

Analisi non lineare di serie temporali

Ludovico Iannello e Federico Fattorini

8 settembre 2023

1 Introduzione

Lo scopo di questo lavoro è quello di esplorare varie misure e metodologie per l'analisi di serie temporali. In particolare, si pone attenzione sul metodo del *Delay Reconstruction*, che consiste nel trasformare una singola serie scalare in una serie di vettori (embedding), da cui è poi possibile estrarre informazioni sulla dinamica del sistema, se presente. Parte del lavoro sarà dedicata ad illustrare metodi per trovare i giusti parametri per l'embedding, facendo uso della mutua informazione e della statistica dei False Nearest Neighbours. Successivamente, viene calcolato il massimo esponente di Lyapunov e la dimensione di correlazione della traiettoria ricostruita, utilizzando l'algoritmo proposto da "Rosenstein et al." [4].

La tecnica del Delay Reconstruction è stata effettuata su una serie prodotta da un sistema dinamico, studiato per la prima volta da Rössler. Infine, la tecnica per il calcolo dell'esponente di Lyapunov e della dimensione di correlazione è stata provata sia in assenza che in presenza di rumore sovrapposto al segnale.

2 Delay reconstruction

Data una serie di misure scalari x_n , una delay reconstruction in m dimensioni si definisce come l'insieme dei vettori \vec{x}_n :

$$\vec{x}_n = (x_{n-(m-1)\tau}, x_{n-(m-2)\tau}, \dots, x_{n-\tau}, x_n)$$

dove τ è la differenza tra gli indici dei sample, che può essere espressa in unità temporali moltiplicando per il *sample time* Δt . τ si definisce come il *lag* o il *delay time*, mentre m rappresenta la *dimensione di embedding*.

Di conseguenza, data una certa serie temporale composta da N misure, il particolare embedding è espresso dalla seguente matrice $M \times m$ (con $M = N - (m - 1)\tau$):

$$\begin{pmatrix} x_1 & x_{1+\tau} & x_{1+2\tau} & \dots & x_{1+(m-2)\tau} & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & x_{2+2\tau} & \dots & x_{2+(m-2)\tau} & x_{2+(m-1)\tau} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_M & x_{M+\tau} & x_{M+2\tau} & \dots & x_{M+(m-2)\tau} & x_{M+(m-1)\tau} \end{pmatrix}$$

L'idea alla base della delay reconstruction è che in generale, dato un sistema n -dimensionale, saranno necessarie n variabili indipendenti per descriverlo [3]. Tali variabili non sono univocamente determinate, ma si richiede che siano diffeomorficamente equivalenti, cioè che rappresentino in modo continuo ed univoco gli stati dell'attrattore. Tale proprietà risulta essere valida per un

generico vettore di delay (con la giusta dimensione) in virtù del teorema di Takens, il quale asserisce che una mappa di delay di dimensione $m = 2D + 1$ è un embedding di un manifold compatto A con dimensione D se la funzione di misura $s : A \rightarrow \mathcal{R}$ è C^2 e accoppia tutti i gradi di libertà del sistema. In tal senso, un qualsiasi attrattore può essere quindi descritto dal set di variabili $x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}$, a patto che m sia maggiore o uguale di $2D + 1$, dove D è la dimensione dell'attrattore.

3 Scelta del delay temporale

Ci sono diversi approcci per la determinazione del delay temporale (τ). Il più intuitivo è quello di calcolare la funzione di autocorrelazione (ACF) della serie temporale, data dalla seguente equazione:

$$\mathcal{C}(\tau) = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \langle x \rangle)(x_{i+\tau} - \langle x \rangle) \quad (1)$$

dove $\langle x \rangle$ rappresenta la media aritmetica. La scelta di τ è determinata dalla durata dopo la quale x_n e $x_{n+\tau}$ diventano scorrelati. In linea di principio la funzione di autocorrelazione dovrebbe essere zero per tutti i delay uguali o maggiori di τ , tuttavia soddisfare questa condizione porterebbe al limite di τ molto grande, con il conseguente problema di dover scartare un alto numero di misure. Di conseguenza, una regola pratica ragionevole è quella di scegliere il tempo in cui l'autocorrelazione decade di $1/e$.

Un'obiezione a questa procedura è che essa si basa su statistiche lineari, senza tenere conto delle correlazioni non lineari. Pertanto un metodo migliore per determinare il τ ottimale è utilizzando la *time delayed mutual information* ($I(\tau)$). La mutua informazione è una misura standard che indica quante informazioni si possono ottenere con la misurazione di una serie temporale, ad esempio x_n , che è influenzata da un'altra misurazione effettuata in un altro momento $x_{n+\tau}$, campionata dopo un intervallo di tempo τ . In altre parole è una misura della dipendenza reciproca di due misure di una serie temporale.

Nella pratica, sull'intervallo esplorato dalla serie, si crea un istogramma per la distribuzione di probabilità dei dati. Indicando con p_i la probabilità che il segnale assuma un valore all'interno dell' i -esimo bin dell'istogramma e con $p_{ij}(\tau)$ la probabilità che x_n sia nel bin i e $x_{n+\tau}$ sia nel bin j , $I(\tau)$ è data da:

$$I(\tau) = \sum_{ij} p_{ij}(\tau) \log(p_{ij}(\tau)) - 2 \sum_i p_i \log(p_i) \quad (2)$$

Il primo minimo di $I(\tau)$ rappresenta il delay al quale si ha la minore dipendenza tra x_n e $x_{n+\tau}$; ovvero, al quale $x_{n+\tau}$ aggiunge la massima informazione alla conoscenza che si ha da x_n .

4 Metodo dei falsi vicini più vicini

Il calcolo della dimensione di embedding (m) ottimale è fatto utilizzando la statistica dei *False Nearest Neighbours* (FNN)[2]. Questo metodo si basa sul presupposto che due punti che si trovano in prossimità l'uno con l'altro con l'appropriata dimensione di embedding devono rimanere vicini considerando dimensioni maggiori. Tuttavia, se m è abbastanza piccola, punti che in realtà sono distanti potrebbero sembrare vicini a causa di un'eccessiva riduzione della dimensionalità; tali punti sono noti come *False Nearest Neighbours*.

La statistica da studiare è la seguente: per ogni vettore di delay ottenuto dalla serie temporale, si considera il primo vicino in m dimensioni e si calcola il rapporto delle distanze euclidee tra questi due punti in $m + 1$ dimensioni e in m dimensioni. Se questo rapporto è maggiore di una soglia r , i due vettori si considerano falsi vicini. Inoltre, indicando con σ la deviazione standard dei punti della serie, le coppie la cui distanza iniziale è già maggiore di $\frac{\sigma}{r}$ per definizione non possono essere falsi vicini, poiché, in media, le variazioni dei dati non sono grandi a sufficienza per garantire che il rapporto le due distanze in dimensioni consecutive sia maggiore di quello di soglia r . Se con $k(n)$ si indica l'indice del primo vicino dell' n -esimo punto, si ha dunque la seguente statistica:

$$X_{fnn}(r) = \frac{\sum_{n=1}^{N-m-1} \Theta\left(\frac{\|\vec{x}_n^{(m+1)} - \vec{x}_{k(n)}^{(m+1)}\|}{\|\vec{x}_n^{(m)} - \vec{x}_{k(n)}^{(m)}\|} - r\right) \Theta\left(\frac{\sigma}{r} - \|\vec{x}_n^{(m)} - \vec{x}_{k(n)}^{(m)}\|\right)}{\sum_{n=1}^{N-m-1} \Theta\left(\frac{\sigma}{r} - \|\vec{x}_n^{(m)} - \vec{x}_{k(n)}^{(m)}\|\right)} \quad (3)$$

Viene dunque studiato il rapporto dei False Nearest Neighbours X_{fnn} in funzione di r per varie dimensioni di embedding optando per quella che annulla il rapporto da una certa soglia in poi, in modo che la delay reconstruction non consideri un numero rilevante di falsi vicini. Il range su cui far variare le soglie deve essere ampio abbastanza da consentire una divergenza esponenziale dovuta al caos deterministico. Tuttavia i risultati in generale dipenderanno dalla scelta del delay τ , quindi è conveniente prima determinare il delay migliore per poi applicare il metodo dei False Nearest Neighbours.

5 Calcolo del maggiore esponente di Lyapunov

Esistono vari metodi per il calcolo del più grande esponente di Lyapunov di un sistema dinamico a partire da una serie temporale; in questo lavoro, si utilizza il metodo illustrato da Rosenstein [4]. Inizialmente vengono costruiti i vari vettori di delay a partire dalla serie temporale, i quali costituiranno i nuovi stati dello spazio delle fasi. Per ciascuno di essi (indicato con \vec{x}_j), viene trovato il primo vicino tra gli altri vettori di delay $\vec{x}_{\hat{j}}$:

$$d_j(0) = \min_{\vec{x}_{\hat{j}}} \|\vec{x}_j - \vec{x}_{\hat{j}}\| = \|\vec{x}_j - \vec{x}_{k(j)}\|$$

dove $d_j(0)$ è distanza iniziale del vettore \vec{x}_j dal suo primo vicino, e $\|\cdot\|$ è la norma euclidea. Inoltre, come ulteriore condizione, si impone che i primi vicini debbano avere una separazione temporale maggiore del periodo medio della serie, in modo da non considerare valori "consecutivi" della serie come primi vicini.

Si assuma adesso che i primi vicini si allontanino con un rate di espansione dato dal massimo esponente di Lyapunov λ_{max} :

$$d_j(i) \simeq d_j(0) \exp \lambda_{max}(i\Delta t)$$

e, prendendo il logaritmo:

$$\ln d_j(i) \simeq \ln d_j(0) + \lambda_{max}(i\Delta t)$$

Infine mediando su tutti i \vec{x}_j si ottiene un'unica retta:

$$\langle \ln d_j(i) \rangle = \langle \ln d_j(0) \rangle + \lambda_{max} i \Delta t$$

Eseguendo un fit dei minimi quadrati, è possibile ricavare quindi il valore del massimo esponente di Lyapunov a partire dal coefficiente angolare.

6 Dimensione di correlazione

Gli attrattori strani dei sistemi dinamici caotici sono caratterizzati da una dimensionalità frattale o dimensione di Hausdorff calcolata comunemente con gli algoritmi di Box Counting. Grassberg e Procaccia in [1] presentano una misura per la "stranezza" degli attrattori che è strettamente correlata alla dimensione frattale. La misura si ottiene considerando le correlazioni tra punti di una serie sull'attrattore, nel caso in analisi si considerano i vettori di delay ottenuti dalla serie temporale. Data la dimensione di embedding m , la somma di correlazione è definita come:

$$\mathcal{C}_m(r) = \frac{2}{M(M-1)} \sum_{i \neq k} \Theta[r - \|\vec{x}_i - \vec{x}_k\|] \quad (4)$$

$\mathcal{C}_m(r)$ è la frazione di coppie di punti che sono separate da una distanza minore ad r . Per un sistema dinamico che presenta comportamento caotico vale $\mathcal{C}_m(r) \propto r^D$ e l'esponente D è correlato alla dimensione di Hausdorff dell'attrattore. In particolare vale $D \leq D_H$, anche se sperimentalmente il valore di D può essere considerato come una stima della dimensione frattale.

7 Attrattore di Rössler

Per quanto riguarda i dati presi in considerazione, si è scelto di analizzare una singola serie temporale ottenuta da un sample della variabile x ricavata integrando le equazioni della dinamica del sistema di Rössler. Le equazioni di seguito riportate sono state integrate per un tempo $T = 200$ s utilizzando il metodo di Runge-Kutta all'ordine 4, con passo di integrazione $\Delta t = 0.001$ s.

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \quad (5)$$

Sono stati fissati i parametri $a = 0.15, b = 0.2, c = 10.0$. Per tali valori dei parametri il valore teorico del massimo esponente di Lyapunov è $\lambda_{max} = 0.090$ [4].

Di seguito, in figura 1 viene riportato il plot tridimensionale della soluzione ottenuta considerando come condizione iniziale $(x_0, y_0, z_0) = (1, 1, 1)$.

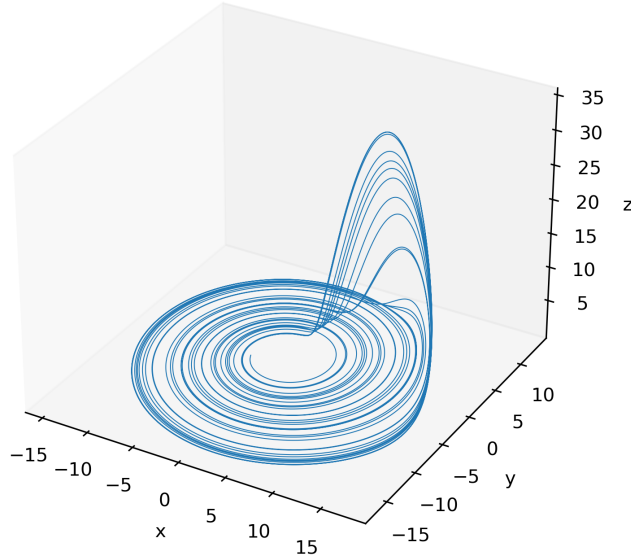


Figura 1: Attrattore di Rössler

8 Spazio delle fasi ricostruito

Al fine di dimostrare intuitivamente come i vettori di delay possano fornire una buona descrizione dello spazio delle fasi, è utile osservare il piano $x_t - x_{t+\tau}$, che corrisponde a rappresentare le prime due colonne della matrice di embedding, una in funzione dell'altra. La serie temporale che verrà analizzata in tutte le sezioni seguenti è stata ottenuta eseguendo una misura ogni 100 ($\Delta t_{sample} = 0.1$ s) a partire dalla serie ottenuta integrando il sistema di Rössler, per un totale di $N = 2000$ misure. Si applica dunque il metodo della ricostruzione dello spazio delle fasi ed il risultato è mostrato in figura 2, da cui, confrontando con la proiezione della dinamica sul piano x-y, risulta evidente la presenza di strutture derivanti dall'attrattore originale.

Infine si noti come la ricostruzione dipenda fortemente dalla scelta del delay temporale. In figura 2 vengono riportati 3 plot per differenti valori di τ . Appare evidente che per valori di τ piccoli i punti risultino molto correlati e pertanto il grafico risulta schiacciato sulla bisettrice, mentre per valori di τ maggiori emerge la struttura derivante dall'attrattore.

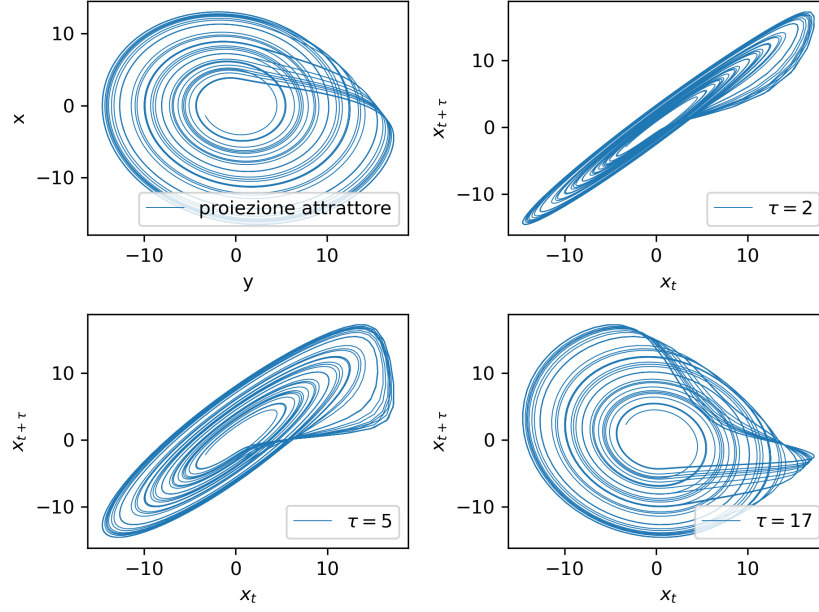


Figura 2: La prima figura a sinistra rappresenta la proiezione della dinamica nello spazio delle fasi sul piano x - y , le restanti figure rappresentano lo spazio delle fasi ricostruito da x_t e $x_{t+\tau}$ per differenti valori di τ .

Fissando $\tau = 17$ (valore determinato nella sezione 9.1 tale metodologia è stata poi applicata allo stesso sistema, ma in presenza di un rumore gaussiano sovrapposto ai dati. In figura 8 si rappresenta lo stesso piano x_t - $x_{t+\tau}$ al variare dell'intensità del rumore sovrapposto. Tale intensità è stata valutata tramite il *Signal to Noise Ratio* (SNR), definito come il rapporto tra la deviazione standard della serie e quella del solo rumore. In particolare, si mostrano i risultati per SNR= 1, 10, 100 e 1000, ovvero per rumore di intensità decrescente. Risulta evidente come la struttura dell'attrattore rimanga riconoscibile fino ad un SNR= 10, per poi venire completamente mascherata dal rumore per valori inferiori.

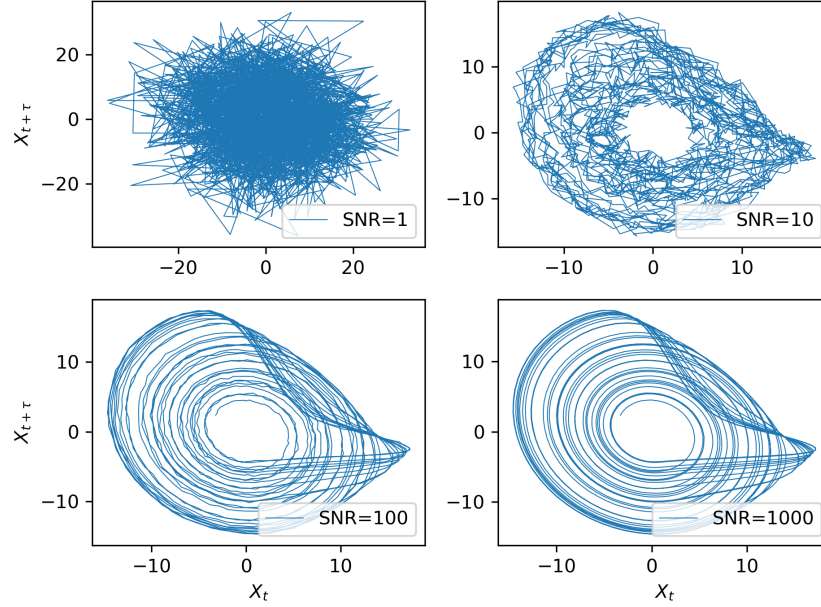


Figura 3: Spazio delle fasi ricostruito da x_t e $x_{t+\tau}$, per diversi valori SNR.

9 Determinazione dei migliori parametri di embedding

Non tutti i valori di τ e m garantiscono una corretta ricostruzione dello spazio delle fasi: in questa sezione, si mostrano quindi le varie problematiche derivanti da scelte non ottimali dei parametri di embedding e come i metodi descritti nelle sezioni 3 e 4, forniscano indicazioni per ottimizzare tali scelte.

9.1 Delay

Come visto dalla ricostruzione dello spazio delle fasi, la scelta del τ influisce notevolmente sul risultato dell'embedding. In particolare, τ troppo piccoli danno come risultato vettori di delay fortemente correlati. Per ovviare tale problema, è possibile fare ricorso alle misure di autocorrelazione e della mutua informazione.

In figura 4 si riporta quindi il grafico dell'autocorrelazione per la serie generata dal sistema di Rossler. Il valore del τ per cui l'autocorrelazione decresce di un fattore $\frac{1}{e}$ risulta essere $\tau = 11$, che fornisce una prima stima del giusto delay temporale.

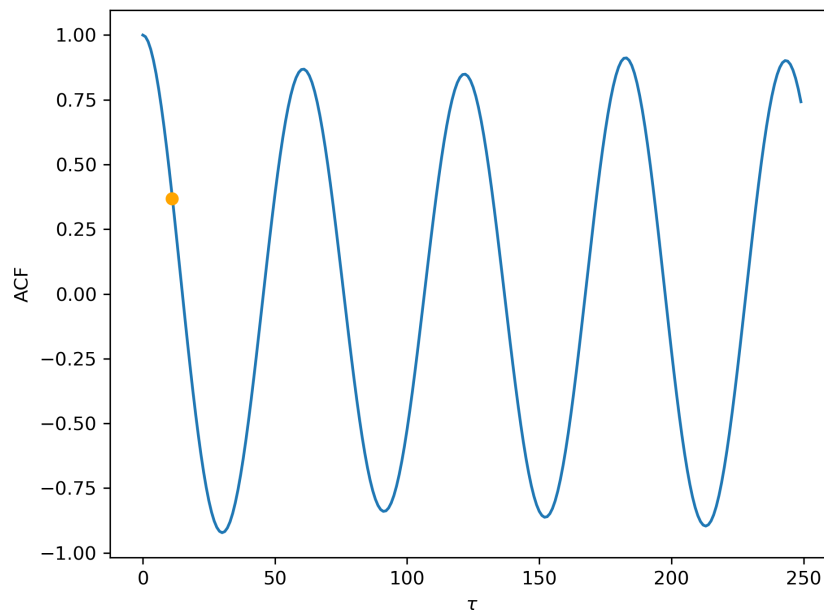


Figura 4: Grafico dell'autocorrelazione in funzione di τ ; si evidenzia anche il valore per cui l'autocorrelazione decresce di un fattore $\frac{1}{e}$.

Infine in figura 5 si riporta il grafico per la mutua informazione. Il primo minimo risulta essere $\tau = 17$.

Per le misure successive verrà utilizzato come τ ottimale quello ottenuto dalla mutua informazione, in quanto garantisce scorrelazione anche per ordini superiori a quello lineare.

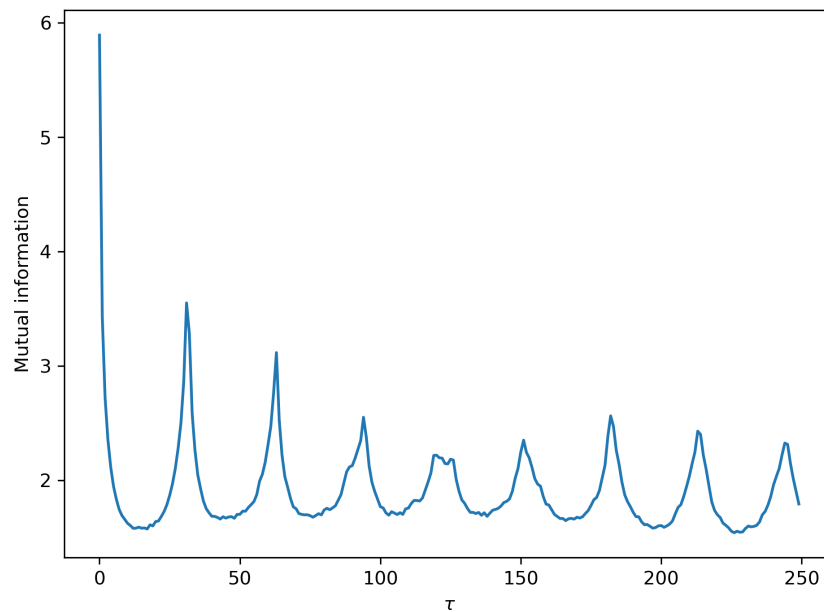


Figura 5: Grafico della mutua informazione in funzione di τ .

9.2 False Nearest Neighbours

Una volta determinato il miglior valore per il delay τ si determina la migliore dimensione di embedding m per mezzo del metodo dei False Nearest Neighbours al variare della dimensione e fissando $\tau = 17$. In figura 6 viene riportato il grafico del rapporto dei false nearest neighbours X_{fnn} in funzione di r . Appare evidente come per tutti gli $m \geq 3$ il rapporto vada a 0 per tutti gli r maggiori di una certa soglia, pertanto si può concludere che i valori della dimensione di embedding ottimali sono tutte le dimensioni $m \geq 3$.

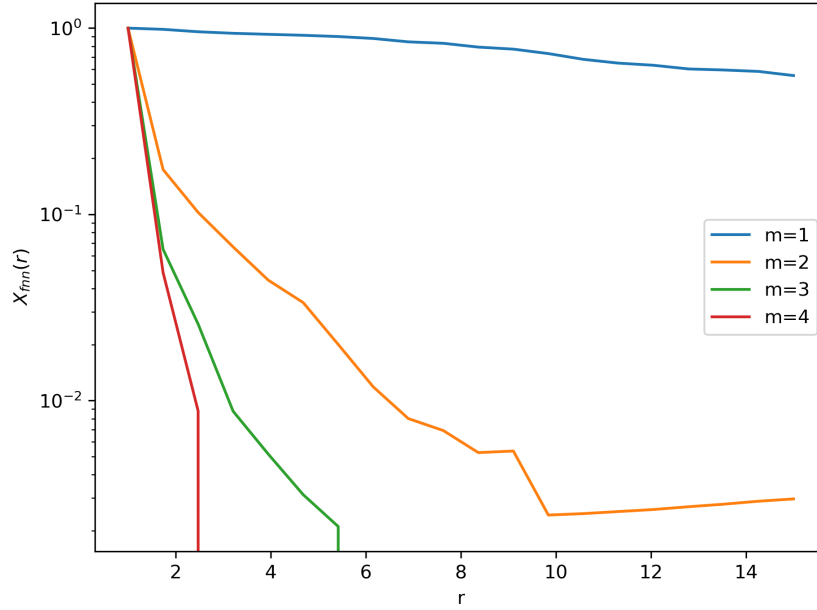


Figura 6: Grafico del rapporto X_{fnn} in scala logaritmica in funzione della distanza r ; le varie curve rappresentano il rapporto al variare della dimensione di embedding m

10 Calcolo dell'esponente di Lyapunov

10.1 Determinazione del periodo della serie

Come descritto in 5, per l'algoritmo di Rosenstein è necessario considerare come primi vicini solo vettori di delay distanti temporalmente almeno un periodo della serie temporale. Ovviamente, per segnali non periodici tale misura non può essere calcolata in modo diretto, ma facendo uso della trasformata di Fourier.

Si procede quindi eseguendo una FFT della serie temporale, il cui modulo viene riportato in figura 7. Da tale misura è stato poi possibile calcolare la frequenza caratteristica, considerando come tale la frequenza media della densità spettrale di potenza, ovvero del modulo quadro della FFT; si ottiene quindi $\hat{f} \simeq 0.16$ Hz, da cui il periodo della serie risulta essere $T \simeq 6.25$ s. Di conseguenza, nell'esecuzione dell'algoritmo si considereranno primi vicini solo vettori vicini spazialmente ma distanti temporalmente più di 6.25 s gli uni dagli altri.

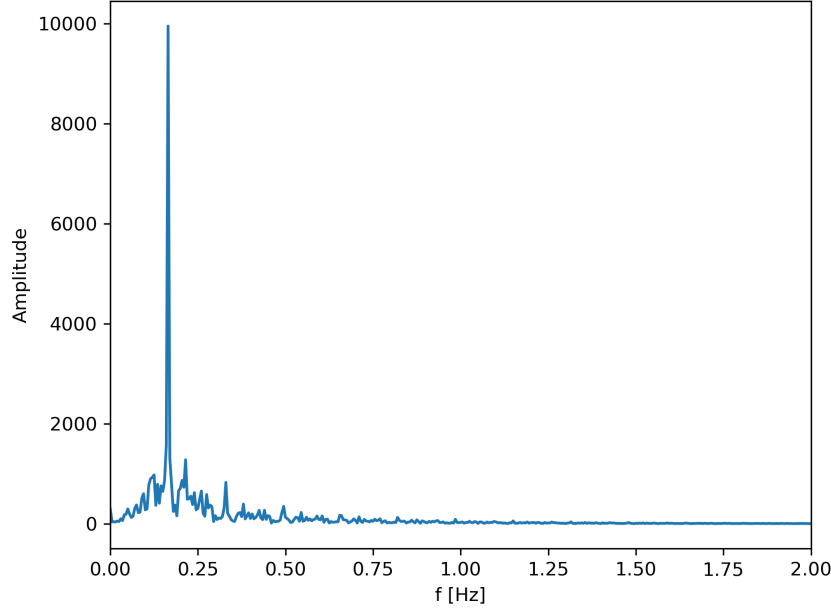


Figura 7: Modulo della FFT della serie temporale.

10.2 Fit per il calcolo dell'esponente di Lyapunov

Una volta ottenuto il periodo della serie, si hanno tutti i parametri necessari per usare la procedura descritta da Rosenstein. Si calcolano quindi le evoluzioni temporali per le distanze tra i primi vicini tra i vettori di delay, al variare della dimensione di embedding m e fissando $\tau = 17$. Una volta mediate le distanze su tutti i vettori di delay, si ottengono quindi i grafici per $d(t)$ al variare di m . Come si può notare dalla figura 8, per $m \geq 2$ l'andamento di $\langle \log d(t) \rangle$ è caratterizzato da un regime lineare seguito da un plateau. Si esegue quindi un fit con una retta, considerando solo il range lineare, ottenendo le stime per gli esponenti di Lyapunov riportate in figura. Si nota come tutti i valori ottenuti per $m > 3$ siano uguali tra loro e al valore teorico per l'esponente di Lyapunov dell'attrattore di Rössler, pari a $\lambda_{max} = 0.09$ (la cui corrispondente retta è stata riportata con una linea nera tratteggiata).

Per quanto riguarda le discrepanze per $m = 1$ e $m = 2$, queste sono da attribuirsi alla non corretta scelta della dimensione di embedding, come osservato dalla False Nearest Neighbours. In particolare, per $m = 1$ l'andamento della $d(t)$ è completamente diverso rispetto agli altri valori di m .

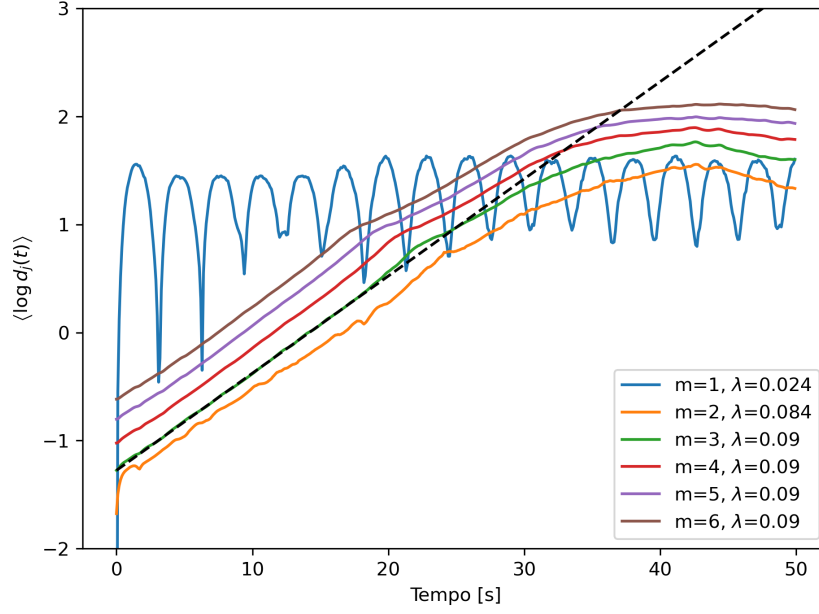


Figura 8: Grafici per $\langle \log d(t) \rangle$ al variare di m con i rispettivi esponenti di Lyapunov ottenuti tramite fit. Per confronto, si riporta la retta calcolata con il valore teorico di λ_{max}

10.3 Fit per il calcolo del l'esponente di Lyapunov in presenza di rumore

Si ripete la procedura descritta sopra sulla stessa serie temporale, ma in presenza di rumore gaussiano sovrapposto, per diversi valori del SNR. Si nota dalla figura 9 come il regime a plateau sia raggiunto per t minori con l'aumentare dell'intensità del rumore, fino al punto di coprire tutto l'intervallo temporale. Allo stesso tempo, si osserva una diminuzione della pendenza della retta nel regime lineare, che porta a valori per l'esponente di Lyapunov discordanti con quello teorico per valori bassi del SNR, per qualsiasi valore della dimensione di embedding.

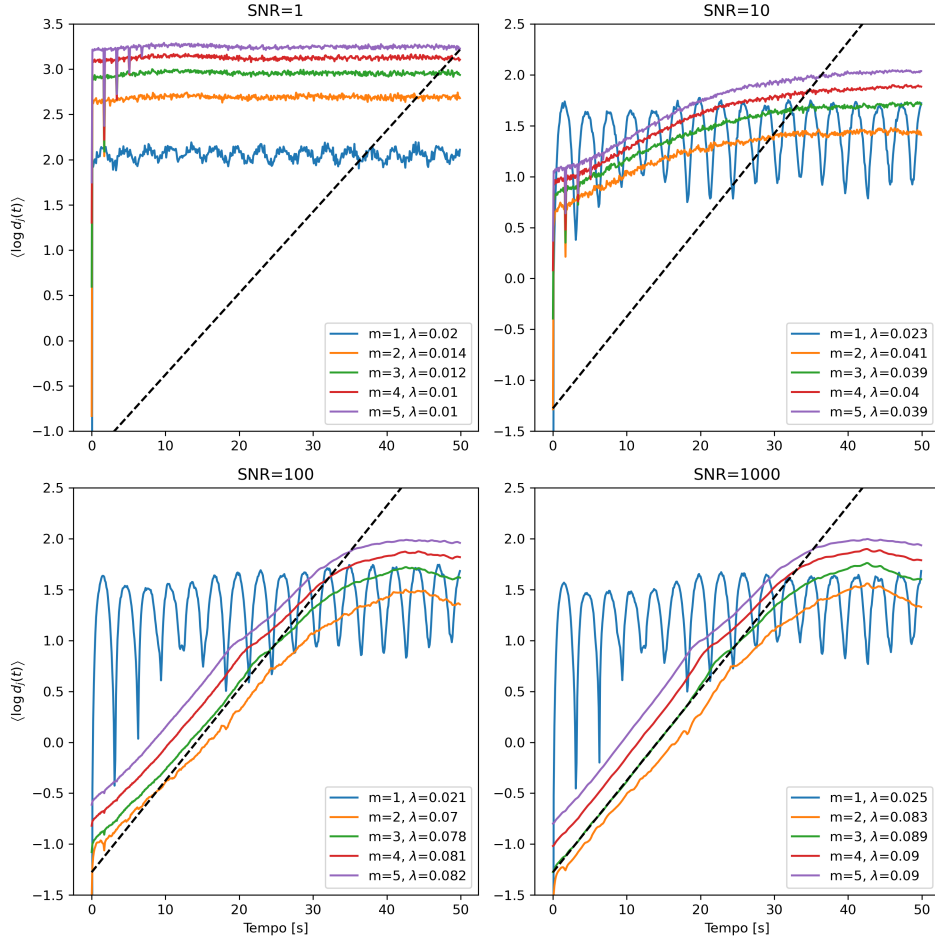


Figura 9: Grafici per $\langle \log d(t) \rangle$ al variare di m con 4 valori differenti di SNR, con i rispettivi esponenti di Lyapunov ottenuti tramite fit.

11 Fit per la dimensione di correlazione

Infine come ultima analisi sulla serie temporale è stata effettuata la misura descritta nella sezione 6, per la determinazione della dimensione di correlazione. In figura 10 viene riportato il grafico della somma di correlazione in funzione della distanza r . Come descritto ci si aspetta che per una serie caotica la relazione sia una legge di potenza: $\mathcal{C}(r) \propto r^D$; pertanto viene effettuato un fit lineare considerando i logaritmi delle due quantità. La funzione di fit utilizzata è dunque $y = D \cdot x + q$ e la dimensione di correlazione sarà proprio il coefficiente angolare della retta. La dimensione di Hausdorff per l'attrattore di Rossler è stimata essere $D_H \approx 2$, quindi effettivamente il valore ottenuto per la dimensione di correlazione risulta $D = 1.82 \leq D_H$.

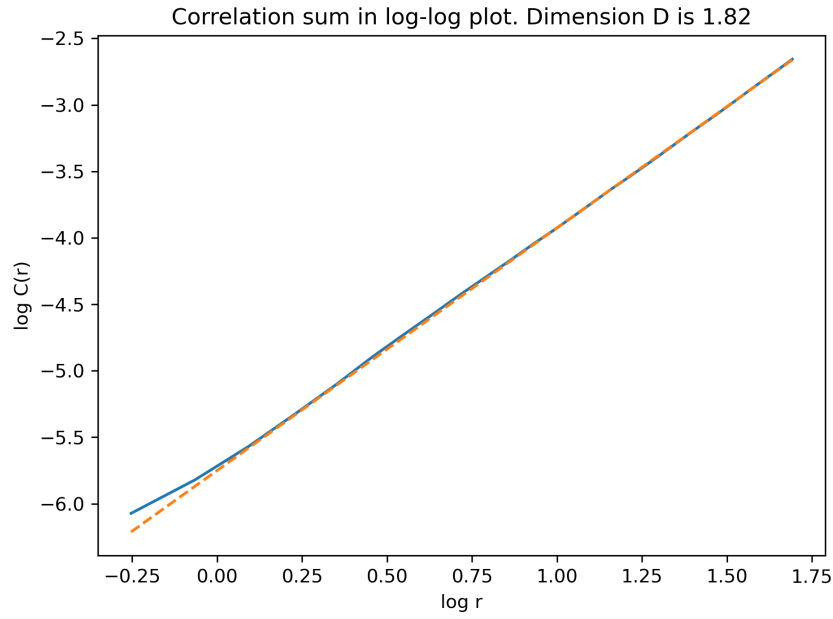


Figura 10: Grafico del logaritmo della somma di correlazione in funzione del logaritmo della distanza r in blu e grafico della retta fittata (in arancione).

Lo stesso metodo è stato applicato nel caso della serie con rumore sovrapposto con $SNR = 1$; il grafico ottenuto è riportato in figura 11. Come previsto dalla teoria, in presenza di rumore il grafico in logaritmica non presenta il comportamento lineare previsto invece per le serie generate da sistemi dinamici.

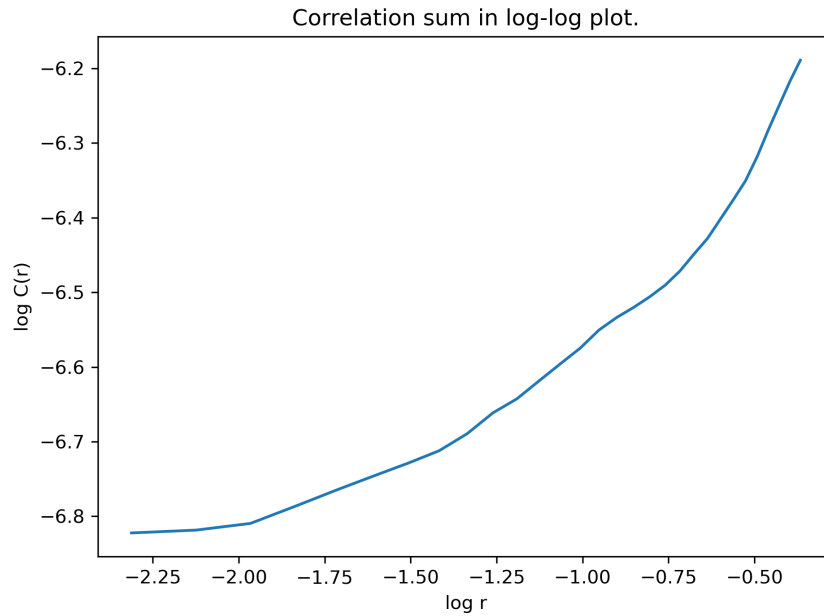


Figura 11: Grafico del logaritmo della somma di correlazione in funzione del logaritmo della distanza r nel caso di segnale rumoroso con $SNR = 1$

Riferimenti bibliografici

- [1] Peter Grassberger and Itamar Procaccia. Characterization of strange attractors. *Phys. Rev. Lett.*, 50:346–349, Jan 1983.
- [2] Holger Kantz and Thomas Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, 2 edition, 2003.
- [3] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45:712–716, Sep 1980.
- [4] Michael T. Rosenstein, James J. Collins, and Carlo J. De Luca. A practical method for calculating largest lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1):117–134, 1993.