Regressione Polinomiale

Nel lavoro che segue ci proponiamo di descrivere alcune curve di adattamento con il metodo dei *minimi quadrati* e di fornire un metodo iterativo per generalizzare tali funzioni a polinomi di grado n. Esamineremo il caso semplificato di misure (x_i) non affette da incertezze.

Spesso possiamo esprimere una variabile, y, come polinomiale di una seconda variabile x:

$$y = A + Bx + Cx^2 + ... + Zx^n \tag{1}$$

Supponiamo per esempio di avere una polinomiale di forma quadratica,

$$y = A + Bx + Cx^2 \tag{2}$$

nota una serie di valori $(x_i,\ y_i)$, $i = 1,..,N$ per ogni x_i il valore y_i si ottiene dalla (2) dove A,B,C sono ancora incognite. Assumiamo che le misure degli y_i seguano distribuzioni normali (ognuna centrata sul proprio valor medio Y_i e tutte con la stessa deviazione standard σ_y). Possiamo quindi calcolare la probabilità di ottenere i valori y_1,...,y_n come:

$$P(y_1,...,y_n) \propto e^{-\frac{\chi^2}{2}}$$

dove

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - A - Bx_i - Cx_i^2)^2}{\sigma_y^2}$$

La miglior stima per A,B,C è data da quei valori per cui $P(y_1,...,y_n)$ è massima, o χ^2 è minimo. Differenziamo quindi χ^2 rispetto a A,B,C :

$$\frac{\partial \chi^2}{\partial A} = \frac{1}{\sigma_y^2} \sum_{i=1}^{N} \left(2A - 2y_i + 2Bx_i + 2Cx_i^2\right)$$

$$\frac{\partial \chi^2}{\partial B} = \frac{1}{\sigma_y^2} \sum_{i=1}^{N} \left(2Bx_i^2 - 2y_ix_i + 2Ax_i + 2Cx_i^3\right)$$

$$\frac{\partial\chi^2}{\partial C}=\frac{1}{\sigma_y^2}\sum_{i=1}^{N}\left(2Cx_i^4-2y_ix_i^2+2Ax_i^2+2Bx_i^3\right)$$

ponendo uguale a zero, omettendo gli estremi di sommatoria

$$\begin{aligned}\sum y_i &= AN+B\sum x_i+C\sum x_i^2\\ \sum x_iy_i &= A\sum x_i+B\sum x_i^2+C\sum x_i^3\\ \sum x_i^2y_i &= A\sum x_i^2+B\sum x_i^3+C\sum x_i^4\end{aligned}$$

si tratta si un sistema 3x3 che in forma matriciale diventa

$$\begin{bmatrix} N & \sum x_i & \sum x_i^2\\ \sum x_i & \sum x_i^2 & \sum x_i^3\\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4\end{bmatrix}\begin{bmatrix}A\\B\\C\end{bmatrix}=\begin{bmatrix}\sum y_i\\ \sum x_iy_i\\ \sum x_i^2y_i\end{bmatrix}$$

che è del tipo

$$\mathcal{A}\qquad X\qquad=\qquad\mathcal{B}$$

e si può risolvere diversi modi tra cui (quello matriciale premoltiplicando per l'inversa di $\mathcal{A}$)

$$X=\mathcal{A}^{-1}\mathcal{B}$$

Trovati i valori di A,B,C si sostituiscono in (2) e otteniamo una curva interpolante di grado 2.

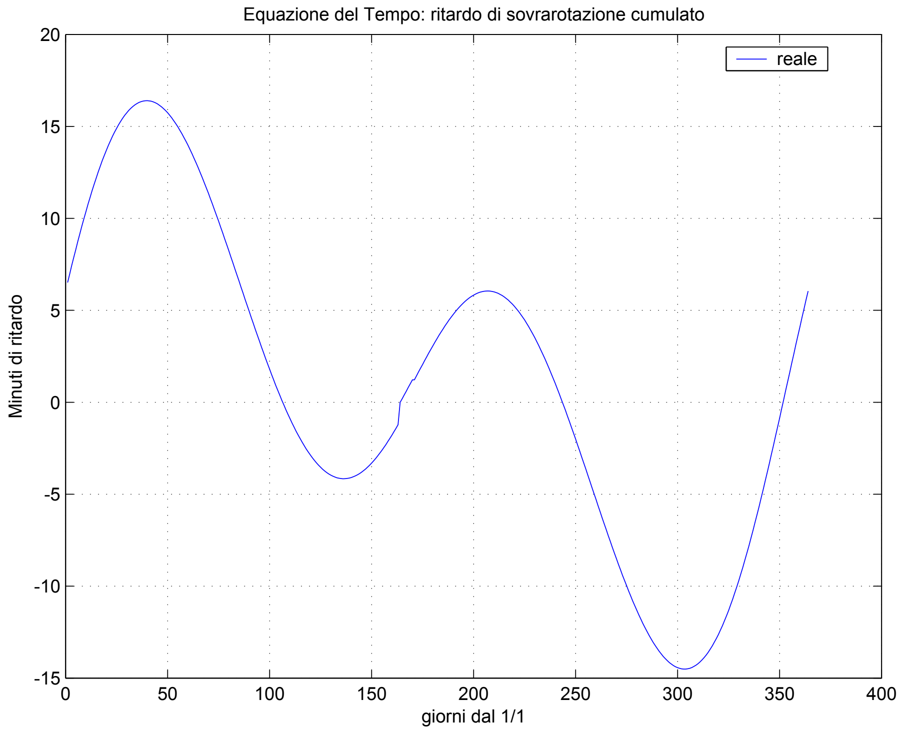
In generale volendo ricavare una polinomiale di grado n avremo $n+1$ equazioni in $n+1$ incognite, la matrice $\mathcal{A}$ sarà $(n+1)_\times(n+1)$, le matrici $\mathcal{B}$ e X saranno $(n+1)_\times(1)$. Le componenti di tali matrici saranno:

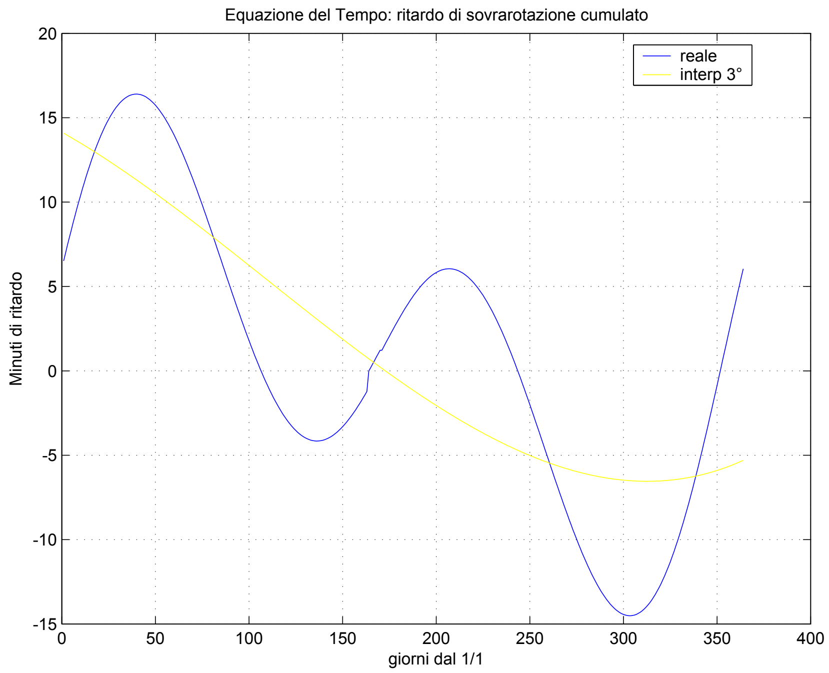
$$\begin{bmatrix} N & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^n\\ \sum x_i & \sum x_i^2 & \cdots & \cdots & \sum x_i^{n+1}\\ \sum x_i^2 & \sum x_i^3 & \cdots & \cdots & \sum x_i^{n+2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \cdots & \sum x_i^{2n}\end{bmatrix}\begin{bmatrix}A\\B\\C\\ \vdots\\ Z\end{bmatrix}=\begin{bmatrix}\sum y_i\\ \sum x_iy_i\\ \sum x_i^2y_i\\ \vdots\\ \sum x_i^ny_i\end{bmatrix}$$

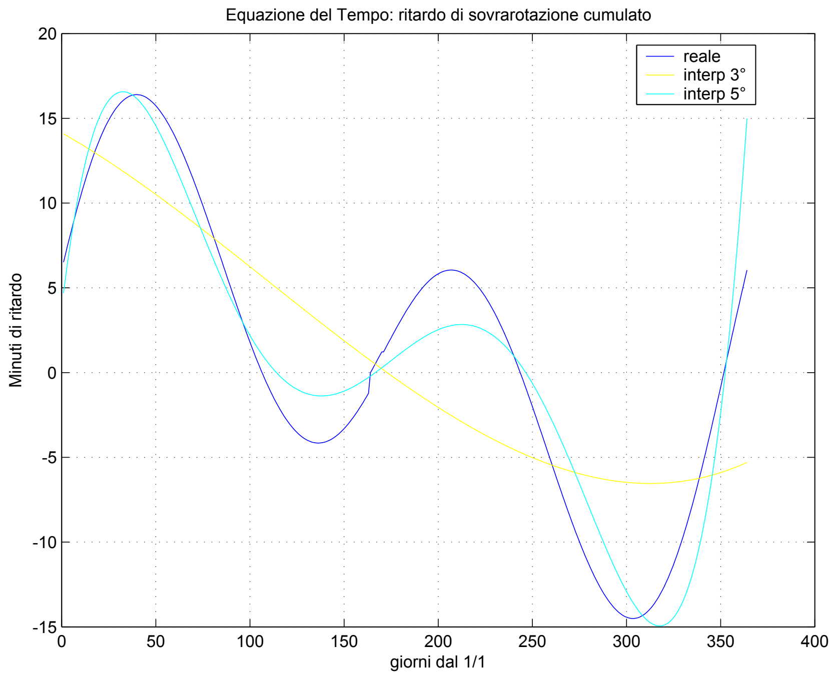
Risolvendo tale sistema si ottengono i valori A,B,C,...,Z da sostituire in (1), questa equazione rappresenta la regressione polinomiale di adattamento ai dati (x_i,y_i).

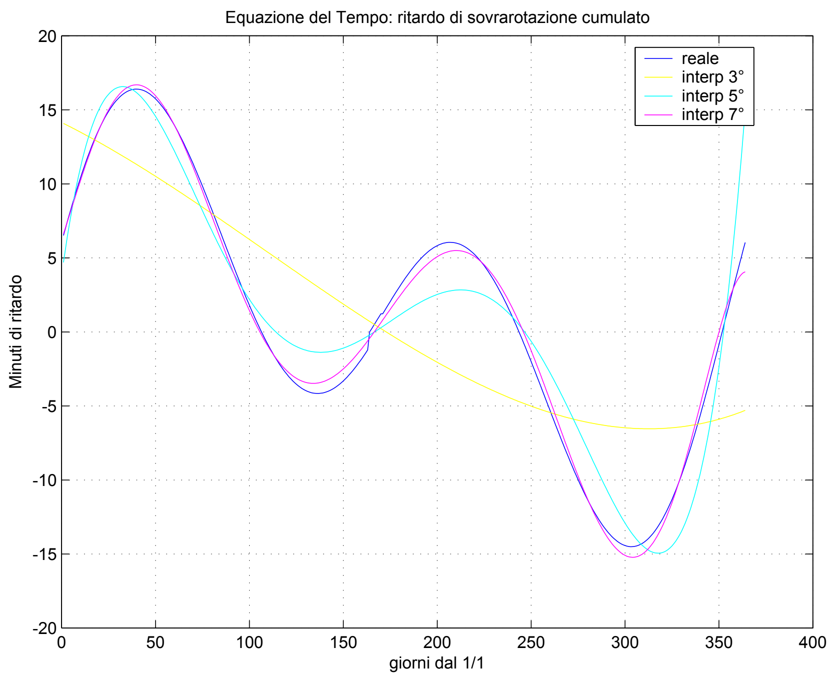
Vediamo ora alcuni esempi grafici:

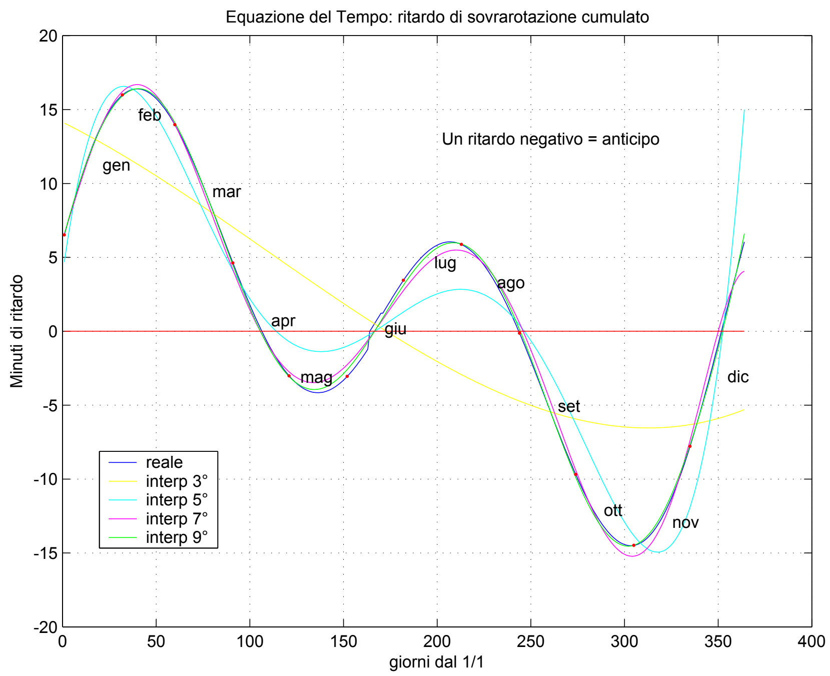
ecco i dati da interpolare

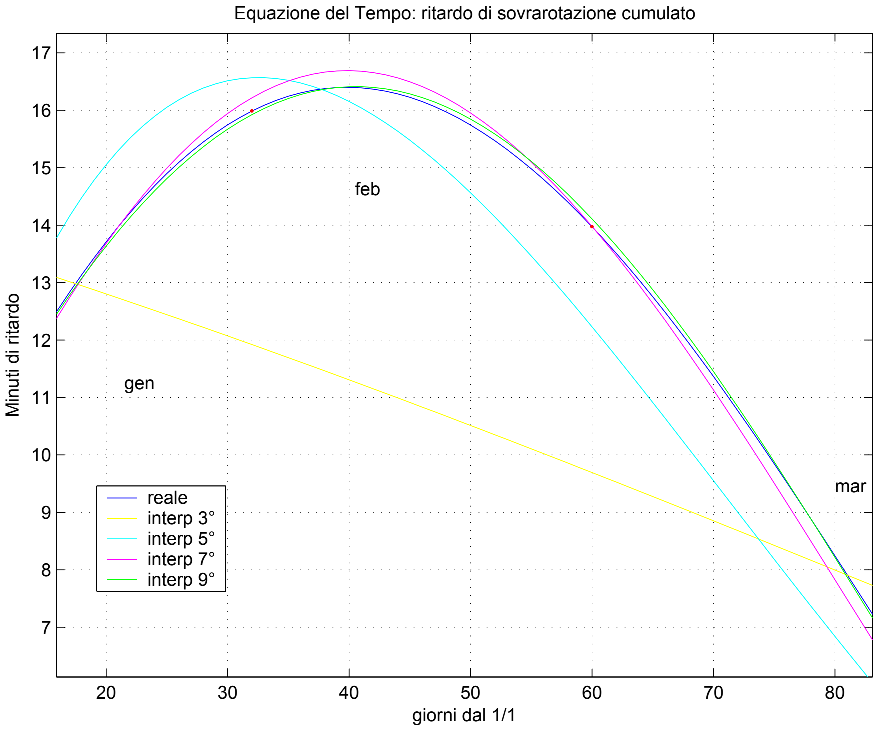



Equazione del Tempo: ritardo di sovrarotazione cumulato
reale
interp 3°
Minuti di ritardo
giorni dal 1/1
20
15
10
5
0
-5
-10
-15
0
50
100
150
200
250
300
350
400


Equazione del Tempo: ritardo di sovrarotazione cumulato
reale
interp 3°
interp 5°
Minuti di ritardo
giorni dal 1/1
20
15
10
5
0
-5
-10
-15
0
50
100
150
200
250
300
350
400


Equazione del Tempo: ritardo di sovrarotazione cumulato
Minuti di ritardo
giorni dal 1/1
reale
interp 3°
interp 5°
interp 7°


Equazione del Tempo: ritardo di sovrarotazione cumulato
Un ritardo negativo = anticipo
Minuti di ritardo
giorni dal 1/1
gen
feb
mar
apr
mag
giu
lug
ago
set
ott
nov
dic
reale
interp 3°
interp 5°
interp 7°
interp 9°



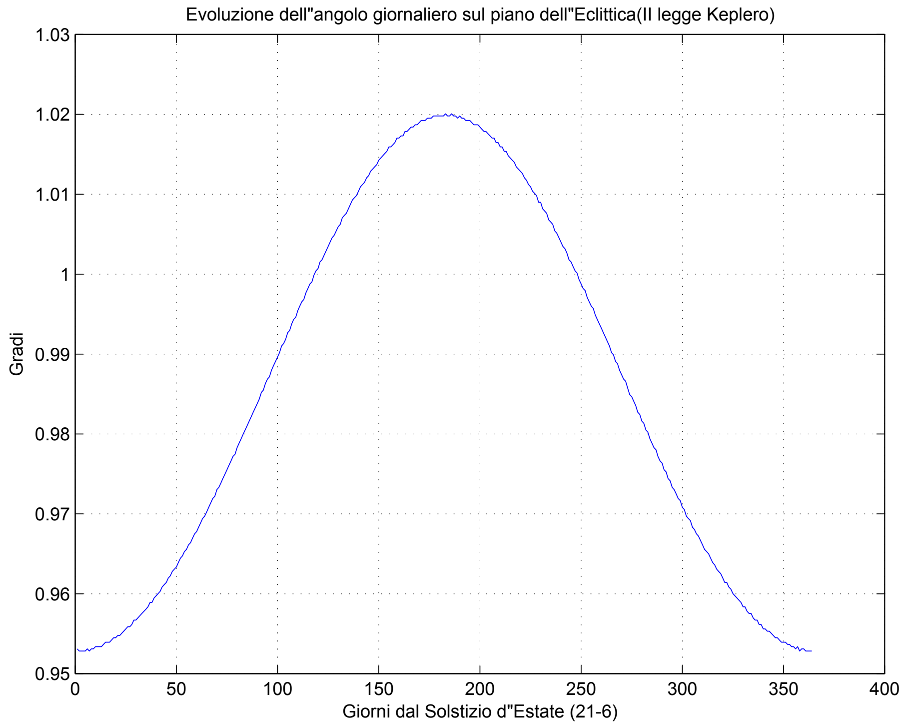
Considerando l'interpolazione di 9^o grado abbiamo:

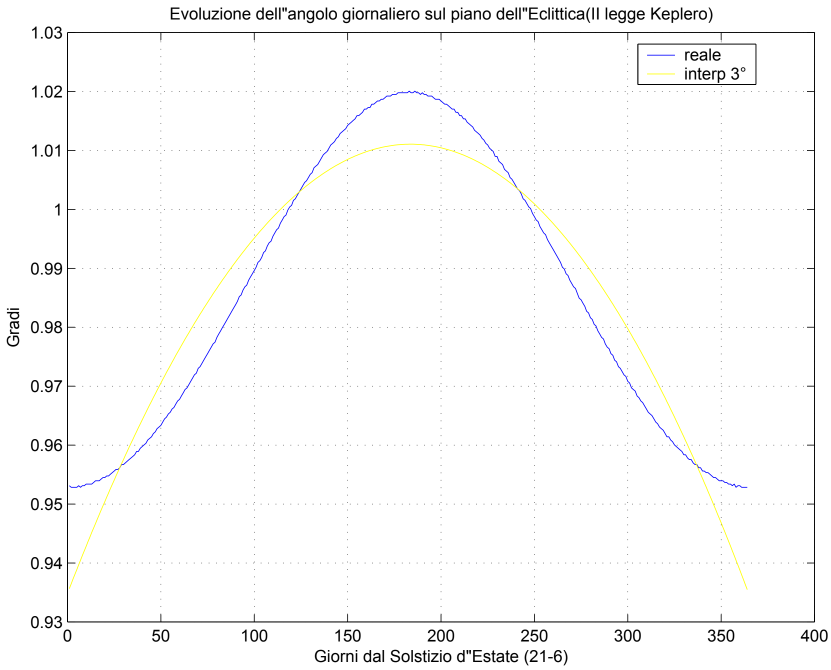
$$y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5 + Gx^6 + Hx^7 + Ix^8 + Jx^9$$

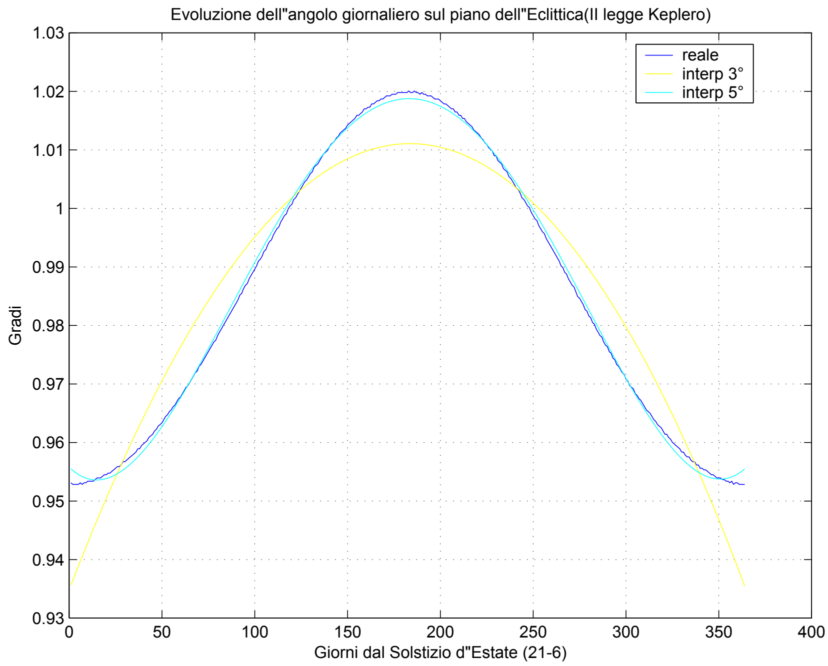
dove

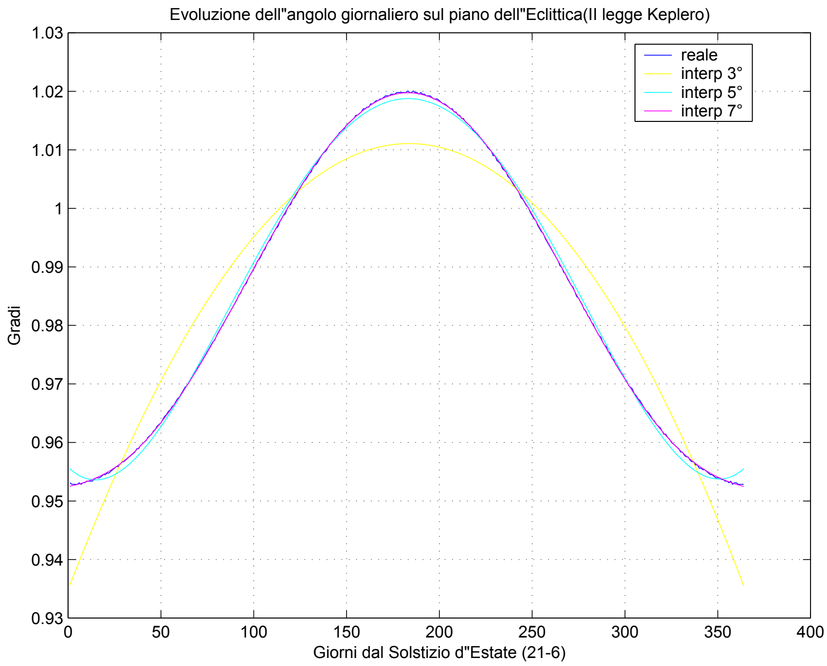
- $A = 6.064440205052961e + 000$ $\quad B = 4.883440707490081e - 001$
- $C = -5.697999782569241e - 003$ $\quad D = 3.686031917027322e - 005$
- $E = -1.825141488165372e - 006$ $\quad F = 2.847366603955792e - 008$
- $G = 2.847366603955792e - 008$ $\quad H = 6.164566509312907e - 013$
- $I = -1.002026362553123e - 015$ $\quad J = 6.447688942780012e - 019$

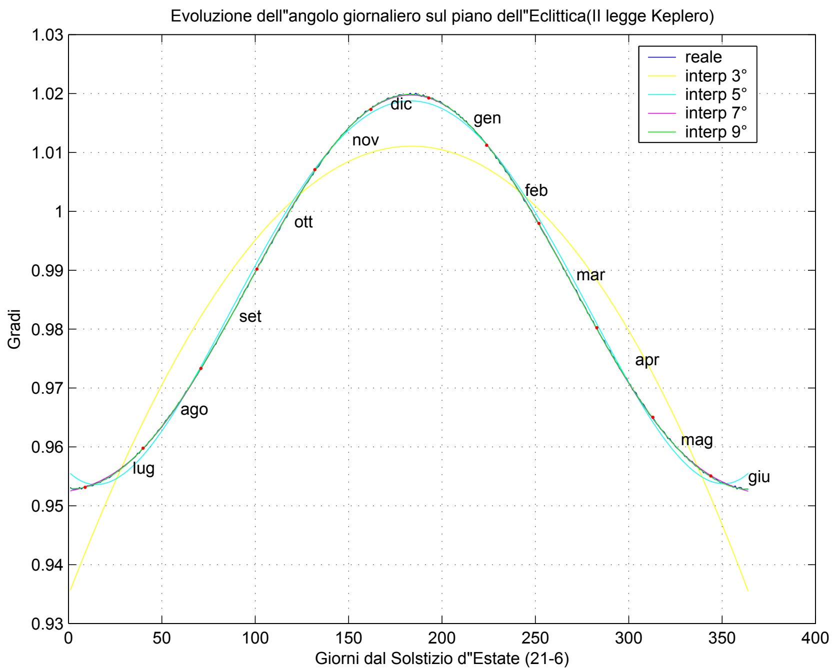
Vediamo ancora un esempio: ecco i dati da interpolare

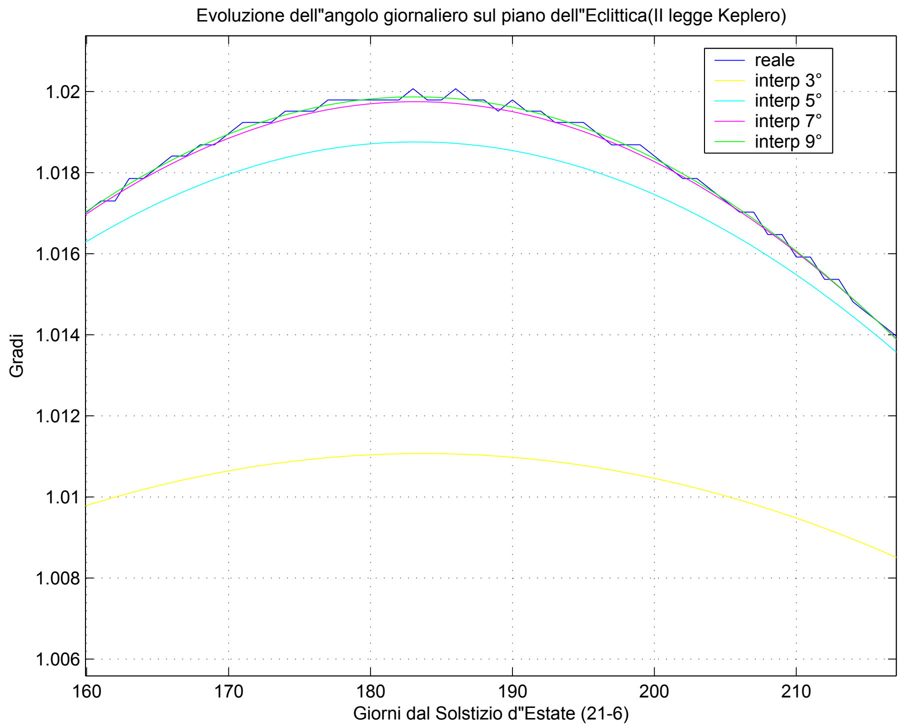



Evoluzione dell"angolo giornaliero sul piano dell"Eclittica(II legge Keplero)
Gradi
Giorni dal Solstizio d"Estate (21-6)
reale
interp 3°


Evoluzione dell"angolo giornaliero sul piano dell"Eclittica(II legge Keplero)
Gradi
Giorni dal Solstizio d"Estate (21-6)
reale
interp 3°
interp 5°


Evoluzione dell"angolo giornaliero sul piano dell"Eclittica(II legge Keplero)
reale
interp 3°
interp 5°
interp 7°
Gradi
Giorni dal Solstizio d"Estate (21-6)


Evoluzione dell"angolo giornaliero sul piano dell"Eclittica(II legge Keplero)
reale
interp 3°
interp 5°
interp 7°
interp 9°
lug
ago
set
ott
nov
dic
gen
feb
mar
apr
mag
giu
Gradi
Giorni dal Solstizio d"Estate (21-6)



Considerando l'interpolazione di 9^o grado abbiamo:

$$y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5 + Gx^6 + Hx^7 + Ix^8 + Jx^9$$

dove

- $A = 9.529548138234532e - 001$ $\quad B = -2.914134529419243e - 005$
- $C = 6.024863751008525e - 006$ $\quad D = -3.861119068915286e - 008$
- $E = 4.481461313154966e - 010$ $\quad F = -4.046453276618234e - 012$
- $G = 1.652769989011343e - 014$ $\quad H = -3.149617237237741e - 017$
- $I = 2.520047095173427e - 020v$ $\quad J = -4.268785104527807e - 024$

*Diego **Alberto**, Aprile 2005*