

Pro e contro delle RNA

Pro:

- flessibilità: le RNA sono approssimatori universali;
- aggiornabilità sequenziale: la stima dei pesi della rete può essere aggiornata man mano che arriva nuova informazione;
- le RNA sono adatte ad essere implementate su macchine per il calcolo parallelo;

Pro e contro delle RNA

Contro:

- *"data hungry"* e *overfitting*: la stima dei pesi richiede numerosità consistenti;
- difficoltà di stima: presenza di molti minimi locali nella funzione obiettivo;
- interpretabilità: grosse difficoltà ad interpretare la relazione fra x ed y ;
- inferenza: non ci sono metodi statistici per il calcolo degli errori di stima associati ai parametri, per la scelta del numero di neuroni nascosti e per il pruning della rete;



Reti con apprendimento supervisionato

Reti neurali *con un livello nascosto* sono in grado di risolvere problemi molto più complessi:

- regressione non lineare
- classificazione con la presenza di due o più rette discriminanti
- classificazione con criteri non lineari



Reti con apprendimento supervisionato

Si tratta di problemi **difficilmente affrontabili con tecniche tradizionali** e per i quali le RNA sanno dare contributi importanti.

Reti con apprendimento supervisionato

RNA per la regressione non lineare

La regressione non lineare si ha quando Y è una variabile quantitativa che si assume dipendente da s variabili esogene X_1, X_2, \dots, X_s attraverso una funzione g **non lineare**.

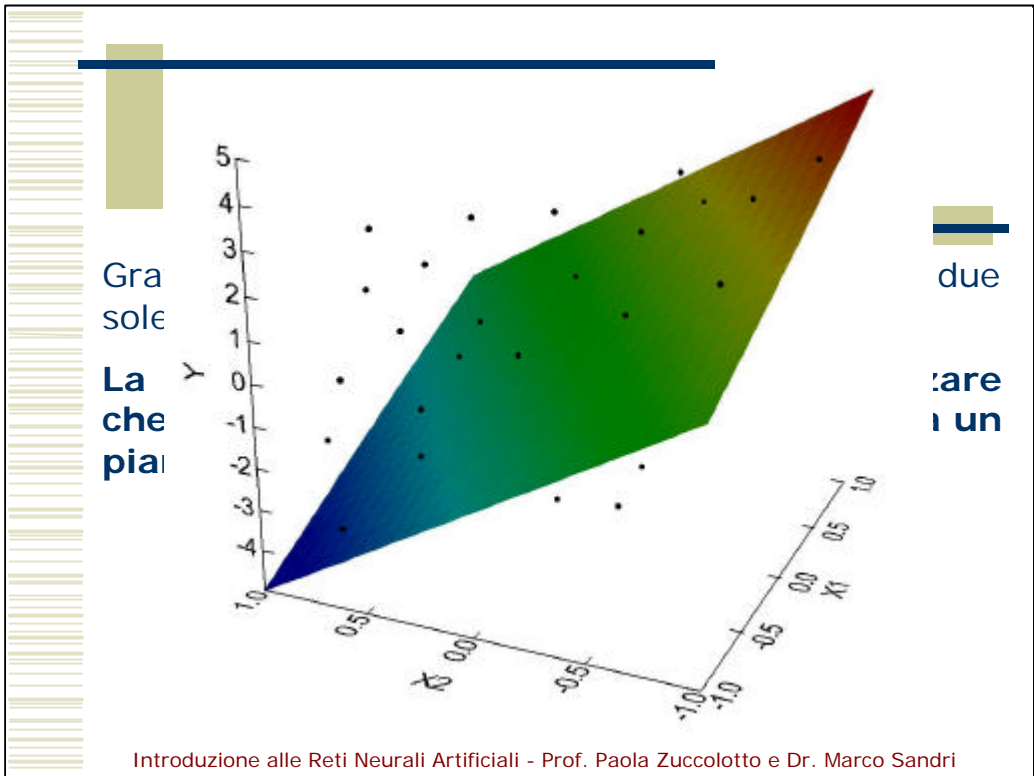
$$\hat{Y} = g(X_1, X_2, \dots, X_s)$$

Reti con apprendimento supervisionato

Si può pensare sempre all'esempio della catena di ipermercati che desidera prevedere il fatturato di un nuovo punto vendita sulla base di tre variabili di contesto relative all'area in cui esso sorgerà.

$$\hat{Y} = g(X_1, X_2, X_3)$$

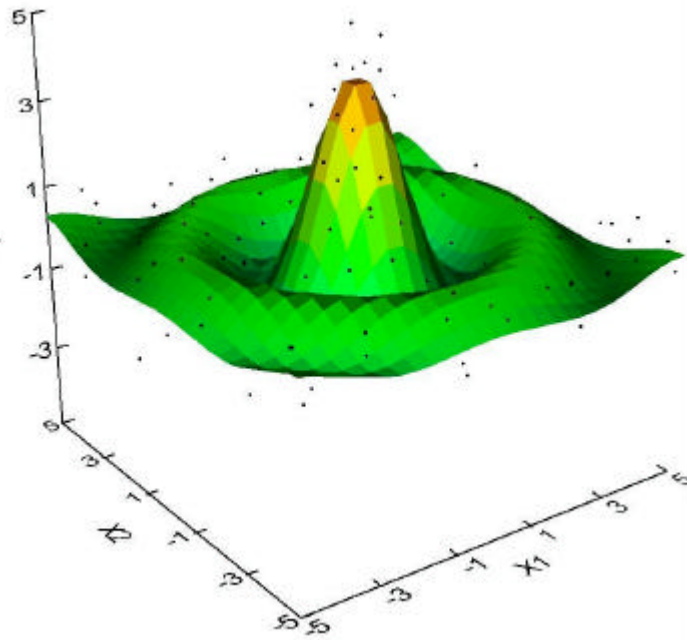
Prima avevamo supposto che g fosse una funzione lineare, ma in realtà si potrebbe ipotizzare che il meccanismo che lega le variabili sia *più complesso*.



Introduzione alle Reti Neurali Artificiali - Prof. Paola Zuccolotto e Dr. Marco Sandri

Costo

J



ti
la



Reti con apprendimento supervisionato

La complessità della funzione di regressione è definita dalla forma che si imprime alle *propagation rule* ed alle funzioni di attivazione.

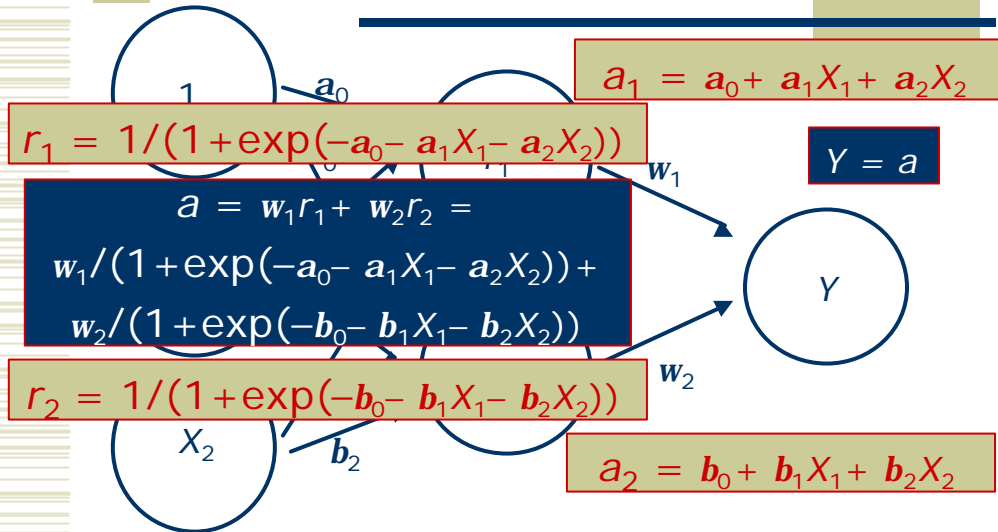
Ovviamente sono possibili un infinito numero di forme.

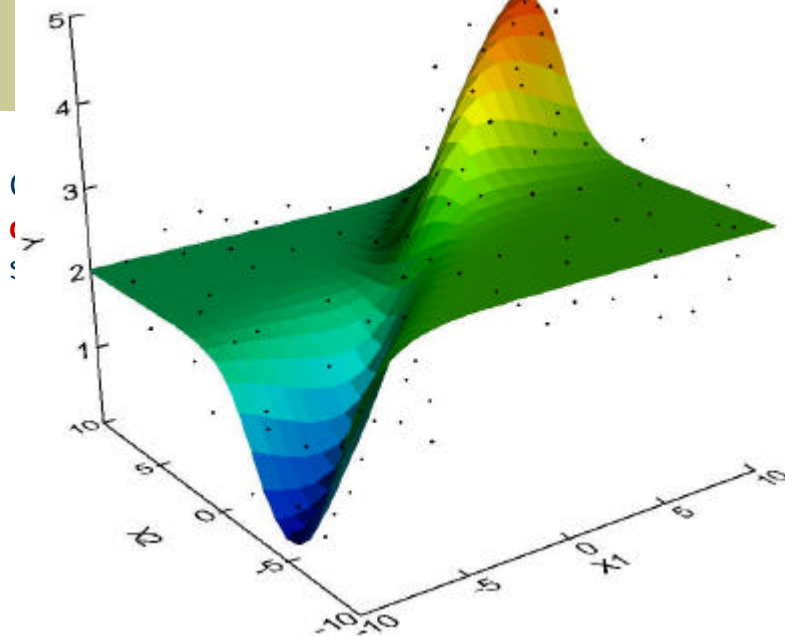
Reti con apprendimento supervisionato

Vediamo cosa accade con due variabili esogene e una comune architettura di RNA:

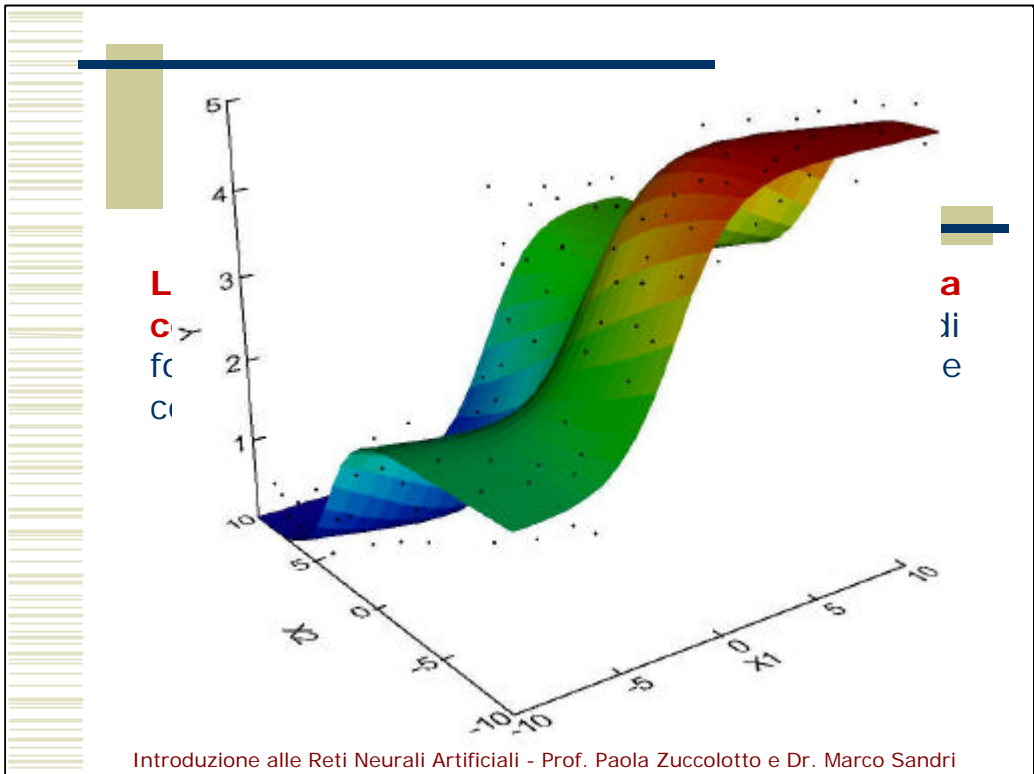
- **Un livello nascosto composto da due nodi ($l=1, m=2$)**
- ***Propagation rule* lineari sia nel livello nascosto che nel livello di *output***
- **funzione di attivazione sigmoide nel livello nascosto e identità nel livello di *output*.**

Reti con apprendimento supervisionato





Introduzione alle Reti Neurali Artificiali - Prof. Paola Zuccolotto e Dr. Marco Sandri



Introduzione alle Reti Neurali Artificiali - Prof. Paola Zuccolotto e Dr. Marco Sandri



Reti con apprendimento supervisionato

Naturalmente quando le variabili esogene sono più di due non è più possibile rappresentare graficamente la funzione di regressione.

Inoltre per architetture più complesse cresce anche la complessità delle funzioni e la varietà di forme che esse possono assumere.

Aver esaminato a fondo dei casi semplici ci è servito unicamente per avere un'idea di come opera la rete nei casi complessi, che sono quelli effettivamente interessanti dal punto di vista operativo.

Introduzione alle Reti Neurali Artificiali - Prof. Paola Zuccolotto e Dr. Marco Sandri

Reti con apprendimento supervisionato

Quindi in generale:

- **E' estremamente difficile e comunque non interessa esplicitare la funzione di regressione.**
- **Non interessano i valori assunti dai parametri.**
- **Una volta dettate le regole fondamentali (l , m , *propagation rule*, funzioni di attivazione) la rete diviene una *black box* a cui fornire i nuovi *inputs* per ottenere i corrispondenti *outputs*.**



Reti con apprendimento supervisionato

Importa soltanto sapere che:

- **Agendo sul numero di nodi nel livello nascosto, sulla forma delle *propagation rule* e delle funzioni di attivazione ai vari livelli, si ottengono un gran numero di diverse tipologie di non linearità**
- **Una stessa RNA è in genere in grado di approssimare una grande varietà di funzioni di regressione molto complesse**

Reti con apprendimento supervisionato

Tutto quanto detto per la classificazione con due variabili esogene si estende chiaramente anche al caso di s variabili esogene.

In questo caso si perde la possibilità di rappresentare graficamente i punti e le funzioni discriminanti ed inoltre cresce molto la complessità delle funzioni stesse.

Aver esaminato a fondo dei casi semplici ci è servito unicamente per avere un'idea di come opera la rete nei casi complessi, che sono quelli effettivamente interessanti dal punto di vista operativo.

Introduzione alle Reti Neurali Artificiali - Prof. Paola Zuccolotto e Dr. Marco Sandri

Reti con apprendimento supervisionato

Anche nel caso della classificazione dunque:

- **E' estremamente difficile e comunque non interessa esplicitare la funzione discriminante.**
- **Non interessano i valori assunti dai parametri.**
- **Una volta dettate le regole fondamentali (I , m , *propagation rule*, funzioni di attivazione) la rete diviene una *black box* a cui fornire i nuovi *inputs* per ottenere le corrispondenti **classificazioni.****



Reti con apprendimento supervisionato

Importa soltanto sapere che:

- **Agendo sul numero di nodi nel livello nascosto, sulla forma delle *propagation rule* e delle funzioni di attivazione ai vari livelli, si ottengono diverse tipologie di discriminazione (lineare o non lineare, a una, due o più funzioni discriminanti).**
- **Una RNA è in genere in grado di effettuare classificazioni anche in casi di forme molto complesse.**



Un'osservazione sulla terminologia

Un'ultima osservazione: nel campo delle RNA si utilizza, come abbiamo visto, una terminologia specifica.

Taluni termini sono spesso "di effetto" (si pensi alla parola "apprendimento") ma non devono ingannare, nel senso che si deve sempre tenere presente il loro effettivo significato, per non attribuire alle RNA capacità che vanno oltre quelle di puro e semplice strumento matematico.

Vediamo quindi una breve rassegna di termini tipici delle RNA, associati al loro corrispondente nella terminologia statistica classica.

Un'osservazione sulla terminologia

RNA	Statistica
<i>Inputs</i>	Variabili esogene o indipendenti
<i>Target</i>	Valori osservati
<i>Outputs</i>	Valori teorici (o previsti)
<i>Training set</i>	Campione
Esempio	Unità campionaria
Pesi	Parametri
Apprendimento o Addestramento	Stima dei parametri