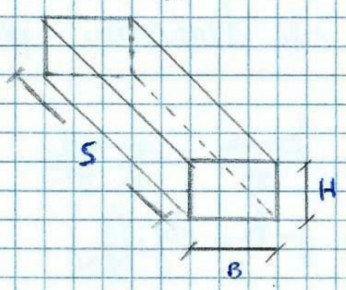
Metodi numerici per la risoluzione di equazioni e sistemi di equazioni non lineari

Consideriamo l'equazione di Manning per il calcolo dell'altezza H dell'acqua in un canale (si deve conoscere la portata con misura sperimentale):

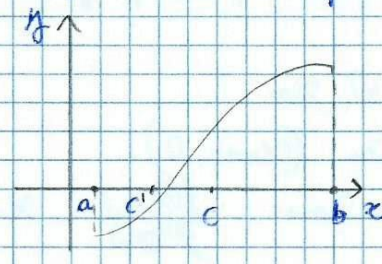


$$Q = \frac{S^{1/2}}{m} \cdot \frac{(B \cdot H)^{5/3}}{(B+2H)^{2/3}}$$

$$\Rightarrow f(H) := \frac{S^{1/2}}{m} \cdot \frac{(BH)^{5/3}}{(B+2H)^{2/3}} - Q = 0$$

Con S pendenza del canale e m parametrizzazione adimensionale dell'attrito del canale. L'unica incognita è H, ma la relazione tra f e H stessa non è lineare. Si tratta quindi di trovare gli zeri della funzione, ovvero le H che la annullano, ma non è disponibile una formula di risoluzione analitica. Tralasciando il metodo grafico, che consente di ottenere visivamente un'approssimazione della soluzione, si presentano due metodi numerici.

Il primo è quello di bisezione, probabilmente già noto al lettore dall'analisi matematica. Si considera un intervallo $[a,b]$ tale che $f(a) \cdot f(b) < 0$. Se f è continua il teorema di esistenza degli zeri garantisce che esista almeno uno zero in $]a,b[$. Si procede poi nel modo seguente, calcolato il punto medio $c = \frac{a+b}{2}$:



•) se $f(a) \cdot f(c) < 0 \Rightarrow$ nuovo intervallo $[a,c] \Rightarrow$ nuovo punto medio $c' = \frac{a+c}{2}$;

••) se $f(a) \cdot f(c) > 0$ (ovvero $f(c) \cdot f(b) < 0$) $\Rightarrow$ nuovo intervallo $[c,b] \Rightarrow$ nuovo punto medio $c' = \frac{c+b}{2}$.

In altre parole si controlla come $f(c)$ condiziona il segno dei prodotti con gli estremi degli intervalli di partenza. In figura sono rappresentati i casi $f(a) \cdot f(c) < 0$ e $f(c') \cdot f(\ ) < 0$, rispettivamente prima e seconda iterazione del processo. Quest'ultimo termina (si ripete infatti il procedimento ad ogni nuovo punto medio) quando l'intervallo ottenuto è più piccolo di una certa ε prefissata (tolleranza).

Data una tolleranza ε è possibile determinare il numero n di iterazioni che è necessario eseguire per raggiungerla:

$$I_0 = b-a \Rightarrow I_1 = \frac{b-a}{2} \Rightarrow I_2 = \frac{b-a}{2^2} \Rightarrow \dots \Rightarrow I_n = \frac{b-a}{2^n}$$

$$\Rightarrow \frac{b-a}{2^n} \le \varepsilon \Rightarrow \frac{b-a}{2^n} \simeq \varepsilon \Rightarrow n = \log_2\left(\frac{b-a}{\varepsilon}\right)$$

Possiamo altresì calcolare il numero di iterazioni per giungere ad una cifra significativa:

$$|x^{(k)} - \alpha| = \frac{|x^{(j)} - \alpha|}{10}, \quad x^{(k)} = \frac{a^k + b^k}{2^k} \Rightarrow \frac{b-a}{2^k} = \frac{b-a}{10 \cdot 2^j}$$

$$\Rightarrow 2^j \cdot 10 = 2^k \Rightarrow 2^{j-k} = 10^{-1} \Rightarrow j-k = \log_2\left(\frac{1}{10}\right)$$

$$\Rightarrow k-j = \log_2(10) \simeq 3{,}32$$

Il metodo di bisezione è convergente (giunge sempre alla soluzione), ma è piuttosto lento: servono ben 3,32 iterazioni per ottenere ogni cifra significativa del risultato.

Più rapido, benché non convergente, è il metodo di Newton-Raphson. Consiste nell'uso delle tangenti. Si deve scegliere un punto iniziale x_0 più vicino possibile allo zero della funzione. Si genera quindi una successione di valori partendo dalla retta tangente alla funzione in $(x_0, f(x_0))$ e che interseca l'asse delle x in x_1. In generale si considera la tangente in $(x_n, f(x_n))$ e la sua intersezione con x in x_{n+1} conoscendo la derivata prima in $(x_n, f(x_n))$, cioè $f'(x_n)$:

$$\begin{cases} f(x_n) = m x_n + q \\ m = f'(x_n) \end{cases} \Rightarrow f(x_n) = f'(x_n) \cdot x_n + q \Rightarrow q = f(x_n) - f'(x_n) \cdot x_n$$

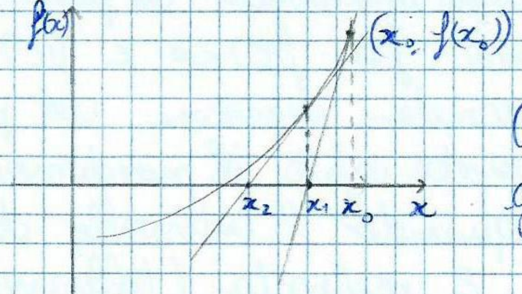
$$\Rightarrow y = f'(x_n)\, x + f(x_n) - f'(x_n) \cdot x_n$$

L'intersezione con l'asse x è in $(x_{n+1}, 0)$:

$$0 = f'(x_n) \cdot x_{n+1} + f(x_n) - f'(x_n) \cdot x_n \Rightarrow x_{n+1} + \frac{f(x_n)}{f'(x_n)} - x_n = 0$$

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad n = 0, 1, \dots$$

Si impone una condizione di arresto:



$$|x_{n+1} - x_n| \le \varepsilon_1$$

Con funzione molto pendente si aggiungono due condizioni:

$$|f(x_{n+1}) - f(x_n)| \le \varepsilon_2, \quad |f(x_n)| \le \varepsilon_3$$

Se la funzione è molto pendente si trovano infatti x_i non sufficientemente vicini allo zero della funzione.

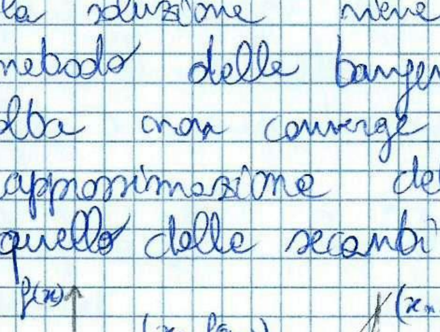
Oltre alla prima basta imporre una delle altre due condizioni (due su tre devono essere verificate).
Un metodo si dice convergente se $\lim_{n\to\infty} x_n = \alpha$ tale che $f(\alpha)=0$. Quello di Newton-Raphson, come già evidenziato, non è sempre convergente, può divergere. Tuttavia se converge ha ordine di convergenza quadratico, pari a 2.
Approfondiamo meglio questo concetto. Un metodo iterativo per equazioni non lineari è detto di ordine p se $\exists\, C>0$ tale che:

$$\underbrace{|x_{n+1}-\alpha|}_{\text{errore all'iterazione } n+1} \le C\underbrace{|x_n-\alpha|}_{\text{errore all'iterazione } n}^{p} \quad \forall n>\bar{n}$$

Misura della velocità di convergenza:

- $p=1 \Rightarrow$ convergenza lineare:
$$|x_{n+1}-\alpha| \le C\,|x^n-\alpha|^1$$
- $p\in(1,2) \Rightarrow$ convergenza superlineare;
- $p=2 \Rightarrow$ convergenza quadratica:
$$|x_{n+1}-\alpha| \le C\,|x^n-\alpha|^2$$

Con $p=2$, se a n si ha $\frac{1}{100}$ di ε, a $n+1$ si avrà $\frac{1}{10000}$ di ε: la soluzione viene raggiunta molto velocemente! Il metodo delle tangenti possiede tale velocità, ma talvolta non converge (oscilla).
Un'approssimazione del metodo delle tangenti è data da quello delle secanti:



$$\begin{cases} f(x_{n-1}) = m x_{n-1} + q \\ f(x_n) = m x_n + q \end{cases}$$

$$\Rightarrow \begin{cases} q = f(x_n) - m x_n & (2) \\ f(x_{n-1}) = m x_{n-1} + f(x_n) - m x_n & (1) \end{cases}$$

$$(1)\quad m(x_{n+1}-x_n) = f(x_{n-1}) - f(x_n) \Rightarrow m = \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}$$

$$\Rightarrow q = f(x_n) - x_n \cdot \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}$$

$$\Rightarrow y = \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}\,x + f(x_n) - x_n \cdot \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}$$

È l'equazione della retta in figura. Ad essa appartiene il punto $(x_{n+1}, 0)$:

$$0 = \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}\,x_{n+1} + f(x_n) - x_n \cdot \frac{f(x_{n-1})-f(x_n)}{x_{n-1}-x_n}$$

Moltiplicando per il reciproco del coefficiente angolare:

$$x_{n+1} = x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \qquad n = 0, 1, \ldots$$

Il metodo delle secanti è superlineare di ordine $p = \frac{1+\sqrt{5}}{2}$, quindi inferiore a quello del metodo delle tangenti. Presenta altresì il problema della convergenza.

L'espressione della successione può anche esplicitata usando lo sviluppo in serie di Taylor. Ad esempio per il metodo delle tangenti:

$$0 = f(\alpha) = f(x) + f'(x) \cdot (\alpha - x) + O((\alpha - x)^2) \Rightarrow \alpha = x - \frac{f(x)}{f'(x)} \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Si trascurano i termini di ordine uguale o superiore al secondo.

Per un sistema di equazioni non lineari definiamo dei vettori. Uno è quello delle incognite $\mathbb{X} = (x_1, x_2, \ldots, x_m)$:

$$\begin{cases} f_1(x_1, x_2, \ldots, x_m) = 0 \\ f_2(x_1, x_2, \ldots x_m) = 0 \\ \vdots \\ f_N(x_1, x_2, \ldots, x_m) = 0 \end{cases} \Rightarrow \mathbb{F}(\mathbb{X}) = \begin{pmatrix} f_1(\mathbb{X}) \\ f_2(\mathbb{X}) \\ \vdots \\ f_N(\mathbb{X}) \end{pmatrix}$$

Cerchiamo α tale che $F(\alpha) = 0$. Per esprimere l'equazione generica del sistema dobbiamo estendere il concetto di derivata nelle N ed M dimensioni in cui è definito il sistema:

f' se $f: \mathbb{R} \to \mathbb{R}$ DERIVATA

∇f se $f: \mathbb{R}^N \to \mathbb{R}$ GRADIENTE

$\mathbb{J}f$ se $f: \mathbb{R}^N \to \mathbb{R}^M$ JACOBIANO

Quindi si passa da $(n = 0, 1, \ldots)$:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{a} \quad \mathbb{X}_{n+1} = \mathbb{X}_n - (\mathbb{J}\mathbb{F}(\mathbb{X}_n))^{-1} \cdot \mathbb{F}(\mathbb{X}_n)$$

Anziché calcolare l'inversa dello jacobiano si preferisce definire $\delta_n = (\mathbb{J}\mathbb{F}(\mathbb{X}_n))^{-1} \cdot \mathbb{F}(\mathbb{X}_n)$, da cui:

$$\mathbb{X}_{n+1} = \mathbb{X}_n - \delta_n \Rightarrow \mathbb{F}(\mathbb{X}_n) = \mathbb{J}\mathbb{F}(\mathbb{X}_n) \cdot \delta_n$$

Si tratta di un sistema lineare, $\mathbb{b} = \mathbb{A}\mathbb{X}$. L'incognita è δ_n. Come condizione si considera la norma del vettore:

$$\|\mathbb{X}_{n+1} - \mathbb{X}_n\| \leq \varepsilon$$