PRESENTAZIONE ZANETTI

Punto 1

Si sono presi in considerazione i terremoti avvenuti nel sud della California avvenuti tra il 1 Gennaio 1982 e il 2011.

Il dataset preso in considerazione contiene 7 colonne:

* La prima contenente l’indice del terremoto.
* La seconda contiene il trigger che corrisponde all’indice dell’evento che ha provocato il terremoto preso in considerazione.
* La terza il tempo in secondi a cui è avvenuto il terremoto a partire dal 1 Gennaio 1982.
* La quarta è la magnitudo del terremoto in una scala tra [0,8]. (La **magnitudo** è una misura indiretta dell'[energia meccanica](https://it.wikipedia.org/wiki/Energia_meccanica) sprigionata da un [evento sismico](https://it.wikipedia.org/wiki/Terremoto) all'[ipocentro](https://it.wikipedia.org/wiki/Ipocentro), basandosi sull'ampiezza delle [onde sismiche](https://it.wikipedia.org/wiki/Onda_sismica) registrate dai [sismografi](https://it.wikipedia.org/wiki/Sismografo) in superficie. )
* La quinta, sesta e settima rappresentano le coordinate cartesiane dell’ipocentro del terremoto espresse in metri.

I terremoti sono stati riordinati rispetto al tempo, in quanto si è notato che per un caso l’indice del terremoto non seguiva l’ordine temporale.

Al fine di visualizzare graficamente i terremoti, si sono convertite le coordinate dell’ipocentro da cartesiane (X,Y,Z) a coordinate geografiche (longitude,latitude,depth). La conversione è stata effettuata usando il pacchetto PyProj. In particolare, la superficie terrestre viene approssimata come un ellissoide con specifiche date dal sistema WGS84 (vedi figura).

La posizione rispetto all’ellissoide è data dalla

* Longitudine che misura l’angolo rotazionale tra il meridiano zero e il punto misurato il cui valore è compreso tra (-180°,180°).
* Latitudine che misura quanto vicini si è ai poli o all’equatore lungo un meridiano, il cui valore è compreso tra (-90°,90°) dove 0° è l’equatore.
* L’altezza ellissoidale che misura l’altezza del punto sopra o sotto l’ellissoide di riferimento lungo la sua normale.

(vedi figura)

Le equazioni che permettono la conversione da coordinate geografiche a cartesiane sono riportare in formula. In particolare, l’asse X punta dal centro della terra verso il punto con longitudine 0 a latitudine 0, l’asse Y punto dal centro della terra verso il punto con longitudine 90 a latitudine 0 a l’asse Z punta verso il polo Nord.

Punto 2

Si è deciso di visualizzare i terremoti nello spazio e nel tempo tramite un’animazione tramite il pacchetto Animation di matplotlib.

Per prima cosa ad ogni terremoto è stato associato il mese e l’anno a cui questo è avvenuto e questi sono stati filtrati per una magnitudine maggiore o uguale di un valore fissato, in modo da alleggerire il costo computazionale e facilitare la visualizzazione.

In ogni frame dell’animazione sono stati rappresentati tutti i terremoti avvenuti in un fissato mese ed anno. Si è creata una mappa bidimensionale sulla quale i terremoti sono visualizzati con uno scatter plot. La posizione dei terremoti è stata rappresentata con la rispettiva latitudine e longitudine, mentre la profondità è stata rappresentata tramite una scala di colori, che va dal bianco (il meno profondo) al nero (il più profondo). Si è infine rappresentata la magnitudo variando il diametro del disco che localizza il terremoto, dove a maggiori diametri corrisponde una magnitudo maggiore.

Inserire tree plot.

Punto

La sismicità è un complesso fenomeno spazio-temporale che obbedisce a semplici e generali regole che governano la statistica della loro frequenza. In particolare, ci si aspetta che la distribuzione temporale e spaziale dei terremoti segua una power law distribution, ovvero che la frequenza di una variabile t sia proporzionale ad una potenza della variabile stessa (vedi formula).

In particolare, la power law è rappresentata come una retta in log-log scale.

Punto 3

Si è calcolata la distribuzione dei waiting time per terremoti di magnitudo maggiore o uguale a un valore m fissato che va da 2 a 5. Dato un m si è calcolato il waiting time come differenza fra il tempo associato a un terremoto e il tempo del terremoto precedente (fondamentale: i dati sono ordinati rispetto al tempo).

Per prima cosa si è visualizzato l’istogramma delle frequenze normalizzate contro i waiting time, utilizzando un binnaggio fisso proporzionale alla radice del numero di eventi. Per migliorare la visualizzazione lo stesso istogramma è stato rappresentato per punti in log-log scale. Notiamo un rumore nella coda della distribuzione che è dato dal fatto che gli eventi con un lungo waiting time sono poco probabili, dunque i conteggi sono bassi, e quindi in ogni bin vi sono molti meno eventi rispetto al resto della distribuzione. Con il rumore abbiamo pochi dati nella parte per noi più significativa della curva, mentre un numero eccessivo di dati nella parte per noi insignificante (cioè abbiamo pochi bin nella parte iniziale e molti nella parte finale).

Supposto che il comportamento della distribuzione sia come una power law bisognerebbe fittare i dati nell’istogramma in log-log scale. La pendenza della retta rappresenta l’esponente della distribuzione. Ma la presenza del rumore provoca un errore nella stima dei parametri del fit (nel caso della retta, con il rumore ne sottostimiamo la pendenza, la retta si abbassa).

Al fine di eliminare il rumore è necessario ampliare la larghezza dei bin nella parte terminale della distribuzione. Per adattare la grandezza dei bin alla rispettiva parte della distribuzione si è utilizzato un binnaggio logaritmico. Questa tecnica è frequentemente utilizzata nel caso di power law. Si osserva come in questo istogramma i bin per waiting time elevati sono più larghi, raggruppando più punti rispetto a prima, e quindi si vede che nell’istogramma in log-log scale per punti che il rumore viene eliminato e quindi la quantità di informazione è regolare in tutta la distribuzione.

Si nota a questo punto una pancia nella distribuzione. Intuiamo che la prima parte si comporta come una power law, mentre per la coda, poiché gli eventi con meno probabilità decadono più velocemente di una power law, si nota un decadimento più rapido (probabilmente esponenziale).

Se si vuole quindi ottenere una stima dell’esponente della power law, è necessario fissare un lower bound entro il quale effettuare il fit lineare. Questo lower bound può essere fissato ad occhio, ma noi abbiamo cercato di automatizzare la procedura. Si è quindi proceduto a calcolare la pendenza della retta a partire dal primo punto fino al punto i-esimo per tutti gli i fino alla fine del vettore (dove i punti rappresentano i bin). Si è ottenuto il grafico delle pendenze in funzione del lower bound scelto e si è scelto come lower bound quello relativo alla zona di stabilità della pendenza (in grafico il numero di bin scelto in modo crescente).

(Come si è scelto? Si è calcolata la retta interpolante per ogni set di 5 punti del grafico delle pendenze in modo continuo (dall’1 al 10, dal 2 al 11 etc…) e si è preso come lower bound il quinto punto corrispondente al set con pendenza minore). (sceglie estremo destro)

Infine si è quindi stimata la slope facendo un’interpolazione lineare dell’istogramma tagliata al lower bound. Si osserva che il valore ottenuto nel caso di m=3 è circa -0.70.

L’intera procedura è stata effettuata per i valori di m compresi tra 2 e 5 e i valori dell’esponente risultato tutti compresi tra 0.70 e 1.

Gli istogrammi con i punti sono stati rappresentati per tutti i differenti valori di m in uno stesso grafico ottenendo il plot in figura.

Punto 4

Si è calcolata la distribuzione delle distanze tra gli ipocentri dei terremoti in coordinate cartesiane tra un terremoto e il successivo per terremoti di magnitudo uguale o superiore a un m fissato (tra 2 e 5). Siccome si ipotizza che anche le distanze siano distribuite come una power law, si è ripetuta una procedura del tutto analoga al caso della distribuzione dei waiting time (ma perché è così brutto???? Ha una pancia strana). (Scorrere le slide velocemente) Anche in questo caso si è eseguita la procedura per m che va da 2 a 5 e i risultati sono stati plottati tutti in uno stesso grafico.

Punto 5

Si è considerata la distribuzione dei waiting time per terremoti di magnitudo m o superiore, che sono separati da una distanza minore di un valore R. Per un dato m e un dato R si è effettuata un’analisi del tutto analoga a quella precedente. In particolare, si è fissato il valore di m tra 2 e 5 e successivamente si sono filtrati i dati, prendendo i terremoti lontani al massimo quanto la metà del valor medio, il valor medio e il doppio del valor medio di tutti i terremoti considerati. Anche in questo caso i grafici degli istogrammi per diversi valori di m ed R sono stati plottati in un’unica figura.

Punto

Una power law distribution è l’unica distribuzione che rimane invariata da qualunque scala la si osserva (con un riscalamento dei dati). Quindi cambiando la scala con la quale misuriamo la variabile x della distribuzione, la forma della stessa rimane invariata a meno di una costante moltiplicativa. Riscalando quindi per una grandezza caratteristica i diversi dataset relativi alla stessa distribuzione, si osserva il collasso delle differenti curve su un’unica curva. Il riscalamento si ottiene dividendo le posizioni dei bin (x) per la grandezza caratteristica e moltiplicando le frequenze normalizzate (y) per la grandezza caratteristica.

Punto 6

Si sono riscalati i dati dei waiting time relativi agli istogrammi per diversi m utilizzando come grandezza caratteristica il tempo medio. Si osserva come le curve collassino molto bene sovrapponendosi.

Si sono riscalati i dati delle distanze relativi agli istogrammi per diversi m utilizzando come grandezza caratteristica la distanza massima. Si osserva come le curve non varino significativamente rispetto alle curve non riscalate, che risultavano già approssimativamente sovrapposte (????).

Infine, si sono riscalati i dati dei waiting time per differenti m, filtrati per distanze minori di R, utilizzando come lunghezza caratteristica il tempo medio moltiplicato per la distanza massima. Si osserva come le curve collassino.

CONCLUSIONI