FATTI DI ANALISI NUMERICA

GENERALITATE

- (**Numeri rappresentabili**) In un metodo numerico i numeri rappresentabili oltre allo zero sono tutti e soli quelli che si possono scrivere in una assegnata base B con un numero finito n di cifre e con un esponente, più precisamente come
$$\pm 0.c_1 c_2 \cdots c_n \times B^e$$
con il significato di $\pm B^e \cdot \sum_{i=1}^{n} c_i \cdot B^{-i}$, dove $c_1, \ldots, c_n$ sono le cifre, ovvero degli interi compresi tra 0 e $B-1$.

- (**Teorema di rappresentazione in base**) Per ogni numero reale $x \neq 0$ esistono unici un intero p ed una successione $\{d_i\}_{i \geq 1}$ tali che $0 \leq d_i \leq B-1$, $d_1 \neq 0$ e $\forall k > 0 \quad \exists j \geq k$ t.c. $d_j \neq B-1$ per i quali quindi
$$x = \text{sgn}\,(x) B^p \sum_{i=1}^{\infty} d_i B^{-i}$$

- (**Errore inerente**) Assegnata una funzione $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ vorremmo calcolare $f(x)$ per un assegnato valore di $x \in \Omega$. Purtroppo dobbiamo accontentarci di calcolare $f(\tilde{x})$, dove $\tilde{x} \in \mathcal{F}^n \cap \Omega$ è una n-upla di numeri macchina tali che $\tilde{x}_i = x_i(1+\varepsilon_i)$, dove ε_i sono gli errori di rappresentazione tali che $\mid \varepsilon_i \mid < u$. Ancora prima di iniziare abbiamo a che fare con l'errore relativo
$$\varepsilon_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)}$$
definito se $f(x) \neq 0$

- (**Errore algoritmico**) Supponiamo che $f : \mathbb{R} \to \mathbb{R}$ sia razionale. Vogliamo calcolare il valore di $f(\tilde{x})$ eseguendo un'opportuna sequenza di operazioni aritmetiche ciascuna delle quali introduce potenzialmente un errore locale. La funzione calcolata in verità sarà qualcosa di diverso da $f(\tilde{x})$, in generale che indichiamo con $\varphi(\tilde{x})$. Definiamo quindi l'errore algoritmico
$$\varepsilon_{alg} = \frac{\varphi(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$$
generato dall'accumularsi degli errori locali relativi a ciascuna operazione aritmetica eseguita in floating point

- (**Errore analitico**) Nel caso di una funzione non razionale $g(x) : \Omega \to \mathbb{R}$ dobbiamo selezionare una funzione razionale $f(x)$ che ben approssimi $g(x)$. Definiamo quindi l'errore analitico definito da
$$\varepsilon_{an} = \frac{f(x) - g(x)}{g(x)}$$

- (**Errore totale**)
$$\varepsilon_{tot} = \frac{\varphi(\tilde{x}) - g(x)}{g(x)}$$
esprime di quanto il valore effettivamente calcolato $\varphi(\tilde{x})$ si discosta dal valore $f(x)$ che avremmo voluto calcolare. $\varepsilon_{tot\cdot} = \varepsilon_{in} + \varepsilon_{alg} + \varepsilon_{an}$

- (**Analisi dell'errore inerente**) Se $f : \mathbb{R} \to \mathbb{R}$ è almeno $\mathcal{C}^2([x, \tilde{x}])$. Considerando l'errore di rappresentazione $\delta_x = \frac{\tilde{x}-x}{x}$ si ricava $\varepsilon_{in.} = \delta_x \frac{xf'(x)}{f(x)}$. La quantità $\delta_x \frac{xf'(x)}{f(x)}$ viene detta coefficiente di amplificazione. Nel caso di $f : \mathbb{R}^n \to \mathbb{R}$ vale una formula analoga per l'errore inerente. Infatti detto $x = (x_i), \delta_{x_i} = \frac{\tilde{x}_i - x_i}{x_i}$ risulta $\varepsilon_{in.} = \sum_{i=1}^n \delta_{x_i} C_i$, dove $C_i = \frac{x_i \frac{\partial f(x)}{\partial x_i}}{f(x)}$ sono i coefficienti di amplificazione rispetto alla variabile x_i

- (**Coefficienti di amplificazione delle operazioni aritmetiche**)

Operazione	C_1	C_2
moltiplicazione	1	1
divisione	1	-1
addizione	$\frac{x_1}{x_1+x_2}$	$\frac{x_2}{x_1+x_2}$
sottrazione	$\frac{x_1}{x_1-x_2}$	$-\frac{x_2}{x_1-x_2}$