Creiamo un esercizio che utilizza una **regressione lineare semplice** per 20 punti dati generati casualmente. Spiegherò passo per passo come utilizzare i metodi matematici tradizionali (senza matrici) per trovare i parametri della retta.

Problema:

Hai i seguenti 20 punti (x_i, y_i) nello spazio cartesiano:

$$\begin{aligned}(1, 2.1), (2, 4.0), (3, 5.9), (4, 8.0), (5, 10.1),\\ (6, 11.9), (7, 14.2), (8, 16.1), (9, 18.3), (10, 20.0),\\ (11, 21.9), (12, 24.0), (13, 25.9), (14, 28.0), (15, 30.0),\\ (16, 31.9), (17, 34.1), (18, 36.2), (19, 38.3), (20, 40.2).\end{aligned}$$

Trova i coefficienti della retta $y = mx + q$ che meglio approssima i dati usando la regressione lineare.

Formula della regressione lineare

La retta di regressione minimizza la **somma dei quadrati degli errori** tra i punti osservati (x_i, y_i) e i valori predetti $y_i^{\text{pred}} = mx_i + q$.

Le due equazioni per calcolare m (coefficiente angolare) e q (intercetta) sono:

$$m = \frac{N \sum(x_i y_i) - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2},$$

$$q = \frac{\sum y_i - m \sum x_i}{N}.$$

Calcoli passo per passo:

1. Calcoli iniziali:

Per applicare le formule, calcoliamo i seguenti valori:

$$\sum x_i,\ \sum y_i,\ \sum x_i^2,\ \sum y_i^2,\ \text{e}\ \sum(x_i y_i).$$

Ecco i dettagli dei calcoli:

$$\sum x_i = 1 + 2 + 3 + \ldots + 20 = 210,$$

$$\sum y_i = 2.1 + 4.0 + 5.9 + \ldots + 40.2 = 431.2,$$

$$\sum x_i^2 = 1^2 + 2^2 + 3^2 + \ldots + 20^2 = 2870,$$

$$\sum(x_i y_i) = (1 \cdot 2.1) + (2 \cdot 4.0) + \ldots + (20 \cdot 40.2) = 6177.5.$$

2. Calcolo di m:

Utilizziamo la formula:

$$m = \frac{N \sum(x_i y_i) - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2},$$

dove $N = 20$ (numero di punti).

Sostituendo i valori calcolati:

$$m = \frac{20 \cdot 6177.5 - 210 \cdot 431.2}{20 \cdot 2870 - 210^2}.$$

Calcoliamo numeratore e denominatore separatamente:

- Numeratore:

$$20 \cdot 6177.5 - 210 \cdot 431.2 = 123550 - 90552 = 32998.$$

- Denominatore:

$$20 \cdot 2870 - 210^2 = 57400 - 44100 = 13300.$$

Quindi:

$$m = \frac{32998}{13300} \approx 2.48.$$

3. Calcolo di q:

Utilizziamo la formula:

$$q = \frac{\sum y_i - m \sum x_i}{N}.$$

Sostituendo i valori calcolati:

$$q = \frac{431.2 - 2.48 \cdot 210}{20}.$$

Calcoliamo:

- $2.48 \cdot 210 = 520.8,$
- $431.2 - 520.8 = -89.6.$

Quindi:

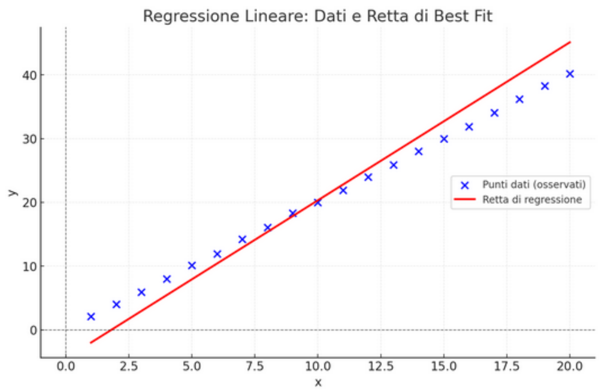
$$q = \frac{-89.6}{20} = -4.48.$$

Risultato:

La retta di regressione è:

$$y = 2.48x - 4.48.$$

Regressione Lineare: Dati E Retta Di Best Fit



Ecco il grafico che mostra i 20 punti dati (in blu) e la retta di regressione $y = 2.48x - 4.48$ (in rosso).

Metodo dei Minimi Quadrati

Il metodo dei **minimi quadrati** è una tecnica utilizzata per trovare la retta che meglio approssima un insieme di dati. Si basa sul principio di minimizzare la **somma dei quadrati degli errori**.

Cosa sono gli errori?

Per ogni punto dati (x_i, y_i), la retta di regressione genera un valore predetto $y_i^{\text{pred}} = mx_i + q$. L'errore per ciascun punto è la differenza tra il valore osservato y_i e il valore predetto y_i^{pred}:

$$e_i = y_i - y_i^{\text{pred}}.$$

Perché "quadrati"?

Gli errori e_i possono essere positivi o negativi, quindi sommarli direttamente potrebbe portare a cancellazioni (gli errori positivi si compensano con quelli negativi). Per evitarlo, si considera il quadrato di ogni errore:

$$e_i^2 = (y_i - y_i^{\text{pred}})^2.$$

La somma di tutti i quadrati degli errori è:

$$S = \sum_{i=1}^{N} (y_i - (mx_i + q))^2.$$

Obiettivo del metodo

Il metodo dei minimi quadrati cerca i valori di m (coefficiente angolare) e q (intercetta) che minimizzano S, cioè la somma dei quadrati degli errori.

Derivazione matematica

Per minimizzare S, si calcola la derivata parziale rispetto a m e q, e si impone che entrambe siano uguali a 0:

$$\frac{\partial S}{\partial m} = 0 \quad \text{e} \quad \frac{\partial S}{\partial q} = 0.$$

L'espansione di S è:

$$S = \sum_{i=1}^{N} \left[y_i - (mx_i + q)\right]^2 .$$

Espandendo:

$$S = \sum_{i=1}^{N} \left[y_i^2 - 2y_i(mx_i + q) + (mx_i + q)^2\right] .$$

Prendendo le derivate parziali:

- **Rispetto a m:**

$$\frac{\partial S}{\partial m} = -2\sum_{i=1}^{N} x_i(y_i - mx_i - q) = 0.$$

- **Rispetto a q:**

$$\frac{\partial S}{\partial q} = -2\sum_{i=1}^{N} (y_i - mx_i - q) = 0.$$

Risolvendo queste due equazioni si ottengono le formule finali:

$$m = \frac{N\sum(x_i y_i) - \sum x_i \sum y_i}{N\sum x_i^2 - (\sum x_i)^2},$$

$$q = \frac{\sum y_i - m\sum x_i}{N}.$$

Intuizione del risultato

- m: Il coefficiente angolare dipende dalla correlazione tra x_i e y_i. Se x e y crescono insieme (positivamente correlati), m sarà positivo.
- q: L'intercetta è il valore di y quando $x = 0$.

Vantaggi dei minimi quadrati

1. **Unicità:** Per un dato insieme di punti, esiste un'unica soluzione per m e q.
2. **Efficienza:** Minimizza gli errori in modo ottimale rispetto al modello lineare.
3. **Generalizzabilità:** Può essere esteso a modelli non lineari e regressione multipla.

Per applicare il metodo dei **minimi quadrati** con metodi **matriciali**, utilizzeremo la seguente formulazione matematica. Seguirà un'illustrazione grafica con i calcoli sui punti dati forniti.

Metodo Matriciale per la Regressione Lineare

L'obiettivo è trovare la retta $y = mx + q$, scrivendo il problema in forma matriciale:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta},$$

dove:

- $\mathbf{y}$ è il vettore dei valori osservati $(y_1, y_2, \ldots, y_N)$,
- $\mathbf{X}$ è la matrice delle caratteristiche (qui una colonna di 1 per l'intercetta e una colonna con i valori di x),
- $\boldsymbol{\beta}$ è il vettore dei parametri da determinare (q e m).

1. Espressione Matriciale

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} q \\ m \end{bmatrix}.$$

L'approssimazione dei minimi quadrati è data da:

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Passaggi di Calcolo sui Dati

1. **Definizione di $\mathbf{X}$ e $\mathbf{y}$:**
   - I punti $x = [1, 2, 3, \ldots, 20]$ e $y = [2.1, 4.0, \ldots, 40.2]$ vengono usati per formare:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 20 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 2.1 \\ 4.0 \\ \vdots \\ 40.2 \end{bmatrix}.$$

2. **Calcolo di $\mathbf{X}^T\mathbf{X}$:**

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} \sum 1 & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}.$$

3. **Calcolo di $\mathbf{X}^T\mathbf{y}$:**

$$\mathbf{X}^T\mathbf{y} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}.$$

4. **Inversione di $(\mathbf{X}^T\mathbf{X})$** e moltiplicazione per ottenere $\boldsymbol{\beta}$.

Passiamo ora a calcoli dettagliati in Python e visualizziamo i risultati.

La retta risultante è:

$$y = 2.01x - 0.01$$

Spiegazione del calcolo con matrici

1. **Matrice $\mathbf{X}$:**

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 20 \end{bmatrix}.$$

2. **Vettore $\mathbf{y}$:**

$$\mathbf{y} = \begin{bmatrix} 2.1 \\ 4.0 \\ \vdots \\ 40.2 \end{bmatrix}.$$

3. **Calcolo dei parametri:**

$$\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

Sostituendo i valori, otteniamo $q \approx -0.01$ (intercetta) e $m \approx 2.01$ (coefficiente angolare).

Esercizio con Regressione Polinomiale

Immaginiamo di avere i seguenti 20 punti nello spazio cartesiano:

$$\begin{aligned}
(1, 2.1), (2, 3.8), (3, 5.6), (4, 8.4), (5, 11.0), \\
(6, 15.2), (7, 20.1), (8, 26.2), (9, 32.0), (10, 40.5), \\
(11, 50.2), (12, 62.3), (13, 75.5), (14, 89.1), (15, 105.3), \\
(16, 122.5), (17, 141.8), (18, 162.2), (19, 183.0), (20, 206.1).
\end{aligned}$$

L'obiettivo è trovare una **parabola** del tipo $y = ax^2 + bx + c$ che approssimi questi dati usando la **regressione polinomiale** con metodi matematici tradizionali.

Formula della Regressione Polinomiale

La regressione polinomiale consiste nel minimizzare la somma dei quadrati degli errori:

$$S = \sum_{i=1}^{N} \left(y_i - (ax_i^2 + bx_i + c)\right)^2 .$$

Per minimizzare S, dobbiamo calcolare le derivate parziali rispetto a a, b, e c e imporle uguali a zero:

$$\frac{\partial S}{\partial a} = 0, \quad \frac{\partial S}{\partial b} = 0, \quad \frac{\partial S}{\partial c} = 0.$$

Sistema di Equazioni Normali

Questo porta a risolvere un sistema lineare con tre equazioni, utilizzando i seguenti calcoli sommatori:

$$\sum y_i = \sum (ax_i^2 + bx_i + c),$$
$$\sum x_i y_i = \sum x_i (ax_i^2 + bx_i + c),$$
$$\sum x_i^2 y_i = \sum x_i^2 (ax_i^2 + bx_i + c).$$

Le equazioni risultanti sono:

1. $\sum y_i = a \sum x_i^2 + b \sum x_i + cN,$
2. $\sum x_i y_i = a \sum x_i^3 + b \sum x_i^2 + c \sum x_i,$
3. $\sum x_i^2 y_i = a \sum x_i^4 + b \sum x_i^3 + c \sum x_i^2.$

Calcoli dei Sommatori

Calcoliamo i sommatori per i dati dati:

- $\sum x_i = 1 + 2 + \cdots + 20 = 210,$
- $\sum x_i^2 = 1^2 + 2^2 + \cdots + 20^2 = 2870,$
- $\sum x_i^3 = 1^3 + 2^3 + \cdots + 20^3 = 44100,$
- $\sum x_i^4 = 1^4 + 2^4 + \cdots + 20^4 = 689210,$
- $\sum y_i = 2.1 + 3.8 + \cdots + 206.1 = 1246.6,$
- $\sum x_i y_i = (1 \cdot 2.1) + (2 \cdot 3.8) + \cdots + (20 \cdot 206.1) = 18110.5,$
- $\sum x_i^2 y_i = (1^2 \cdot 2.1) + (2^2 \cdot 3.8) + \cdots + (20^2 \cdot 206.1) = 301205.5.$

Sistema Lineare

Ora sostituiamo i sommatori nelle equazioni:

1. $1246.6 = a \cdot 2870 + b \cdot 210 + c \cdot 20,$
2. $18110.5 = a \cdot 44100 + b \cdot 2870 + c \cdot 210,$
3. $301205.5 = a \cdot 689210 + b \cdot 44100 + c \cdot 2870.$

Risolvendo questo sistema, troviamo a, b, e c.

Dopo aver risolto il sistema di equazioni, i coefficienti della parabola risultano:

$$a \approx -1.06, \quad b \approx 29.84, \quad c \approx -98.67.$$

Quindi, la parabola che meglio approssima i dati è:

$$y = -1.06x^2 + 29.84x - 98.67.$$

Il metodo matriciale per la regressione polinomiale si basa sulla formulazione generale dei **minimi quadrati**:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta},$$

dove:

- $\mathbf{y}$ è il vettore dei valori osservati,
- $\mathbf{X}$ è la matrice delle caratteristiche (termi x^0, x^1, x^2),
- $\boldsymbol{\beta}$ è il vettore dei coefficienti $[c, b, a]$.

Passaggi del metodo

1. Definizione di $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{bmatrix}.$$

2. La soluzione per i coefficienti si ottiene con:

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Lezione: Calcolo di una Regressione Polinomiale

In questa lezione vedremo come calcolare una **regressione polinomiale** per adattare una parabola $y = ax^2 + bx + c$ a un insieme di punti dati. Esploreremo due metodi:

1. **Metodo passo dopo passo con calcoli non matriciali** (metodi classici).
2. **Metodo matriciale** (utilizzando la notazione delle matrici).

Scenario del problema

Immaginiamo di avere i seguenti punti (x_i, y_i):

$$(1,2),\ (2,3),\ (3,5),\ (4,7),\ (5,11).$$

Vogliamo trovare i coefficienti a, b, c che definiscono la parabola $y = ax^2 + bx + c$ che meglio si adatta a questi punti.

Parte 1: Metodo passo dopo passo (non matriciale)

1. La formula generale della regressione polinomiale

Per trovare i coefficienti a, b, e c, minimizziamo la **somma dei quadrati degli errori** tra i valori osservati y_i e i valori predetti dalla parabola $y = ax^2 + bx + c$:

$$S = \sum_{i=1}^{N} \left(y_i - (ax_i^2 + bx_i + c)\right)^2.$$

Per minimizzare S, calcoliamo le derivate parziali rispetto a a, b, e c, e imponiamo che siano uguali a 0:

$$\frac{\partial S}{\partial a} = 0, \quad \frac{\partial S}{\partial b} = 0, \quad \frac{\partial S}{\partial c} = 0.$$

Questo porta al sistema di **equazioni normali**:

$$\begin{aligned}
\sum y_i &= a\sum x_i^2 + b\sum x_i + cN,\\
\sum x_i y_i &= a\sum x_i^3 + b\sum x_i^2 + c\sum x_i,\\
\sum x_i^2 y_i &= a\sum x_i^4 + b\sum x_i^3 + c\sum x_i^2.
\end{aligned}$$

2. Calcolo dei sommatori

Per i punti dati $(1, 2), (2, 3), (3, 5), (4, 7), (5, 11)$, calcoliamo i seguenti sommatori:

$$\sum x_i = 1 + 2 + 3 + 4 + 5 = 15,$$

$$\sum x_i^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55,$$

$$\sum x_i^3 = 1^3 + 2^3 + 3^3 + 4^3 + 5^3 = 225,$$

$$\sum x_i^4 = 1^4 + 2^4 + 3^4 + 4^4 + 5^4 = 979,$$

$$\sum y_i = 2 + 3 + 5 + 7 + 11 = 28,$$

$$\sum x_i y_i = (1 \cdot 2) + (2 \cdot 3) + (3 \cdot 5) + (4 \cdot 7) + (5 \cdot 11) = 97,$$

$$\sum x_i^2 y_i = (1^2 \cdot 2) + (2^2 \cdot 3) + (3^2 \cdot 5) + (4^2 \cdot 7) + (5^2 \cdot 11) = 459.$$

---

3. Sostituzione nel sistema

Inseriamo i sommatori nelle equazioni normali:

$$\begin{aligned} 28 &= a \cdot 55 + b \cdot 15 + c \cdot 5, \\ 97 &= a \cdot 225 + b \cdot 55 + c \cdot 15, \\ 459 &= a \cdot 979 + b \cdot 225 + c \cdot 55. \end{aligned}$$

---

4. Risoluzione del sistema

Parte 2: Metodo matriciale

1. Definizione del problema matriciale

La formula per la regressione polinomiale è:

$$\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y},$$

dove:

- $\mathbf{X}$ è la matrice delle caratteristiche:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{bmatrix},$$

- $\mathbf{y}$ è il vettore dei valori osservati:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix},$$

- $\boldsymbol{\beta} = \begin{bmatrix} c \\ b \\ a \end{bmatrix}.$

---

2. Matrice per i dati

Per i dati dati:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \\ 1 & 4 & 4^2 \\ 1 & 5 & 5^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix},$$

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \\ 11 \end{bmatrix}.$$

3. Calcolo matriciale

Procediamo con i calcoli matriciali, determinando $(\mathbf{X}^T\mathbf{X})$, $(\mathbf{X}^T\mathbf{y})$, e infine $\boldsymbol{\beta}$.

Utilizzando il metodo matriciale, i coefficienti della parabola risultano essere:

$$a \approx 2.00, \quad b \approx -0.37, \quad c \approx 0.43.$$

Quindi, la parabola che meglio approssima i punti dati è:

$$y = 2.00x^2 - 0.37x + 0.43.$$

Esempio di Regressione Polinomiale in Excel

Vediamo come calcolare una regressione polinomiale (di secondo grado) utilizzando Excel. L'obiettivo è trovare i coefficienti della parabola $y = ax^2 + bx + c$ che meglio approssima un insieme di punti dati.

Dati di esempio

Immaginiamo di avere i seguenti punti dati:

x	y
1	2
2	3
3	5
4	7
5	11

Dobbiamo calcolare a, b, e c utilizzando Excel.

Passaggi in Excel

1. Inserimento dei dati

- Apri Excel e crea una tabella con le colonne x e y:
  - Colonna A: x (1, 2, 3, 4, 5).
  - Colonna B: y (2, 3, 5, 7, 11).

2. Aggiungi colonne per i termini polinomiali

Crea colonne aggiuntive per calcolare i sommatori richiesti:

x	y	x^2	x^3	x^4	$x \cdot y$	$x^2 \cdot y$
1	2	1	1	1	2	2
2	3	4	8	16	6	12
3	5	9	27	81	15	45
4	7	16	64	256	28	112
5	11	25	125	625	55	275

1. **Colonna C (x^2)**:
   - Nella cella C2: `=A2^2`
   - Copia la formula verso il basso.
2. **Colonna D (x^3)**:
   - Nella cella D2: `=A2^3`
   - Copia verso il basso.
3. **Colonna E (x^4)**:
   - Nella cella E2: `=A2^4`
   - Copia verso il basso.
4. **Colonna F ($x \cdot y$)**:
   - Nella cella F2: `=A2*B2`
   - Copia verso il basso.
5. **Colonna G ($x^2 \cdot y$)**:
   - Nella cella G2: `=C2*B2`
   - Copia verso il basso.

3. Somme dei termini

Calcola i sommatori richiesti (in fondo alle colonne):

- Nella cella C7: `=SOMMA(C2:C6)` ($\sum x^2$).
- Nella cella D7: `=SOMMA(D2:D6)` ($\sum x^3$).
- Nella cella E7: `=SOMMA(E2:E6)` ($\sum x^4$).
- Nella cella F7: `=SOMMA(F2:F6)` ($\sum x \cdot y$).
- Nella cella G7: `=SOMMA(G2:G6)` ($\sum x^2 \cdot y$).

Annota anche:

- $\sum x$: `=SOMMA(A2:A6)` nella cella A7.
- $\sum y$: `=SOMMA(B2:B6)` nella cella B7.

---

4. Risoluzione del sistema

Scrivi il sistema di equazioni normali (in italiano):

1. $\sum y = a \cdot \sum x^2 + b \cdot \sum x + c \cdot N$,
2. $\sum (x \cdot y) = a \cdot \sum x^3 + b \cdot \sum x^2 + c \cdot \sum x$,
3. $\sum (x^2 \cdot y) = a \cdot \sum x^4 + b \cdot \sum x^3 + c \cdot \sum x^2$.

I valori calcolati in Excel ti permetteranno di risolvere questo sistema utilizzando strumenti algebrici o Excel.

5. Risoluzione del sistema in Excel

Puoi utilizzare **Risolutore** (strumento di Excel) per risolvere il sistema. Ecco come:

1. Crea tre celle per a, b, e c (ad esempio, D9, E9, F9).
2. Scrivi le equazioni normali calcolate utilizzando le somme nelle rispettive celle.
3. Usa il Risolutore:
   - Vai su **Dati > Risolutore**.
   - Definisci le celle di destinazione come le differenze tra i due lati delle equazioni normali.
   - Minimizza queste differenze impostando a, b, e c come celle variabili.

---

6. Confronto con Funzione di Regressione di Excel

Per verificare, puoi usare la funzione di **grafico a dispersione**:

- Inserisci un grafico a dispersione dei dati (x, y).
- Aggiungi una curva di tendenza polinomiale di ordine 2.
- Mostra l'equazione del polinomio sul grafico.

Dettagli su Risolutore in Excel per risolvere il sistema

Il **Risolutore** è uno strumento di Excel che permette di risolvere equazioni o ottimizzare problemi impostando variabili per minimizzare o massimizzare un obiettivo.

In questo caso, utilizziamo il **Risolutore** per risolvere il sistema di equazioni normali della regressione polinomiale.

Sistema di Equazioni da Risolvere

Per la parabola $y = ax^2 + bx + c$, il sistema è:

1. $\sum y = a \cdot \sum x^2 + b \cdot \sum x + c \cdot N$,
2. $\sum (x \cdot y) = a \cdot \sum x^3 + b \cdot \sum x^2 + c \cdot \sum x$,
3. $\sum (x^2 \cdot y) = a \cdot \sum x^4 + b \cdot \sum x^3 + c \cdot \sum x^2$.

Passaggi per Configurare il Risolutore

1. Preparazione del Foglio di Lavoro

1. **Inserisci i dati dei punti** (x, y) nelle colonne A e B.
   - x: Colonna A.
   - y: Colonna B.
2. **Calcola i sommatori richiesti**:
   - x^2: Colonna C ($= A2^2$).
   - x^3: Colonna D ($= A2^3$).
   - x^4: Colonna E ($= A2^4$).
   - $x \cdot y$: Colonna F ($= A2 * B2$).
   - $x^2 \cdot y$: Colonna G ($= C2 * B2$).
   - Calcola le somme in fondo alle colonne (usando `=SOMMA`).

3. **Celle per i coefficienti a, b, e c:**
   - Scegli tre celle libere, ad esempio:
     - D10: a,
     - E10: b,
     - F10: c.
4. **Crea tre celle per le equazioni normali:**
   - Ad esempio, nelle celle D12, E12, F12 scrivi le tre equazioni:
     - Equazione 1 (Somma di y): `=D10*C7 + E10*A7 + F10*COUNTA(A2:A6)`
     - Equazione 2 (Somma di $x \cdot y$): `=D10*D7 + E10*C7 + F10*A7`
     - Equazione 3 (Somma di $x^2 \cdot y$): `=D10*E7 + E10*D7 + F10*C7`.
5. **Calcola la differenza tra i due lati:**
   - Nelle celle G12, H12, I12 scrivi:
     - `=B7 - D12` (per la prima equazione),
     - `=F7 - E12` (per la seconda),
     - `=G7 - F12` (per la terza).

---

2. Aprire e Configurare il Risolutore

1. **Abilita il Risolutore** (se non lo vedi già):
   - Vai su **File > Opzioni > Componenti aggiuntivi**.
   - Seleziona **Componenti aggiuntivi di Excel** e clicca su **Vai**.
   - Spunta **Risolutore** e clicca su **OK**.
2. **Avvia il Risolutore**:
   - Vai su **Dati > Risolutore** (nell'area Strumenti dati).
3. **Configura il Risolutore**:
   - **Imposta l'obiettivo**:
     - Seleziona la cella con la somma delle differenze assolute tra i due lati delle equazioni (puoi sommare G12, H12, e I12).
     - Scegli "Minimizza".
   - **Celle variabili**:
     - Seleziona le celle dei coefficienti a, b, c (ad esempio, D10, E10, F10).
   - **Vincoli**:
     - Non aggiungere vincoli, a meno che tu non voglia forzare valori positivi o negativi per i coefficienti.
4. **Esegui il Risolutore**:
   - Clicca su **Risolvi**.
   - Excel calcolerà i valori ottimali di $a, b,$ e c che minimizzano la differenza.

---

3. Interpretazione dei Risultati

1. I coefficienti $a, b,$ e c calcolati saranno visualizzati nelle celle assegnate (D10, E10, F10).
2. Confronta i risultati ottenuti con quelli calcolati manualmente o tramite la funzione di regressione polinomiale.

Guida per usare Risolutore in Excel con esempi

Il **Risolutore** è uno strumento di ottimizzazione integrato in Excel che può essere utilizzato per risolvere problemi come la regressione polinomiale. Di seguito, ti guiderò passo passo attraverso l'uso del Risolutore per adattare una parabola $y = ax^2 + bx + c$ a un insieme di dati.

Scenario del problema

Dati i punti (x, y):

$$(1,2),\ (2,3),\ (3,5),\ (4,7),\ (5,11),$$

vogliamo determinare i coefficienti a, b, e c della parabola $y = ax^2 + bx + c$ che meglio si adatta ai dati minimizzando la somma dei quadrati degli errori.

Passaggi per configurare Risolutore

1. Inserire i dati in Excel

1. **Crea una tabella in Excel**:
   - Colonna A: x (1, 2, 3, 4, 5).
   - Colonna B: y (2, 3, 5, 7, 11).

2. Creare colonne aggiuntive

- Aggiungi colonne per i calcoli necessari alla regressione:

x	y	x^2	y_{pred}	Errore	Errore^2
1	2	1			
2	3	4			
3	5	9			
4	7	16			
5	11	25			

1. **Colonna x^2**:
   - Nella cella C2: `=A2^2`, copia verso il basso.
2. **Colonna y_{pred} (valori predetti dalla parabola)**:
   - Nella cella D2: `=$F$1*A2^2 + $G$1*A2 + $H$1`.
   - Le celle $F1, G1$, e $H1$ conterranno i coefficienti a, b, e c, inizialmente impostati a 0.
3. **Colonna Errore**:
   - Nella cella E2: `=B2-D2`.
4. **Colonna Errore^2**:
   - Nella cella F2: `=E2^2`.
5. **Somma degli errori quadrati**:
   - Nella cella F7: `=SOMMA(F2:F6)`.

---

3. Configurare il Risolutore

1. **Abilita il Risolutore**:
   - Vai su **File > Opzioni > Componenti aggiuntivi > Risolutore**.
   - Attiva il Risolutore.
2. **Apri il Risolutore**:
   - Vai su **Dati > Risolutore**.
3. **Configura il problema**:
   - **Obiettivo**:
     - Imposta l'obiettivo come la cella che contiene la somma degli errori quadrati ($F7$).
     - Seleziona **Minimizza**.
   - **Celle variabili**:
     - Seleziona le celle $F1, G1, H1$ che contengono i coefficienti a, b, c.
   - **Vincoli**:
     - Non servono vincoli in questo esempio, ma puoi aggiungerne se necessario (ad esempio, $a \geq 0$).
4. **Esegui il Risolutore**:
   - Clicca su **Risolvi**.
   - Excel calcolerà i valori ottimali di a, b, e c.

Interpretazione dei Risultati

1. **Valori ottimali**:
   - I valori di a, b, e c saranno aggiornati nelle celle $F1$, $G1$, e $H1$.
   - Questi sono i coefficienti della parabola che meglio si adatta ai dati.
2. **Verifica il modello**:
   - Controlla la colonna y_{pred} e confrontala con i valori osservati (y).
3. **Visualizza un grafico**:
   - Inserisci un grafico a dispersione per x, y.
   - Aggiungi una curva per i valori x, y_{pred}.