

I segnali biologici sono caratterizzati da un'ampia variabilità, a volte presentano delle caratteristiche dinamiche (biforcazioni) non facilmente predicibili con gli studi classici fatti tramite la media e la varianza del segnale (un buon esempio è lo scompenso cardiaco).

Quindi introduciamo il concetto di serie temporale. (wikipedia: una serie storica (o temporale) si definisce come un insieme di variabili casuali ordinate rispetto al tempo, ed esprime la dinamica di un certo fenomeno nel tempo. Le serie storiche vengono studiate sia per interpretare un fenomeno, individuando componenti di trend, di ciclicità, di stagionalità e/o di accidentalità, sia per prevedere il suo andamento futuro.)

Quindi non avremo più un modello parametrico, bensì un segnale di cui non conosciamo gli ingressi, ma del quale ci interessa sapere: **la stima delle caratteristiche invarianti (cosa sono le "certezze del nostro sistema"), la ricostruzione dello stato, capire se esiste determinismo, trovare la dimensione frattale, trovare l'esponente di Lyapunov e ridurre il rumore.**

Qual è lo scopo di andare ad analizzare questo tipo di oggetti (serie temporali)? Voglio trovare delle caratteristiche statistiche e trovare un modo per simulare e predire il mio segnale.

Partiamo dal presupposto che il nostro sistema abbia N variabili di stato, tali che $\dot{x}(t) = f(x(t))$, non abbiamo conoscenze a priori del sistema. Misurare tutte le variabili non è possibile, in più la misura ha un costo. A volte addirittura non conosciamo il numero delle variabili di stato.

Come risolviamo?

Andiamo a campionare la variabile y in tempo $t_0, y(\tau), y(2\tau), \dots, y((N-1)\tau)$, questa analisi avrà un numero finito (N) di campioni, e sarà affetta da un rumore.

Lorentz ha scoperto che un sistema caotico può essere identificato come un attrattore, i sistemi attrattivi sono caratterizzati dalla dimensione frattale e dall'esponente di Lyapunov. Come si ricostruisce un sistema frattale? **Manè-Takens**

Definisco il vettore delle uscite ritardate $S = \begin{bmatrix} y(t) & y(t+\tau) & \dots & y(t+(n-1)\tau) \end{bmatrix}$ la ricostruzione dello stato si ottiene: $x(t) = O^{-1}S^T \rightarrow O = \begin{bmatrix} c^T & A^T c^T & \dots & (A^T)^{n-1} c^T \end{bmatrix}^T$

O = Observability matrix. N : è la dimensione di embedding dell'attrattore A : se la funzione G ha una corrispondenza uno a uno in A i.e. $x' \neq x''$, allora $G(x') \neq G(x'')$.

Embedding: La tecnica del misurare una serie temporale come una sequenza di punti in uno spazio n -dimensionale viene detta *time-lag embedding*. $(y(t), y(t-\tau), \dots, y(t-(m-1)\tau))$:
 $m = N =$ embedding dimension, τ è l'embedding lag.

Il segnale ricostruito ha un esponente di Lyapunov uguale a quello del sistema reale, infatti questo è un parametro invariante.

Una condizione necessaria è che $m > d$. Se m è un embedding nello spazio vettore delle coordinate ritardate non c'è sovrapposizione. (**Manè Takens**). Se supponiamo $m > 2d$ (sufficiente non necessario), quindi $d < m \leq 2d$. Negli esperimenti reali d non è nota a priori.

Dimensione m :

- FN: False Neighbors: la percentuale di punti vicini (distanza < soglia) è calcolata in m e $m+1$. Se due punti vicini sono in m ma non in $m+1$, e si ripete incrementando m finché non si trova una dimensione corretta.

Come si fa a trovare la **dimensione tau**: ci sono dei valori di tau che sono migliori degli altri. Ci sono dei metodi empirici che possono essere

$$d_{KY} = k + \frac{L_1 + L_2 + \dots + L_k}{|L_{k+1}|}$$

- Il primo zero dell'autocorrelazione
- Primo minimo della funzione di *Mutua Informazione*

Determinismo\Caos:

- **Test delle ipotesi** basato sui dati surrogati: Devo rigettare l'ipotesi nulla, cioè devo annullare l'ipotesi che il processo si generato da un processo stocastico stazionario. Noi vogliamo che la serie è inconsistente con l'ipotesi di linearità. **Dati surrogati**: Serie di dati random che condividono delle proprietà lineari con la serie originale. Per vedere se posso rigettare l'ipotesi mi baso sulla funzione di autocorrelazione o sulla mutua informazione

$$\gamma_{xx}(\tau) = E[x_n^* \cdot x_{n+\tau}]$$

$$I(\tau) = \sum_{n=1}^N P(x_n, x_{n+\tau}) \cdot \ln \frac{P(x_n, x_{n+\tau})}{P(x_n) \cdot P(x_{n+\tau})}$$

In particolare ci interessa conoscere la percentuale di variazione della mutua informazione

$$\bar{I}_{avr}(\tau) = \frac{1}{M} \sum_{i=1}^M I_{si}(\tau)$$

Se c'è una ampia variazione della mutua informazione, allora il processo sarà deterministico.

$$I_d(\tau) = I_{or}(\tau) - \bar{I}_{avr}(\tau)$$

Esempio: **HVR**: possiamo rifiutare l'hp nulla, quindi possiamo definirlo come un segnale casuale.

$$I_m = \frac{1}{T} \sum_{\tau=1}^T I_{or}(\tau)$$

$$\Delta I_{\%}(\tau) = \frac{I_d(\tau)}{I_m}$$

1. Dimensione frattale:

Dimensione d dell'attrattore! Abbiamo un problema, questa caratteristica non è nota a priori. Calcoliamo la

$$C_m(r) = \frac{\# \text{pairs}(y_m(i), y_m(j)) \text{ t.c. } \|(y_m(i) - y_m(j))\| < r}{\# \text{pairs}(y_m(i), y_m(j))}$$

funzione di correlazione

per vari valori di

r. Per r piccole vale la relazione $\log C(r) = d \log r + \alpha$, dà una stima di d.

2. Esponente di Lyapunov?

A cosa serve? L1 caratterizza la dinamica del sistema. È anche possibile calcolare l'intero spettro degli esponenti L1, L2, ... Lm.

Si consideri un vettore $S_m(i)$ molto vicino ad un altro vettore $S_m(j)$, così vicino che

$$\Delta(0) = |S_m(i+t) - S_m(j+t)| \text{ sia infinitesimale } \Delta(t) \sim \exp(Lt) \text{ da qua calcoliamo L.}$$

Kaplan e Yorke hanno stimato la dimensione frattale d come:

$$d_L = k + \frac{S_k}{|L_{k+1}|} \text{ where } k = \max\{m | S_m \geq 0\}$$

Dove k è il massimo intero tale che la somma dei k esponenti maggiori sia non negativa.

Rumore

Devo eliminare il rumore da una serie temporale, chiaramente non posso usare i filtri semplici che usavo nei metodi parametrici. La prima cosa che faccio è stimare i con il teorema di Takens. Si trova la **matrice della covarianza**: i k autovalori principali rappresentano la potenza del segnale utile. → Metodo di Schreiber

Applicazioni: Posso notare che per quanto riguarda l'analisi HRV, i segnali di un cuore trapiantato seguono delle caratteristiche molto meno caotiche rispetto a quelle di un cuore trapiantato.

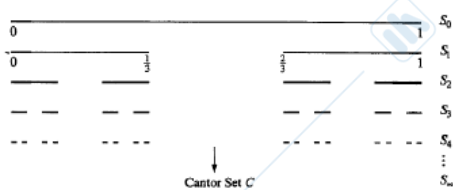
Frattalità:

-Concetto di self-similarity, un processo può avere delle caratteristiche che si ripetono su scale diverse, è come se il suo comportamento fosse in qualche modo coerente con le varie scale di se stesso. La dimensione di questo tipo di sistemi non è intera bensì frazionaria = sistemi frattali.

Possiamo pensare concretamente che esistono delle dimensioni che tutti conosciamo, come la dimensione = 0, quella del punto, 2 della retta e 3 del cubetto.

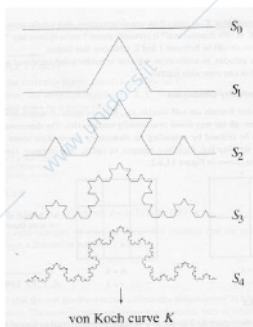
Esistono quindi anche dei set frattali famosi.

- Il primo è quello di **Cantor**



$C = 1/3$, il sistema viene man mano diviso in $1/3$ di se stesso

La lunghezza del set S è 0, infatti al crescere di k , quindi del numero di divisione, i puntini saranno sempre più piccoli e separati quindi arriveranno a scomparire. È facile immaginare che il frattale di Cantor abbia quindi una dimensione compresa tra 0 e 1.



Un set è topologicamente un set di Cantor se: S è totalmente disconnesso, S non contiene punti isolati = tutti i punti hanno nella loro vicinanza un altro punto generato da S .

- Il secondo è quello di **Von Kock:**

Lunghezza infinita. $S \rightarrow \infty$ Per $L \rightarrow \infty$. Infatti $L(k+1) = 4/3 L(k)$. K è compresa tra 1 e 2.

Come si calcola la dimensione frattale?

$$D_H = \sup \{ d \in \mathbb{R}_+ : m_d(X) = \infty \}$$

$$D_C = \lim_{\rho \rightarrow 0} \frac{\log(N(\rho))}{\log(1/\rho)}$$

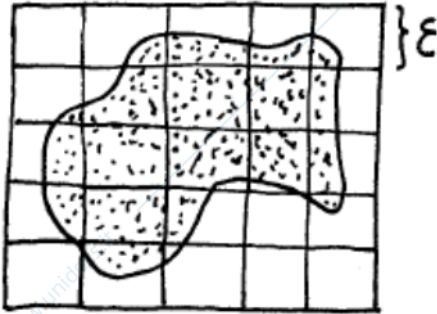
N sfere di diametro ρ . Si può dimostrare che $D_H \leq D_C$

In un frattale la dimensione di Hausdorff dimension è strettamente più alta della dimensione topologica $D_H > D_T$

Hausdorff dimension. Per ogni set di X valori reali esiste $D(H)$ tale che $d > D(H)$ e uguale a 0 per $d > D(H)$

- d_B : box counting dimension
- d_C : Correlation dimension
- d_L : Lyapunov dimension

1- d_B : box counting dimension



S appartiene a \mathbb{R}^n , contenuto in un cubo H con dimensione n , che dividiamo in H piccoli pezzettini di dimensione ε .

$T(\varepsilon)$ è proporzionale a $(1/\varepsilon)^n$. $N(\varepsilon)$ è il numero di pezzetti che contengono almeno un punto di S . $N(\varepsilon)$ segue la seguente legge:

$$N(\varepsilon) = \gamma \left(\frac{1}{\varepsilon} \right)^{d_B} \quad \text{that is} \quad \log N(\varepsilon) = \log \gamma + d_B \log(1/\varepsilon)$$

$$d_B = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}$$

2- d_C : Correlation dimension

$$C(r) = \lim_{t \rightarrow \infty} \frac{\#\text{coppie}(x(i), x(j)) \text{ t.c. } \|x(i) - x(j)\| < r}{\#\text{coppie}(x(i), x(j))}$$

$$d_C = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}$$

Vedi prima

3 - d_L : Lyapunov dimension

Traiettoria per cui $L_1 \geq L_2 \geq \dots \geq L_n$.

Per ogni $m \leq n$, $\exp(S_m) = \exp(L_1 + L_2 + \dots + L_m)$, è la del tasso di espansione ($S_k > 0$), o di contrazione ($S_k < 0$)
DL esisterà se si calcola come:

$$d_L = k + \frac{S_k}{|L_{k+1}|} \quad \text{where } k = \max\{m \mid S_m \geq 0\}$$

→ Per quanto riguarda i sistemi caotici, spesso ma non sempre i sistemi caotici hanno dimensione frattale.

→ **Parametric diagram:** Fissata una valore c appartenente ad M (a point belonging to the Mandelbrot set) there is a set B of starting times $z(0)$ which generates limited trajectories (certainly B is nonempty, since contains at least $z(0) = 0$.)

Cosa possiamo ottenere da una serie temporale?

→ Una delle caratteristiche è la self-similarità, un processo stocastico si definisce self-similar, con un esponente di self-similarity $H > 0$ se:

- Per ogni stretching factor $\lambda > 0$, il processo scalato $\lambda^{-H} X_{\lambda t}$, soddisfa questa condizione $\{\lambda^{-H} X_{\lambda t}\} \stackrel{d}{=} \{X_t\}$.

Per un processo di self-similarity i sono 3 caratteristiche fondamentali:

- $\rho_k = ck^{-\lambda}$, $\lambda = 2H-1$
- Power spectrum follows a power law: $1/f^\beta$, $\beta > 0$, $\beta = 2H+1$
- Variance tends to zero fastly
- $\rho_k^{(m)} = \rho_k$

Un processo si definisce self similar se $0 < H < 1$. In particolare

$H=0$: rumore bianco antipersistenza

$H=0.5$: moto browniano random

$H=1$ segnale differenziabile, persistenza

FBm= fractional Brownian moto = $B_H(t)$

FGn=fractional Gaussian noise $Y_H(t) = B_H(t) - B_H(t-1)$.

Come faccio a calcolare H?

1. Detrended fluctuation analysis (DFA)

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2}$$

- $y_n(k)$ is the regression line
- n is the length of the observation interval
- in case of self-similarity

$$F(n) \sim n^\gamma$$

$\gamma = 1$ is typical of series with **1/f power-law** in power spectrum

$\gamma = 1.5$ characterizes the ordinary diffusion (Brownian motion)

$\gamma = 0.5$ white noise

$\gamma \geq 1$ indicates a correlation different from the power-law

Considero degli slot temporali e faccio una sorta di interpolazione (cerco la linea di regressione), e la sottraggo con il valore reale del segnale. Ricorda che γ e H sono la stessa cosa!

Esempio: camminata \rightarrow nella camminata si nota che se si invecchia si perdono le caratteristiche frattali. Infatti in un soggetto giovane H è quasi ad 1, mentre in un soggetto vecchio la H è a 0.5.

In un soggetto normale la DFA nella camminata normale è $H=1.10$, mentre in un soggetto a cui si da un pacing $\alpha=0.32-0.08$

HR: in un soggetto normale H è circa ad 1, mentre in un soggetto in coma $H=1.55$, in un soggetto con scompenso $H=1.26$, in un soggetto con infarto $H=0.99$.

2. Rescaled range analysis: