^2X_t = LLX_t = X_{t-2}$, $L^3X_t = X_{t-3}$, ecc.
- Espressioni del tipo

$$a_0X_t + a_1X_{t-1} + \ldots + a_kX_{t-k} = a_0X_t + a_1LX_t + a_2L^2X_t + \ldots a_kL^kX_t$$

si possono abbreviare come

$$p(L)X_t = (a_0 + a_1L + a_2L^2 + \ldots a_kL^k)X_t.$$

Modelli AR

I modelli autoregressivi generalizzano l'equazione lineare con smorzamento.

- L'equazione

"$Y = \alpha X + \text{Residuo}$"

$$X_t = \alpha X_{t-1} + W_t$$

è una regressione lineare semplice di X_t rispetto a X_{t-1}.

Modelli AR

I modelli autoregressivi generalizzano l'equazione lineare con smorzamento.

- L'equazione

$$X_t = \alpha X_{t-1} + W_t$$

  è una regressione lineare semplice di X_t rispetto a X_{t-1}.

- Si estende ad una regressione lineare multipla su $\underline{p \geq 1}$ istanti precedenti.

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + W_t$$

Modelli AR

I modelli autoregressivi generalizzano l'equazione lineare con smorzamento.

- L'equazione
$$X_t = \alpha X_{t-1} + W_t$$
è una regressione lineare semplice di X_t rispetto a X_{t-1}.
- Si estende ad una regressione lineare multipla su $p \geq 1$ istanti precedenti.
- Dato $p \geq 0$, un processo $(X_t)_{t \in \mathcal{T}}$ è detto AR(p) (autoregressivo di ordine p) se esistono parametri $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$ tali che, per ogni $t \in \mathcal{T}$ (tale che $t - p \in \mathcal{T}$) si abbia

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + W_t.$$

$(\mathcal{T} = \mathbb{Z}$ o $\mathcal{T} = \mathbb{N})$

Notazione compatta

Possiamo scrivere la ricorsione

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + W_t.$$

in modo compatto:

$$\boxed{p(L)X_t = W_t,} \qquad AR(p)$$

dove $\tilde{p}(L)$ è il polinomio formale nella variabile L dato da

$$\boxed{\tilde{p}(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \ldots - \alpha_p L^p} = 1 - \sum_{i=1}^{p} \alpha_i L^i.$$

Media mobile

$\boxed{\mathcal{T} = \mathbb{Z}}$

La media mobile su una finestra temporale sinistra di ampiezza $q \geq 1$, trasforma un processo $(Z_t)_{t \in \mathcal{T}}$ con le medie

$$\bar{Z}_t = \frac{1}{q} \sum_{i=0}^{q-1} Z_{t-i}.$$

- È caso particolare di *convoluzione* $Z * g$ tra il processo e il filtro

$$g(t) = \begin{cases} \frac{1}{q} & \text{se } i = 0, 1, \ldots, (q-1) \\ 0 & \text{altrimenti.} \end{cases}$$

$$Z * g(t) = \sum_{i \in \mathbb{Z}} Z_{t-i}\, g(i)$$

Convoluzione e stazionarietà

Data una qualsiasi sia g (nota e fissata), se il processo Z è stazionario (in senso lato o anche in senso stretto), anche $Z * g$ lo è (nello stesso senso).

- Infatti, la funzione di media è

$$\mathbb{E}\left[(Z * g)_t\right] = \mathbb{E}\left[\sum_i Z_{t-i} g(i)\right] = \sum_i \mathbb{E}\left[Z_{t-i}\right] g(i) = m \sum_i g(i),$$

avendo indicato con $m = \mathbb{E}\left[Z_s\right]$.

Convoluzione e stazionarietà

Data una qualsiasi sia g (nota e fissata), se il processo Z è stazionario (in senso lato o anche in senso stretto), anche $Z * g$ lo è (nello stesso senso).

- Infatti, la funzione di media è

$$\mathbb{E}\left[(Z * g)_t\right] = \mathbb{E}\left[\sum_i Z_{t-i} g(i)\right] = \sum_i \mathbb{E}\left[Z_{t-i}\right] g(i) = \boxed{m \sum_i g(i),}$$

avendo indicato con $m = \mathbb{E}\left[Z_s\right]$.

- La funzione di autocovariaza è, usando la bilinearità,

$$\begin{aligned} C(s,t) &= \mathrm{Cov}\left(\underbrace{(Z * g)_s}, \underbrace{(Z * g)_t}\right) = \sum_i \sum_j g(i) g(j) \underbrace{\mathrm{Cov}\left(Z_{s-i}, Z_{t-j}\right)}_{= C(0,(s-i-t+j))} \\ &= \sum_i \sum_j g(i) g(j) C((t-s) + (i-j)) \end{aligned}$$

che dipende da $t - s$ solamente.

Modelli MA

Dato $q \geq 0$, un processo $(X_t)_{t \in \mathcal{T}}$ è detto MA(q) (a media mobile di ordine q) se esistono parametri $\beta_1, \beta_2, \ldots, \beta_q \in \mathbb{R}$ tali che, per ogni $t \in \mathcal{T}$ (tale che $t - q \in \mathcal{T}$) si abbia

$$X_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q} = W * g(t)$$

- Un processo a media mobile MA(q) è semplicemente del tipo $W * g$, dove g è dato dai coefficienti 1, β_1, β_2, $\ldots$, β_q (e nullo altrove). In particolare, X è **gaussiano e stazionario**.

Modelli MA

Dato $q \geq 0$, un processo $(X_t)_{t \in \mathcal{T}}$ è detto MA(p) (a media mobile di ordine q) se esistono parametri $\beta_1, \beta_2, \ldots, \beta_q \in \mathbb{R}$ tali che, per ogni $t \in \mathcal{T}$ (tale che $t - q \in \mathcal{T}$) si abbia

$$X_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}.$$

- Un processo a media mobile MA(q) è semplicemente del tipo $W * g$, dove g è dato dai coefficienti 1, β_1, β_2, ..., β_q (e nullo altrove). In particolare, X è **gaussiano e stazionario**.
- Con il polinomio dell'operatore ritardo riscriviamo

$$X_t = q(L) W_t,$$

dove

$$q(L) = 1 + \beta_1 L + \ldots + \beta_q L^q = 1 + \sum_{j=1}^{q} \beta_q L^j.$$

Integrazione discreta

$$S_t = S_{t-1} + W_t \Rightarrow S_t - S_{t-1} = W_t$$

Consideriamo l'operazione di derivazione discreta: la derivata di un processo $(X_t)_{t\in\mathcal{T}}$ a tempi discreti è

$$\boxed{X_t - X_{t-1} = (1-L)X_t,}$$

per $t \geq 1$.

- Iterando, si trova

$$\boxed{(1-L)^d X_t = \sum_{i=0}^{d} \binom{d}{i} (-1)^i L^i X_t.}$$

es $(1-L)^2 X_t = X_t - 2X_{t-1} + X_{t-2}$

Integrazione discreta

Consideriamo l'operazione di derivazione discreta: la derivata di un processo $(X_t)_{t\in\mathcal{T}}$ a tempi discreti è

$$X_t - X_{t-1} = (1-L)X_t,$$

per $t \geq 1$.

- Iterando, si trova

$$(1-L)^d X_t = \sum_{i=0}^{d} \binom{d}{i} (-1)^i L^i X_t.$$

- La formula è una convoluzione $X * g$: se X è stazionario, lo è anche ogni derivata discreta di qualsiasi ordine d.

L'integrazione discreta è l'inversa della derivata discreta: X è l'integrale discreto di Y se vale $(1-L)X = Y$, e similmente per gli integrali iterati.

- la passeggiata aleatoria gaussiana, $S_t = S_{t-1} + W_t$, ossia $(1-L)S_t = W_t$ è l'integrale di W_t.

L'integrazione discreta è l'inversa della derivata discreta: X è l'integrale discreto di Y se vale $(1-L)X = Y$, e similmente per gli integrali iterati.

- la passeggiata aleatoria gaussiana, $S_t = S_{t-1} + W_t$, ossia $(1-L)S_t = W_t$ è l'integrale di W_t.
- l'integrazione discreta non mantiene la stazionarietà di un processo.

X è $\underline{I(d)}$ di Y se $(1-L)^d X = Y$

integrazione

Definizione generale ARIMA

Dati $p, d, q \geq 0$, un processo $(X_t)_{t \in \mathcal{T}}$ è detto ARIMA(p, d, q) se esistono parametri $(\alpha_i)_{i=1}^p$, $(\beta_j)_{j=1}^q$ reali tali che, per ogni $t \in \mathcal{T}$ (tale che $t - d - p$ e $t - q \in \mathcal{T}$), posto

$$Y_t = (1 - L)^d X_t$$

valga

$$Y_t = \underbrace{\sum_{i=1}^p \alpha_i Y_{t-i}}_{AR(p)} + W_t + \underbrace{\sum_{j=1}^q \beta_j W_{t-j}}_{MA(q)}.$$

- Usando i polinomi

$$p(L) = 1 - \sum_{i=1}^p \alpha_i L^i, \quad \text{e} \quad q(L) = 1 + \sum_{j=1}^q \beta_j L^j$$

si può scrivere in forma compatta:

$$\boxed{p(L)(1 - L)^d X_t = q(L) W_t.}$$

Esempi

- il rumore bianco gaussiano è ARIMA$(0, 0, 0)$,

$X_t = W_t$ ok

Esempi

- il rumore bianco gaussiano è ARIMA$(0, 0, 0)$,
- la passeggiata aleatoria è ARIMA$(0, 1, 0)$

$S_t - S_{t-1} = W_t$

è anche ARIMA$(1,0,0)$ con $\underline{\underline{\alpha_1 = 1}}$

$(1-L) S_t = W_t$

$\underline{\underline{d=1}}$ $\quad p=0, q=0$

Esempi

- il rumore bianco gaussiano è ARIMA$(0,0,0)$,
- la passeggiata aleatoria è ARIMA$(0,1,0)$
- l'equazione lineare con smorzamento è ARIMA$(1,0,0)$.

$$X_t = \alpha X_{t-1} + W_t$$

$$(1 - \alpha L) X_t = W_t \qquad \underline{\underline{\alpha_1 = \alpha}}$$

Section 4

Proprietà dei processi ARIMA

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)
3. come stimare i parametri sulla base delle osservazioni.

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)
3. come stimare i parametri sulla base delle osservazioni.

- Consideriamo un processo $(X_t)_{t \in \mathcal{T}}$ ARIMA(p, d, q) con parametri $(\alpha_i)_{i=1}^p$ e $(\beta_j)_{j=1}^q$.

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L)W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L)W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

- Nel caso generale, l'idea è "risolvere" l'equazione del modello

$$p(L)(1-L)^d X_t = q(L)W_t, \quad \text{da cui} \quad X_t = \frac{q(L)}{p(L)(1-L)^d} W_t,$$

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L)W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

- Nel caso generale, l'idea è "risolvere" l'equazione del modello

$$p(L)(1-L)^d X_t = q(L)W_t, \quad \text{da cui} \quad X_t = \frac{q(L)}{p(L)(1-L)^d} W_t,$$

- Per dare senso alla scrittura, sviluppiamo in serie

$$\frac{q(z)}{p(z)(1-z)^d} = \sum_{k=0}^{\infty} b_k z^k, \quad \text{così} \quad X_t = \sum_{k=0}^{\infty} b_k L^k W_t.$$

Criterio di stazionarietà

Il problema sta nella convergenza della serie, che dipende dalla crescita dei coefficienti b_k al tendere di $k \to \infty$ e in ultima analisi agli zeri (complessi) del denominatore $p(z)(1-z)^d$.

- Dati (p, d, q) e coefficienti $(\alpha_i)_{i=1}^p$, $(\beta_j)_{j=1}^q$, posto

$$p(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i,$$

allora esiste un modello ARIMA(p, d, q) stazionario con tali coefficienti se e solo se $d = 0$ e tutte le radici complesse di $p(z)$ hanno modulo $|z| > 1$, ossia

$$\text{se } z \in \mathbb{C} \text{ è tale che } p(z) = 0, \text{ allora } |z| > 1.$$

Criterio di stazionarietà

Il problema sta nella convergenza della serie, che dipende dalla crescita dei coefficienti b_k al tendere di $k \to \infty$ e in ultima analisi agli zeri (complessi) del denominatore $p(z)(1-z)^d$.

- Dati (p, d, q) e coefficienti $(\alpha_i)_{i=1}^p$, $(\beta_j)_{j=1}^q$, posto

$$p(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i,$$

  allora esiste un modello ARIMA(p, d, q) stazionario con tali coefficienti se e solo se $d = 0$ e tutte le radici complesse di $p(z)$ hanno modulo $|z| > 1$, ossia

$$\text{se } z \in \mathbb{C} \text{ è tale che } p(z) = 0, \text{ allora } |z| > 1.$$

- Nel caso dell'equazione lineare con smorzamento si trova $|\alpha| < 1$.

Autocovarianza

I processi ARIMA hanno media nulla e l'equazione permette anche di ottenere una formula ricorsiva (le *equazioni di Yule-Walker*) per la funzione di autocovarianza:

$$C(0,t) = \sum_{i=1}^{p} \alpha_i C(0, t-i), \quad \text{se } t > q.$$

- Nel caso ARIMA$(1,0,0)$ si ha

$$C(0,t) = \alpha_1 C(0, t-1) = \ldots = \alpha_1^t C(0,0).$$

Autocovarianza

I processi ARIMA hanno media nulla e l'equazione permette anche di ottenere una formula ricorsiva (le *equazioni di Yule-Walker*) per la funzione di autocovarianza:

$$C(0,t) = \sum_{i=1}^{p} \alpha_i C(0, t-i), \quad \text{se } t > q.$$

- Nel caso ARIMA$(1,0,0)$ si ha

$$C(0,t) = \alpha_1 C(0, t-1) = \ldots = \alpha_1^t C(0,0).$$

- Nel caso ARIMA$(0,0,q)$ si ha

$$C(0,t) = 0 \quad \text{se } t > q.$$

Dimostrazione

Stima dei parametri

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.

Stima dei parametri

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.
- Il metodo si riduce alla minimizzazione dei residui quadratici.

Stima dei parametri

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.
- Il metodo si riduce alla minimizzazione dei residui quadratici.
- Si può fare lo stesso per un ARIMA generale, ma la nozione di residuo va chiarita.

Partendo dall'equazione definente

$$p(L)(1-L)^d X_t = q(L) W_t$$

si ricava una equazione ricorsiva per il rumore bianco gaussiano W_t:

$$W_t = -\sum_{j=1}^{q} \beta_j W_{t-j} + p(L)(1-L)^d X_t$$

- Avendo osservato $X_t = x_t$, definiamo ricorsivamente i residui

$$w_t = -\sum_{j=1}^{q} \beta_j w_{t-j} + p(L)(1-L)^d x_t,$$

Partendo dall'equazione definente

$$p(L)(1-L)^d X_t = q(L) W_t$$

si ricava una equazione ricorsiva per il rumore bianco gaussiano W_t:

$$W_t = -\sum_{j=1}^{q} \beta_j W_{t-j} + p(L)(1-L)^d X_t$$

- Avendo osservato $X_t = x_t$, definiamo ricorsivamente i residui

$$w_t = -\sum_{j=1}^{q} \beta_j w_{t-j} + p(L)(1-L)^d x_t,$$

- La stima di massima verosimiglianza si ottiene minimizzando la somma dei quadrati:

$$(\alpha_{\mathrm{MLE}}, \beta_{\mathrm{MLE}}) \in \arg\min_{\alpha,\beta} \sum_t w_t^2$$

Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.

Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.
- Oltre alle stime puntuali forniscono stime delle deviazoni standard associate ai parametri.

Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.
- Oltre alle stime puntuali forniscono stime delle deviazoni standard associate ai parametri.
- Inoltre con la funzione `forecast()` si ottengono previsioni (basate sulla stima ottenuta) dei valori futuri.

Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.

Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.
- vi sono indicatori particolari (AIC, BIC) che "guidano" opportunamente nella scelta.

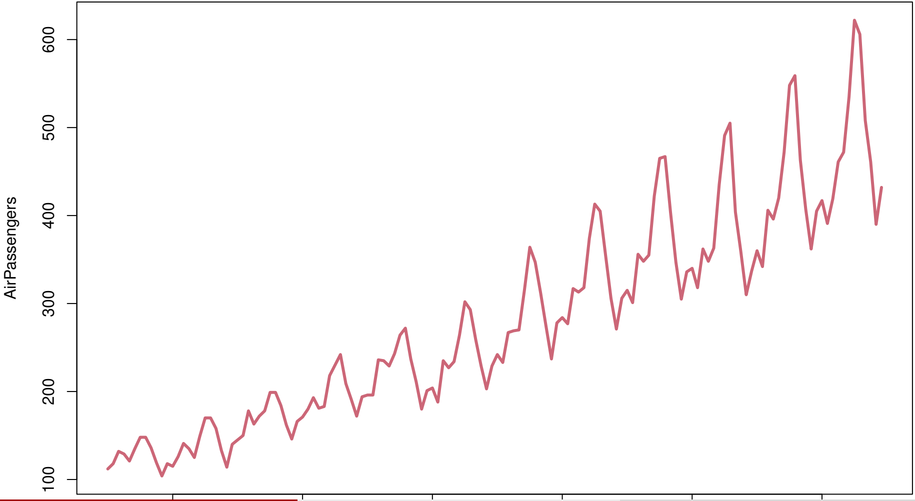
Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.
- vi sono indicatori particolari (AIC, BIC) che "guidano" opportunamente nella scelta.
- In R il comando `auto.arima()` valuta automaticamente il migliore modello che aderisce alle osservazioni sulla base di tali indicatori.

Un esempio

Consideriamo il dataset precaricato in R `AirPassengers`.



```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

## Series: AirPassengers
## ARIMA(2,1,1)(0,1,0)[12]
##
## Coefficients:
##          ar1     ar2      ma1
##       0.5960  0.2143  -0.9819
## s.e.  0.0888  0.0880   0.0292
##
## sigma^2 = 132.3:  log likelihood = -504.92
## AIC=1017.85   AICc=1018.17   BIC=1029.35
##
## Training set error measures:
##                    ME     RMSE     MAE       MPE     MAPE
## Training set 1.342299 10.84619 7.86754 0.4206976 2.800458 (
```

La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2, 1, 1)(0, 1, 0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.

La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2,1,1)(0,1,0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).

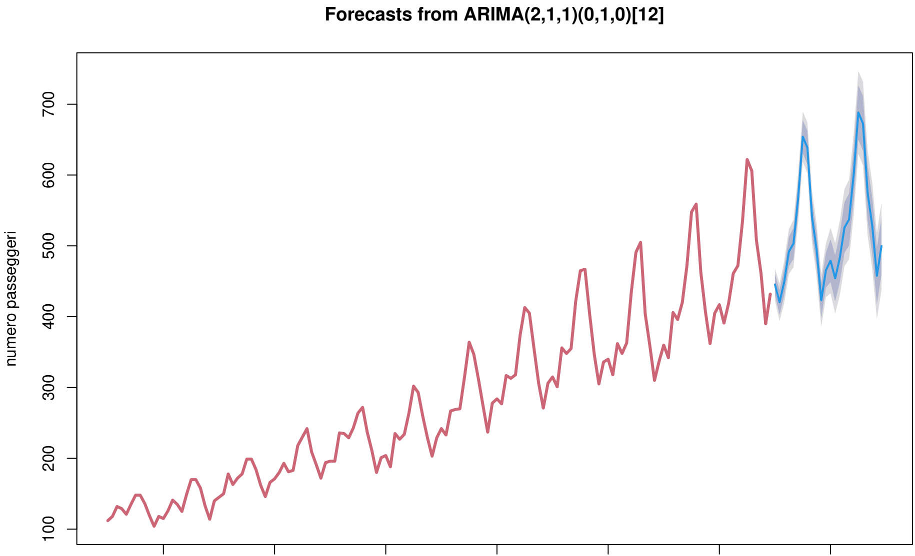
La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2, 1, 1)(0, 1, 0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).
- In generale si può ricorrere alla funzione di autocorrelazione empirica o ad analisi spettrale per determinare eventuali stagionalità e il loro periodo.

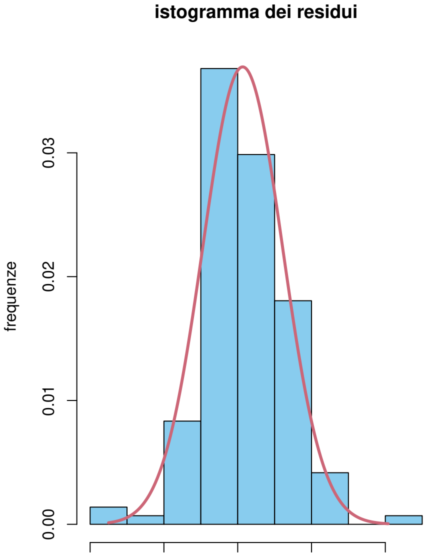
La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2,1,1)(0,1,0)$ (la seconda tripla si riferisce alla stagionalità).

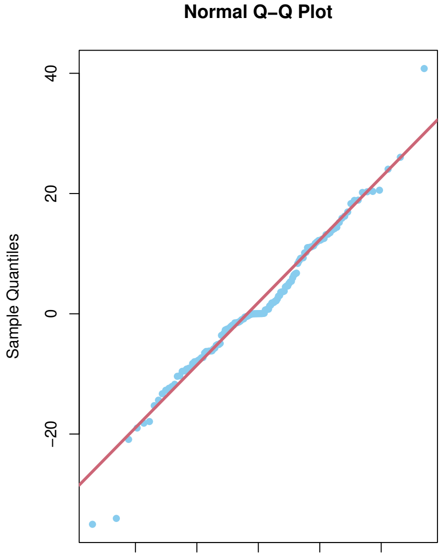
- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).
- In generale si può ricorrere alla funzione di autocorrelazione empirica o ad analisi spettrale per determinare eventuali stagionalità e il loro periodo.
- Un comando automatico è invece `findfrequency()`.

Previsione


Forecasts from ARIMA(2,1,1)(0,1,0)[12]
numero passeggeri
100
200
300
400
500
600
700

Residui


istogramma dei residui
frequenze
0.00
0.01
0.02
0.03


Normal Q–Q Plot
Sample Quantiles
−20
0
20
40