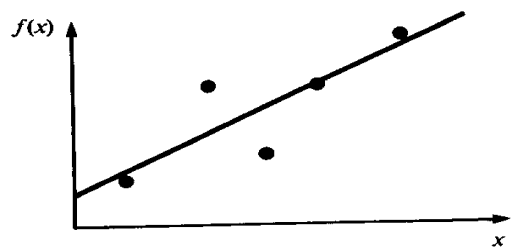


INTERPOLAZIONE E REGRESSIONE

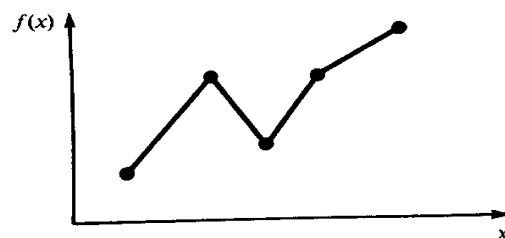
Nella pratica ingegneristica accade sovente che si conosca il valore approssimato di una grandezza variabile con continuità solo in corrispondenza di alcuni punti (ad esempio attraverso operazioni di misura). Si pone allora il problema di determinare il valore della grandezza nei punti intermedi a partire dai dati noti. Allo scopo, è possibile ricavare, mediante una tecnica numerica detta **regressione**, la funzione che, pur non passando necessariamente per i punti noti, approssima meglio il loro comportamento secondo un criterio che verrà definito di seguito.

Si noti che generalmente i dati di partenza sono affetti da errore (o *rumore*) significativo pertanto, in questi casi, non ha senso imporre alla funzione approssimante di passare attraverso di essi.

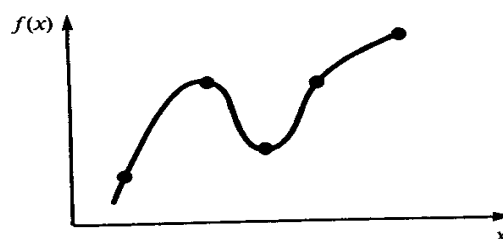
Una situazione differente è quella di approssimare una funzione molto complicata con una funzione più semplice che l'approssimi al meglio. In questo caso è possibile valutare la funzione complessa in un certo numero di punti nell'intervallo di interesse e quindi ricavare, attraverso una tecnica detta di **interpolazione**, la funzione più semplice che passa attraverso i punti determinati. Poiché in questo caso nei punti assegnati si conoscono i valori **esatti** della funzione, è adesso richiesto alla funzione approssimante di assumere in quei punti esattamente gli stessi valori (naturalmente, nei punti intermedi le due funzioni assumeranno generalmente valori differenti).



(a)



(b)



(c)

Figura IV.1 Tre diversi tentativi di determinare la curva migliore che passa per cinque punti: (a) regressione ai minimi quadrati, (b) interpolazione lineare e (c) interpolazione curvilinea.

Esempi di applicazione delle tecniche di regressione e interpolazione

- Sono note attraverso misure o prove sperimentali i valori di una grandezza in corrispondenza di determinati punti attraverso una **tabella**. Viene richiesto il valore della grandezza in punti diversi da quelli noti con due obiettivi diversi:
 1. **Analisi delle tendenze**: l'andamento dei dati noti viene utilizzato per predire il valore nei punti incogniti.
Se i dati noti sono esatti o misurati con grande precisione la funzione approssimante è costruita per interpolazione, in caso contrario per regressione. Se si richiede di determinare la grandezza di interesse per valori delle variabili indipendenti esterni (risp. interni) rispetto al range dei dati disponibili si parla di **estrapolazione** (risp. **interpolazione**).
 1. **Caratterizzazione di un sistema reale**: Allo scopo di caratterizzare un sistema reale si introduce un modello matematico predefinito i cui parametri non sono noti. Questi ultimi vengono determinati in modo che il modello approssimi al meglio i dati sperimentali ottenuti attraverso misure sul sistema reale per un insieme discreto di valori delle variabili di ingresso.
 2. **Verifica delle ipotesi**: Nel caso in cui siano stati proposti diversi modelli alternativi, i dati sperimentali vengono utilizzati per determinare quello che meglio rappresenta il comportamento di un dato sistema reale.
- La regressione e l'interpolazione assumono grande importanza anche negli altri metodi numerici: ad esempio, nella integrazione definita, nella soluzione di equazioni differenziali, nell'approssimazione di funzioni matematiche complesse.

Nozioni matematiche di base: elementi di statistica

Sia y una variabile casuale (in altre parole il suo valore non sia ben determinato, ma ottenuto per esempio attraverso misure sperimentali). Sia $\{y_i, i=1:n\}$ un insieme di dati rappresentativi della variabile y ottenuti attraverso una serie di prove.

Per rappresentare i risultati di una serie di misure di una grandezza definiamo la **distribuzione della densità di probabilità**: dal punto di vista matematico, è la

funzione $f(y) = \lim_{n \rightarrow \infty} \frac{n_y}{n}$ definita come limite del rapporto tra “ n_y ”, il numero di volte in cui la variabile random ha assunto il valore y , e “ n ”, il numero totale di esperimenti effettuati, quando quest’ultimo valore tende a infinito.

Usualmente i dati tendono ad assumere una distribuzione dalla caratteristica forma a campana (linea continua in figura) detta *distribuzione normale*.

Se il numero n dei campioni è limitato, si può approssimare la distribuzione dei dati attraverso un istogramma. Allo scopo, si suddivide il range dei valori riportati in intervalli di uguale ampiezza. In ascisse si riportano i valori assunti dalla variabile e in ordinata il numero di misure che ricadono in ogni intervallo:

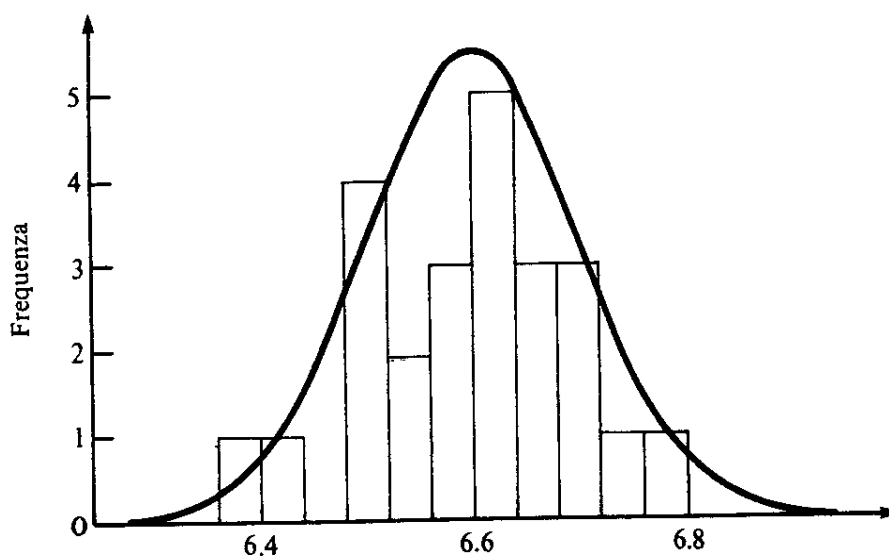


Figura IV.2 Istogramma utilizzato per visualizzare la distribuzione dei dati. Al crescere del loro numero, l’istogramma può tendere alla curva regolare a forma di campana nota come distribuzione normale.

Nell'analisi statistica, vengono definiti alcuni parametri di carattere globale che permettono di descrivere in modo sintetico il comportamento di una distribuzione.

- I parametri più utilizzati per definire il valore centrale di una distribuzione sono la **media** e **il valore quadratico medio** definiti rispettivamente come:

$$\bar{y}_t = \int_{-\infty}^{\infty} yf(y)dy \quad \text{e} \quad \overline{y_t^2} = \int_{-\infty}^{\infty} y^2 f(y)dy$$

Se non si conosce la densità di distribuzione, si ottiene una stima

- della media dividendo per n la somma dei valori riportati per y in n

esperimenti: $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

- del valore quadratico medio dividendo per n la somma dei quadrati dei

valori riportati per y in n esperimenti: $\overline{y^2} = \frac{1}{n} \sum_{i=1}^n y_i^2$.

- La misura più comune della dispersione dei dati rispetto alla media è ottenuta attraverso la **deviazione standard** $\sigma_t(y)$, ottenuta attraverso la radice quadrata della **varianza** $\sigma_t^2(y) = \overline{y_t^2} - \bar{y}_t^2$.

Per stimare la deviazione standard si usa la **deviazione standard campione**

$$S^2(y) = \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{dove} \quad \sigma_y = \sqrt{\frac{S^2(y)}{n-1}}$$

Il quadrato della deviazione standard

campione prende il nome di **varianza campione** $\sigma_y^2 = \frac{S^2(y)}{n-1}$.

La divisione per $n-1$ è giustificata dal fatto che non ha senso definire la dispersione per un singolo dato.

- Per quantificare la dispersione dei dati si può utilizzare un ulteriore parametro, che ha lo stesso significato dell'errore relativo percentuale, detto **coefficiente di variazione (c.v.)**, definito come il rapporto tra la deviazione standard e la media vera moltiplicato per 100 (per poter essere espresso in termini percentuali) e

approssimato dal rapporto: $c.v. = \frac{\sigma_y}{y} 100\%$

Esempio:

Determinare i principali parametri statistici relativi ai dati riportati in tabella.

Tabella IV.1 Misure del coefficiente di espansione termica di un metallo da costruzione ($\times 10^{-6}$ cm/cm °C).

6.495	6.625	6.635	6.655
6.665	6.515	6.625	6.775
6.755	6.615	6.575	6.555
6.565	6.435	6.395	6.655
6.595	6.715	6.485	6.605
6.505	6.555	6.715	6.685

Media :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{24} \sum_{i=1}^{24} y_i = \frac{158.400}{24} = 6.6$$

Deviazione Standard Campione:

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{0.217000}{23}} = 0.097133$$

Varianza Campione:

$$\sigma_y^2 = 0.009435$$

Coefficiente di Variazione:

$$c.v. = \frac{\sigma_y}{\bar{y}} 100\% = \frac{0.097133}{6.6} 100\% = 1.47\%$$

Tabella IV.2 Calcolo per la costruzione dell'istogramma e la determinazione dei parametri statistici relativi alla misura del coefficiente di espansione termica.

i	y_i	$(y_i - \bar{y})^2$	Frequenza	Intervallo	
				Limite inferiore	Limite superiore
1	6.395	0.042025	1	6.36	6.40
2	6.435	0.027225	1	6.40	6.44
3	6.485	0.013225	4	6.48	6.52
4	6.495	0.011025			
5	6.505	0.009025			
6	6.515	0.007225			
7	6.555	0.002025	2	6.52	6.56
8	6.555	0.002025			
9	6.565	0.001225	3	6.56	6.60
10	6.575	0.000625			
11	6.595	0.000025			
12	6.605	0.000025	5	6.60	6.64
13	6.615	0.000225			
14	6.625	0.000625			
15	6.625	0.000625			
16	6.635	0.001225			
17	6.655	0.003025	3	6.64	6.68
18	6.655	0.003025			
19	6.665	0.004225			
20	6.685	0.007225	3	6.68	6.72
21	6.715	0.013225			
22	6.715	0.013225	1	6.72	6.76
23	6.755	0.024025			
24	6.775	0.030625	1	6.76	6.80

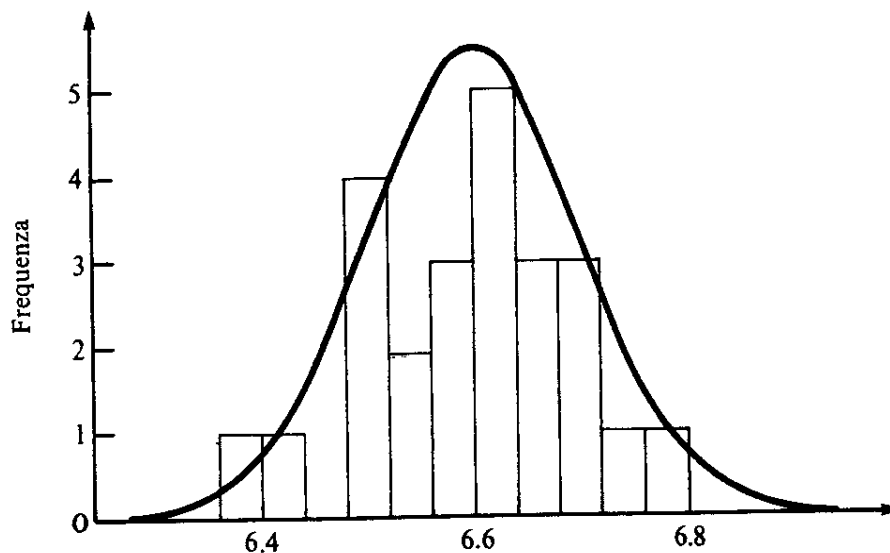


Figura IV.2 Istogramma utilizzato per visualizzare la distribuzione dei dati. Al crescere del loro numero, l'istogramma può tendere alla curva regolare a forma di campana nota come distribuzione normale.

Regressione ai minimi quadrati

Si consideri un sistema fisico o ingegneristico, e si assuma, per semplicità, che abbia un unico ingresso ed un'unica uscita (per esempio il valore di una grandezza)

Sia nota attraverso misure sperimentali l'uscita del sistema per un numero discreto di valori assunti dall'ingresso e, attraverso questa informazione, si supponga di voler stimare l'uscita corrispondente ad un diverso insieme di valori dell'ingresso.

Se i dati relativi alle misure sono affetti da un errore sensibile, la loro approssimazione attraverso interpolazione polinomiale si dimostra inadeguata per la stima dei valori intermedi.

In questi casi la migliore strategia per l'approssimazione dei dati è quella di cercare una funzione approssimante che si adatti all'andamento generale dei dati senza passare per nessuno in particolare. Si parla, allora, di regressione.

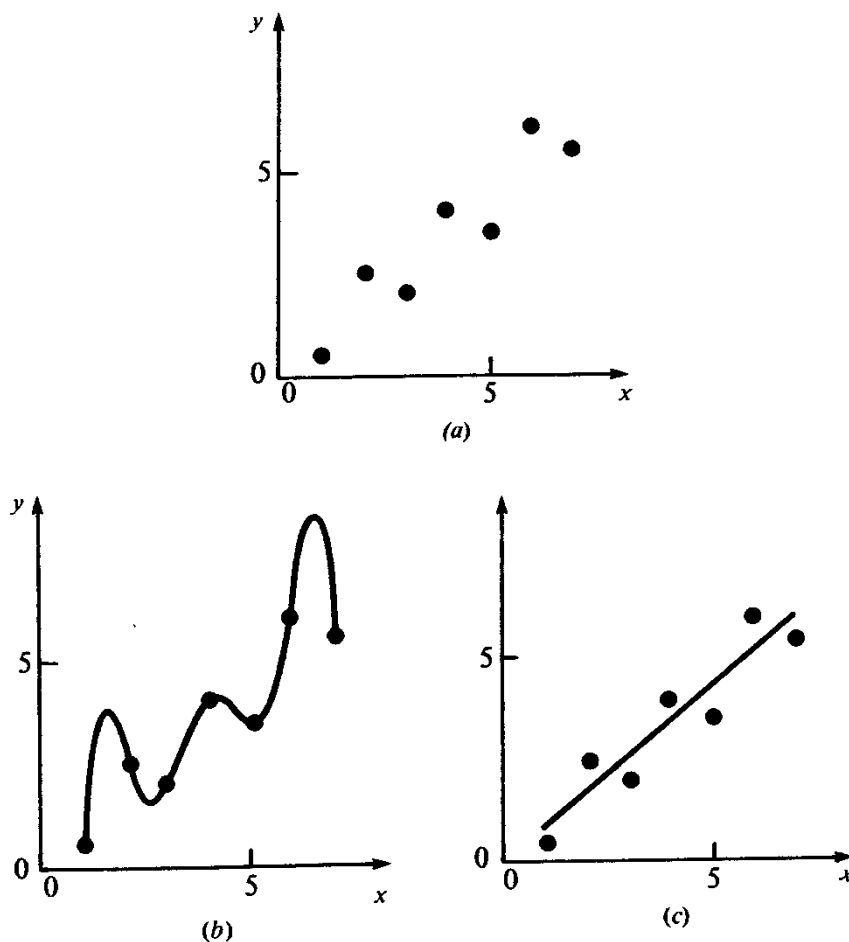


Figura 10.1 (a) Dati caratterizzati da errore significativo. (b) Approssimazione polinomiale oscillante al di fuori del campo di valori in cui sono contenuti i dati. (c) Risultato più soddisfacente ottenuto con una regressione lineare che soddisfa il criterio dei minimi quadrati.

REGRESSIONE LINEARE AI MINIMI QUADRATI

Il modo più semplice di effettuare una regressione è quello di approssimare un insieme di punti definiti da coppie di dati (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) attraverso una linea retta di equazione:

$$y = a_0 + a_1 x + E,$$

dove a_0 ed a_1 rappresentano l'intercetta e la pendenza della retta, mentre E è l'errore o residuo tra il modello e il dato osservato vale a dire la differenza tra il vero valore y_i corrispondente al dato x_i e l'approssimazione $a_0 + a_1 x_i$ data dall'equazione lineare:

$$E = y - (a_0 + a_1 x)$$

Criteri per la valutazione dell'approssimazione ottima

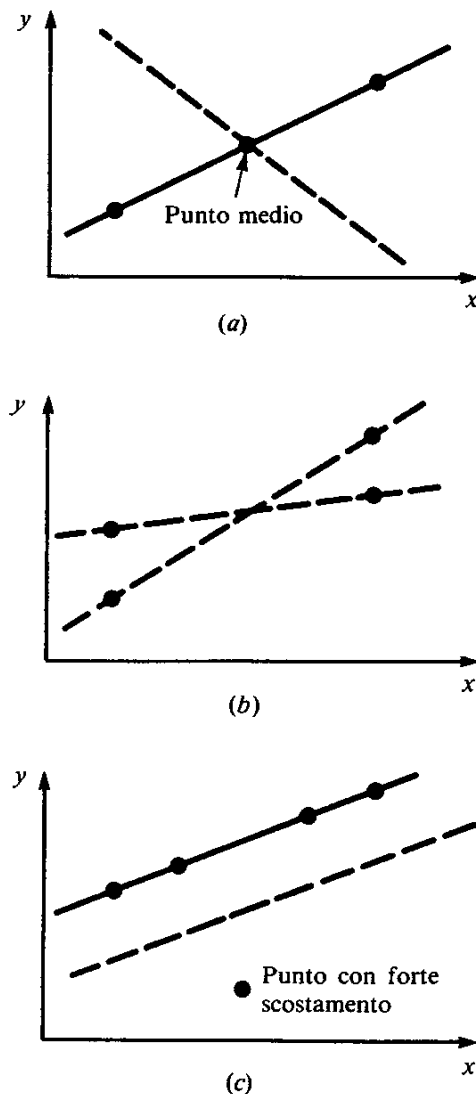


Figura 10.2 Esempi di alcuni criteri di valutazione della bontà di un'approssimazione che non sono adatti nel caso della regressione lineare: (a) minima somma dei residui; (b) minima somma dei valori assoluti dei residui e (c) minimo errore massimo di ogni singolo punto.

Ci si pone ora il problema di scegliere i coefficienti della retta così da approssimare i dati nella maniera migliore.

- Un prima strategia potrebbe essere quella di minimizzare la somma degli errori:

$$\sum_{i=1}^n E_i = \sum_{i=1}^n (y_i - a_0 - a_1 x_i).$$

Questa strategia risulta inadeguata in situazioni come quelle rappresentate in fig.10.2a. Con due soli punti, ovviamente la migliore approssimazione è costituita dalla retta che passa per i punti. Viceversa col criterio definito qualunque retta che passi attraverso il punto intermedio (tranne la retta verticale) produce errore complessivo nullo: gli errori di segno opposto si compensano.

- Si può pensare di rimediare a questo inconveniente minimizzando la somma dei

valori assoluti degli errori: $\sum_{i=1}^n |E_i| = \sum_{i=1}^n |(y_i - a_0 - a_1 x_i)|$. Tuttavia, come risulta

dalla figura 10.2b, anche questo criterio appare inadeguato. Infatti, qualsiasi retta compresa tra le due tratteggiate minimizza la funzione di errore globale sopra definita.

- Si potrebbe allora pensare di utilizzare una strategia che va sotto il nome di *minimax*: scegliere la retta che minimizza la massima distanza di ogni singolo punto dalla linea. La fig. 10.2c dimostra come anche questo criterio risulti adeguato in quanto porta ad attribuire un peso notevole ai punti isolati che, nella realtà, possono essere ottenuti attraverso misure particolarmente imprecise.
- Tutte le situazioni degeneri illustrate possono essere superate utilizzando un approccio nel quale viene richiesto di minimizzare la somma dei quadrati degli

errori $S_r = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2$. Tra i maggiori pregi di questa tecnica compare quello di determinare una soluzione unica.

Determinazione dei parametri della retta ai minimi quadrati

Determiniamo ora i coefficienti a_0 e a_1 che minimizzano la funzione S_r , somma dei quadrati degli errori. Annullando le derivate parziali, risulta:

$$\frac{\partial S_r}{\partial a_0} = -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i) = 0 \quad \rightarrow \quad n a_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$\frac{\partial S_r}{\partial a_1} = -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i) x_i = 0 \quad \rightarrow \quad a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i$$

Il sistema di due equazioni in due incognite può essere facilmente risolto:

$$a_1 = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$

Esempio: Determinare la retta che meglio approssima i dati nelle prime 2 colonne

Tabella 10.1 Analisi dell'errore nel caso di approssimazione lineare.

x_i	y_i	$(y_i - \bar{y})^2$	$(y_i - a_0 - a_1 x_i)^2$
1	0.5	8.5765	0.1687
2	2.5	0.8622	0.5625
3	2.0	2.0408	0.3473
4	4.0	0.3265	0.3265
5	3.5	0.0051	0.5896
6	6.0	6.6122	0.7972
7	5.5	4.2908	0.1993
Σ	24	22.7143	2.9911

$$n = 7; \sum_{i=1}^n y_i x_i = 119.5; \sum_{i=1}^n x_i^2 = 140; \sum_{i=1}^n x_i = 28;$$

$$\bar{x} = \frac{28}{7} = 4; \sum_{i=1}^n y_i = 24; \bar{y} = \frac{24}{7} = 3.428571429$$

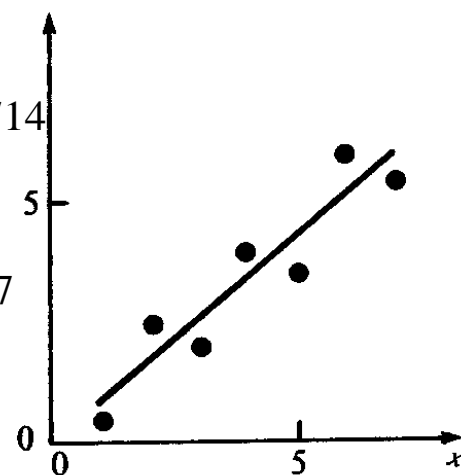
da cui si ottiene:

$$a_1 = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} = \frac{7 * 119.5 - 28 * 24}{7 * 140 - (28)^2} = 0.839285714$$

$$a_0 = \bar{y} - a_1 \bar{x} = 3.428571429 - 0.839285714 * 4 = 0.07142857$$

pertanto l'approssimazione lineare ai minimi quadrati è data

dalla retta: $y = 0.07142857 + 0.839285714 * x$



(c)

Quantificazione dell'errore nella regressione lineare

Qualsiasi retta con coefficienti a_0 , a_1 differenti da quelli ottenuti attraverso le equazioni trovate dà origine ad una somma degli errori quadratici superiore rispetto a quella minima. Pertanto la retta determinata è, nel criterio assunto, la migliore.

Esistono altre proprietà della regressione lineare. Consideriamo il modo in cui sono

stati calcolati i residui:
$$S_r = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2 \quad (1)$$

Si può notare la somiglianza tra questa espressione e quella dell'errore quadratico medio.

Questa ultima quantità rappresenta il quadrato della differenza tra i dati e la stima della media, il punto attorno al quale i dati tendono a raggrupparsi.

Nella (1) i residui rappresentano il quadrato della distanza misurata in direzione verticale tra i dati e un'altra rappresentazione della zona centrale dei dati, vale a dire la retta data dalla regressione (vedi figura 10.3)

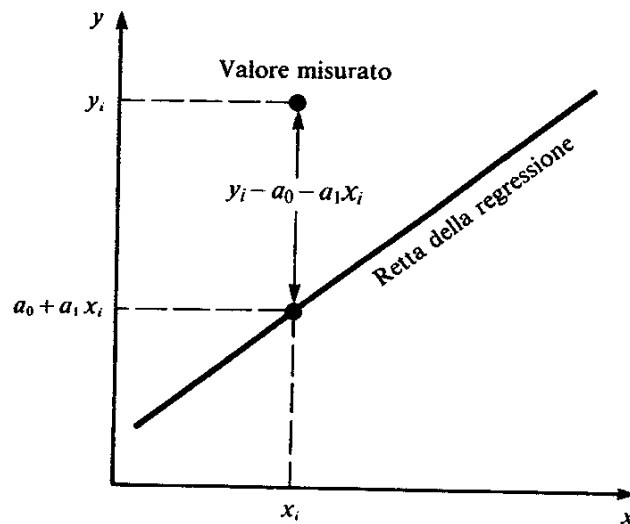


Figura 10.3 Nella regressione lineare il residuo rappresenta il quadrato della distanza verticale tra il punto e la retta.

L'analogia è maggiormente rilevabile nei casi in cui la dispersione dei punti sui due lati della retta è della stessa entità lungo tutto il campo di variabilità dei dati. e nei casi in cui la dispersione dei punti attorno alla retta è descrivibile attraverso la distribuzione normale.

Se queste condizioni sono verificate si può dimostrare che la regressione ai minimi quadrati fornisce la stima migliore (cioè la più verosimile) dei parametri a_0 e a_1 .

Inoltre in queste condizioni è possibile determinare la deviazione standard della retta

definita dalla regressione come:
$$\sigma_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

dove $\sigma_{y/x}$ è detto **errore standard della stima**.

Si noti che al denominatore compare $n-2$ perché per calcolare S_r sono state utilizzate due quantità a_0 e a_1 e così sono stati persi due gradi di libertà (con 2 soli punti l'errore quadratico è nullo e non esiste dispersione)

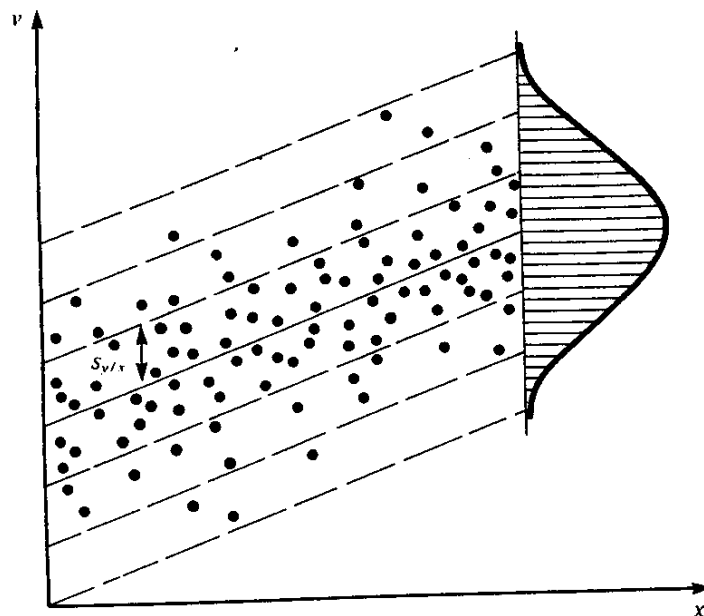


Figura 10.4 Rappresentazione della dispersione dei dati attorno a una retta approssimante.

L'errore standard della stima, come la deviazione standard, quantifica la dispersione dei dati attorno alla linea di regressione come mostrato dalla fig. 10.4.

Questi concetti possono essere utilizzati per valutare la bontà dell'approssimazione ottenuta.

A questo scopo calcoliamo la somma dei quadrati delle differenze tra la variabile dipendente e il suo valore medio e chiamiamo S_t questa quantità:

$$S_t = \sum_{i=1}^n (y_i - \bar{y})^2$$

La differenza tra le due quantità S_r e S_t quantifica la riduzione dell'errore ottenuta assumendo come modello una linea retta anziché una quantità costante. Normalizzando rispetto a S_t si definisce il **coefficiente di correlazione** r e il

coefficiente di determinazione r^2 : $r^2 = \frac{S_t - S_r}{S_t}$

Nel caso di un'approssimazione ideale (la retta passa per tutti i punti) risulta $S_r=0$ e $r^2=1$. In questo caso la retta di regressione rappresenta perfettamente l'andamento variabile dei dati. Se invece $r^2=0$ l'approssimazione con la retta non porta alcun vantaggio rispetto a quella col solo valore medio.

Esempio:

Stimare gli errori per l'approssimazione lineare ai minimi quadrati dell'esempio precedente.

La deviazione standard campione risulta:
$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{22.7143}{7-1}} = 1.9457$$

l'errore standard della stima:
$$\sigma_{y/x} = \sqrt{\frac{S_r^2}{n-2}} = \sqrt{\frac{2.9911}{7-2}} = 0.7735$$

Pertanto poiché $\sigma_y > \sigma_{y/x}$ il modello dato dalla regressione lineare risulta efficace. Per valutare il miglioramento calcoliamo il coefficiente di determinazione e quello di

correlazione
$$r^2 = \frac{S_t - S_r}{S_t} = \frac{22.7143 - 2.9911}{22.7143} = 0.868; \quad r = \sqrt{0.868} = 0.932$$

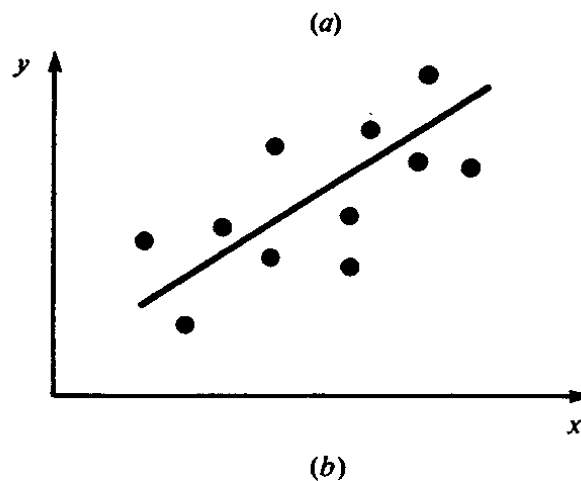
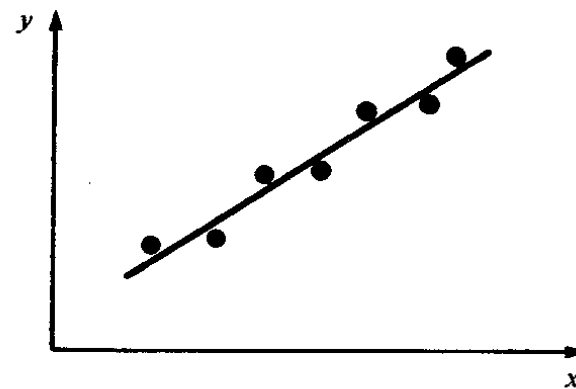


Figura 10.5 Esempi di regressione lineare con errori residui (a) piccoli e (b) grandi.

IPOTESI ASSUNTE SULLA REGRESSIONE LINEARE

Vale la pena di sottolineare alcune ipotesi statistiche che sono state assunte nella discussione relativa all'approccio lineare dei minimi quadrati:

- Le variabili indipendenti x hanno valore fisso, non sono variabili casuali e sono misurate senza errore
- Le variabili dipendenti y sono tra loro indipendenti e sono tutte caratterizzate dalla stessa varianza.
- Le y corrispondenti a una data x devono essere disposte secondo la distribuzione normale.

Queste ipotesi influiscono sulla derivazione della regressione e sul suo uso.

LINEARIZZAZIONE DI RELAZIONI NON LINEARI

La regressione lineare permette di scegliere la migliore retta che approssima un insieme di dati. Il presupposto è però che la relazione tra la variabile dipendente e quella indipendente sia lineare. Ciò non sempre si verifica.

La figura 10.8 mostra un insieme di dati che richiedono in maniera evidente un'approssimazione di tipo curvilinea e pertanto la regressione lineare risulta scarsamente efficace. Si potrebbe pensare di utilizzare una regressione polinomiale come in fig. 10.8b e, questo argomento sarà discusso nel seguito.

In alternativa è talvolta possibile eseguire una trasformazione dei dati tale da renderli compatibili con la regressione lineare.

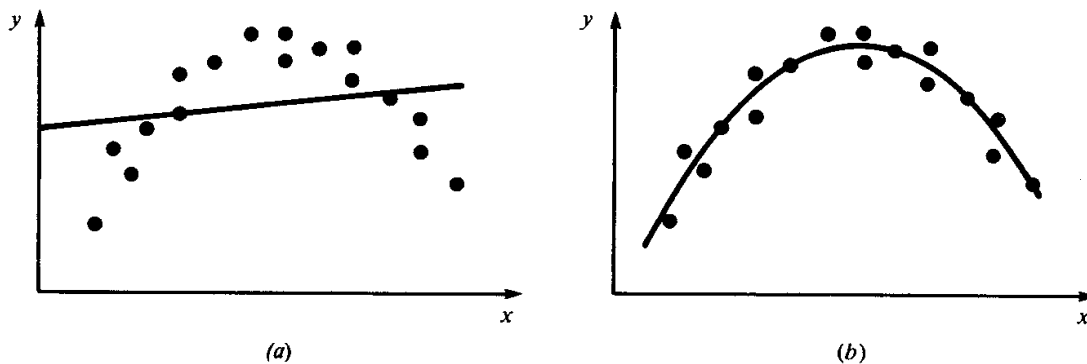


Figura 10.8 (a) Dati che non si adattano alla regressione lineare ai minimi quadrati. (b) Indicazione che un'approssimazione per mezzo di una parabola è preferibile.

Si consideri ad esempio il modello esponenziale: $y = a_1 e^{b_1 x}$ con a_1 e b_1 costanti.

Il modello è peraltro molto diffuso in ingegneria in quanto rappresenta l'integrale generale dell'equazione differenziale del primo ordine a coefficienti costanti.

Un altro esempio di modello non lineare è rappresentato dalla funzione potenza

$$y = a_2 x^{b_2} \text{ con } a_2 \text{ e } b_2 \text{ coefficienti costanti.}$$

Un terzo esempio è rappresentato dalla funzione a saturazione della velocità di

crescita:

$$y = a_3 \frac{x}{b_3 + x} \text{ con } a_3 \text{ e } b_3 \text{ coefficienti costanti.}$$

Per trattare dati che soddisfano a questi modelli, si eseguono delle manipolazioni tali da rendere lineari le funzioni sopra descritte.

Con riferimento al **modello esponenziale** è ad esempio possibile applicare la funzione di logaritmo naturale al primo e secondo membro. In tal modo si ottiene un legame lineare tra $\ln y$ e x : $\ln y = \ln(a_1) + b_1 x$

Per la **funzione potenza** la linearizzazione si effettua passando ai logaritmi in base 10. Si ottiene in tal modo un legame lineare tra $\text{Log}(y)$ e $\text{Log}(x)$:

$$\text{Log}(y) = \text{Log}(a_2) + b_2 \text{Log}(x)$$

La **funzione a saturazione di velocità** si trasforma invece facendone il reciproco:

$$\frac{1}{y} = \frac{b_3}{a_3} \frac{1}{x} + \frac{1}{a_3}. \text{ Pertanto la variabile } y^{-1} \text{ è legata linearmente alla variabile } x^{-1}$$

Nella loro forma trasformata i dati vengono utilizzati per costruire la retta di regressione lineare. I valori dei parametri calcolati possono poi essere riportati nelle funzioni originali per costruire una funzione approssimante dei dati di partenza.

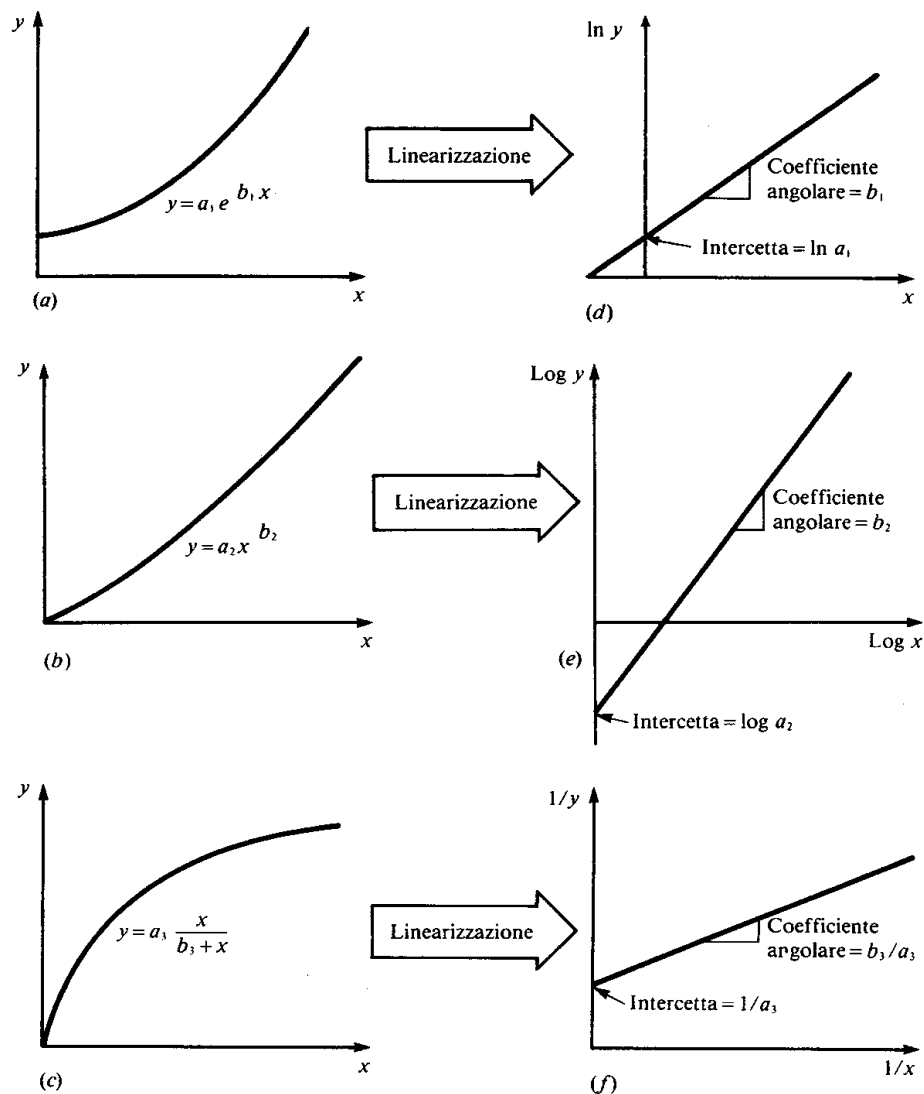


Figura 10.9 (a) Funzione esponenziale, (b) funzione potenza e (c) funzione a saturazione della velocità di crescita. Le parti (d), (e) e (f) rappresentano versioni linearizzate di queste funzioni che derivano da semplici trasformazioni.

Esempio: Trasformazione della funzione $y=a_2 x^{b_2}$

Ricerca la funzione di tipo potenza che meglio approssima i dati di tabella. Eseguite una trasformazione logaritmica dei dati.

Tabella 10.3 Dati per l'approssimazione con la funzione potenza.

x	y	$\text{Log } x$	$\text{Log } y$
1	0.5	0	-0.301
2	1.7	0.301	0.226
3	3.4	0.477	0.534
4	5.7	0.602	0.753
5	8.4	0.699	0.922

La figura 10.10(a) mostra i dati nella forma originale, mentre la figura 10.10(b) mostra i dati dopo la trasformazione logaritmica.

Eseguendo la regressione lineare dei dati trasformati si ottiene la funzione

$$\text{Log } y = 1.75 \text{ Log } x - 0.300$$

L'intercetta $\text{Log } a_2$ è uguale a -0.300; pertanto $a_2 = 10^{-0.3} = 0.5$.

La pendenza è $b_2 = 1.75$. L'equazione definitiva è pertanto $y = 0.5 x^{1.75}$

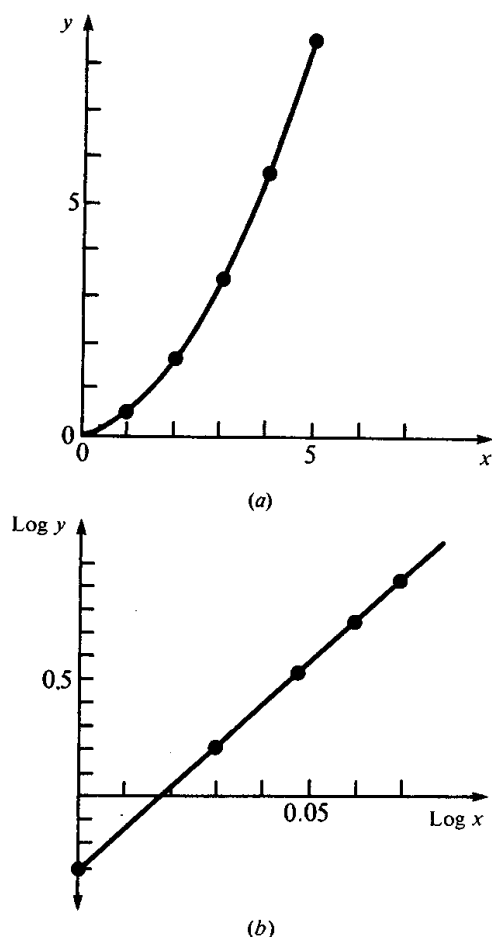


Figura 10.10 (a) Grafico dei dati non trasformati con la funzione potenza che li approssima. (b) Grafico dei dati trasformati utilizzati per determinare i coefficienti della funzione potenza.

REGRESSIONE POLINOMIALE

Come detto alcuni dati sperimentali non si sposano bene con un modello approssimante di tipo lineare. In questi casi è ragionevole utilizzare per l'approssimazione una curva polinomiale di ordine superiore.

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m$$

In questo caso la funzione da minimizzare diventa:

$$S_r = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n \left(y_i - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_m x_i^m \right)^2$$

Seguendo la procedura già vista il sistema di equazioni che permette di determinare le espressioni dei coefficienti si può ottenere eguagliando a zero le derivate rispetto ai coefficienti stessi della S_r .

$$\begin{aligned} \frac{\partial S_r}{\partial a_0} &= -2 \sum_{i=1}^n \left(y_i - a_0 - a_1 x_i - \dots - a_m x_i^m \right) = 0 \\ \frac{\partial S_r}{\partial a_1} &= -2 \sum_{i=1}^n \left(y_i - a_0 - a_1 x_i - \dots - a_m x_i^m \right) x_i = 0 \\ \frac{\partial S_r}{\partial a_m} &= -2 \sum_{i=1}^n \left(y_i - a_0 - a_1 x_i - \dots - a_m x_i^m \right) x_i^m = 0 \end{aligned} \quad (2)$$

Come si vede, si è pervenuti ad un sistema di $m+1$ equazioni lineari che va risolto in funzione degli $m+1$ coefficienti incogniti $a_j, j=0, m$

Rispetto a quanto detto per la regressione lineare, va aggiornata la definizione dell'errore standard della stima:

$$\sigma_{y/x} = \sqrt{\frac{S_r}{n - (m + 1)}}$$

Esempio: Utilizzando la regressione polinomiale approssimare i dati che compaiono nelle prime due colonne della tabella 10.4

Tabella 10.4

x_i	y_i	$(y_i - \bar{y})^2$	$(y_i - a_0 - a_1 x_i - a_2 x_i^2)^2$
0	2.1	544.44	0.14332
1	7.7	314.47	1.00286
2	13.6	140.03	1.08158
3	27.2	3.12	0.80491
4	40.9	239.22	0.61951
5	61.1	1272.11	0.09439
Σ	152.6	2513.39	3.74657

Dai dati si ricava

$$m = 2; n = 6; \sum_{i=1}^n x_i = 15; \sum_{i=1}^n x_i^4 = 979; \sum_{i=1}^n y_i x_i^2 = 2488.8;$$

$$\sum_{i=1}^n y_i = 152.6; \sum_{i=1}^n y_i x_i = 585.6; \sum_{i=1}^n x_i^2 = 55;$$

$$\sum_{i=1}^n x_i^3 = 225; \bar{x} = 2.5; \sum_{i=1}^n y_i = 152.6; \bar{y} = 25.433$$

Sostituendo questi valori nel sistema (2) si ottiene: $y = a_0 + a_1 x + a_2 x^2$
 con $a_0 = 2.47857; a_1 = 2.35929; a_2 = 1.86071;$

I coefficienti di determinazione e di correlazione sono $r^2 = 0.99851; r = 0.99925$

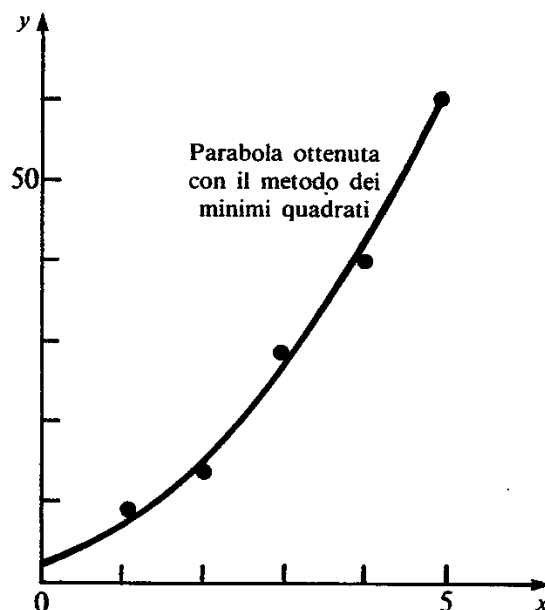


Figura 10.11 Approssimazione per mezzo di una funzione polinomiale del secondo ordine.

INTERPOLAZIONE

Quando viene richiesto di **stimare valori intermedi tra dati noti con precisione** è possibile ricorrere ad una tecnica numerica detta **Interpolazione polinomiale**.

L'espressione generale di una funzione polinomiale di ordine n è:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad (1)$$

Se sono assegnati $n+1$ punti esiste una e una sola polinomiale di ordine n (o inferiore) che passa per tutti i punti.

Ad esempio, esiste una sola retta che congiunge 2 punti,
una sola parabola che congiunge tre punti, etc.

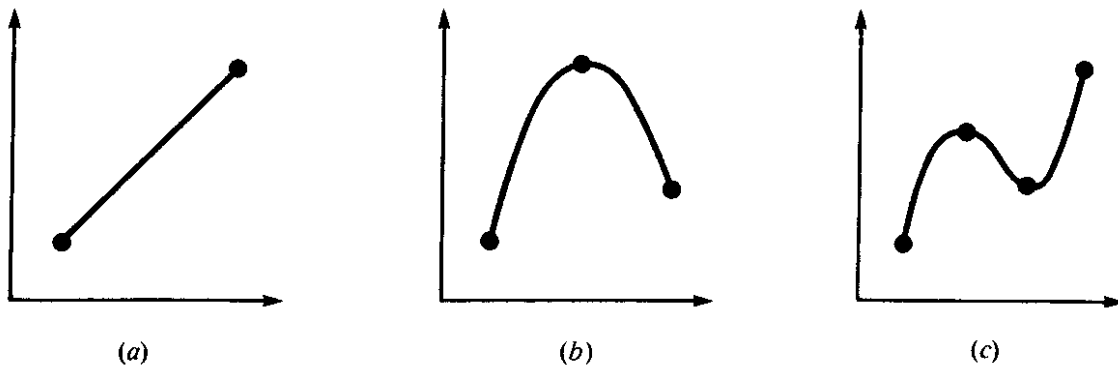


Figura 11.1 Esempi di polinomiali interpolanti: (a) del primo ordine (lineare) che connette due punti; (b) del secondo ordine (quadratica o parabolica) che connette tre punti e (c) del terzo ordine (cubica) che connette quattro punti.

Sebbene esista una sola polinomiale di ordine n -mo che passa per $n+1$ punti, le forme matematiche che essa può assumere sono molteplici.

In questa parte del corso vedremo come è possibile determinare i coefficienti nel caso di due forme polinomiali alternative note col nome di **formula di Newton** e **formula di Lagrange**.

Formula di interpolazione di Newton alle differenze divise Interpolazione lineare

La più semplice procedura di interpolazione consiste nel collegare due punti con una retta (si parla in tal caso di **interpolazione lineare**).

Usando i teoremi dei triangoli simili si ottiene la formula d'interpolazione lineare:

$$\frac{f_1(x) - f(x_0)}{x - x_0} = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad \rightarrow \quad f_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) \quad (2)$$

dove la notazione $f_1(x)$ indica che la funzione interpolante è del primo ordine. Si noti che il coefficiente angolare della retta è anche l'approssimazione alle differenze finite della derivata prima della $f(x)$

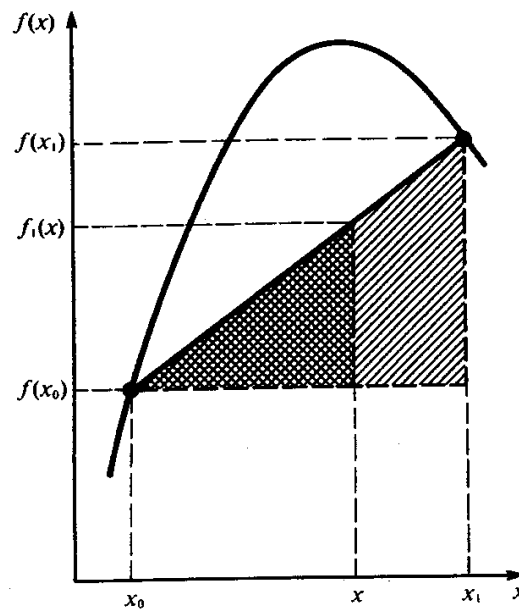


Figura 11.2 Rappresentazione grafica dell'interpolazione lineare. Le aree tratteggiate indicano i triangoli simili che vengono usati per ricavare la formula d'interpolazione lineare (11.2).

Esempio: Si determini il logaritmo naturale di 2 (valore esatto:0.69314718) attraverso l'interpolazione lineare prendendo come estremi dell'intervallo

a) $\ln 1=0$ e $\ln 6=1.7917595$

b) $\ln 1=0$ e $\ln 4=1.3862944$

a) Applicando la formula (2) si ottiene:

$$f_1(2) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) = f(0) + \frac{f(6) - f(1)}{6 - 1}(2 - 1) =$$

$$= 0 + \frac{f(6) - f(1)}{6 - 1} * 1 = \frac{1.7917595 - 0}{5} = 0.35835190$$

con un errore percentuale $\epsilon_r=48.3\%$

b) Sempre attraverso la formula (2) si ricava:

$$f_1(2) = f(0) + \frac{f(4) - f(1)}{4 - 1}(2 - 1) = 0 + \frac{f(4) - f(1)}{4 - 1} * 1 = \frac{1.3862944 - 0}{3} = 0.46209813$$

con un errore percentuale sceso a $\epsilon_r=33.3\%$

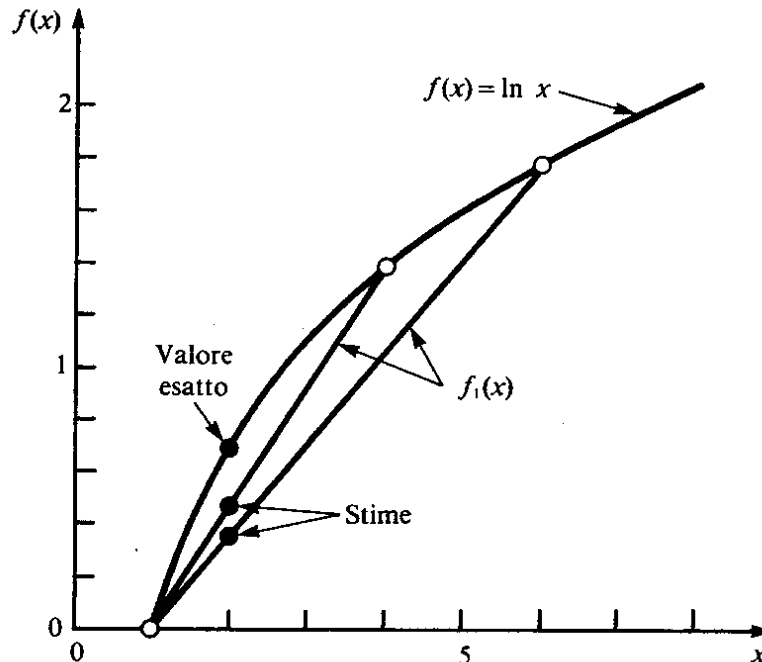


Figura 11.3 Due interpolazioni lineari per stimare $\ln 2$. Osservate come l'intervallo più piccolo consenta un'approssimazione migliore.

Formula di interpolazione di Newton alle differenze divise Interpolazione quadratica

Come si è visto nell'esempio precedente, gli errori ottenuti nel risultato erano alquanto marcati. Il motivo di questo comportamento è legato al fatto che la curva logaritmica viene approssimata in modo molto approssimativo da una retta.

Se il valore della curva logaritmica è disponibile in un terzo punto, è possibile ottenere una stima decisamente migliore. In questo caso come funzione approssimante può essere scelto un polinomio del secondo ordine. Tra le molte espressioni possibili, risulta particolarmente comoda la seguente:

$$f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) \quad (3)$$

che con opportune manipolazioni si dimostra essere equivalente all'espressione generale (1): $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$.

I coefficienti della funzione possono essere determinati in maniera semplice:

- Per calcolare b_0 si può sostituire $x=x_0$ nella (3). In tal modo si ottiene:

$$b_0 = f(x_0) \quad (4)$$

- Per calcolare b_1 si può sostituire $x=x_1$ nella (3). In tal modo si ottiene:

$$b_1 = \frac{f(x_1) - f(x_0)}{(x_1 - x_0)} \quad (5)$$

- Infine per il calcolo di b_2 attraverso la (3), la (4) e la (5) si ottiene:

$$b_2 = \frac{\frac{f(x_2) - f(x_1)}{(x_2 - x_1)} - \frac{f(x_1) - f(x_0)}{(x_1 - x_0)}}{(x_2 - x_0)} \quad (6)$$

N.B.

- Il termine b_1 rappresenta ancora il coefficiente angolare della retta che congiunge i punti x_0 e x_1 .
- Il termine b_2 tiene conto della curvatura del secondo ordine ed assomiglia all'approssimazione della derivata seconda.
- La (3) ricorda uno sviluppo approssimato in serie di Taylor.

Esempio: Si determini il logaritmo naturale di 2 (valore esatto:0.69314718) attraverso l'interpolazione quadratica. I punti noti sono:

$$\ln 1 = 0$$

$$\ln 4 = 1.3862944$$

$$\ln 6 = 1.7917595$$

Soluzione:

Dalle formule precedenti si ottiene:

$$b_0 = f(x_0) = 0$$

$$b_1 = \frac{f(x_1) - f(x_0)}{(x_1 - x_0)} = \frac{1.3862944 - 0}{(4 - 1)} = 0.46209813$$

$$b_2 = \frac{\frac{f(x_2) - f(x_1)}{(x_2 - x_1)} - \frac{f(x_1) - f(x_0)}{(x_1 - x_0)}}{(x_2 - x_0)} = \frac{\frac{1.7917595 - 1.3862944}{(6 - 4)} - \frac{1.3862944 - 0}{(4 - 1)}}{(6 - 1)} = \frac{0.20273255 - 0.46209813}{5} = -0.051873116$$

Da cui sostituendo questi valori si ottiene:

$$f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) = 0 + 0.46209813(x - 1) - 0.051873116(x - 1)(x - 4)$$

che, per $x=2$, restituisce $f_2(2) = 0.56584436$

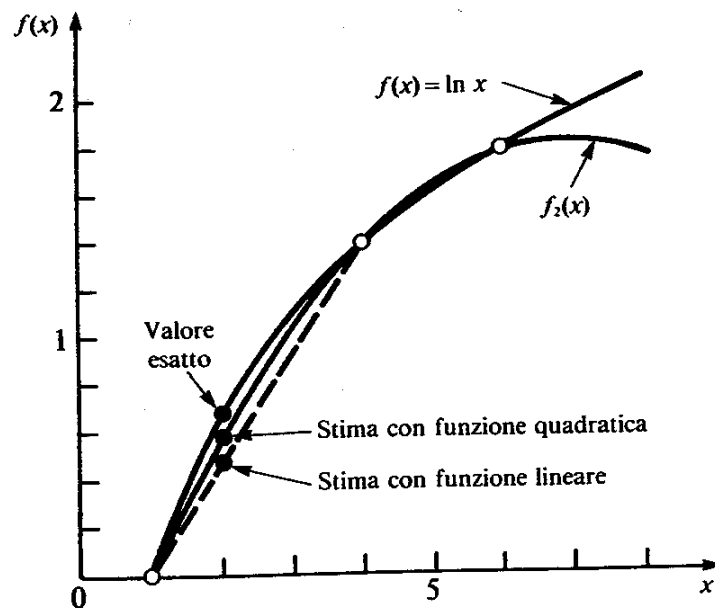


figura 11.4 Uso dell'interpolazione quadratica per stimare $\ln 2$. Per confronto è stata inclusa anche l'interpolazione lineare per x che va da 1 a 4.

Espressione generale della formula di Newton

L'analisi condotta può essere estesa al caso di una funzione polinomiale di ordine n -mo che deve passare per $n+1$ punti assegnati:

$$f_n(x) = b_0 + b_1(x - x_0) + \dots + b_n(x - x_0)(x - x_1)\dots(x - x_{n-1}) \quad (7)$$

I coefficienti b_0, b_1, \dots, b_n possono essere determinati come nei casi precedenti attraverso le differenze divise (rappresentate nel seguito con le parentesi quadre)

$$\begin{aligned} b_0 &= f(x_0) \\ b_1 &= f[x_1, x_0] \\ b_2 &= f[x_2, x_1, x_0] \\ &\vdots \\ b_n &= f[x_n, x_2, x_1, x_0] \end{aligned} \quad (8)$$

Ad esempio, la prima differenza divisa viene rappresentata come:

$$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{(x_i - x_j)}$$

La seconda differenza divisa è data da:

$$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{(x_i - x_k)}$$

La n -sima differenza divisa è data da

$$f[x_n, x_{n-1}, \dots, x_1, x_0] = \frac{f[x_n, x_{n-1}, \dots, x_1] - f[x_{n-1}, \dots, x_1, x_0]}{(x_n - x_0)}$$

I coefficienti b_j possono essere calcolati tramite le differenze divise applicate in maniera ricorsiva.

i	x_i	$f(x_i)$	prima	seconda	terza
0	x_0	$f(x_0)$	$f[x_1, x_0]$	$f[x_2, x_1, x_0]$	$f[x_3, x_2, x_1, x_0]$
1	x_1	$f(x_1)$	$f[x_2, x_1]$	$f[x_3, x_2, x_1]$	
2	x_2	$f(x_2)$	$f[x_3, x_2]$		
3	x_3	$f(x_3)$			

Figura 11.5 Rappresentazione grafica della natura ricorsiva delle differenze finite divise utilizzate nella formula di Newton.

Risulta infine:

$$\begin{aligned} f_2(x) &= f(x_0) + f[x_1, x_0](x - x_0) + \dots \\ &+ f[x_n, x_{n-1}, \dots, x_1, x_0](x - x_0)(x - x_1)\dots(x - x_{n-1}) \end{aligned}$$

Esempio: Determinare $\ln(2)$ utilizzando un polinomio del terzo ordine. Si assumano come punti noti: $\ln 1=0$, $\ln 4=1.3862944$, $\ln 6=1.7917595$, $\ln 5=1.6094379$. La funzione polinomiale del terzo ordine risulta:

$$f_3(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) + b_3(x - x_0)(x - x_1)(x - x_2)$$

Le prime differenze divise assumono i seguenti valori:

$$f[x_1, x_0] = \frac{f(x_1) - f(x_0)}{(x_1 - x_0)} = \frac{1.3862944 - 0}{(4 - 1)} = 0.46209813$$

$$f[x_2, x_1] = \frac{f(x_2) - f(x_1)}{(x_2 - x_1)} = \frac{1.7917595 - 1.3862944}{(6 - 4)} = 0.20273255$$

$$f[x_3, x_2] = \frac{f(x_3) - f(x_2)}{(x_3 - x_2)} = \frac{1.6094379 - 1.7917595}{(5 - 6)} = 0.18232160$$

Le seconde differenze divise valgono:

$$f[x_2, x_1, x_0] = \frac{0.20273255 - 0.46209813}{(6 - 1)} = -0.051873116$$

$$f[x_3, x_2, x_1] = \frac{0.18232160 - 0.20273255}{(5 - 6)} = -0.020410950$$

La terza differenza divisa vale:

$$f[x_3, x_2, x_1, x_0] = \frac{-0.020410950 + 0.051873116}{(5 - 1)} = 0.0078655415$$

Le espressioni di $f(x_0)$, $f[x_1, x_0]$, $f[x_2, x_1, x_0]$, $f[x_3, x_2, x_1, x_0]$ coincidono coi coefficienti b_0 , b_1 , b_2 , b_3 . Il polinomio approssimante di Newton del terzo ordine è pertanto:

$$f_3(x) = 0 + 0.46209813(x - 1) - 0.051873116(x - 1)(x - 4) + 0.0078655415(x - 1)(x - 4)(x - 6)$$

che valutato per $x=2$ fornisce $f(2)=0.62876869$ con un errore relativo $\varepsilon_r=9.3\%$

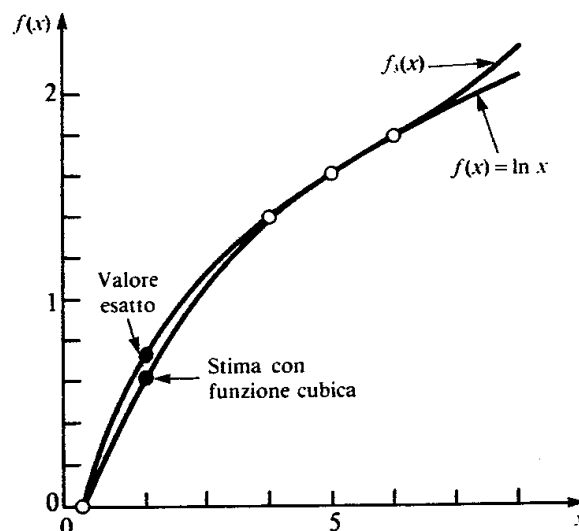


Figura 11.6 Uso dell'interpolazione cubica per stimare $\ln 2$.

ANALISI DELL'ERRORE PER LA FORMULA DI NEWTON

La formula di Newton richiama alla mente uno sviluppo della funzione in serie di Taylor nel quale le derivate analitiche sono state approssimate numericamente. Utilizzando le tecniche cui si è fatto in precedenza ricorso per la serie di Taylor è allora possibile definire l'errore di troncamento anche per la formula di Newton. Allo scopo si ricorda che l'errore di troncamento della serie di Taylor è uguale a:

$$R_n = f^{n+1}(\xi) / (n+1)! (x_{i+1} - x_i)^{n+1} \text{ dove } \xi \in [x_i, x_{i+1}]$$

Si definisce allora l'errore di troncamento per la formula di Newton come:

$$R_n = f^{n+1}(\xi) / (n+1)! (x - x_0)(x - x_1) \dots (x - x_n) \text{ dove } \xi \in [x_0, x_n]$$

Perché la formula sia utilizzabile la funzione approssimata dovrebbe essere nota e derivabile, condizioni queste che in genere non sono soddisfatte. Si può allora utilizzare la formula alternativa nella quale compare la differenza finita di ordine $n+1$ e che è valida sotto condizioni molto meno restrittive:

$$R_n = f[x, x_n, x_{n-1}, \dots, x_0] / (n+1)! (x - x_0)(x - x_1) \dots (x - x_n) \quad (9)$$

Si noti che siccome nell'espressione (9) compare l'incognita $f(x)$, essa non può essere utilizzata per determinare la stima dell'errore. D'altra parte se è noto il valore della funzione da approssimare in un ulteriore punto $f(x_{n+1})$, la stima dell'errore può essere ottenuta attraverso l'espressione:

$$R_n = f[x_{n+1}, x_n, x_{n-1}, \dots, x_0] (x - x_0)(x - x_1) \dots (x - x_n) \quad (10)$$

Esempio: Tramite l'espressione (10) stimare l'errore commesso nell'esempio precedente nel quale si è utilizzata una funzione polinom. del secondo ordine come funzione approssimante. Utilizzare come punto ulteriore $f(x_3) = f(5) = 1.6094379$.

Soluzione: Tramite la polinomiale del secondo ordine si era ottenuto $f(2) = 0.56584436$ con un errore

$$E_t = 0.69314718 - 0.56584436 = 0.127302835$$

Attraverso la (10) si ricava:

$$R_n = f[x_3, x_2, x_1, x_0] (x - x_0)(x - x_1)(x - x_2) = 0.0078655415(x - 1)(x - 4)(x - 6)$$

Che valutata per $x=2$ restituisce: $R_n = 0.0078655415(2 - 1)(2 - 4)(2 - 6) = 0.062924332$
Dello stesso ordine di grandezza dell'errore effettivo.

Caratteristiche dell' algoritmo per la formula di Newton

E' possibile rimarcare tre caratteristiche della formula di Newton che la rendano particolarmente adeguata per l'implementazione su un calcolatore:

- Le versioni di ordine superiore della formula possono essere ottenute per iterazione aggiungendo un solo termine alla formula di grado inferiore. Questa caratteristica risulta vantaggiosa quando l'ordine della funzione polinomiale approssimante non è noto a priori, ma viene ottenuto durante la fase di calcolo attraverso la valutazione dell'errore approssimato relativo commesso. Il confronto dell'errore stimato con una tolleranza prefissata rappresenta il criterio di terminazione dell'algoritmo.
- Le differenze divise utilizzate per determinare i coefficienti della funzione interpolante possono essere calcolate attraverso una formula iterativa: differenze divise di ordine inferiore vengono utilizzate per ottenere le differenze divise di ordine immediatamente superiore.
- L'espressione per la stima dell'errore commesso nell'interpolazione viene ricavata utilizzando le stesse differenze divise calcolate per la determinazione dei coefficienti del polinomio interpolante. Questa fase richiede pertanto un costo computazionale marginale.

FORTRAN

```

DIMENSION FX(10,10),X(10)
READ(5,1)N
1 FORMAT(15)
DO 140 I=1,N
  READ(5,2)X(I),FX(I,1)
2 FORMAT(2F10.0)
140 CONTINUE
  N=N-1
  DO 200 J=1,N
    K=J+1
    NP=N-J
    DO 190 I=1,NP
      FX(I,K)=(FX(I+1,J)-FX(I,J))/(X(I+J)-X(I))
190 CONTINUE
200 CONTINUE
    DO 230 J=1,N
      WRITE(6,3)FX(1,J)
3 FORMAT(' ',F10.3)
230 CONTINUE
      READ(5,2)XI
      FA=1.
      Y=0.
      DO 340 J=1,N
        Y=Y+FX(1,J)*FA
        WRITE(6,3)Y
        FA=FA*(XI-X(J))
        IF(J.GE.N)GO TO 350
        JP=J+1
        ER=FA*FX(1,JP)
        WRITE(6,3)ER
340 CONTINUE
350 STOP
      END
  
```

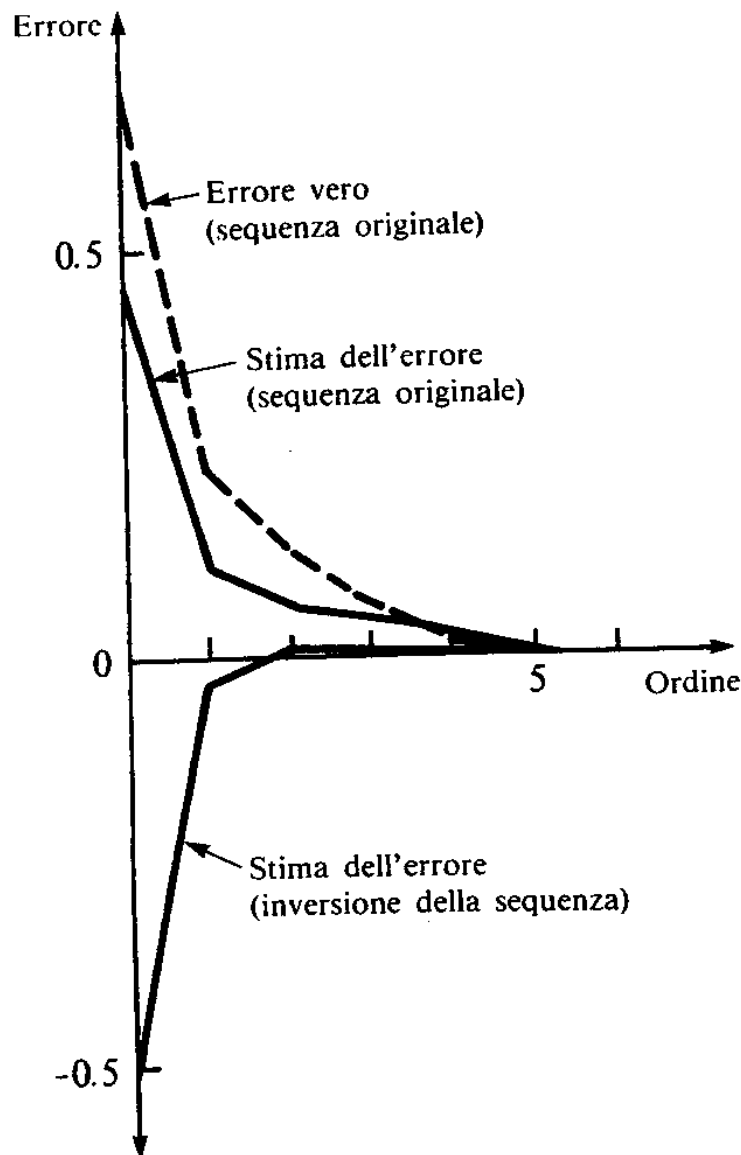
BASIC

```

100 DIM FX(10,10),X(10)
110 INPUT N
120 FOR I = 1 TO N
130 INPUT X(I),FX(I,1)
140 NEXT I
150 FOR J = 1 TO N - 1
160 K = J + 1
170 FOR I = 1 TO N - J
180 FX(I,K) = (FX(I + 1,J) - FX(I
,J)) / (X(I + J) - X(I))
190 NEXT I
200 NEXT J
210 FOR J = 1 TO N
220 PRINT FX(1,J)
230 NEXT J
240 INPUT XI
250 FA = 1
260 Y = 0
270 FOR J = 1 TO N
280 Y = Y + FX(1,J) * FA
290 PRINT Y
300 FA = FA * (XI - X(J))
310 IF J > = N THEN 350
320 ER = FA * FX(1,J + 1)
330 PRINT ER
340 NEXT J
350 END
  
```

N = numero di punti
 X(I) = vettore della variabile indipendente
 FX(I, 1) = matrice la cui prima colonna contiene la variabile dipendente
 (Loop che calcola le differenze finite divise)
 (Visualizzazione dei coefficienti della polinomiale interpolante)
 XI = valore della variabile indipendente per la quale si richiede l'interpolazione
 (Loop che calcola interpolazione degli ordini da 0 a n-1)
 ER = Stima dell'errore

Figura 11.7 Programma per la formula di Newton.



x	$f(x) = \ln x$
1	4
4	1.3862944
6	1.7917595
5	1.6094379
3	1.0986123
1.5	0.40546511
2.5	0.91629073
3.5	1.2527630

Fig.[11.8] Errore relativo percentuale associato all'approssimazione di $\ln 2$ in funzione dell'ordine della polinomiale interpolante.

Si noti dalla figura come

- un'approssimazione del quinto ordine fornisca una soluzione adeguata.
- le curve dell'errore stimato e dell'errore effettivo risultino simili
- I risultati migliorano decisamente quando nella sequenza dei valori noti (vedi tabella) utilizzati per costruire la polinomiale interpolante intervengono punti più vicini a quello incognito ($x=2$).
- A conferma di ciò, se si inverte l'ordine progressivo dei punti assunti per costruire il polinomio interpolante i risultati migliorano decisamente (già col polinomio di ordine 2 l'errore è estremamente contenuto $\epsilon_t < 2\%$).

FORMULA DI LAGRANGE

La formula di interpolazione di Lagrange non è null'altro che una riformulazione della formula di Newton nella quale si evita il calcolo delle differenze divise.

In modo compatto risulta:

$$f_n(x) = \sum_{i=0}^n L_i(x) f(x_i) \quad \text{con } L_i = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}$$

Nel caso lineare risulta: $f_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1)$

Il polinomio del secondo ordine è invece:

$$f_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f(x_0) + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f(x_1) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f(x_2)$$

Anche in questo caso è possibile ottenere una stima dell'errore:

$$\mathfrak{R}_n = f[x, x_n, x_{n-1}, \dots, x_0] \prod_{i=0}^n (x - x_i)$$

DEDUZIONE DELLA FORMULA DI LAGRANGE DA QUELLA DI NEWTON NEL CASO DEL PRIMO ORDINE

La formula di Newton del primo ordine è $f_1(x) = f(x_0) + f[x_1, x_0](x - x_0)$ (2)

Le differenze divise possono essere espresse in una forma alternativa (detta simmetrica). Ad esempio per quella del primo ordine si ha:

$$f[x_1, x_0] = \frac{f(x_1) - f(x_0)}{(x_1 - x_0)} = \frac{f(x_1)}{(x_1 - x_0)} + \frac{f(x_0)}{(x_0 - x_1)}$$

Sostituendo nella (2), raccogliendo i termini simili e semplificando si ha:

$$f_1(x) = f(x_0) + \frac{(x - x_0)}{(x_1 - x_0)} f(x_1) + \frac{(x - x_0)}{(x_0 - x_1)} f(x_0) = \frac{(x - x_1)}{(x_0 - x_1)} f(x_0) + \frac{(x - x_0)}{(x_1 - x_0)} f(x_1)$$

Esempio: Utilizzando le formule di Lagrange del primo e del secondo ordine determinare il $\ln(2)$.

Si assumano noti $\ln(1)=0$, $\ln(4)=1.3862944$, $\ln(6)=1.7917595$

Soluzione: Nel caso lineare la formula di Lagrange porta al seguente risultato:

$$f_1(2) = \frac{x-x_1}{x_0-x_1} f(1) + \frac{x-x_0}{x_1-x_0} f(4) = \frac{2-4}{1-4} 0 + \frac{2-1}{4-1} 1.3862944 = 0.4620981$$

In modo simile utilizzando l'approssimazione polinomiale del secondo ordine :

$$\begin{aligned} f_2(2) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f(1) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f(4) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f(6) = \\ &= \frac{(2-4)(2-6)}{(1-4)(1-6)} 0 + \frac{(2-1)(2-6)}{(4-1)(4-6)} 1.3862944 + \frac{(2-1)(2-4)}{(6-1)(6-4)} 1.7917595 = 0.565844 \end{aligned}$$

Questi risultati corrispondono (a meno degli errori di arrotondamento) con quelli ottenuti attraverso la formula di Newton, pari rispettivamente per i casi lineare e quadratico a 0.46209813 e a 0.56584436.

VANTAGGI DELLA FORMULA DI NEWTON E VANTAGGI DELLA FORMULA DI LAGRANGE

- Nei casi in cui l'ordine della polinomiale non è noto a priori risulta più vantaggioso utilizzare l'algorithmo di Newton perché
 1. consente di analizzare il comportamento di polinomiali di ordine diverso senza ridondanze computazionali
 2. è facile introdurre nei calcoli la stima dell'errore numerico perché allo scopo si richiede la conoscenza delle differenze divise già calcolate per la definizione dei coefficienti della polinomiale.
- D'altro canto, quando l'ordine della polinomiale è noto a priori è preferibile utilizzare la formula di Lagrange perché
 1. è più facile da programmare
 2. è meno soggetta ad errori di arrotondamento

Interpolazione con dati uniformemente distanziati

Nonostante il fatto che le formule di Newton e Lagrange valgano per dati distribuiti in modo generici, vale la pena di ricordare le **tecniche che permettono di interpolare dati distanziati l'uno dall'altro dal medesimo intervallo** che, a parte l'interesse storico (prima dell'avvento dei calcolatori queste erano in effetti gli unici strumenti matematici che consentissero la definizione di funzioni interpolanti), risultano utili nel ricavare formule di integrazione numerica.

Con dati uniformemente distanziati le variabili indipendenti, disposte in ordine crescente, assumono i seguenti valori: $x_n = x_0 + n h$, dove h è l'ampiezza del passo costante tra i dati.

Le differenze divise possono essere espresse in forma compatta: ad esempio la seconda differenza divisa in avanti è:

$$f[x_2, x_1, x_0] = \frac{\frac{f(x_0) - f(x_1)}{x_0 - x_1} - \frac{f(x_1) - f(x_2)}{x_1 - x_2}}{(x_0 - x_2)} = \frac{f(x_0) - 2f(x_1) + f(x_2)}{2h^2} = \frac{\Delta^2 f(x_0)}{2!h^2}$$

dove con $\Delta^2 f(x_0)$ si è denotata la seconda differenza in avanti.

L'espressione precedente si ricava come caso particolare da quella generale:

$$f[x_0, x_1, \dots, x_n] = \frac{\Delta^n f(x_0)}{n!h^n},$$

attraverso la quale possiamo ricavare la formula di Newton per dati equi-distanziati come (R_n è definita dalle espressioni 9-10 precedenti):

$$f_n(x) = f(x_0) + \frac{\Delta f(x_0)}{h}(x - x_0) + \dots + \frac{\Delta^n f(x_0)}{n!h^n}(x - x_0)(x - x_0 - h)\dots(x - x_0 - (n-1)h) + R_n$$

che, posto $\alpha = (x - x_0)/h$, conduce con una ulteriore semplificazione alla **formula in avanti di Newton—Gregory**, utile per ricavare formule di integrazione

$$f_n(x) = f(x_0) + \Delta f(x_0)\alpha + \dots + \frac{\Delta^n f(x_0)}{n!}\alpha(\alpha - 1)\dots(\alpha - n + 1) + R_n$$

Dove:

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1} \alpha(\alpha - 1)\dots(\alpha - n)$$

ESTRAPOLAZIONE

Il procedimento tramite il quale si determina il valore di $f(x)$ per x esterno all'intervallo dei dati noti prende il nome di estrapolazione.

Precedentemente si è affermato che l'interpolazione è maggiormente accurata quando il punto da stimare occupa una posizione centrale rispetto all'insieme dei dati noti. Nel caso dell'estrapolazione questa condizione ovviamente non viene verificata e, pertanto, l'errore associato può essere notevole.

Allo scopo si osservi la figura riportata di seguito. Siccome si cerca di predire un valore di $f(x)$ esterno rispetto alla regione dei dati noti, può accadere che la funzione polinomiale diverga ampiamente rispetto alla funzione da approssimare. Per questi motivi i risultati ottenuti tramite processi di estrapolazione vanno valutati con grande attenzione e senso critico prima di essere utilizzati per prendere delle decisioni.

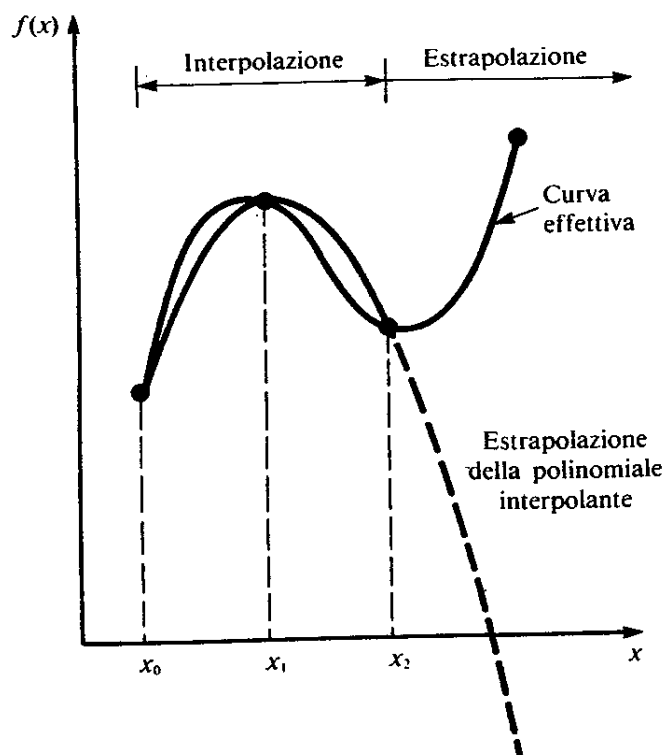


Figura 11.11 Possibile divergenza dei valori estrapolati. L'estrapolazione è basata sulla parabola che passa per i primi tre punti.

INTERPOLAZIONE MEDIANTE FUNZIONI SPLINE

Esistono dei casi pratici nei quali l'interpolazione di $n+1$ punti attraverso una funzione polinomiale di ordine n (come proposto nei metodi precedenti) fornisce risultati inadeguati.

È allora possibile fare ricorso ad un approccio alternativo attraverso le cosiddette funzioni spline: interpolare la funzione raccordando diverse polinomiali di ordine inferiore ciascuna definita all'interno di un sottoinsieme dei dati disponibili.

Per esempio, le curve del terzo ordine utilizzate per collegare i punti dati presi a coppie vengono dette **spline cubiche**.

Il vantaggio di questo approccio è di consentire il raccordo senza discontinuità tra le diverse funzioni spline.

Si potrebbe pensare che un'approssimazione tramite diverse funzioni del terzo ordine sia meno precisa di una ottenuta tramite una unica funzione di ordine superiore. Ciò non è sempre vero, ad esempio, quando la funzione interpolata presenti brusche variazioni all'interno della regione di interesse.

Si vede in particolare come le funzioni polinomiali interpolanti di ordine elevato tendano ad oscillare in modo ampio in presenza di brusche variazioni di $f(x)$.

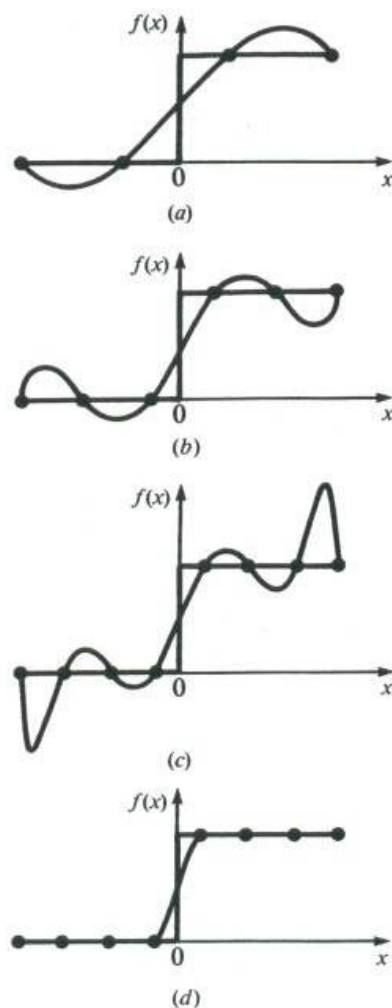


Figura 11.12 Rappresentazione visiva di una situazione in cui le spline danno risultati migliori delle polinomiali interpolanti di ordine superiore. La funzione da approssimare mostra una brusca variazione per $x = 0$. (a), (b) e (c) indicano che una brusca variazione induce oscillazioni nelle polinomiali interpolanti. Le spline cubiche (d), al contrario, forniscono un'approssimazione molto più fedele in quanto sono curve di basso ordine (terzo) raccordate senza punti angolosi.

SPLINE LINEARI

La funzione polinomiale più semplice che permette di collegare due punti è ovviamente la retta. Se intendiamo collegare tra di loro senza discontinuità un gruppo di n punti presi nell'ordine delle x crescenti si possono utilizzare ovviamente $(n-1)$ linee rette del tipo:

$$f(x)=f(x_i)+m_i (x-x_i) \text{ per } x \in [x_i, x_{i+1}], \text{ con } m_i=(f(x_{i+1})-f(x_i))/(x_{i+1}-x_i)$$

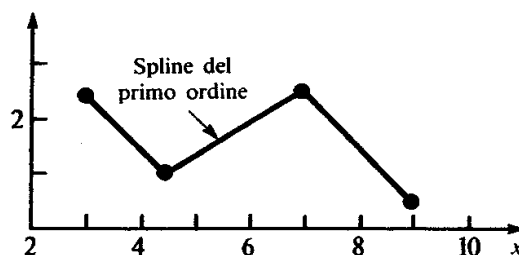
Esempio: Attraverso la conoscenza dei dati forniti in tabella determinare il valore della funzione per $x=5$.

Soluzione Utilizzando i dati forniti si ricava il coefficiente angolare della retta interpolante: $m=(2.5-1.0)/(7.0-4.5)=0.60$

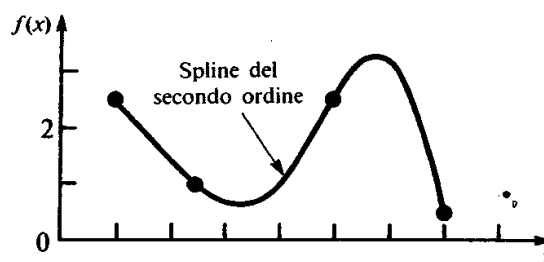
Da cui la funzione interpolante vale:

$$f(x)=1.0+0.60*(x-4.5) \rightarrow f(5)=1.0+0.6*0.5=1.3$$

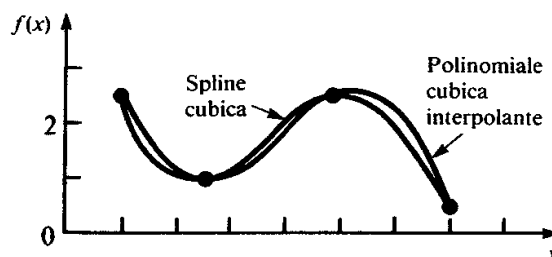
x	$f(x)$
3.0	2.5
4.5	1.0
7.0	2.5
9.0	0.5



(a)



(b)



(c)

Figura 11.14 Interpolazione di quattro punti mediante spline: (a) spline lineare, (b) spline quadratica e (c) spline cubica con, sovrapposta, una polinomiale interpolante cubica.

SPLINES QUADRATICHE

Il principale svantaggio delle splines lineari è la presenza di punti angolosi laddove splines differenti vengono raccordate (i punti di raccordo sono detti **nodi**).

Questo implica che nei nodi la derivata prima della funzione approssimante è discontinua. Per essere certi che la derivata di ordine m-simo sia continua è necessario utilizzare splines di ordine m+1 o superiore.

Nella pratica le splines più utilizzate sono quelle cubiche che consentono di ottenere una interpolante che non presenta discontinuità nella derivata prima e seconda.

Siccome l'illustrazione del concetto di interpolazione per le splines cubiche risulta alquanto complesso, lo si presenterà nel caso di splines quadratiche per le quali è la sola derivata prima a mantenersi continua in corrispondenza dei nodi.

In questo caso le coppie consecutive di nodi vengono raccordate tramite una polinomiale del secondo ordine del tipo:

$$f(x) = a_i x^2 + b_i x + c_i \text{ per } x \in [x_i, x_{i+1}]$$

Nel caso di n+1 punti da raccordare i coefficienti incogniti sono $3 \cdot n$ e pertanto per determinarli occorre scrivere $3 \cdot n$ equazioni. Le condizioni da imporre sono:

1. I valori delle n funzioni interpolanti devono essere uguali tra loro e a quelli della $f(x)$ in corrispondenza dei nodi interni.

Questo si traduce nelle seguenti $2 \cdot (n-1)$ equazioni:

$$\begin{aligned} f(x_{i-1}) &= a_{i-1} (x_{i-1})^2 + b_{i-1} x_{i-1} + c_{i-1} & \text{per } i=1:(n-1) \\ f(x_i) &= a_i (x_i)^2 + b_i x_i + c_i \end{aligned}$$

2. La prima e l'ultima equazione devono passare per i corrispondenti punti esterni (2 equazioni):

$$\begin{aligned} f(x_0) &= a_0 (x_0)^2 + b_0 x_0 + c_0 \\ f(x_n) &= a_n (x_n)^2 + b_n x_n + c_n \end{aligned}$$

3. Le derivate prime di ciascuna coppia di splines adiacenti devono assumere lo stesso valore nel nodo comune (n-1 equazioni)

$$2 a_{i-1} x_i + b_{i-1} = 2 a_i x_i + b_i \text{ per } i=1:(n-1)$$

In totale abbiamo $2n+n-1=3n-1$ equazioni.

4. Tra le infinite possibili scelte per l'ultima equazione si impone che la derivata seconda si annulli nel primo punto (che corrisponde a unire i primi due punti con una retta): $a_1=0$

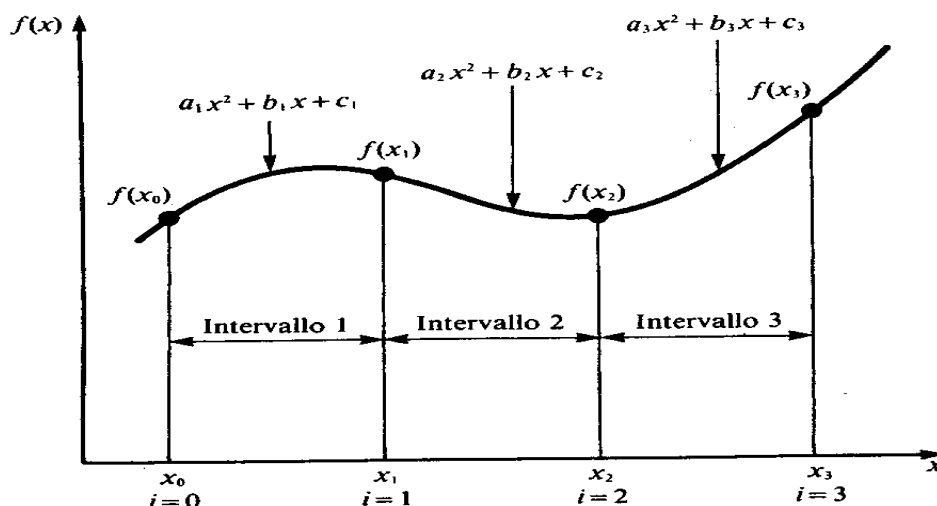


Figura 11.15 Notazione usata per ricavare splines quadratiche. Osservate che ci sono n intervalli e $n + 1$ punti. L'esempio mostrato si riferisce al caso di $n = 3$.

Esempio: Utilizzando i dati dell'esempio precedente stimare il valore corrispondente al punto $x=5$ utilizzando un'interpolazione con splines quadratiche.

I 4 punti noti sono $(3,2.5)$, $(4.5,1.0)$, $(7.0,2.5)$, $(9.0,0.5)$. Si avranno 3 splines e 9 incognite. Le equazioni da scrivere sono:

L'imposizione dei valori noti nei punti interni comporta $2 \cdot (3-1) = 4$ equazioni

$$20.25 a_1 + 4.5 b_1 + c_1 = 1.0$$

$$20.25 a_2 + 4.5 b_2 + c_2 = 1.0$$

$$49.00 a_2 + 7.0 b_2 + c_2 = 2.5$$

$$49.00 a_3 + 7.0 b_3 + c_3 = 2.5$$

L'imposizione dei valori noti nei due punti esterni comporta due equazioni

$$9.00 a_1 + 3.0 b_1 + c_1 = 2.5$$

$$81.00 a_3 + 9.0 b_3 + c_3 = 0.5$$

La continuità della derivata prima comporta altre $(3-1) = 2$ equazioni

$$9.00 a_1 + b_1 = 9.00 a_2 + b_2$$

$$14.00 a_1 + b_1 = 14.00 a_2 + b_2$$

L'annullamento della derivata seconda nella prima spline implica $a_1 = 0$

Risolvendo il sistema si perviene alle seguenti equazioni:

$$f_1(x) = -x + 5.5 \quad 3.0 \leq x \leq 4.5$$

$$f_2(x) = 0.64 x^2 - 6.76 x + 18.46 \quad 4.5 \leq x \leq 7.0$$

$$f_3(x) = -1.6 x^2 + 24.6 x - 91.3 \quad 7.0 \leq x \leq 9.0$$

$$\text{Pertanto in } x=5 \text{ risulta } f_2(5) = 0.64 \cdot 5^2 - 6.76 \cdot 5 + 18.46 = 0.66$$

Dalla figura si evidenziano due aspetti negativi: i primi due punti sono interpolati tramite una retta. La spline nell'ultimo intervallo arriva troppo in alto

SPLINES CUBICHE

Le splines cubiche connettono i nodi tramite funzioni polinomiali del terzo ordine:

$$f(x) = a_i x^3 + b_i x^2 + c_i x + d_i \text{ per } x \in [x_i, x_{i+1}]$$

Nel caso di $n+1$ punti noti ci sono n intervalli e conseguentemente $4*n$ incognite.

Le $4*n$ condizioni da imporre sono:

1. I valori delle n funzioni interpolanti devono essere uguali tra loro e a quelli della $f(x)$ in corrispondenza dei nodi interni.

Questo si traduce nelle seguenti $2*(n-1)$ equazioni:

$$\begin{aligned} f(x_{i-1}) &= a_{i-1} (x_{i-1})^3 + b_{i-1} (x_{i-1})^2 + c_{i-1} x_{i-1} + d_{i-1} & \text{per } i=1:(n-1) \\ f(x_i) &= a_i (x_i)^3 + b_i (x_i)^2 + c_i x_i + d_i \end{aligned}$$

2. La prima e l'ultima equazione devono passare per i corrispondenti punti esterni (2 equazioni):

$$\begin{aligned} f(x_0) &= a_0 (x_0)^3 + b_0 (x_0)^2 + c_0 x_0 + d_0 \\ f(x_n) &= a_n (x_n)^3 + b_n (x_n)^2 + c_n x_n + d_n \end{aligned}$$

3. Le derivate prime di ciascuna coppia di spline adiacenti devono assumere lo stesso valore nel nodo comune ($n-1$ equazioni)

$$3 a_{i-1} (x_i)^2 + 2 b_{i-1} x_i + c_{i-1} = 3 a_i (x_i)^2 + 2 b_i x_i + c_i \text{ per } i=1:(n-1)$$

4. Le derivate seconde di ciascuna coppia di spline adiacenti devono assumere lo stesso valore nel nodo comune ($n-1$ equazioni)

$$6 a_{i-1} x_i + 2 b_{i-1} = 6 a_i x_i + 2 b_i \text{ per } i=1:(n-1)$$

5. Le derivate seconde si devono annullare nei nodi esterni (il che corrisponde a unire i primi due punti e gli ultimi 2 con una retta) (2 equazioni)

$$\begin{aligned} 6 a_0 x_0 + 2 b_0 &= 0 \\ 6 a_n x_n + 2 b_n &= 0 \end{aligned}$$

In alternativa, se il problema da risolvere richiede una derivata seconda differente nei nodi esterni, il termine noto viene aggiustato in modo da soddisfare la richiesta.

In totale abbiamo $2n+2*(n-1)+2=4*n$ equazioni.

Sviluppando le equazioni precedenti si perviene a delle espressioni delle spline cubiche nelle quali gli unici valori incogniti sono le derivate seconde agli estremi di ciascun sottointervallo:

$$f_i(x) = \frac{f''(x_{i-1})}{6(x_i - x_{i-1})}(x_i - x)^3 + \frac{f''(x_i)}{6(x_i - x_{i-1})}(x - x_{i-1})^3 + \left[\frac{f(x_{i-1})}{(x_i - x_{i-1})} - \frac{f''(x_{i-1})(x_i - x_{i-1})}{6} \right](x_i - x) + \left[\frac{f(x_i)}{(x_i - x_{i-1})} - \frac{f''(x_i)(x_i - x_{i-1})}{6} \right](x - x_{i-1}) \quad (1)$$

Gli n-1 valori incogniti delle derivate seconde nei nodi interni possono essere calcolati attraverso le seguenti n-1 equazioni

$$(x_i - x_{i-1})f''(x_{i-1}) + 2(x_{i+1} - x_{i-1})f''(x_i) + (x_{i+1} - x_i)f''(x_{i+1}) = \frac{6}{(x_{i+1} - x_i)}[f(x_{i+1}) - f(x_i)] + \frac{6}{(x_i - x_{i-1})}[f(x_{i-1}) - f(x_i)] \quad (2)$$

Esempio: Si utilizzino le spline cubiche per interpolare i 4 punti dell'esempio precedente: (3,2.5), (4.5,1.0),(7.0,2.5),(9.0,0.5). Si stimi il valore per $x=5$. Sostituendo questi valori nelle n-2=2 equazioni (2) si ottiene per il primo nodo interno:

$$(4.5 - 3)f''(3) + 2(7 - 3)f''(4.5) + (7 - 4.5)f''(7) = \frac{6}{(7 - 4.5)}[2.5 - 1] + \frac{6}{(4.5 - 3)}[2.5 - 1]$$

Le condizioni naturali (annullamento della derivata seconda nei nodi esterni) comportano $f''(3)=0$ e $f''(9)=0$. Pertanto l'eq. precedente diventa:

$$2(7 - 3)f''(4.5) + (7 - 4.5)f''(7) = \frac{6}{(7 - 4.5)}[2.5 - 1] + \frac{6}{(4.5 - 3)}[2.5 - 1]$$

$$8f''(4.5) + 2.5f''(7) = 9.6$$

In modo analogo per il secondo nodo interno risulta: $3.5f''(4.5) + 9f''(7) = -9.6$

La soluzione del sistema è pertanto $f''(4.5)=1.74545$; $f''(7.5)=-1.74545$;

Sostituendo questi valori nelle espressioni (1) per i tre intervalli si ottiene:

$$f_1(x) = 0.193939(x - 3)^3 + 1.666667(4.5 - x) + 0.230303(x - 3)$$

$$f_2(x) = 0.116364(7 - x)^3 - 0.116364(x - 4.5)^3 - 0.327273(7 - x) + 1.727273(x - 4.5)$$

$$f_3(x) = -0.145455(9 - x)^3 + 1.831818(9 - x) + 0.25(x - 7)$$

Infine

$$f_2(5) = 0.116364(7 - 5)^3 - 0.116364(5 - 4.5)^3 - 0.327273(7 - 5) + 1.727273(5 - 4.5) = 1.1255$$