MINIMI QUADRATI NON LINEARI

Supponiamo di avere una $y = y(t, a)$ e di avere alcune misure sperimentali $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ della variabile dipendente y in corrispondenza di certi istanti $t_1, t_2, \dots, t_m$.

Cerchiamo di trovare per quale valore $\bar{a}$ del parametro a si ha il minimo della funzione:

$$\phi(a) = \sum_{i=1}^{m} \left[y(t_i, a) - \bar{y}_i \right]^2$$

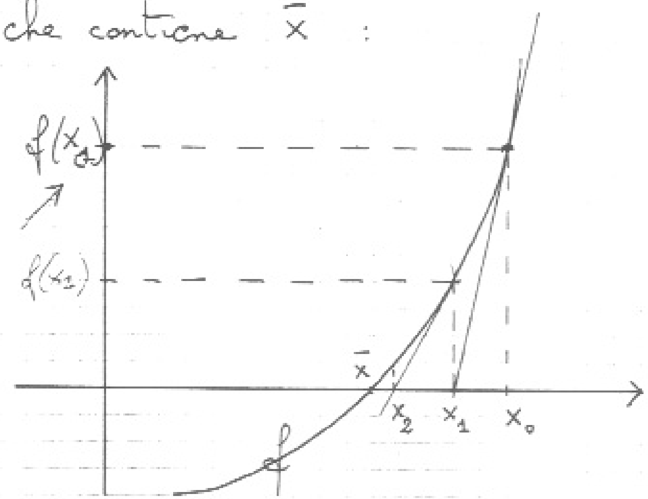
In corrispondenza di $\bar{a}$ avremo il migliore adattamento della funzione $y = y(t, a)$ ai dati sperimentali.

Se $\bar{a}$ è un punto di minimo per $\phi(a)$, avremo $\phi'(\bar{a}) = 0$.
Cercheremo la soluzione dell'equazione $\phi'(a) = 0$ con il metodo iterativo delle tangenti (o di Newton):

$$\bar{a} = \lim_{k \to +\infty} a_k \quad \text{dove} \quad a_{k+1} = a_k - \frac{\phi'(a_k)}{\phi''(a_k)}$$

con a_0 arbitrario (ma da scegliere vicino ad $\bar{a}$).

Ricordiamo il metodo di Newton per trovare uno zero $\bar{x}$ di una funzione $f(x)$ convessa e crescente (per esempio) su un intervallo che contiene $\bar{x}$:



x_2, per esempio, è lo zero della tangente nel punto $(x_1, f(x_1))$ che ha equazione:

$$y - f(x_1) = f'(x_1) \cdot (x - x_1).$$

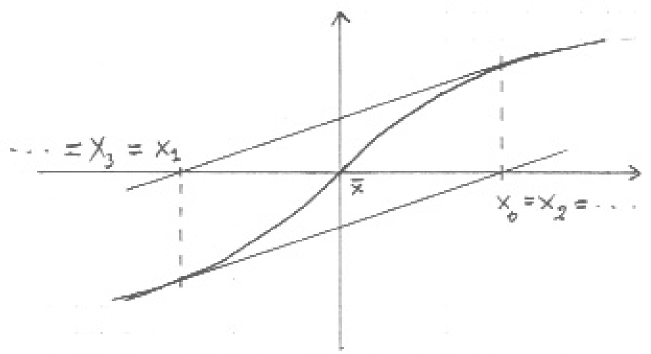
Ponendo $y = 0$ e risolvendo rispetto a $x = x_2$

$$x_2 - x_1 = -\frac{f(x_1)}{f'(x_1)} \quad \text{da cui:}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}; \quad \text{in generale} \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Osservazioni.

1. Il metodo può non convergere:



2. Quando converge in generale è molto veloce (più di un metodo tipo bisezione o peggio tabulazione della funzione in molti punti dell'intervallo, etc...).

3. Nel nostro caso l'equazione da risolvere non è $f(x)=0$ ma $\phi'(a)=0$ quindi la formula precedente diventa:

$$a_{k+1} = a_k - \frac{\phi'(a_k)}{\phi''(a_k)}$$

Occorre quindi saper valutare ϕ' e ϕ'' e generalizzare il tutto al caso "a" vettore.

Continuiamo prima nel caso a scalare:

$$\phi(a) = \sum_{i=1}^{m} \left[y(t_i, a) - \bar{y}_i \right]^2$$

$$\phi'(a) = 2 \sum_{i=1}^{m} \left[y(t_i, a) - \bar{y}_i \right] \cdot \frac{\partial y}{\partial a}(t_i, a)$$

$$\phi''(a) = 2 \sum_{i=1}^{m} \left[\frac{\partial y}{\partial a}(t_i, a) \right]^2 + $$

$$+ 2 \sum_{i=1}^{m} \left[y(t_i, a) - \bar{y}_i \right] \cdot \frac{\partial^2 y}{\partial a^2}(t_i, a) .$$

Per valori di a già abbastanza buoni, i fattori $\left[y(t_i, a) - \bar{y}_i \right]$ sono piccoli e la seconda sommatoria può essere trascurabile.

Allora l'algoritmo diventa:

$$\boxed{a_{k+1} = a_k - \frac{\cancel{2} \sum_{i=1}^{m} \left[y(t_i, a_k) - \bar{y}_i \right] \cdot \frac{\partial y}{\partial a}(t_i, a_k)}{\cancel{2} \sum_{i=1}^{m} \left[\frac{\partial y}{\partial a}(t_i, a_k) \right]^2}}$$

con a_0 "vicino" alla soluzione $\bar{a}$.

Se non converge, provare a cambiare a_0.

Il problema successivo è l'ottimizzazione del parametro quando $y(t,a)$ non è data esplicitamente, ma come soluzione di un problema ai valori iniziali:

$$(*)\begin{cases} y' = f(y,t,a) \\ y(t_0) = y_0 \end{cases}$$

nel quale l'equazione differenziale non sia integrabile in forma esatta, altrimenti si ricade nel caso precedente, ma solo con metodi approssimati tipo Runge-Kutta.

L'algoritmo illustrato:

$$a_{k+1} = a_k - \frac{\sum_{i=1}^{m}\left[y(t_i,a_k)-\bar{y}_i\right]\cdot\frac{\partial y}{\partial a}(t_i,a_k)}{\sum_{i=1}^{m}\left[\frac{\partial y}{\partial a}(t_i,a_k)\right]^2}$$

è ancora valido, ma richiede il calcolo, a ogni iterazione (di indice k), delle successioni: $y(t_i,a_k)$ $\quad i=1,\dots,m$

e $\frac{\partial y}{\partial a}(t_i,a_k)$ $\quad i=1,\dots,m$

dove $y(t,a)$ è definita dalla $(*)$.

Avendo come nuova incognita la $\frac{\partial y}{\partial a}(t,a)$ occorre trovare anche per lei una equazione differenziale che la definisca, che, messa in sistema con quella che definisce $y(t,a)$, sarà poi integrata con Runge-Kutta.

$$\frac{\partial\left(\frac{\partial y}{\partial a}\right)}{\partial t} = \frac{\partial^2 y}{\partial t\,\partial a} = \text{ per il teor. di Schwartz: }$$

$$= \frac{\partial^2 y}{\partial a\,\partial t} = \frac{\partial\left(\frac{\partial y}{\partial t}\right)}{\partial a} = (*)$$

ma $\frac{\partial y}{\partial t}$ è data dall'equazione di partenza, che è meglio scrivere per esteso:

$$\frac{\partial y}{\partial t}(t,a) = f(y(t,a),t,a) \text{ , quindi:}$$

$$(*) = \frac{\partial}{\partial a} f(y(t,a),t,a) =$$

$$= \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial a} + \frac{\partial f}{\partial a}$$

Riassumendo:

$$\frac{\partial}{\partial t}\left(\frac{\partial y}{\partial a}\right) = \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial a} + \frac{\partial f}{\partial a} \qquad \text{con } \frac{\partial y}{\partial a}(t_0,a_k) = 0$$

La condizione iniziale è giustificata dal fatto che poniamo: $\forall a \quad y(t_0, a) = y_0$

Il sistema da integrare sarà quindi:

$$\begin{cases} \dfrac{\partial y(t, a_k)}{\partial t} = f(y(t, a_k), t, a_k) \\ \dfrac{\partial z(t, a_k)}{\partial t} = \dfrac{\partial f}{\partial y}(y(t, a_k), t, a_k) \cdot z(t, a_k) + \dfrac{\partial f}{\partial a}(\text{idem}) \\ y(t_0, a_k) = y_0 \\ z(t_0, a_k) = 0 \end{cases}$$

dove $z(t, a)$ rappresenta $\frac{\partial y}{\partial a}(t, a)$.

Esempio

(1) $y' = t + a e^{y \cdot t} \qquad y(0) = 0$

$$f(y, t, a) = t + a e^{y \cdot t}$$

$$\frac{\partial f}{\partial y} = a t e^{y \cdot t} \qquad \frac{\partial f}{\partial a} = e^{y \cdot t}$$

L'equazione ausiliaria in $z(t, a)$ è:

(2) $z' = a t e^{y \cdot t} z + e^{y \cdot t} \qquad z(0) = 0$

$\boxed{\text{FACOLTATIVO}}$

Un'altra generalizzazione possibile è nel caso di una funzione di più parametri:

$$y = y(t, a_1, \dots, a_n)$$

o $y = y(t, \mathbf{a}) \qquad \mathbf{a}$ vettore $= (a_1, \dots a_n)$

Dovendo minimizzare la funzione

$$\phi(a_1, \dots, a_n) = \sum_{i=1}^{m} \left[y(t_i, a_1, \dots, a_n) - \bar{y}_i \right]^2$$

dobbiamo imporre $\nabla \phi(a_1, \dots, a_n) = \vec{0}$

Il metodo di Newton si generalizza a questa equazione vettoriale:

$$\mathbf{a}^{(k+1)} = \mathbf{a}^{(k)} + S^{(k)} \qquad \mathbf{a}^{(0)} \text{ arbitrario}$$

dove $S^{(k)}$ è soluzione dell'equazione vettoriale:

$$H\phi(\mathbf{a}^{k}) \cdot S^{(k)} = -\nabla \phi(\mathbf{a}^{k})$$

(che generalizza: $\quad s_k = -\dfrac{\phi'(a_k)}{\phi''(a_k)}$)

approssimabile con:

$$\left[J(\mathbf{a}^{(k)}) \right]^T \cdot J(\mathbf{a}^{(k)}) \cdot S^{(k)} = -\left[J(\mathbf{a}^{(k)}) \right]^T Y(\mathbf{a}^{(k)})$$

FACOLTATIVO

Nella pagina precedente

$$Y(\mathbf{a}) = \begin{bmatrix} y(t_1, \mathbf{a}) - \bar{y}_1 \\ \cdots\cdots \\ \cdots\cdots \\ y(t_m, \mathbf{a}) - \bar{y}_m \end{bmatrix}$$

$J(\mathbf{a})$ è la Jacobiana di Y, cioè, in pratica:

$$\begin{pmatrix} \dfrac{\partial y(t_1, \mathbf{a})}{\partial a_1} & \cdots\cdots & \dfrac{\partial y(t_1, \mathbf{a})}{\partial a_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial y(t_m, \mathbf{a})}{\partial a_1} & \cdots\cdots & \dfrac{\partial y(t_m, \mathbf{a})}{\partial a_n} \end{pmatrix}$$

$\left[J(\mathbf{a})\right]^T$ è la sua trasposta.

il gradiente di $\phi(\mathbf{a})$ è: $\nabla\phi(\mathbf{a}) = 2\left[J(\mathbf{a})\right]^T \cdot Y(\mathbf{a})$

mentre $H\phi$ è la matrice Hessiana di $\phi(a_1, \dots, a_n)$

cioè, per $n=2$ $\begin{pmatrix} \dfrac{\partial^2\phi}{\partial a_1^2} & \dfrac{\partial^2\phi}{\partial a_2 \partial a_1} \\ \dfrac{\partial^2\phi}{\partial a_1 \partial a_2} & \dfrac{\partial^2\phi}{\partial a_2^2} \end{pmatrix}$ e si ha:

$$H\phi \approx 2 J^T \cdot J$$

FACOLTATIVO

Invece di dimostrare le formule precedenti osserviamo solo che per $n=1$ (un solo parametro) si riducono a quelle di pagina 4. Infatti:

$-\left[J(\mathbf{a}^k)\right]^T \cdot Y(\mathbf{a}^k) =$ se $\mathbf{a}^k = a_k$ con a unico param.:

$$= -\left[\frac{\partial y}{\partial a}(t_1, a_k), \ \dots, \frac{\partial y}{\partial a}(t_m, a_k)\right] \cdot \begin{bmatrix} y(t_1, a_k) - \bar{y}_1 \\ \cdots\cdots \\ \cdots\cdots \\ y(t_m, a_k) - \bar{y}_m \end{bmatrix} =$$

$$= -\sum_{i=1}^{m} \frac{\partial y}{\partial a}(t_i, a_k) \cdot \left[y(t_i, a_k) - \bar{y}_i\right].$$

mentre

$$\left[J(\mathbf{a}^k)\right]^T \cdot J(\mathbf{a}^k) = \left[\frac{\partial y}{\partial a}(t_1, a_k), \dots, \frac{\partial y}{\partial a}(t_m, a_k)\right] \begin{bmatrix} \frac{\partial y}{\partial a}(t_1, a_k) \\ \vdots \\ \frac{\partial y}{\partial a}(t_m, a_k) \end{bmatrix}$$

$$= \sum_{i=1}^{m} \left[\frac{\partial y}{\partial a}(t_i, a_k)\right]^2$$

e ritroviamo il numeratore e il denominatore della formula in basso a pag. 4.

Integrazione numerica equazioni differenziali (11)

$$\begin{cases} y' = f(t,y) \\ y(t_0) = y_0 \end{cases} \iff y'(t) = f(t, y(t)) \quad \forall t \in U(t_0)$$

integrando fra t_0 e t:

$$\int_{t_0}^{t} y'(u)\,du = \int_{t_0}^{t} f(u, y(u))\,du$$

$$y(t) = y(t_0) + \int_{t_0}^{t} f(u, y(u))\,du \qquad \text{ma } y(t_0) = y_0$$

$$\boxed{y(t) = y_0 + \int_{t_0}^{t} f(u, y(u))\,du}$$ Equazione di Volterra

Approssimando l'integrale con la regola dei rettangoli:

$$y(t_0 + h) \approx y_1 = y_0 + h \cdot f(t_0, y(t_0)) = y_0 + h f(t_0, y_0)$$

da cui la formula ricorsiva (di Eulero)

$$(t_n = t_0 + n \cdot h) \quad \boxed{y_{n+1} = y_n + h \cdot f(t_n, y_n)} \quad \text{(imprecisa!)}$$

Formule di Runge-Kutta (2° ordine = Heun) (12)

$$y(t) = y_0 + \int_{t_0}^{t} f(u, y(u))\,du$$

Approssimando l'integr. con regola trapezi:

$$y(t_0 + h) \approx \tilde{y}_1 = y_0 + \frac{h}{2}\left[f(t_0, y(t_0)) + f(t_0 + h, y(t_0 + h))\right]$$

è l'incognita che stiamo cercando: la approssimiamo con Eulero!

$$\tilde{y}_1 \approx y_1 = y_0 + \frac{h}{2}\Big[\underbrace{f(t_0, y_0)}_{k_0} + \underbrace{f(t_1, y_0 + h \cdot \underbrace{f(t_0, y_0)}_{k_0})}_{k_1}\Big]$$

riassumendo:

$$\left.\begin{aligned} y_1 &= y_0 + \frac{h}{2}\left[k_0 + k_1\right] \\ \text{con:} \quad k_0 &= f(t_0, y_0) \\ k_1 &= f(t_1, y_0 + h \cdot k_0) \end{aligned}\right\} \text{formula di Heun} \quad (t_1 = t_0 + h)$$

Riassunto formule numeriche:

problema: $y'(t) = f(t, y(t)) \qquad y(t_0) = y_0$

equivalente a $\qquad y(t) = y_0 + \int_{t_0}^{t} f(u, y(u))\, du$

posto $t = t_1 = t_0 + h$, approssimando l'integrale in vari modi:

① $y(t_1) \approx y_1 = y_0 + h \cdot f(t_0, y_0)$ (Eulero ⊛)

② $y(t_1) \approx y_1 = y_0 + \frac{h}{2}\left[k_0 + k_1\right]$ (Heun)

$k_0 = f(t_0, y_0) \qquad k_1 = f(t_1, y_0 + h k_0)$

Aggiungiamo le formule di Runge Kutta del 4° ordine:

③ $y(t_1) \approx y_1 = y_0 + \frac{h}{6}\left[k_0 + 2k_1 + 2k_2 + k_3\right]$ (R-K-4)

$k_0 = f(t_0, y_0) \qquad k_1 = f\left(t_0 + \frac{h}{2}, y_0 + \frac{h}{2} \cdot k_0\right)$

$k_2 = f\left(t_0 + \frac{h}{2}, y_0 + \frac{h}{2} \cdot k_1\right) \qquad k_3 = f\left(t_0 + h, y_0 + h \cdot k_2\right)$

⊛ stabilità per $\frac{\partial f}{\partial y} < 0$

Eulero e Heun per un sistema di 2 equadiff. 14

$$\begin{cases} y' = f(t, y, z) & y(t_0) = y_0 \\ z' = g(t, y, z) & z(t_0) = z_0 \end{cases}$$

analogamente al caso di una equazione:

$$\begin{cases} y(t_0+h) = y_0 + \int_{t_0}^{t_0+h} f(u, y(u), z(u))\,du \\ z(t_0+h) = z_0 + \int_{t_0}^{t_0+h} g(u, y(u), z(u))\,du \end{cases}$$

$$\left.\begin{cases} y(t_0+h) \approx \tilde{y}_1 = y_0 + h \cdot f(t_0, y_0, z_0) \\ z(t_0+h) \approx \tilde{z}_1 = z_0 + h \cdot g(t_0, y_0, z_0) \end{cases}\right\} \underline{\text{Eulero}}$$

Esempio $y' = z$ $\quad z' = -y$ $\quad y(0) = 0$ $\quad z(0) = 1$

(soluz. esatta $y(t) = \operatorname{sen} t$ $\quad z(t) = \cos(t)$)

$$\begin{cases} \tilde{y}_1 = y_0 + h \cdot z_0 & y_0 = 0 \\ \tilde{z}_1 = z_0 - h \cdot y_0 & z_0 = 1 \end{cases}$$

oppure, più precisamente:

$$y(t_0+h) \approx y_1 = y_0 + \frac{h}{2}\left[f(t_0, y_0, z_0) + f(t_1, \tilde{y}_1, \tilde{z}_1)\right]$$

$t_1 = t_0 + h$

$$z(t_0+h) \approx z_1 = z_0 + \frac{h}{2}\left[g(t_0, y_0, z_0) + g(t_1, \tilde{y}_1, \tilde{z}_1)\right]$$

(Segue: Euler-Heun per sistemi) 15

e posto:

$k_{0y} = f(t_0, y_0, z_0)$ $\qquad k_{1y} = f(t_1, y_0 + h \cdot k_{0y}, z_0 + h \cdot k_{0z})$

$k_{0z} = g(t_0, y_0, z_0)$ $\qquad k_{1z} = g(t_1, y_0 + h k_{0y}, z_0 + h \cdot k_{0z})$

si ottiene:

$$\left.\begin{cases} y_1 = y_0 + \frac{h}{2}\left[k_{0y} + k_{1y}\right] \\ z_1 = z_0 + \frac{h}{2}\left[k_{0z} + k_{1z}\right] \end{cases}\right\} \underline{\text{Heun}}$$

Esempio

$$\begin{cases} y' = z & y(0) = 0 \\ z' = -y & z(0) = 1 \end{cases}$$

(soluz. esatta $\begin{cases} y(t) = \operatorname{sen} t \\ z(t) = \cos t \end{cases}$)

$f(t, y, z) = z$ $\qquad g(t, y, z) = -y$ $\qquad y_0 = 0$ $\quad z_0 = 1$

$y_0 = 0$	$k_{0y} = z_0$	$k_{1y} = z_0 + h \cdot k_{0z}$	$y_1 = y_0 + \frac{h}{2}\left[k_{0y} + k_{1y}\right]$
$z_0 = 1$	$k_{0z} = -y_0$	$k_{1z} = -y_0 - h\,k_{0y}$	$z_1 = z_0 + \frac{h}{2}\left[k_{0z} + k_{1z}\right]$
1	2	3	4

I numeri mostrano l'ordine in cui si devono definire le varie celle di Excel.