Probabilità e Processi Stocastici (455AA)

Lezione 19

Dario Trevisan

30/11/2023

Section 1

Cenni alla teoria delle code

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer...

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer...
- L'obiettivo è individuare strategie per migliorare l'esperienza di chi è in attesa (ridurre i tempi) rendendone più efficiente il servizio.

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer...
- L'obiettivo è individuare strategie per migliorare l'esperienza di chi è in attesa (ridurre i tempi) rendendone più efficiente il servizio.
- La teoria delle code è un campo molto esteso, presentiamo i modelli più semplici come esempi di processi di Markov a salti.

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,
- e infine l'uscita dalla coda quando il compito è svolto

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,
- e infine l'uscita dalla coda quando il compito è svolto
- Una volta introdotto un modello, è di interesse calcolare il tempo medio di attesa, il numero medio di clienti in coda e stimare i parametri di un modello sulla base di quantità osservate nella realtà.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un “processo” di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).
- Formalmente i modelli sono processi di Markov a salti negli stati $E = \mathbb{N}$.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un “processo” di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).
- Formalmente i modelli sono processi di Markov a salti negli stati $E = \mathbb{N}$.
- Lo stato n indica il numero di clienti in attesa o in corso di servizio.

Caso $M/M/0$: il processo di Poisson

Il modello più semplice è il caso in cui non vi siano serventi (oppure si è interessati solo al processo di arrivo dei clienti): il processo è detto *processo di Poisson* di intensità $\lambda > 0$.

Le intensità di salto sono

$$L_{n\to n+1} = \lambda, \quad L_{n\to n} = -\lambda \quad \text{e} \quad L_{n\to k} = 0 \qquad k \neq n,\ k \neq n+1.$$

$$0 \xrightarrow{\lambda} 1 \xrightarrow{\lambda} 2 \xrightarrow{\lambda} 3 \xrightarrow{\lambda} \cdots$$

- Ogni stato è transitorio, non vi sono distribuzioni invarianti.

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp(-t\lambda).$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda \mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp(-t\lambda).$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda\mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

- Calcoliamo quindi

$$\frac{d}{dt} \exp(-t\lambda) \frac{(t\lambda)^n}{n!} = \begin{cases} -\lambda \exp(-t\lambda) & \text{se } n = 0, \\ \frac{nt^{n-1}\lambda^n}{n!} \exp(-t\lambda) - \lambda \frac{(t\lambda)^n}{n!} \exp(-t\lambda) & \text{se } n \geq 1 \end{cases}$$

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp(-t\lambda).$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda \mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

- Calcoliamo quindi

$$\frac{d}{dt} \exp(-t\lambda) \frac{(t\lambda)^n}{n!} = \begin{cases} -\lambda \exp(-t\lambda) & \text{se } n = 0, \\ \frac{nt^{n-1}\lambda^n}{n!} \exp(-t\lambda) - \lambda \frac{(t\lambda)^n}{n!} \exp(-t\lambda) & \text{se } n \geq 1 \end{cases}$$

- Per concludere nel caso $n \geq 1$ basta notare che

$$\frac{nt^{n-1}\lambda^n}{n!} \exp(-t\lambda) = \lambda \frac{(t\lambda)^{n-1}}{(n-1)!} \exp(-t\lambda) = \lambda \mu_{n-1}^t.$$

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.
- La verosimiglianza è

$$L(\Lambda = \lambda; X = \gamma) = \prod_{k=1}^{n} \exp\left(-\lambda t_k\right) \lambda = \lambda^n \exp\left(-\lambda T\right).$$

dove supponiamo noto a priori che $X_0 = x_0$ e $T = \sum_{k=1}^{n} t_k$.

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.
- La verosimiglianza è

$$L(\Lambda = \lambda; X = \gamma) = \prod_{k=1}^{n} \exp(-\lambda t_k)\, \lambda = \lambda^n \exp(-\lambda T)\,.$$

  dove supponiamo noto a priori che $X_0 = x_0$ e $T = \sum_{k=1}^{n} t_k$.

- La stima di massima verosimiglianza si trova annullando la derivata rispetto a λ e vale

$$\frac{n}{\lambda_{MLE}} - T = 0 \quad \text{quindi} \quad \lambda_{MLE} = \frac{n}{T}.$$

Code $M/M/1$

Supponiamo vi sia un solo servente e che il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (ogni cliente sia indipendente dagli altri).

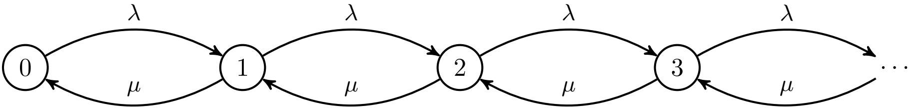
- Per arrivare al modello come un processo di Markov a salti, supponiamo non vi siano arrivi: si salta solo da n verso $n-1$ (se $n \geq 1$) con dei tempi di permanenza esponenziali di parametro μ. Pertanto, se $n \geq 1$,

$$L_{n \to n-1} = \mu.$$

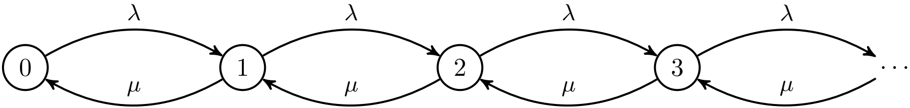
Code $M/M/1$

Supponiamo vi sia un solo servente e che il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (ogni cliente sia indipendente dagli altri).

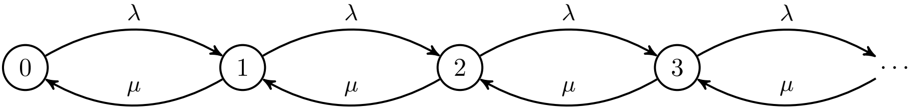
- Per arrivare al modello come un processo di Markov a salti, supponiamo non vi siano arrivi: si salta solo da n verso $n-1$ (se $n \geq 1$) con dei tempi di permanenza esponenziali di parametro μ. Pertanto, se $n \geq 1$,
$$L_{n \to n-1} = \mu.$$
- Aggiungiamo gli arrivi come un processo di Poisson di intensità λ: per $n \geq 0$,
$$L_{n \to n+1} = \lambda,$$
e di conseguenza
$$L_{n \to n} = \begin{cases} -\lambda & \text{se } n = 0 \\ -(\lambda + \mu) & \text{se } n \geq 1, \end{cases}$$



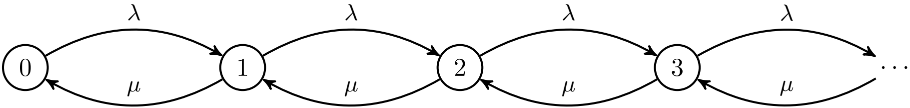
- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.
- Risolvendo l'equazione $\mu L = 0$ (o imponendo il bilancio di flusso) si trova
$$\mu_n \propto \left(\frac{\lambda}{\mu}\right)^n .$$



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.
- Risolvendo l'equazione $\mu L = 0$ (o imponendo il bilancio di flusso) si trova

$$\mu_n \propto \left(\frac{\lambda}{\mu}\right)^n .$$

- Per garantire che μ sia una densità di probabilità, bisogna che

$$\sum_n \left(\frac{\lambda}{\mu}\right)^n < \infty,$$

ossia che $\lambda < \mu$.

- La distribuzione invariante è *geometrica* di parametro $1-\lambda/\mu$, con valor medio
$$\mathbb{E}[N] = \sum_n n\mu_n = \frac{\lambda}{\mu-\lambda},$$

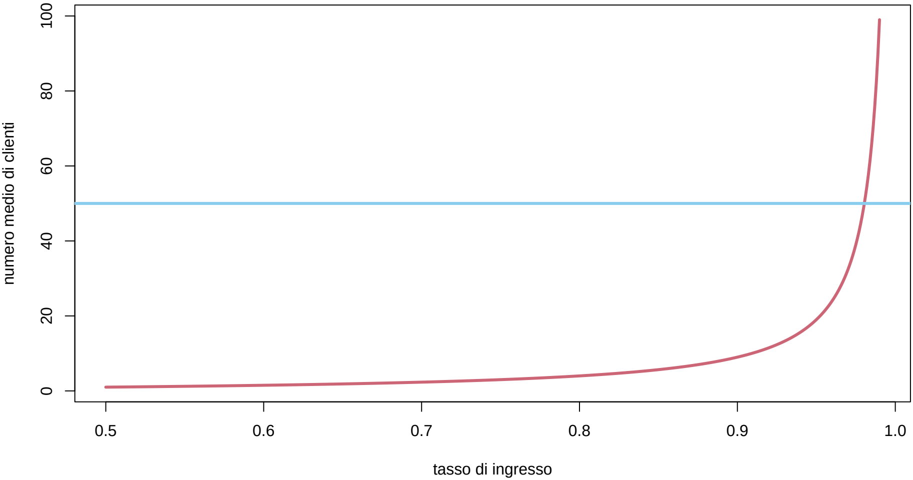


Figure 1: grafico di $\mathbb{E}[N]$ per $\mu = 1$ in funzione di λ (in rosso) e una soglia massima di possibili persone in coda (in azzurro)

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza t_1, t_2, ..., t_ℓ.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza t_1, t_2, ..., t_ℓ.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.
- La verosimiglianza è

$$L(\lambda, \mu; X = \gamma) = \exp\left(-(\lambda + \mu) T\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

  dove γ_+ indica il numero di arrivi osservati in γ (ossia transizioni da uno stato n a $n+1$), mentre γ_- il numero di uscite.

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza $t_1, t_2, \ldots, t_\ell$.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.
- La verosimiglianza è

$$L(\lambda, \mu; X = \gamma) = \exp\left(-(\lambda + \mu) T\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

  dove γ_+ indica il numero di arrivi osservati in γ (ossia transizioni da uno stato n a $n+1$), mentre γ_- il numero di uscite.
- La stima di massima verosimiglianza è

$$\lambda_{MLE} = \frac{\gamma_+}{T}, \quad \mu_{MLE} = \frac{\gamma_-}{T}.$$

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda + \mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda+\mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

- Intepretazione: il tempo in cui la coda è vuota non si può usare per stimare il tasso di uscita.

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda+\mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

- Intepretazione: il tempo in cui la coda è vuota non si può usare per stimare il tasso di uscita.
- *Esempio*: In un intervallo di 10 minuti si osservano 5 persone arrivare alla cassa di un supermercato e 3 persone uscirne. Supponendo che la cassa non sia mai senza lavoro si stimano i parametri $\lambda = 1/2$ persone al minuto, $\mu = 3/10$ persone al minuto. Se invece la cassa è rimasta priva di persone in coda per 4 minuti, si stima $\mu = 3/6 = 1/2$ persone al minuto.

Code $M/M/\infty$

Consideriamo la situazione con un numero arbitrariamente grande, idealmente infinito, di serventi ($M/M/\infty$).

- Il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (e ogni cliente sia indipendente dagli altri).

Code $M/M/\infty$

Consideriamo la situazione con un numero arbitrariamente grande, idealmente infinito, di serventi ($M/M/\infty$).

- Il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (e ogni cliente sia indipendente dagli altri).
- Per arrivare al modello , consideriamo il caso in cui non vi siano arrivi: si osservano salti da n a $n-1$ con dei tempi di permanenza dati dal minimo di n variabili aleatorie esponenziali indipendenti tra loro (la transizione avviene appena il cliente che impega meno tempo tra gli n in servizo lascia la coda).

Esercizio: il minimo di n variabili esponenziali indipendenti, tutte di parametro μ, ha densità esponenziale di parametro $n\mu$.

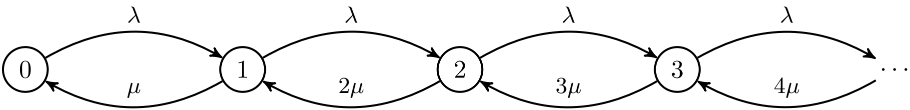
Pertanto, si avrà (se $n \geq 1$)

$$L_{n \to n-1} = n\mu.$$

Nel caso in cui vi siano arrivi con un processo di Poisson di intensità λ, poniamo, per $n \geq 0$,

$$L_{n \to n+1} = \lambda,$$

e di conseguenza



Distribuzione invariante

Come nel caso $M/M/1$, ogni stato è ricorrente.

- La competizione tra il tasso di arrivo λ e di uscita μ, è "smorzata" dal fatto che per n abbastanza grande si ha comunque $\lambda < n\mu$.

Distribuzione invariante

Come nel caso $M/M/1$, ogni stato è ricorrente.

- La competizione tra il tasso di arrivo λ e di uscita μ, è "smorzata" dal fatto che per n abbastanza grande si ha comunque $\lambda < n\mu$.
- Infatti una distribuzione invariante esiste sempre: risolvendo il sistema $\mu L = 0$ si trova una densità di Poisson di parametro λ/μ:

$$\mu_n \propto \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n .$$

Stima dei parametri dalle osservazioni

Come nel caso $M/M/1$, per stimare (λ, μ) sulla base dell'osservazione di un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_\ell)$ con i tempi di permanenza t_1, t_2, $\ldots$, t_ℓ scriviamo la verosimiglianza:

$$L(\lambda, \mu; X = \gamma) \propto \exp\left(-\lambda T - \mu T_\gamma\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

dove $T = \sum_{k=1}^{\ell} t_i$, γ_+ e γ_- sono come nel caso $M/M/1$.

Il termine nuovo è

$$T_\gamma = \sum_{k=1}^{\ell} t_i n_i,$$

(il tempo totale trascorso da tutti i clienti osservati nella coda). La stima di massima verosimiglianza è è

$$\lambda_{MLE} = \frac{\gamma_-}{T}, \quad \mu_{MLE} = \frac{\gamma_-}{T_\gamma}.$$

Il caso $M/M/c$

Il caso $M/M/c$ con $2 \leq c < \infty$ è intermedio tra i gli estremi che abbiamo considerato.

- Una distribuzione invariante esiste se e solo se $\lambda < c\mu$, ma le formule sono meno eleganti.

Section 2

Processi a stati continui

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

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

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,

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

- la **funzione di autocovarianza** del processo

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

- Notiamo che la varianza è

$$C(s,s) = \operatorname{Cov}(X_s, X_s) = \operatorname{Var}(X_s)$$

Funzione di autocorrelazione

A volte si considera anche la funzione

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

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

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

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

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

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 $i, j \in \{1, \ldots, d\}$, 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

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,

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

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,

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

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

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

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

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

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

  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.

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).

Section 3

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

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

- $(W_t)_{t\in\mathcal{T}}$ 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 σ^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:

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

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

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

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

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

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

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

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

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

- 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

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

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

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

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

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

- 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

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

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

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

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:

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

- La varianza vale

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

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

Funzione di media e di autocovarianza

La media di ciascuna S_t è nulla:

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

- La varianza vale

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

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

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

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.

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

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

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,

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 $t \in \mathcal{T}$, $t \geq 1$,

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

- 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:

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

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:

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

e quindi, ripetendo t volte,

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

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} è indipendente da W_t,

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

Condizione di stazionarietà

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

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

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

$$|\alpha| < 1.$$

Funzione di autocovarianza

Funzione di autocorrelazione

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

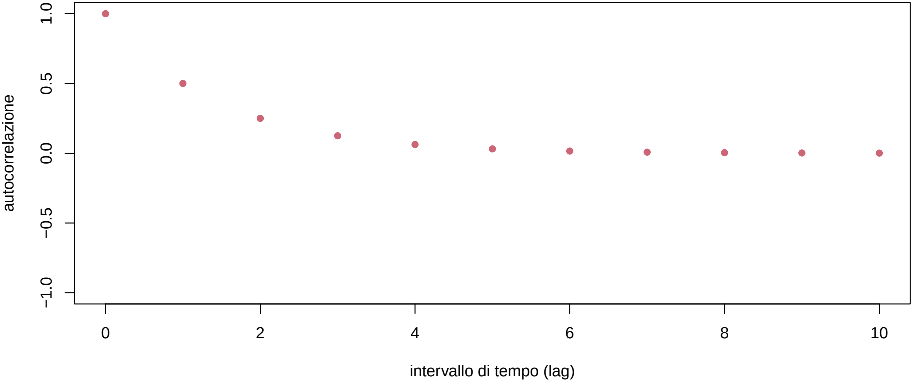


Figure 2: funzione di autocorrelazione dell'equazione lineare con smorzamento per $\alpha = 1/2$

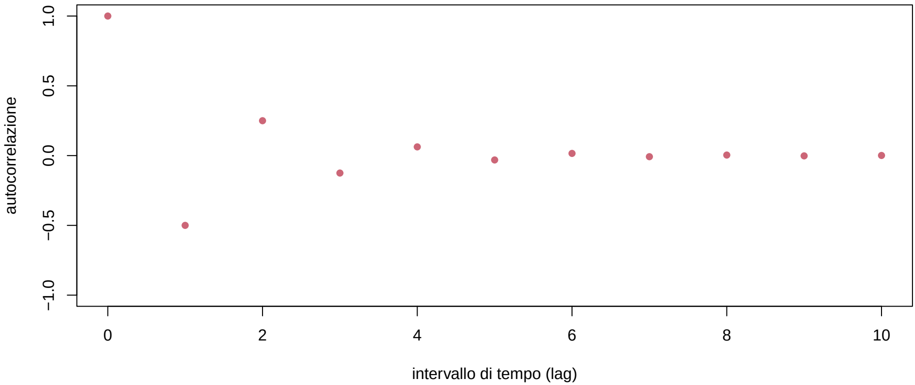


Figure 3: funzione di autocorrelazione dell'equazione lineare con smorzamento per $\alpha = -1/2$

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

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

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"

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

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