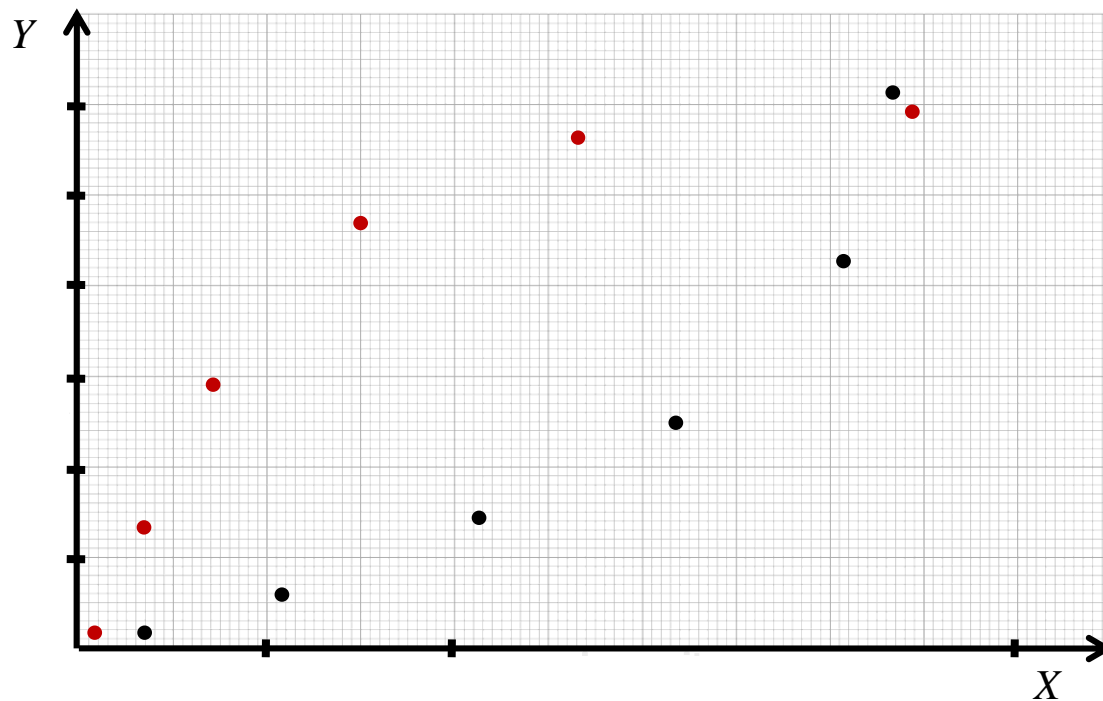


Linearizzazione



Il diagramma di dispersione suggerisce che le funzione di interpolazione dei dati non sono lineari, ma presentano un andamento che in un caso (dots neri) potrebbe essere di tipo esponenziale, mentre nell'altro caso (dots rossi) potrebbe essere di tipo radice.

Quando i dati sperimentali non evidenziano una correlazione di tipo lineare, ma presentano un andamento di tipo esponenziale, polinomiale, etc. è difficile (se non impossibile) ricavare graficamente i parametri caratteristici della curva, che spesso hanno un significato chimico-fisico ben preciso e potrebbero pertanto fornire utilissime informazioni su un certo fenomeno.

In questi casi è conveniente trasformare l'equazione che rappresenta la curva che meglio approssima i dati in quella di una retta effettuando un **cambiamento di variabili (linearizzazione)**

Il procedimento di linearizzazione consiste nell'usare una funzione delle variabili anziché le variabili stesse.

Una classe importante di linearizzazioni è quella legata ad andamenti esponenziali (o logaritmici) e a leggi di potenza. Consideriamo la funzione

$$y = Ae^{Bx}$$

Per ricondurci ad una forma più «comoda», applichiamo una "linearizzazione" della funzione applicando la funzione logaritmo:

$$\ln y = \ln A + Bx$$

Ponendo

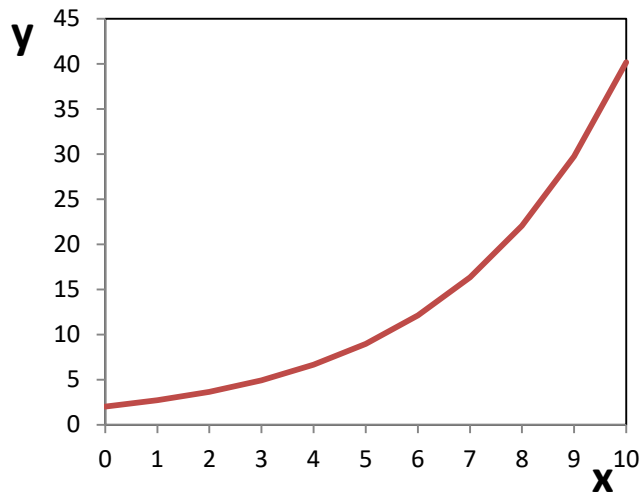
$$Y = \ln y, C = \ln A, X = x$$

Si graficano i dati mediante le nuove variabili funzionali X, Y che hanno così dipendenza lineare.

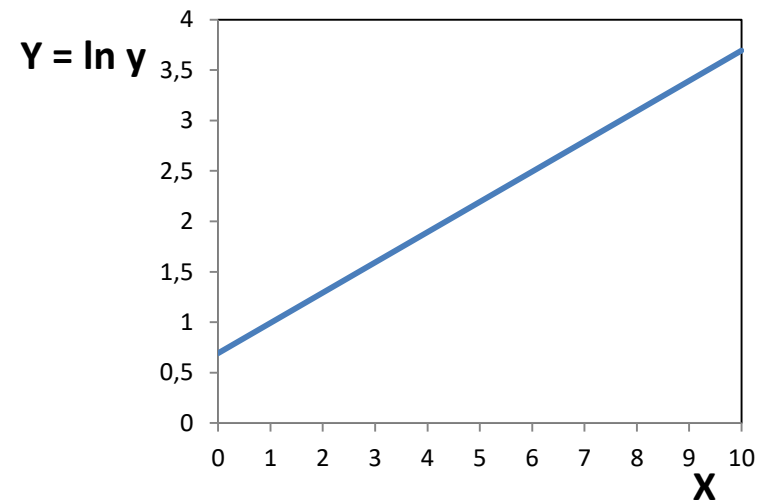
$$Y = C + BX$$

Esempio

$$y = Ae^{Bx} = 2e^{0.3x}$$



$$Y = \ln y = 0.3x + \ln 2 = 0.3x + 0.6391$$



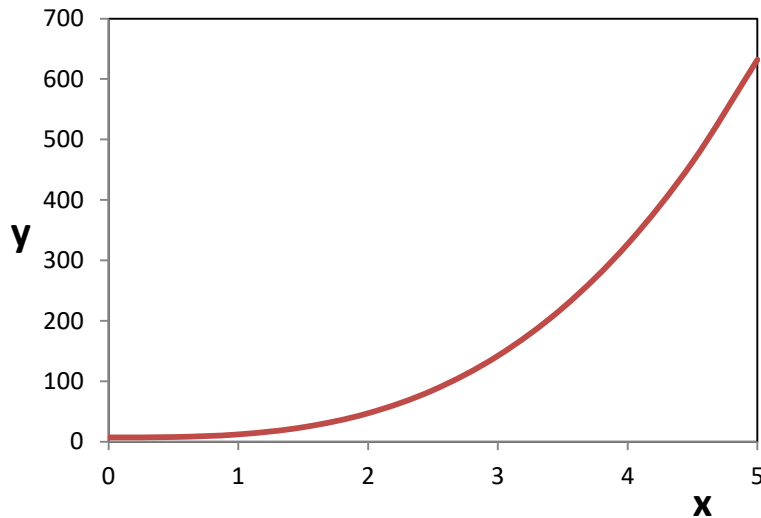
Altri esempi:

$$y = Ax^n + B$$

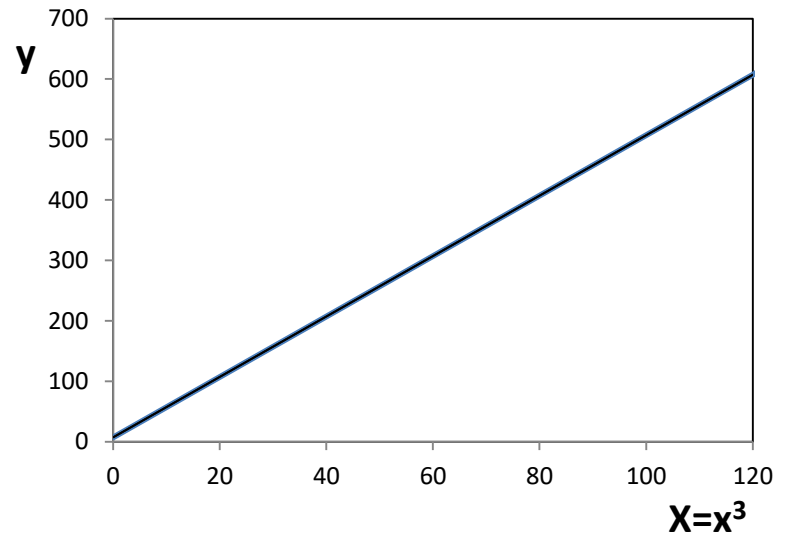
$$X = x^n$$

$$Y = y$$

$$y = 5x^3 + 7$$



$$Y = 5X + 7$$

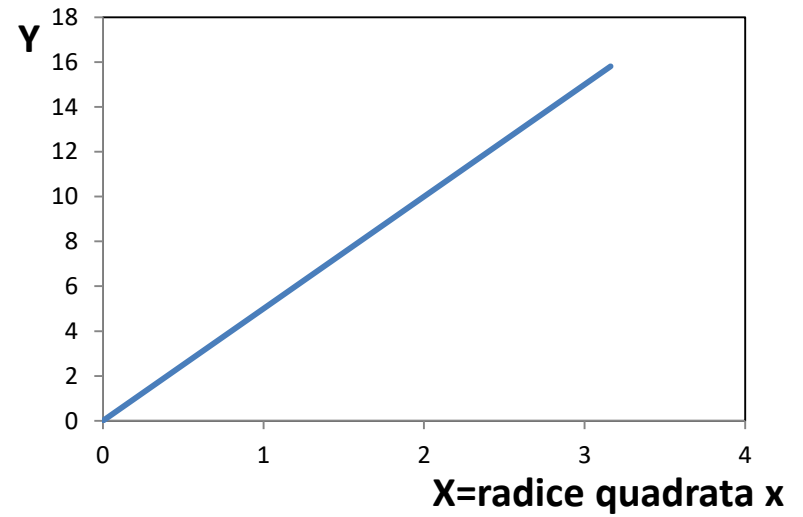
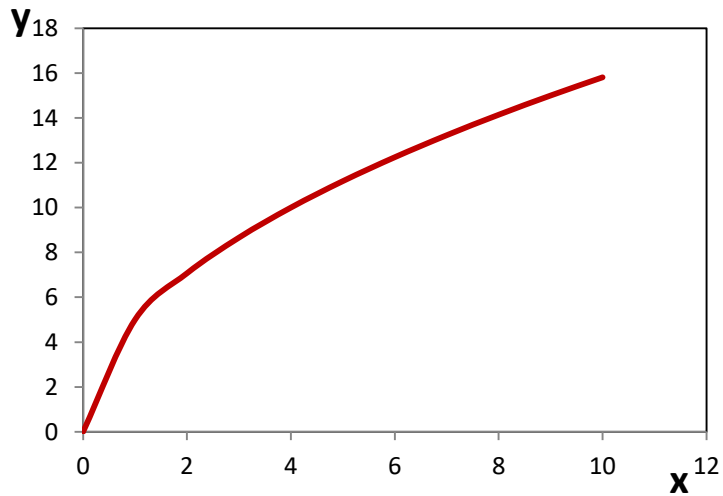


$$y = A\sqrt{x}$$

$$X = \sqrt{x} \quad Y = y$$

$$y = 5\sqrt{x}$$

$$Y = 5X$$



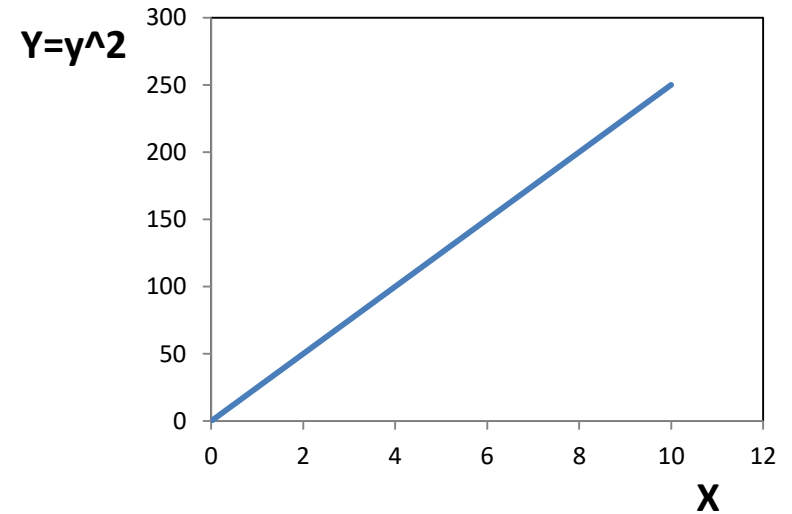
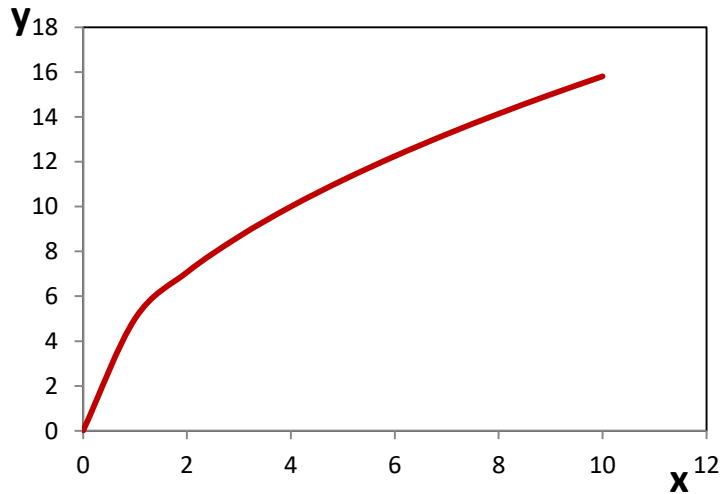
oppure

$$y = A\sqrt{x} \iff y^2 = A^2 x$$

$$X = x \quad Y = y^2$$

$$y = 5\sqrt{x}$$

$$Y = 25X$$

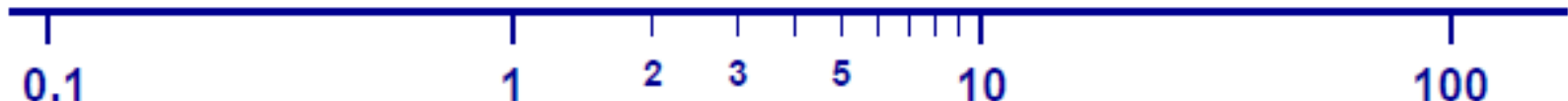


Grafici con scale non lineari

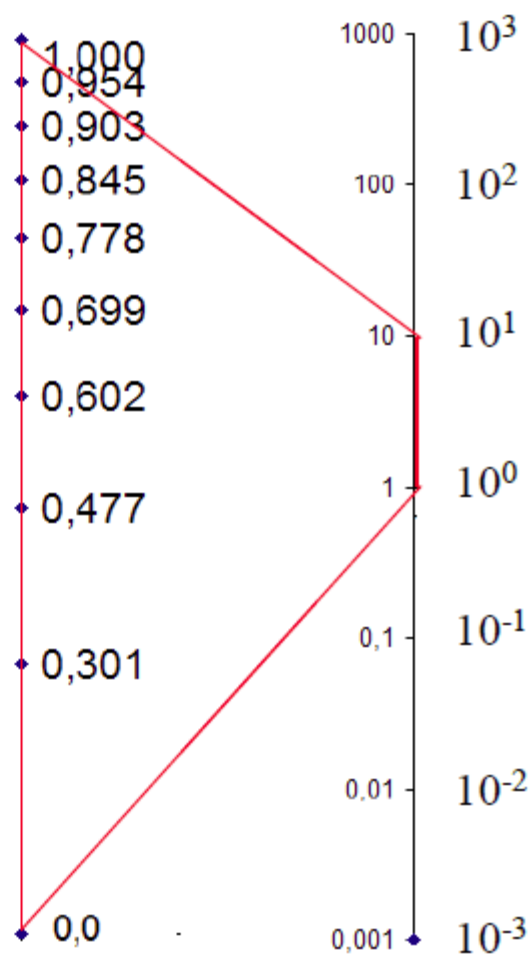
Scale Logaritmiche

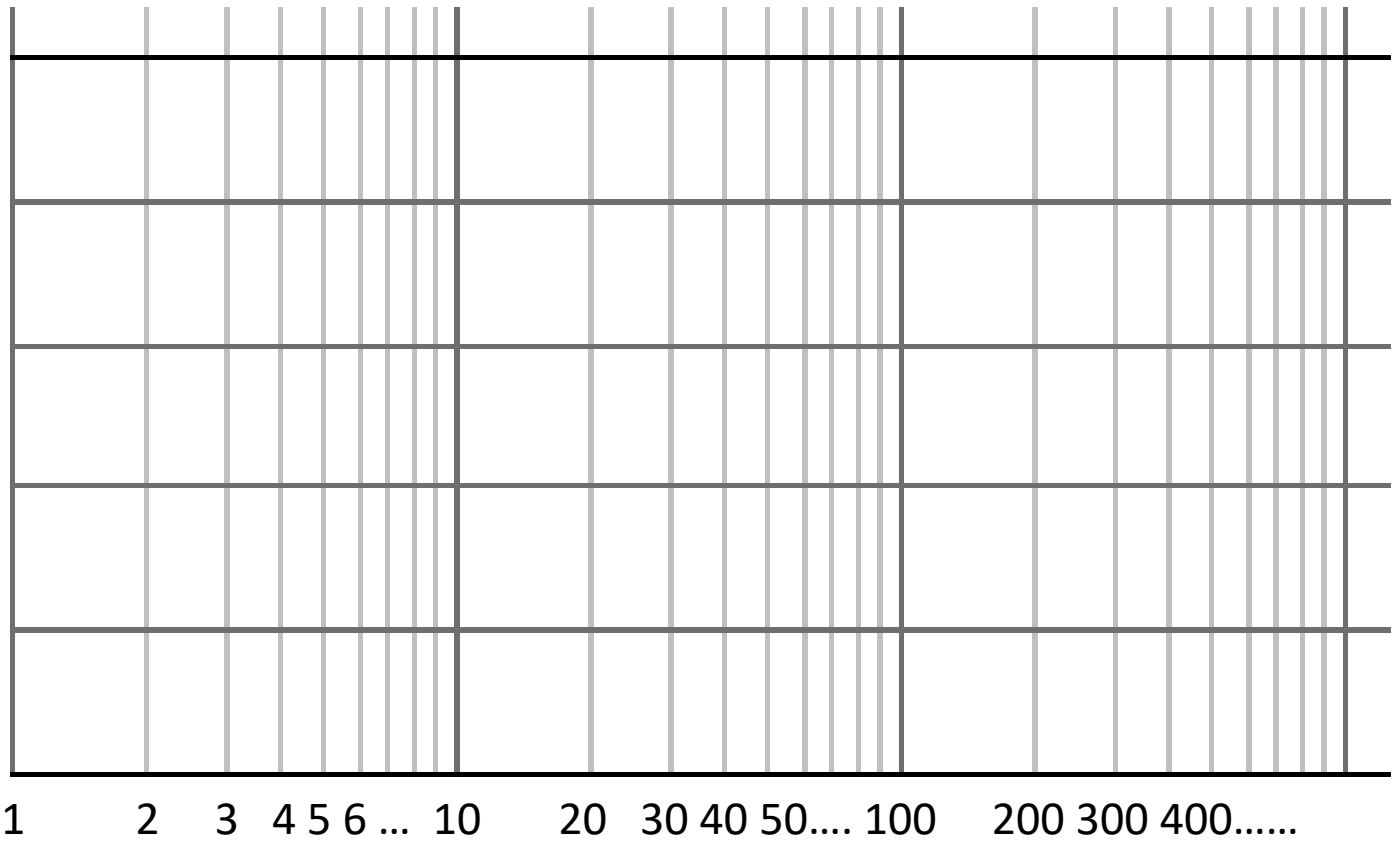
Una scala logaritmica è un asse sul quale sono riportati segmenti proporzionali ai logaritmi dei numeri

- sull'asse prescelto (ad es. asse x) si rappresenta il punto di ascissa $1 = 10^0$
- nella direzione positiva si rappresentano, a distanze uguali fra di loro, i punti di ascissa $10^1, 10^2, 10^3, \dots$
- nella direzione negativa si rappresentano, a distanze uguali fra di loro, i punti di ascissa $10^{-1}, 10^{-2}, 10^{-3}, \dots$
- i valori intermedi tra una potenza di 10 e la successiva (ad.es. 2, 3, \dots 9) sono posizionati ai valori dei rispettivi logaritmi decimali

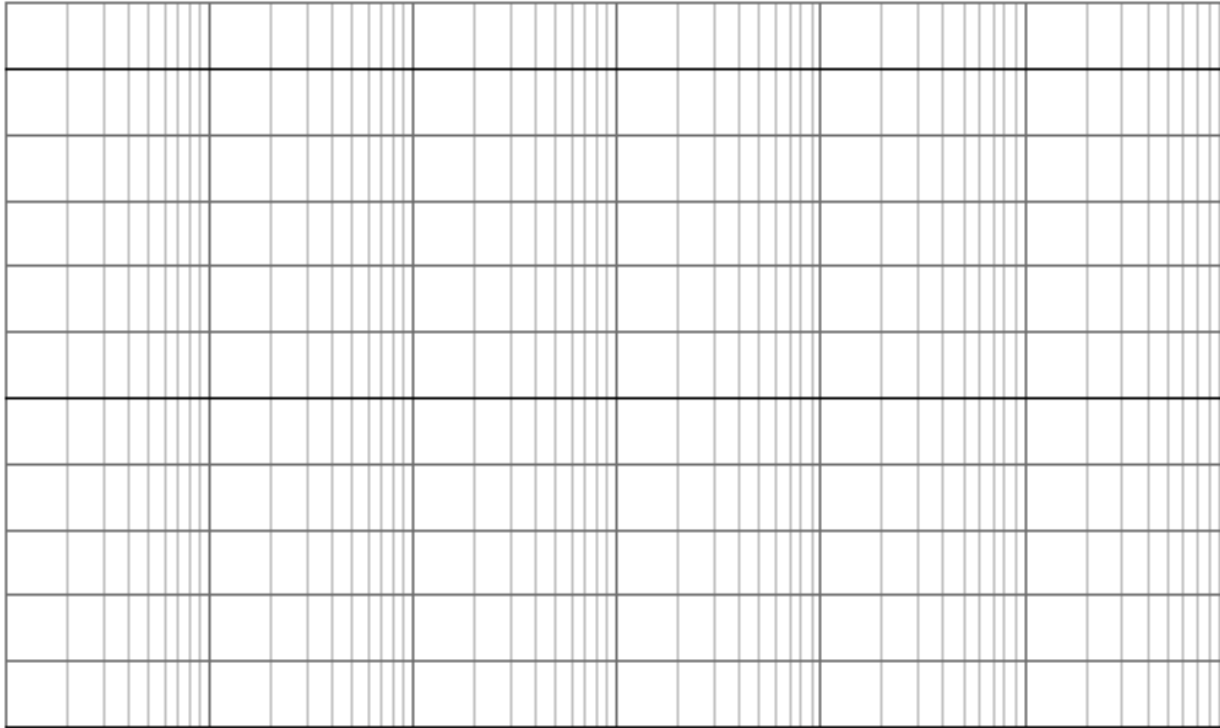


valore	logaritmo
1	0
2	0,30103
3	0,477121
4	0,60206
5	0,69897
6	0,778151
7	0,845098
8	0,90309
9	0,954243
10	1





Carta SemiLogaritmica



scala lineare sull'asse delle ascisse X e scala logaritmica sull'asse delle ordinate Y (o viceversa)

TRASFORMAZIONE DI VARIABILI

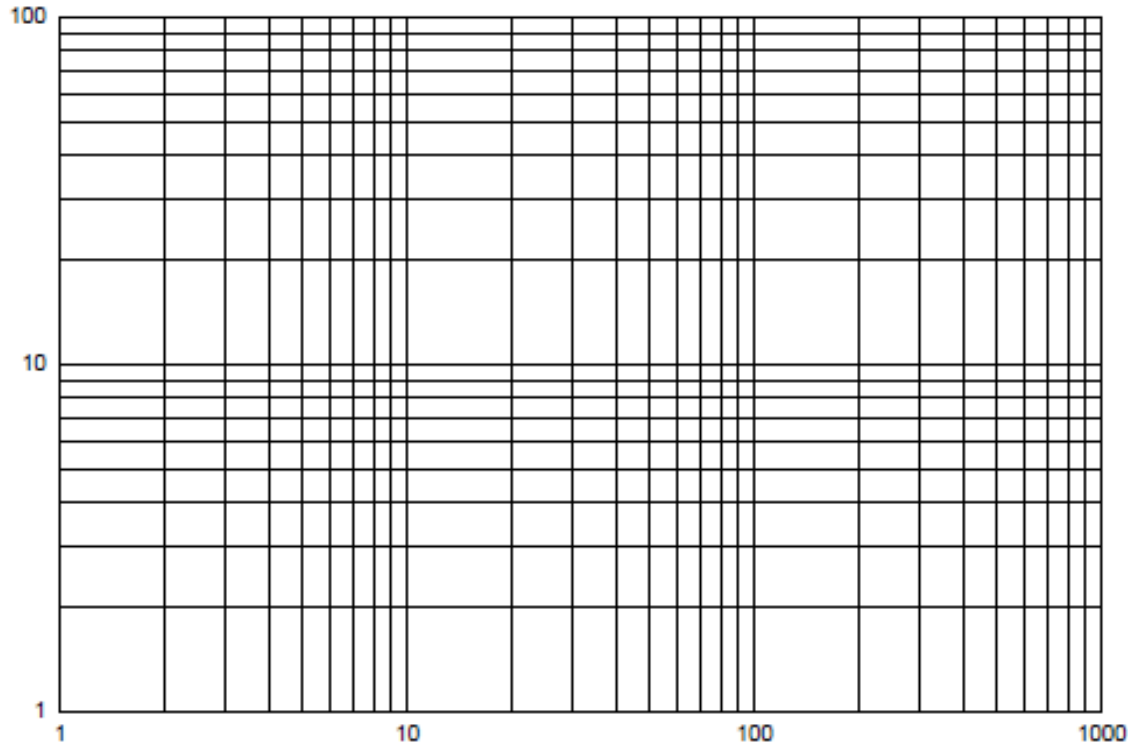
$$X = x \quad Y = \log_{10} y$$

lineare sull'asse X e logaritmica sull'asse Y

$$X = \log_{10} x \quad Y = y$$

logaritmica sull'asse X e lineare sull'asse Y

Carta Logaritmica



scala logaritmica sull'asse delle ascisse X e scala logaritmica sull'asse delle ordinate Y

TRASFORMAZIONE DI VARIABILI

$$X = \log_{10} x \quad Y = \log_{10} y$$

APPLICAZIONI:

- rappresentare misure positive con ordini di grandezza molto diversi fra loro
- linearizzare funzioni esponenziali $y = Ka^x$
scale semilogaritmiche
- linearizzare funzioni potenza $y = Bx^a$
scale logaritmiche

Carta SemiLogaritmica

Data la funzione esponenziale $y = Ka^x$

passando ai logaritmi decimali e utilizzando le proprietà dei logaritmi

$$\underline{\log_{10} y} = \log_{10} Ka^x = \log_{10} K + \log_{10} a^x =$$
$$\underline{\log_{10} K + x \log_{10} a}$$

ponendo $X = x$ $Y = \log_{10} y$

si ha l'equazione di una retta

$$Y = \log_{10} K + X \log_{10} a$$

intercetta

coefficiente angolare

Carta Logaritmica

Data la funzione potenza

$$y = Bx^a$$

passando ai logaritmi decimali e utilizzando le proprietà dei logaritmi

$$\underline{\log_{10} y} = \log_{10} Bx^a = \log_{10} B + \log_{10} x^a =$$
$$\underline{\log_{10} B + a \log_{10} x}$$

ponendo $X = \log_{10} x$ $Y = \log_{10} y$

si ha l'equazione di una retta

$$Y = \log_{10} B + aX$$

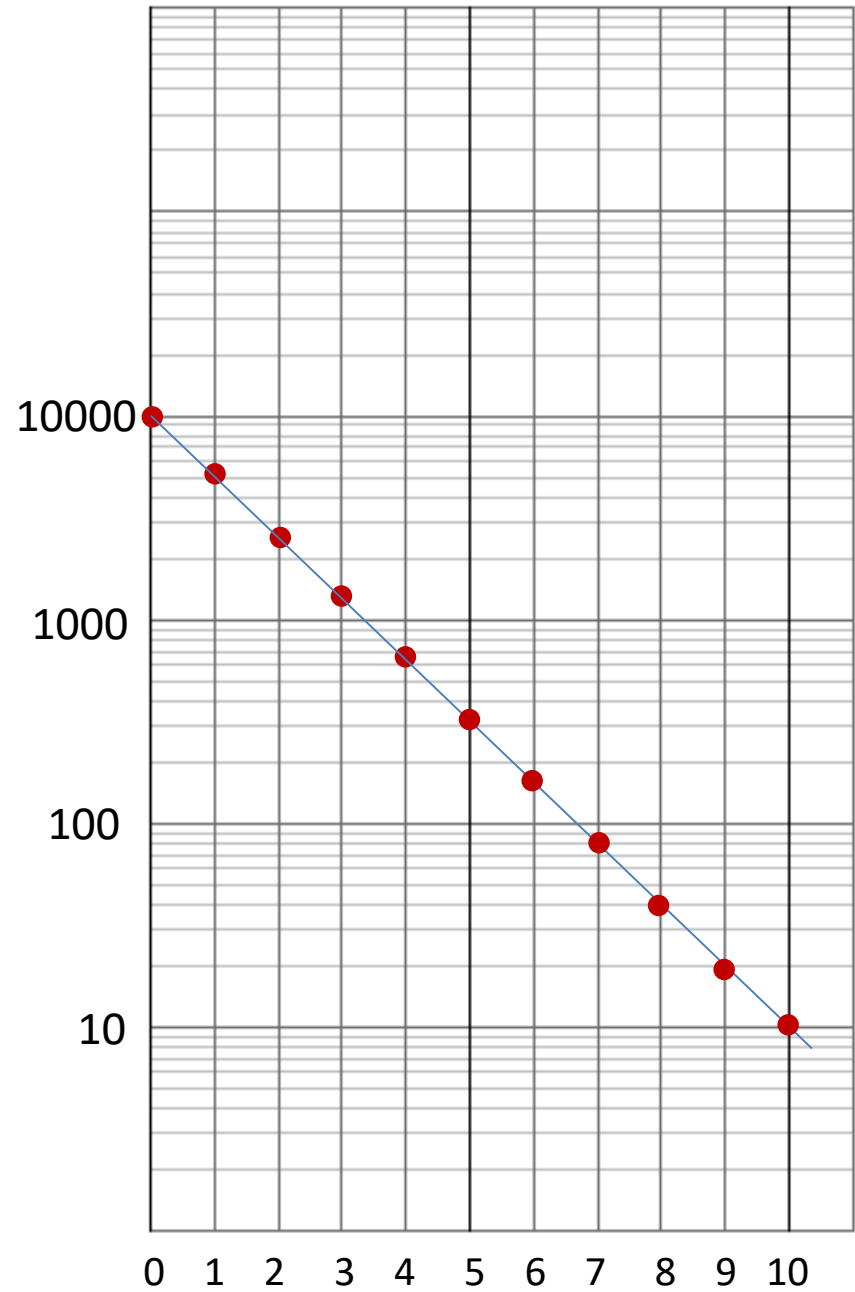
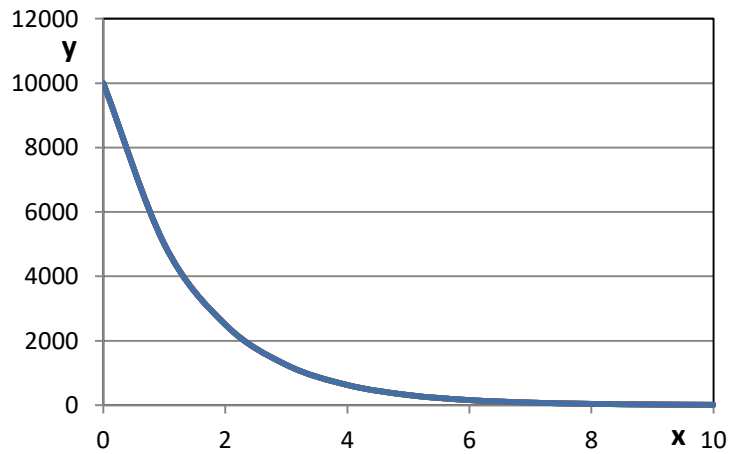
intercetta

coefficiente angolare

Esempio

$$y = 10^4 (0.5)^x$$

0	10000
1	5000
2	2500
3	1250
4	625
5	312
6	156
7	78
8	39
9	20
10	10



Retta dei minimi quadrati

Il metodo della massima e minima pendenza per determinare la retta di best-fit è un metodo grafico che funziona abbastanza bene, ma non esiste alcuna giustificazione teorica della procedura.

Affrontiamo il problema con un metodo più rigoroso.

Consideriamo N coppie di misure (x_i, y_i) di due grandezze x e y fra le quali sappiamo che esiste una relazione di tipo lineare:

$$y = ax + b$$

Supponiamo che l'incertezza su una delle due variabili (ad esempio sulla x) sia trascurabile rispetto a quella dell'altra (y)

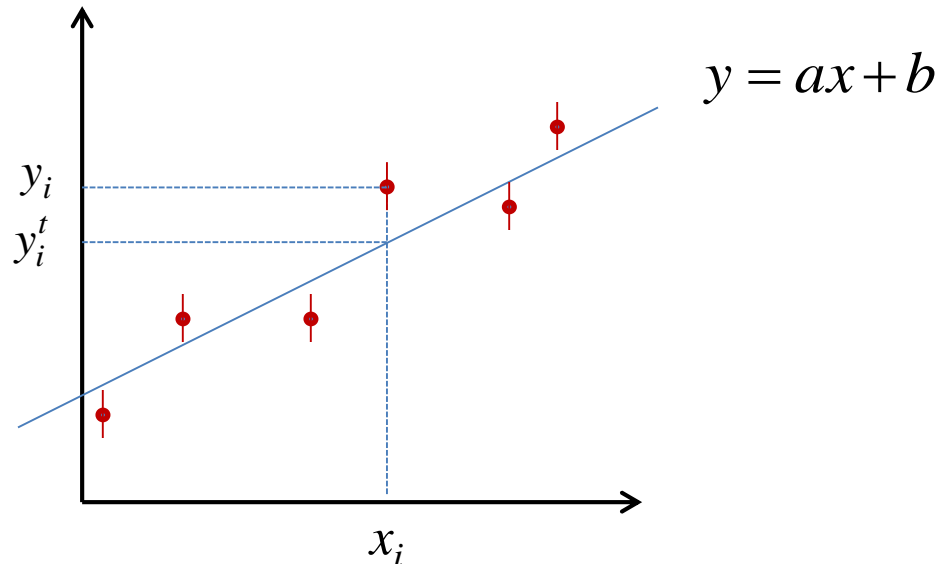


determiniamo i migliori valori di a e b corrispondenti ai dati sperimentali

In corrispondenza di ogni x_i esiste un valore “teorico” y_i^t dato dalla relazione

$$y_i^t = ax_i + b$$

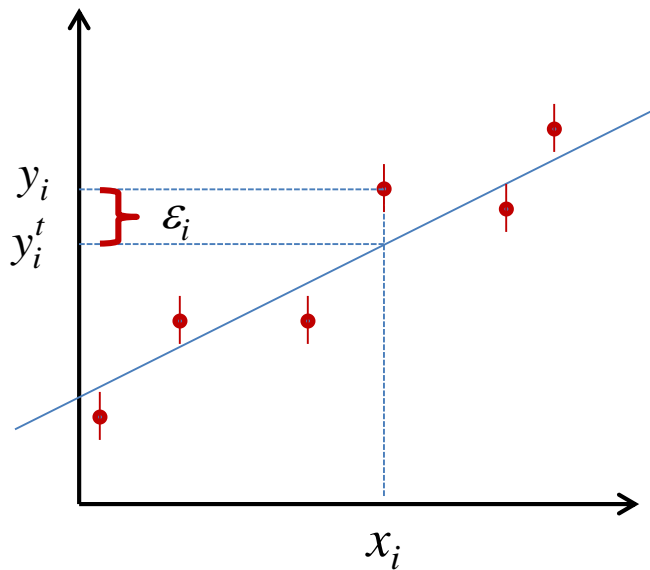
ovvero y_i^t rappresenta l'ordinata del punto di ascissa x_i appartenente alla retta, mentre y_i è quello misurato sperimentalmente in corrispondenza di x_i .



In genere, a causa degli errori di misura, $y_i^t \neq y_i$

e può essere a volte $y_i^t < y_i$ a volte $y_i^t > y_i$

La quantità $\varepsilon_i = y_i - y_i^t$ rappresenta lo scarto del valore sperimentale da quello teorico



$$\varepsilon_i = y_i - y_i^t = y_i - (ax_i + b)$$

$$y_i - ax_i - b$$

Al variare di a e b la retta cambierà pendenza o traslerà e gli scarti assumeranno valori diversi

Osservazione: un grande valore di ε_i non ha di per sè un grande significato . Ciò che conta è il valore assoluto dello scarto rispetto all'errore da cui e' affetto y_i .

La “miglior retta” deve rendere il più possibile piccoli i valori assoluti degli scarti (divisi per i rispettivi errori).

Poichè rendere piccolo uno scarto può renderne grandi altri, allora come “miglior retta” si prende quella che rende minima

la somma dei quadrati degli scarti (*divisi per i rispettivi errori*)

(Principio di Massima Verosimiglianza)

La “miglior retta” è quella che rende minima la quantità

$$z = \sum_{i=1}^N \left(\frac{\varepsilon_i}{\sigma_i} \right)^2 = \sum_{i=1}^N \frac{(y_i - ax_i - b)^2}{\sigma_i^2}$$

dove σ_i è l'incertezza statistica sulla misura y_i .

$z = z(a, b)$, al variare di a e b , assume uno ed un solo minimo in corrispondenza dei valori per cui le derivate parziali rispetto a e b si annullano

$$\frac{\partial z}{\partial a} = \frac{\partial z}{\partial b} = 0$$

$$\begin{aligned}\frac{\partial z}{\partial a} &= \frac{\partial}{\partial a} \left[\sum_{i=1}^N \frac{(y_i - ax_i - b)^2}{\sigma_i^2} \right] = \sum_{i=1}^N \frac{2(y_i - ax_i - b)(-x_i)}{\sigma_i^2} = \\ &-2 \sum_{i=1}^N \frac{x_i(y_i - ax_i - b)}{\sigma_i^2} = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial z}{\partial b} &= \frac{\partial}{\partial b} \left[\sum_{i=1}^N \frac{(y_i - ax_i - b)^2}{\sigma_i^2} \right] = \sum_{i=1}^N \frac{2(y_i - ax_i - b)(-1)}{\sigma_i^2} = \\ &-2 \sum_{i=1}^N \frac{(y_i - ax_i - b)}{\sigma_i^2} = 0\end{aligned}$$

Che equivale ad avere il sistema:

$$\left\{ \begin{array}{l} \sum_{i=1}^N \frac{x_i (y_i - ax_i - b)}{\sigma_i^2} = 0 \\ \sum_{i=1}^N \frac{(y_i - ax_i - b)}{\sigma_i^2} = 0 \end{array} \right.$$

Supponendo che le σ_i siano tutte uguali $\sigma_i = \sigma$

$$\left\{ \begin{array}{l} \sum_{i=1}^N x_i y_i - a \sum_{i=1}^N x_i^2 - b \sum_{i=1}^N x_i = 0 \\ \sum_{i=1}^N y_i - a \sum_{i=1}^N x_i - \sum_{i=1}^N b = 0 \end{array} \right.$$

Sistema di 2 equazioni
in 2 incognite

$$\begin{cases}
 \sum_{i=1}^N x_i y_i - a \sum_{i=1}^N x_i^2 - b \sum_{i=1}^N x_i = 0 \\
 \sum_{i=1}^N y_i - a \sum_{i=1}^N x_i - \sum_{i=1}^N b = 0
 \end{cases}$$

S_{xy} S_{xx} S_x
 S_y S_x bN

$$\begin{cases}
 S_{xy} - a \cdot S_{xx} - b \cdot S_x = 0 \\
 S_y - a \cdot S_x - b \cdot N = 0
 \end{cases}$$

$$\begin{cases}
 a \cdot S_{xx} + b \cdot S_x = S_{xy} \\
 a \cdot S_x - b \cdot N = S_y
 \end{cases}
 \quad \rightarrow$$

Usando la regola di Cramer

$$a = \frac{S_{xy}N - S_x S_y}{S_{xx}N - S_x^2}$$

$$b = \frac{S_{xx}S_y - S_x S_{xy}}{S_{xx}N - S_x^2}$$

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

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

I parametri a e b sono affetti da incertezza (perché lo sono le y_i).

Per la propagazione degli errori statistici:

$$\sigma^2(a) = \sum_{i=1}^N \left(\frac{\partial a}{\partial y_i} \right)^2 \cdot \sigma_{y_i}^2$$

$$\sigma^2(b) = \sum_{i=1}^N \left(\frac{\partial b}{\partial y_i} \right)^2 \cdot \sigma_{y_i}^2$$

$$\frac{\partial a}{\partial y_i} = \frac{Nx_i - \sum_{i=1}^N x_i}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2}$$

Nell'ipotesi che tutte le varianze siano uguali

$$\sigma^2(a) = \sum_{i=1}^N \left(\frac{Nx_i - \sum_{i=1}^N x_i}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2} \right)^2 \cdot \sigma_y^2 =$$

$$\frac{N}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2} \sigma_y^2$$

Analogamente si calcola

$$\sigma^2(b) = \frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2} \sigma_y^2$$

Riassumendo:

$$\sigma^2(a) = \frac{N}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2} \sigma_y^2 = \frac{N}{N^2 \left[\frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N^2} \left(\sum_{i=1}^N x_i \right)^2 \right]} \sigma_y^2 =$$

$$\frac{1}{N \left[\overline{x^2} - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2 \right]} \sigma_y^2 = \frac{1}{\overline{x^2} - (\bar{x})^2} \frac{\sigma_y^2}{N}$$

$$\sigma^2(b) = \frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2} \sigma_y^2 = \frac{\overline{x^2}}{\overline{x^2} - \bar{x}^2} \frac{\sigma_y^2}{N}$$

Queste espressioni indicano che le incertezze sui parametri della retta diminuiscono all'aumentare di N , ma anche della quantità

$$\overline{x^2} - \bar{x}^2$$

denominata “braccio della leva” della estensione delle misure: lo stesso numero di misure sparse su un intervallo più grande determinano i parametri con una maggiore precisione.

Si può inoltre dimostrare che la retta dei minimi quadrati passa per il punto $B(x_B, y_B)$ le cui coordinate sono i baricentri delle misure di x ed y .

$$x_B = \frac{1}{N} \sum_{i=1}^N x_i \qquad y_B = \frac{1}{N} \sum_{i=1}^N y_i$$

Se la σ_y non è nota la si può stimare tramite la dispersione dei punti intorno alla retta:

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^N [y_i - (ax_i + b)]^2}{N-2}}$$

residui ←

È stata così determinata la “migliore retta” che approssima un insieme di dati sperimentali rendendo minima la somma dei quadrati delle distanze, misurate nella direzione della asse y, dei punti dalla retta.

La retta così determinata si dice **retta di regressione** (di y su x)

Avremmo potuto minimizzare la somma dei quadrati delle distanze misurate nella direzione dell'asse x , ovvero la quantità

$$\sum_{i=1}^N \frac{(x_i - a' y_i - b')^2}{\sigma_i^2}$$

La retta così determinata si dice **retta di regressione** (di x su y).

In genere le due rette non coincidono, ma quanto più la distribuzione dei punti è prossima ad essere rettilinea, tanto più le due rette di regressione si avvicinano.

Infine è possibile rendere minima la somma dei quadrati delle distanze misurate ortogonalmente dai punti alla retta.

La retta così determinata si dice **retta di regressione ortogonale**.

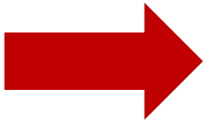
Generalizzazione del metodo dei minimi quadrati

Il metodo può essere applicato in modo molto più generale.

Consideriamo N coppie di misure (x_i, y_i) di due grandezze x e y e supponiamo che l'incertezza su una delle due variabili (ad esempio sulla x) sia trascurabile rispetto a quella dell'altra (y) e che le σ_i siano tutte uguali $\sigma_i = \sigma$

Inoltre la relazione che lega le 2 variabili sia espressa dalla generica funzione

$$y = f(x, c_1, c_2, c_3, c_4, \dots, c_p)$$

 **c_j sono i parametri di cui si vogliono determinare i valori in corrispondenza dei quali la curva meglio si adatta ai dati sperimentali**

Il metodo dei minimi quadrati permette di determinare i valori dei parametri c_j che rendono minima la quantità:

$$z = \sum_{i=1}^N \frac{[y_i - f(x_i, c_1, c_2, \dots)]^2}{\sigma^2}$$

Mediante la risoluzione del sistema di equazioni:

$$\frac{\partial z}{\partial c_1} = \frac{\partial z}{\partial c_2} = \dots = \frac{\partial z}{\partial c_p} = 0$$