

Chapter 1

Introduzione

The number of things you don't know is
one of the things you don't know

Peter Green

Quando si vuole identificare un sistema dinamico una fase importante è stabilire la struttura del modello. Spesso infatti non sono incognite solo le costanti della legge che regola il processo ma non si conoscono nemmeno quali operazioni sui segnali di ingresso e di stato riproducono accettabilmente la dinamica del sistema.

I casi in cui è sufficiente una identificazione parametrica sono i casi in cui si ha già la struttura della legge. Per ottenere la legge (a meno del valore delle costanti) è necessario uno studio del sistema a partire dai principi primi che regolano tutti i fenomeni che sono coinvolti. In definitiva l'identificazione parametrica è sufficiente solo se sono verificate le seguenti condizioni

- è **possibile** ottenere una descrizione del sistema applicando i principi primi delle scienze
- è **vantaggioso** ottenere una descrizione del sistema applicando i principi primi delle scienze (tempo speso, rischio di errori etc)
- ci si accontenta della descrizione ottenuta applicando i principi primi delle scienze (rischio di tralasciare fenomeni importanti)
- la complessità delle leggi che si ottengono applicando i principi primi delle scienze non è tale da renderle inutilizzabili

Quando una di queste condizioni non è soddisfatta è possibile inserire nel processo di identificazione anche la ricerca della struttura del modello. A tale scopo si ipotizza una classe (abbastanza vasta) di strutture di modello e si lascia al processo di identificazione l'onere di scegliere quella che maggiormente si adatta alle misure. I metodi Bayesiani in particolare sono appropriati per la selezione dei modelli perchè forniscono naturalmente (senza analisi aggiuntive) un contesto in cui si quantifica l'incertezza associata all'identificazione. L'approccio bayesiano infatti, non fornisce una singola descrizione del sistema ma fornisce una distribuzione di probabilità associata al set di modelli possibili. Dunque risolvere

il problema di inferenza bayesiana conduce a scelte più ponderate rispetto alla semplice estrazione del modello che si adatta meglio ai dati reali. Analizzando la distribuzione prodotta si può per esempio individuare un modello poco meno accurato del migliore ma preferirlo perchè meno complesso .

Nella seguente trattazione la classe di modelli scelta è quella dei nonlineari, autoregressivi, a media mobile con ingressi esogeni (NARMAX); essa è una popolare classe modelli ingresso-uscita spesso usata nell'identificazione di sistemi non lineari in vari ambiti dell'ingegneria (e non) poichè ha il pregio di produrre leggi compatte ma accurate. Come si vedrà nei successivi capitoli, ciascun modello NARMAX può essere visto come lo sviluppo su una opportuna base polinomiale della legge che regola il sistema. Ciò che distingue un modello da un altro è il numero e l'insieme dei termini presenti nello sviluppo.

La ricerca esaustiva tra tutte le possibili combinazioni di termini NARMAX è tuttavia quasi sempre computazionalmente improponibile , così si è puntato negli ultimi anni a sviluppare metodi di campionamento efficaci che, avvalendosi di una catena di Markov opportunamente costruita, esplorano in maniera non esaustiva l'insieme dei modelli ottenendo comunque statistiche significative.

Chapter 2

Identificazione bayesiana e metodi di campionamento Monte Carlo

2.1 Approccio Bayesiano per l'identificazione

L'inferenza bayesiana è un approccio all'inferenza statistica in cui le probabilità non sono interpretate come frequenze ma piuttosto come livelli di fiducia nel verificarsi di un dato evento. Il nome deriva dal teorema di Bayes, che costituisce il fondamento di questo approccio. Gli statistici bayesiani sostengono che i metodi dell'inferenza bayesiana rappresentano una formalizzazione del metodo scientifico, che normalmente implica la raccolta di dati che avvalorano o confutano una data ipotesi. Queste caratteristiche rendono l'approccio bayesiano un utile ausilio per discriminare tra alternative in conflitto e dunque un ottimo strumento per l'identificazione dei sistemi. Il metodo usa una stima del grado di fiducia in una data ipotesi prima dell'osservazione dei dati al fine di associare un valore numerico al grado di fiducia in quella stessa ipotesi successivamente all'osservazione dei dati. Si supponga di voler identificare un sistema scegliendo il più adatto in un insieme M di modelli. Sia M una variabile aleatoria che rappresenta la legge che regola il vero sistema. L'obiettivo dell'identificazione bayesiana è quello di calcolare per ogni $m \in M$ la probabilità a posteriori

$$P(M = m|Y)$$

dunque complessivamente determinare la distribuzione della probabilità sull'insieme dei modelli condizionatamente alle misure. La probabilità condizionata è stata definita in termini della probabilità congiunta e marginale dei due eventi

$$P(M = m|Y) = P(M = m, Y)P(Y)$$

la definizione corrisponde all'idea ragionevole

$$P(M = m|Y) \propto P(M = m, Y)$$

Avendo per l'inferenza bayesiana è un approccio all'inferenza statistica in cui le probabilità non sono interpretate come frequenze ma piuttosto come livelli

di fiducia nel verificarsi di un dato evento. Il nome deriva dal teorema di Bayes, che costituisce il fondamento di questo approccio. Gli statistici bayesiani sostengono che i metodi dell'inferenza bayesiana rappresentano una formalizzazione del metodo scientifico, che normalmente implica la raccolta di dati che avvalorano o confutano una data ipotesi. Queste caratteristiche rendono l'approccio bayesiano un utile ausilio per discriminare tra alternative in conflitto e dunque un ottimo strumento per l'identificazione dei sistemi. Il metodo usa una stima del grado di fiducia in una data ipotesi prima dell'osservazione dei dati al fine di associare un valore numerico al grado di fiducia in quella stessa ipotesi successivamente all'osservazione dei dati. Si supponga di voler identificare un sistema scegliendo il più adatto in un insieme M di modelli. Sia M una variabile aleatoria che rappresenta la legge che regola il vero sistema. L'obiettivo dell'identificazione bayesiana è quello di calcolare per ogni $m \in M$ la probabilità a posteriori

$$P(M = m|Y)$$

dunque complessivamente determinare la distribuzione della probabilità sull'insieme dei modelli condizionatamente alle misure. La probabilità condizionata è stata definita in termini della probabilità congiunta e marginale dei due eventi

$$P(M = m|Y) = P(M = m, Y)P(Y)$$

la definizione corrisponde all'idea ragionevole

$$P(M = m|Y) \propto P(M = m, Y)$$

Avendo però ristretto l'insieme di supporto su cui è definita la metrica di probabilità, è necessario imporre che valgano ancora gli assiomi della probabilità, in particolare chiamando k la costante di proporzionalità e integrando su tutti i casi possibili si ha

$$P(M = m|Y)dm = k \cdot P(M = m, Y)dm = k \cdot P(Y) := 1$$

da cui si ricava il valore di k

$$k = \frac{1}{P(Y)}$$

ottenendo la definizione di probabilità condizionata. Scambiando i ruoli delle variabili aleatorie si ha anche che $P(Y|M = m) = P(M = m, Y)P(M = m)$ (2.2) dunque mettendo insieme le equazioni (2.1) e la (2.2) si ottiene la formula di Bayes $P(M = m|Y) = P(Y|M = m)P(M = m)P(Y)$ (2.3) usando il teorema della probabilità totale si può inoltre esprimere la costante al denominatore in funzione delle quantità al numeratore ottenendo $P(M = m|Y) = P(Y|M = m)P(M = m)P(Y|M = m)P(M = m)dm$ (2.4) Nei problemi di identificazione l'espressione della probabilità a posteriori è quasi sempre impossibile da valutare analiticamente soprattutto per colpa dell'integrale al denominatore che deve essere calcolato su tutti i possibili modelli; si ricorre quindi a metodi numerici di tipo Monte Carlo basati sul campionamento della distribuzione. Spesso risulta impossibile anche campionare direttamente la distribuzione e se ne deve ottenere una stima accettando o scartando opportunamente i campioni estratti da una distribuzione più semplice.