Appunti di Algebra

http://poisson.dm.unipi.it/~daurizio/Algebra2011.pdf

Jacopo D'Aurizio
elianto84@gmail.com

11 Ottobre 2011

1 Il campo dei numeri complessi

Sia $i \notin \mathbb{R}$ una radice quadrata di -1. Lo spazio vettoriale

$$\mathbb{R}[i] = \left\{a + b\,i : (a,b) \in \mathbb{R}^2\right\}$$

è detto *campo dei numeri complessi* e usualmente denotato con $\mathbb{C}$. Valgono i seguenti fatti:

- $(a+b\,i)(c+d\,i) = (ac-bd)+(ac+bd)\,i$;
- $\frac{1}{a+b\,i} = \frac{1}{a^2+b^2}\,(a-b\,i)$.

Se $z = a+b\,i$, la *parte reale* di z è $\Re(z) = a$, la *parte immaginaria* è $\Im(z) = b$.
Il *coniugio* è un automorfismo dello spazio vettoriale:

$$\bar{z} = \Re(z) - i\cdot\Im(z),$$

mentre il *modulo* è la distanza del punto $(a,b) \in \mathbb{R}^2$ dall'origine:

$$|z| = \sqrt{a^2+b^2} = z\cdot\bar{z}.$$

Estendendo il dominio dell'usuale funzione esponenziale al campo dei numeri complessi, attraverso [1]

$$e^z = \sum_{j=0}^{+\infty}\frac{z^j}{j!},$$

si ha che per ogni numero reale θ vale *l'identità di De Moivre*:

$$e^{i\theta} = \cos(\theta) + i\,\sin(\theta);$$

l'*argomento* di $z \neq 0$, in simboli $\arg z$, è l'unico numero reale θ nell'intervallo $[0,2\pi)$ per cui si ha:

$$z = |z|e^{i\theta}.$$

Si noti che per $z,w \in \mathbb{C}\setminus\{0\}$ si verifica:

$$|z\cdot w| = |z|\cdot|w|, \qquad \arg(z\cdot w) = \arg(z)+\arg(w) \pmod{2\pi}.$$

Si ha inoltre:

$$\cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2}, \qquad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

[1] Notiamo come il parametro dell'esponenziale così definito possa vivere in spazi molto diversi: se z è un numero reale, troviamo una funzione analitica intera; se z è un numero complesso, troviamo una funzione olomorfa e intera; se z è una matrice, troviamo l'esponenziale di matrice, che fornisce le soluzioni dei sistemi dinamici linearizzati; se z è un generico operatore differenziale con spettro limitato, troviamo il flusso integrale.

2 Polinomi

Dato un anello $\mathbb{A}$, commutativo e con identità, l'anello dei polinomi a coefficienti in $\mathbb{A}$, in simboli $\mathbb{A}[x]$, è costituito dagli elementi

$$p(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n, \quad a_i \in \mathbb{A}.$$

Poiché $\mathbb{Z}$ è un anello euclideo, la funzione *grado* rende $\mathbb{Z}[x]$ un anello euclideo: tra i polinomi a coefficienti interi è dunque possibile implementare un algoritmo coerente di divisione con resto. Se ora ξ è *radice* di $p(x)$, ossia realizza $p(\xi) = 0$, si ha

$$(x - \xi) \mid p(x),$$

(il viceversa è banale) e ξ è detta radice *di molteplicità* k se

$$(x - \xi)^k \mid p(x), \quad (x - \xi)^{k+1} \nmid p(x).$$

Un risultato centrale è dovuto a Gauss[2]:

Teorema 2.1 (fondamentale dell'Algebra)**.** *Ogni polinomio a coefficienti reali o complessi di grado k ammette esattamente k radici in $\mathbb{C}$, contate con molteplicità.*

Altri risultati importanti sono:

Lemma 2.2. *Ogni polinomio di grado dispari a coefficienti reali ammette almeno una radice reale.*

Lemma 2.3. *Se $p(x) = a_0 x^n + \ldots + a_n$ è un polinomio a coefficienti interi con una radice razionale $\xi = \pm\frac{a}{b}$, allora $a|a_n$ e $b|a_0$.*

Lemma 2.4. *Se $\xi \in \mathbb{C} \setminus \mathbb{R}$ è radice di $p(x) \in \mathbb{R}[x]$, allora lo è anche $\bar{\xi}$.*

Lemma 2.5. *Ogni polinomio a coefficienti interi coprimo con la sua derivata è privo di radici multiple (ossia con molteplicità maggiore di uno).*

Dimostrazione. Se ξ è radice di molteplicità $k \geq 2$ di $p(x)$, allora $(x - \xi)^k$ divide $p(x)$, ossia:

$$p(x) = (x - \xi)^k q(x),$$

da cui:

$$p'(x) = (x - \xi)^{k-1} \left((x - \xi)q'(x) + k\, q(x)\right),$$

dunque ξ è radice di molteplicità $k - 1$ di $p'(x)$. □

Analizziamo ora le relazioni che intercorrono tra radici e coefficienti. Utilizziamo la notazione

$$[x^k]\, p(x)$$

per denotare il coefficiente del termine x^k nel polinomio $p(x)$; diciamo che un polinomio è *monico* se $[x^{\partial p}]\, p(x) = 1$. Si ha:

Teorema 2.6 (Viète)**.** *Se $p(x) \in \mathbb{C}[x]$ è monico di grado n, con radici $\xi_1, \ldots, \xi_n \in \mathbb{C}$ (eventualmente coincidenti), allora*

$$[x^{n-j}]\, p(x) = (-1)^j \sum_{1 \leq i_1 < \ldots < i_j \leq n} \prod_{h=1}^{j} \xi_{i_j}$$

Dimostrazione. E' sufficiente considerare il coefficiente di x^{n-j} in $p(x) = \prod_{j=1}^{n}(x - \xi_j)$. □

E' ora naturale discutere di *funzioni simmetriche*: una funzione $f : \mathbb{K}^n \to \mathbb{K}$ è detta *simmetrica* se

$$\forall \sigma \in S_n,\ f(x_1, x_2, \ldots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}).$$

Alla categoria appartengono le funzioni simmetriche elementari di n variabili:

$$e_1(x_1, \ldots, x_n) = x_1 + \ldots + x_n, \qquad e_2(x_1, \ldots, x_n) = x_1x_2 + x_1x_3 + \ldots + x_1x_n + x_2x_3 + \ldots + x_{n-1}x_n,$$

[2]Ne omettiamo la dimostrazione, di carattere prettamente analitico. I curiosi sono liberi di indagare sul concetto di *indice di avvolgimento* di una curva e sul *principio di Rouché*.

$$e_k(x_1, \ldots, x_n) = \sum_{1 \le i_1 < \ldots < i_k \le n} x_{i_1} \cdot \ldots \cdot x_{i_k}$$

e le somme di potenze:

$$p_k(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j^k.$$

Le une sono legate alle altre attraverso il seguente

Teorema 2.7 (formule di Newton-Girard)**.**

$$k \cdot e_k = \sum_{j=1}^{k} (-1)^{j-1} e_{k-j}\, p_1.$$

Dimostrazione. Si consideri il polinomio $p(t)$, di grado n, avente termine noto 1 e radici $\frac{1}{x_1}, \ldots, \frac{1}{x_n}$. Per il teorema di Viète si ha:

$$p(t) = \prod_{j=1}^{n} (1 - x_j\, t) = \sum_{k=0}^{n} (-1)^k\, e_k\, t^k,$$

alché, derivando il membro centrale e il membro destro rispetto a t, quindi moltiplicando per t:

$$\sum_{k=0}^{n} (-1)^k\, k e_k\, t^k = -\left(\sum_{j=1}^{n} \frac{x_j\, t}{1 - x_j\, t} \right) \cdot \prod_{j=1}^{n} (1 - x_j\, t).$$

Espandendo ora il fattore sinistro del membro destro in serie di potenze otteniamo:

$$\sum_{k=0}^{n} (-1)^k\, k e_k\, t^k = \left(\sum_{j=1}^{\infty} p_j\, t^j \right) \cdot \left(\sum_{l=0}^{n} (-1)^{l-1}\, e_l\, t^l \right),$$

e la tesi segue dal raffronto del coefficiente di x^k nei due termini. □

Abbiamo dunque modo di esprimere le funzioni simmetriche elementari in termini delle somme di potenze e viceversa, il tutto indipendentemente dal numero di variabili in gioco. Un'altra funzione simmetrica notevole è il *discriminante* di un polinomio: sia $p(z) \in \mathbb{C}[z]$ un polinomio con radici $\xi_1, \ldots, \xi_{\partial p} \in \mathbb{C}$ (eventualmente ripetute se presenti con molteplicità maggiore di uno). La quantità[3]

$$\Delta p = \prod_{1 \le i < j \le \partial p} (\xi_i - \xi_j)^2 = (-1)^{\binom{\partial p}{2}} \prod_{j=1}^{\partial p} p'(\xi_i)$$

è un polinomio simmetrico nelle variabili $\xi_1, \ldots, \xi_{\partial p}$, dunque può essere espresso in termini dei coefficienti di p, in quanto:

Lemma 2.8. *Ogni polinomio simmetrico può essere espressa attraverso somme e prodotti di funzioni simmetriche elementari.*

Dimostrazione. E' sufficiente provare la tesi per polinomi omogenei, ossia somme di monomi di pari grado, quindi procedere per induzione sul grado delle variabili nei monomi. □

Inoltre, è evidente che il discriminante è invariante per traslazioni: se $q(x) = p(x + \tau)$, allora $\Delta p = \Delta q$.

Discutiamo ora di numeri *algebrici*. Un numero $z \in \mathbb{C}$ è detto algebrico su $\mathbb{Q}$ se esiste un polinomio monico $p \in \mathbb{Q}[x]$ per cui si abbia $p(z) = 0$, o, equivalentemente, se esiste un polinomio (non necessariamente monico) $q(x) \in \mathbb{Z}[x]$ per cui si abbia $q(z) = 0$. Se z è algebrico su $\mathbb{Q}$, esiste un unico elemento monico di

[3]La seconda uguaglianza segue dal fatto che, posto $n = \partial p$,

$$p(x) = \prod_{j=1}^{n} (x - \xi_j) \quad \longrightarrow \quad p'(\xi_i) = \prod_{j \ne i} (x - \xi_j).$$

$\mathbb{Q}[x]$ di minimo grado che annulla z: se infatti vi fossero due distinti polinomi monici p_1, p_2 di grado n (minimo) per cui $p_1(z) = 0 = p_2(z)$, z risulterebbe radice di $p_1 - p_2$, di grado strettamente inferiore a n, contravvenendo l'ipotesi di minimalità. Tale elemento è detto *polinomio minimo* di z su $\mathbb{Q}$: le sue radici sono dette *radici coniugate* di z, il suo grado è il *grado algebrico* di z su $\mathbb{Q}$. Ogni polinomio minimo su $\mathbb{Q}$ di grado n è *irriducibile*, ossia non può essere espresso come prodotto di due elementi di $\mathbb{Q}[x]$ di grado inferiore ad n; inoltre è privo di radici multiple.
Esistono diversi criteri per certificare l'irriducibilità di un polinomio su $\mathbb{Q}$:

Lemma 2.9 (Eisenstein)**.** *Se $q(x) \in \mathbb{Z}[x]$ è un polinomio monico di grado n, ed esiste un numero primo p per cui si abbia:*

$$\forall m \in [1, n-1] \quad p \mid [x^m]q(x), \qquad p^2 \nmid q(0),$$

allora $q(x)$ è irriducibile su $\mathbb{Q}$.

Lemma 2.10. *Se $q(x)$ è irriducibile su $\mathbb{K} \supseteq \mathbb{Z}$, lo è anche $q(x+m)$ per ogni intero m.*

Lemma 2.11. *Se nell'immagine di $\mathbb{Z}$ attraverso $q(x) \in \mathbb{Z}[x]$ vi sono più di $2 \cdot \partial q$ numeri primi, $q(x)$ è irriducibile su $\mathbb{Q}$.*

Lemma 2.12. *Se $q(x)$ è irriducibile su $\mathbb{F}_p$ lo è anche su $\mathbb{Q}$.*

Lemma 2.13 (Stickelberger)**.** *Se $q(x) \in \mathbb{F}_p[x]$ è privo di radici multiple, condizione necessaria affinché $q(x)$ sia irriducibile su $\mathbb{F}_p$ è che si abbia $\left(\frac{\Delta q}{p}\right) = (-1)^{1+\partial q}$.*

E' inoltre importante ricordare che:

Lemma 2.14. *Per ogni numero naturale $n \geq 2$, il polinomio minimo su $\mathbb{Q}$ di $z = e^{\frac{2\pi i}{n}}$ (n-esimo polinomio ciclotomico) ha grado $\phi(n)$; inoltre ogni numero della forma z^m con $\gcd(m, n) = 1$ è radice coniugata di z.*

Lemma 2.15. *Per ogni numero naturale n e per ogni coppia (α, β) di numeri algebrici su $\mathbb{Q}$ non nulli,*

$$\alpha^{\frac{1}{n}}, \quad \alpha^n, \quad \alpha \cdot \beta, \quad \frac{\alpha}{\beta}, \quad \alpha + n\,\beta$$

sono numeri algebrici su $\mathbb{Q}$.

Dimostrazione. Supponiamo che α e β abbiano gradi algebrici u e v, e polinomi minimi

$$p_\alpha(x) = x^u - a(x), \qquad p_\beta(x) = x^v - b(x).$$

Sia ora $\mathbb{V}$ lo spazio vettoriale generato dai monomi $\alpha^t\,\beta^s$ con $0 \leq t < u$, $0 \leq s < v$: tale spazio ha dimensione limitata da $u \cdot v$. Ne consegue che, comunque presi $uv + 1$ elementi di $\mathbb{V}$, ne esiste una combinazione lineare nulla. In particolare $\alpha + \beta$ risulta radice di un polinomio a coefficienti in $\mathbb{Q}$ di grado limitato da $u \cdot v$ (prendiamo $1, (\alpha+\beta), (\alpha+\beta)^2, \ldots, (\alpha+\beta)^{uv}$ come elementi di $\mathbb{V}$), e lo stesso vale per $\alpha \cdot \beta$. Per quanto riguarda α^n, si riproduca lo stesso ragionamento sullo spazio vettoriale generato da $1, \alpha, \alpha^2, \ldots, \alpha^{u-1}$. Si noti inoltre che $\frac{1}{\beta}$ è radice del polinomio

$$x^v\, p_\beta\left(\frac{1}{x}\right) \in \mathbb{Q}[x],$$

$\alpha^{\frac{1}{n}}$ è radice del polinomio

$$p_\alpha(x^n) \in \mathbb{Q}[x],$$

$n \cdot \beta$ è radice del polinomio

$$p_\beta\left(\frac{x}{n}\right) \in \mathbb{Q}[x].$$

A questo punto i polinomi minimi delle quantità citate nel lemma vanno identificati tra i fattori irriducibili dei polinomi così costruiti. Si ricordi tuttavia che, se u è coprimo con v, il grado algebrico di $\alpha + \beta$ su $\mathbb{Q}$ è esattamente pari al prodotto $u \cdot v$, il che garantisce l'automatica irriducibilità del polinomio costruito attraverso $\mathbb{V}$. □

Parliamo ora di problemi di problemi di *interpolazione*: ci chiediamo, ad esempio, come determinare un polinomio a coefficienti interi $p(x)$ che realizzi

$$\begin{cases} p(1) &= 13 \\ p(2) &= 27 \\ p(3) &= 40. \end{cases}$$

Per l'euclidicità di $\mathbb{Z}[x]$ abbiamo:

$$(x-1) \,|\, p(x) - 13, \qquad (x-2) \,|\, p(x) - 27, \qquad (x-3) \,|\, p(x) - 40.$$

Condizione necessaria e sufficiente per soddisfare la prima condizione è che si abbia:

$$p(x) = (x-1)p_1(x) + 13,$$

alché la seconda e la terza condizione mutano in:

$$p_1(2) = 14, \qquad 2 \cdot p_1(3) = 27,$$

dunque non vi sono soluzioni in $\mathbb{Z}[x]$, in quanto p_1, polinomio a coefficienti interi, non può valere $\frac{27}{2}$ nel punto 3. Proseguiamo tuttavia nella determinazione delle soluzioni in $\mathbb{Q}[x]$:

$$p_1(x) = (x-2)p_2(x) + 14, \qquad p_2(3) = -\frac{1}{2},$$

$$p_2(x) = (x-3)q(x) - \frac{1}{2}.$$

Abbiamo che tutte le soluzioni in $\mathbb{Q}[x]$ sono della forma:

$$p(x) = 13 + (x-1)\left(14 + (x-2)\left(-\frac{1}{2} + (x-3)q(x)\right)\right),$$

in corrispondenza con le soluzioni $q(x)$ del sistema

$$\begin{cases} q(0) &= 0 \\ q(1) &= 14 \\ q(2) &= 27. \end{cases}$$

attraverso la relazione $q(x) + 13 = p(x+1)$. Saremmo potuto pervenire più rapidamente alla soluzione attraverso una semplice considerazione: se

$$\begin{cases} n_1(x) &= \frac{1}{2}(x-2)(x-3) \\ n_2(x) &= -(x-1)(x-3) \\ n_3(x) &= \frac{1}{2}(x-1)(x-2) \end{cases}$$

si ha che, per $m \in [1,3]$, $n_i(m) = \delta_{im}$, dunque tutte e sole le soluzioni in $\mathbb{Q}[x]$ del sistema iniziale sono della forma:

$$(13n_1(x) + 27n_2(x) + 40n_3(x)) + (x-1)(x-2)(x-3)q(x).$$

Il metodo esposto è sostanzialmente un'equivalente polinomiale del teorema cinese del resto e prende il nome di *interpolazione di Newton-Lagrange*. Vediamo ora un'interessante applicazione di questa tecnica al calcolo del determinante delle matrici di Vandermonde, ricordando preliminarmente un risultato di algebra lineare.

Lemma 2.16 (Cramer)**.** *Se $A \in \mathrm{GL}_n(\mathbb{K})$, la soluzione del sistema lineare*

$$A\,x = (v_1, \ldots, v_n)^T = v \neq 0$$

è data, per ogni $i \in [1,n]$, da

$$x_i = \frac{\det A^{(i)}}{\det A},$$

dove $A^{(i)}$ è la matrice ottenuta da A sostituendo la i-esima colonna con il vettore v.

Dimostrazione. Sia $\{e_1, \ldots, e_n\}$ la base canonica di $\mathbb{F}^n$ e siano f ed f_i le applicazioni lineari associate ad A ed $A^{(i)}$, rispettivamente. Si ha:

$$\begin{cases} f_i : e_i \to v = f\,x, \\ f_i : e_k \to f\,e_k \qquad \forall k \neq i, \end{cases}$$

dunque:

$$\begin{cases} f^{-1}\,f_i : e_i \to x, \\ f^{-1}\,f_i : e_k \to e_k \qquad \forall k \neq i. \end{cases}$$

La matrice associata, nella base canonica, all'applicazione $f^{-1}\,f_i$ ha perciò elementi non nulli unicamente sulla diagonale e sulla i-esima colonna; in particolare si verifica $\det(f^{-1}\,f_i) = x_i$. D'altro canto, per il teorema di Binet si ha $\det(f^{-1}\,f_i) = \frac{\det f_i}{\det f}$, da cui la tesi. □

Sia ora $V(x_0, \ldots, x_k)$ la matrice di Vandermonde $(k+1)\times(k+1)$ per cui si ha $V_{i,j} = x_i^j$ con $i, j \in [0, k]$. V è naturalmente associata ad un problema di interpolazione: supponendo infatti di voler determinare i coefficienti $d^T = (d_0, \ldots, d_k)$ del polinomio $q(x) = d_0 + d_1\,x + \ldots + d_k x^k$ sotto le ipotesi

$$q(x_0) = 1, \quad q(x_i) = 0 \quad \forall i \in [1, k],$$

detta $\{e_0, \ldots, e_k\}$ la base canonica di $\mathbb{F}^{k+1}$, ci troviamo di fronte alla risoluzione del sistema lineare:

$$Vd = e_0.$$

Se gli x_i sono tutti distinti, certamente $V \in \mathrm{GL}_{k+1}(\mathbb{K})$, in quanto una soluzione al problema è data da

$$q(x) = \frac{\prod_{i=1}^k (x_i - x)}{\prod_{i=1}^k (x_i - x_0)},$$

e, in particolare, vale:

$$d_0 = \frac{\prod_{i=1}^k x_i}{\prod_{i=1}^k (x_i - x_0)}.$$

In virtù del lemma precedente abbiamo però:

$$d_0 = \frac{\det V(x_1, \ldots, x_n)}{\det V(x_0, \ldots, x_n)} \prod_{i=1}^k x_i,$$

dunque:

$$\forall n, \quad \frac{\det V(x_0, \ldots, x_n)}{\det V(x_1, \ldots, x_n)} = \prod_{i=1}^k (x_i - x_0)$$

e, per induzione su n,

Teorema 2.17.

$$\det V(x_0, \ldots, x_k) = \prod_{k \geq i > j \geq 0} (x_i - x_j) = \pm \sqrt{\Delta\left(\prod_{i=0}^k (x - x_i)\right)}.$$

È particolarmente interessante interpretare "a rovescio" l'ultima identità: supponiamo che $p \in \mathbb{Q}[x]$ sia un polinomio coprimo con la sua derivata. In tali ipotesi, p ammette ∂p radici complesse distinte $(\zeta_1, \ldots, \zeta_{\partial p})$, e si ha:

$$\Delta p = \det\left(V^T(\zeta_1, \ldots, \zeta_{\partial p}) \cdot V(\zeta_1, \ldots, \zeta_{\partial p})\right);$$

la matrice che figura nel membro destro, che denotiamo con P, è una matrice simmetrica e invertibile, i cui elementi sono somme di potenze[4] delle radici di p:

$$P \in \mathrm{GL}_{\partial p}(\mathbb{Q}), \qquad \forall i, j \in [0, \partial p - 1],\ P_{i,j} = p_{i+j}(\zeta_1, \ldots, \zeta_{\partial p}).$$

Il Teorema (2.17) permette perciò di esplicitare la dipendenza che sussiste tra il discriminante e le funzioni simmetriche elementari delle radici, e, conseguentemente, la dipendenza che sussiste tra il discriminante di un polinomio e l'insieme dei suoi coefficienti.

[4] Convenzionalmente, assumiamo $p_0(\zeta_1, \ldots, \zeta_{\partial p}) = \partial p$.

3 Coefficienti binomiali e successioni per ricorrenza

Il coefficiente binomiale $\binom{n}{k}$ è dato dal numero di possibili scelte di k elementi tra n:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!};$$

per $m \leq 0,\ n < 0$ o $n > m$ si pone, convenzionalmente:

$$\binom{m}{n} = 0.$$

Valgono le seguenti proprietà:

1. $\binom{n}{k} = \binom{n}{n-k}$;
2. $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$;
3. $(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j\, b^{n-j}$;
4. $\sum_{j=0}^{n} \binom{n}{j} = 2^n$;
5. $\sum_{j=0}^{n} \binom{n}{j}(-1)^j = 0$;
6. $\sum_{j=k}^{N} \binom{j}{k} = \binom{N+1}{k+1}$.

Inoltre:

Lemma 3.1 (Chu-Vandermonde)**.**

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Dimostrazione.

$$\sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k} = [x^n]\left(\sum_{k=0}^{n} \binom{n}{k} x^k\right)^2 = [x^n](1+x)^{2n} = \binom{2n}{n}.$$

□

Digerita questa cospicua mole di risultati, ci chiediamo cosa si possa dire della somma delle k-esime potenze dei primi N numeri naturali positivi, ossia di:

$$p_k(N) \doteq \sum_{n=1}^{N} n^k.$$

Leggende narrano che Gauss da giovane fosse uno scolaro particolarmente molesto, tanto che il suo insegnante, per tenerlo buono, gli assegnò un giorno il compito di calcolare la somma dei numeri naturali da 1 a 100, reputando ciò piuttosto tedioso e certamente non alla portata di un bambino. Il piccolo Gauss, dopo appena qualche attimo di macchinazione, esclamò con fierezza: "5050!", lasciando di stucco il suo insegnante, che aveva impiegato tempi ben maggiori per portare a termine il medesimo compito. Quest'ultimo, sopraffatto dall'ammirazione, o forse solo dall'invidia, pretese che il suo alunno palesasse l'artificio diabolico usato per rispondere con tale prontezza. Il piccolo Gauss, senza scomporsi, articolò: "$100+1 = 99+2 = 98+3$, così via per 50 volte, è semplice", quindi riprese a far cagnara.

Il caso $k = 1$ è dunque anteriore al Risorgimento. Per non sfigurare al cospetto di Gauss, consideriamo dapprima il caso $k = 2$. Poiché $n^2 = 2\binom{n}{2} + \binom{n}{1}$,

$$\sum_{n=1}^{N} n^2 = 2\sum_{n=1}^{N} \binom{n}{2} + \sum_{n=1}^{N} \binom{n}{1} = 2\binom{N+1}{3} + \binom{N+1}{2} = \frac{N(N+1)(2N+1)}{6},$$

fatto probabilmente farraginoso da articolare a parole ma sicuramente conciso e di grande effetto. Analogamente, per $k = 3$:

$$\sum_{n=1}^{N} n^3 = 6\sum_{n=1}^{N}\binom{n+1}{3} + \sum_{n=1}^{N}\binom{n}{1} = 6\binom{N+2}{4} + \binom{N+1}{2} = \frac{N^2(N+1)^2}{4} = \left(\sum_{n=1}^{N} n\right)^2.$$

Possiamo dunque asserire che, in generale, $p_k(N)$ è un polinomio in N, a coefficienti razionali[5], di grado $k+1$: i suoi coefficienti possono essere dunque dedotti per interpolazione, anche in virtù del fatto che[6]:

$$p_k(N+1) - p_k(N) = (N+1)^k,$$

in stile gaussiano. Chiamando ora δ l'operatore (di *differenza in avanti*)

$$\delta : p(x) \longrightarrow p(x) - p(x+1),$$

abbiamo che:

- $p \in \mathbb{Z}[x] \longrightarrow \delta p \in \mathbb{Z}[x]$;
- $\delta(pq) - p\cdot(\delta q) - q\cdot(\delta p) + (\delta p)\cdot(\delta q) = 0$;[7]
- $\partial(\delta p) = \partial p - 1$;
- $[x^{\partial p - 1}](\delta p) = -\partial p \cdot [x^{\partial p}](p)$;
- $\delta^{\partial p} p = (-1)^{\partial p}\cdot(\partial p)!\cdot[x^{\partial p}](p)$.

E' ora possibile fornire una risposta ben motivata a tutti i quiz di intelligenza che chiedono: *"Qual è il prossimo numero nella sequenza?"*. Supponiamo di avere una sequenza di 5 numeri interi, ad esempio $(10, 11, 17, 23, 71)$, e supponiamo che questi siano i valori assunti in $(0, 1, 2, 3, 4)$ da un certo polinomio $q(x)$, di grado 4, a coefficienti razionali: è semplice determinare $q(5)$. Infatti δq assume i valori

$$(-1, -6, -6, -48) \text{ su } (0, 1, 2, 3),$$

$\delta^2 q$ assume i valori

$$(5, 0, 42) \text{ su } (0, 1, 2),$$

$\delta^3 q$ assume i valori

$$(5, -42) \text{ su } (0, 1),$$

dunque $(\delta^4 q)(0) = 47$. Ma se q ha grado 4, $\delta^4 q$ è costante, alché, procedendo a ritroso, $\delta^3 q$ assume i valori

$$(5, -42, -89) \text{ su } (0, 1, 2),$$

$\delta^2 q$ assume i valori

$$(5, 0, 42, 131) \text{ su } (0, 1, 2, 3),$$

δq assume i valori

$$(-1, -6, -6, -48, -179) \text{ su } (0, 1, 2, 3, 4),$$

q assume i valori

$$(10, 11, 17, 23, 71, 250) \text{ su } (0, 1, 2, 3, 4, 5),$$

[5] *Accortezza, prego:* p_k ha coefficienti razionali, ma sugli interi assume unicamente valori interi.

[6] Suggeriamo al lettore di provare che $[N^{k+1}]\, p_k(N) = \frac{1}{k+1}$ in quanto:

$$\left|\sum_{n=1}^{N} n^k - \int_0^N x^k\, dx\right| = O(N^k).$$

[7] Ricordiamo che un operatore differenziale D è una *derivazione* se soddisfa:

$$D(f\cdot g) = (D\, f)\cdot g + f\cdot(D\, g),$$

dunque δ non è una derivazione, ma solo a causa del piccolo contributo $(\delta f)(\delta g)$.

dunque $q(5) = 250$ è una risposta perfettamente plausibile per il nostro test di intelligenza. Il metodo esposto per la ricostruzione del valore di un certo polinomio in un punto è detto *metodo delle differenze finite*. Passiamo ora a dibattere di *successioni (definite) per ricorrenza*, in particolare di successioni *autonome* e *ricorrenti lineari*:

$$a_n = b_1 \cdot a_{n-1} + b_2 \cdot a_{n-2} + \ldots + b_k \cdot a_{n-k},$$

cui la successione di Fibonacci è esempio prototipico:

$$\begin{cases} F_0 &= 0 \\ F_1 &= 1 \\ F_n &= F_{n-1} + F_{n-2}. \end{cases}$$

Per tali successioni $\{a_n\}_{n\in\mathbb{N}}$ il polinomio

$$p(x) = x^k - b_1 x^{k-1} - b_2 x_{k-2} - \ldots ... - b_k$$

è detto *polinomio caratteristico*. Successioni di numeri razionali aventi il medesimo polinomio caratteristico $p(x)$ costituiscono uno spazio vettoriale su $\mathbb{Q}$ di dimensione ∂p; nell'ipotesi che $p(x)$ sia irriducibile su $\mathbb{Q}$, con radici distinte $\xi_1, \ldots, \xi_k$, le successioni

$$\{\xi_1^n\}_{n\in\mathbb{N}}, \quad \ldots, \quad \{\xi_k^n\}_{n\in\mathbb{N}}$$

sono linearmente indipendenti con polinomio caratteristico $p(x)$, dunque $\forall n \in \mathbb{N}$ si ha:

$$a_n = c_1\xi_1^n + c_2\xi_2^n + \ldots + c_k\xi_1^k, \quad c_i \in \mathbb{Q}[\xi_1].$$

Nel caso della successione di Fibonacci il polinomio caratteristico è il polinomio minimo su $\mathbb{Q}$ della sezione aurea ($\varphi = \frac{1+\sqrt{5}}{2}$, con radice coniugata $\bar{\varphi} = \frac{1-\sqrt{5}}{2} = 1 - \varphi = -\frac{1}{\varphi}$), dunque:

$$F_n = c_1\varphi^n + c_2(1-\varphi)^n, \quad F_0 = c_1 + c_2 = 0, \quad F_1 = c_1\varphi + c_2(1-\varphi) = 1$$

e si ha il seguente

Lemma 3.2 (Binet)**.**

$$F_n = \frac{1}{\sqrt{5}}\left(\varphi^n - (1-\varphi)^n\right).$$

Presa invece la successione di Lucas:

$$\begin{cases} L_0 &= 2 \\ L_1 &= 1 \\ L_n &= L_{n-1} + L_{n-2}. \end{cases}$$

si ha:

$$L_n = \varphi^n + (1-\varphi)^n,$$

inoltre L_n è combinazione lineare di F_n ed F_{n+1}, così come F_n è combinazione lineare di L_n ed L_{n+1}:

$$\begin{cases} L_n &= -F_n + 2\,F_{n+1} &= F_{n-1} + F_{n+1} \\ F_n &= \frac{1}{5}\left(-L_n + 2\,L_{n+1}\right) &= \frac{1}{5}\left(L_{n-1} + L_{n+1}\right) \end{cases}$$

Per le formule di Binet si ha inoltre:

$$\begin{cases} L_{2k} &= L_k^2 - 2\,(-1)^k & \\ L_{2k+1} &= L_k L_{k+1} - (-1)^k &= L_{k+1}^2 - L_k^2 \\ F_{2k} &= F_k L_k &= F_k(2\,F_{k+1} - F_k) \\ F_{2k+1} &= F_{k+1}^2 + F_k^2 & \end{cases}$$

Un'altra importante identità, facilmente dimostrabile per induzione su k, è la seguente:

$$F_m = F_k\,F_{m-k+1} + F_{k-1}\,F_{m-k}. \tag{1}$$

Con la scelta dei parametri $m = 2n-1, k = n$ si ottiene

$$F_{2n-1} = F_n^2 + F_{n-1}^2,$$

mentre con la scelta dei parametri $m = 2n-1, k = n-1$ si ottiene

$$F_{2n-1} = F_{n-1}F_{n+1} + F_{n-2}F_n.$$

Raffrontando le due identità prodotte possiamo scrivere:

$$(F_n^2 - F_{n-1}F_{n+1}) + (F_{n-1}^2 - F_{n-2}F_n) = 0,$$

fatto che garantisce, per induzione, la validità dell'uguaglianza:

$$F_n^2 - F_{n-1}F_{n+1} = (-1)^{n+1}.$$

In modo del tutto analogo, la scelta dei parametri $m = 2n-1, k = n-r$ porta a concludere:

$$F_n^2 - F_{n-r}F_{n+r} = (-1)^{n+r}\,F_r^2,$$

$$F_n^4 - F_{n-2}F_{n-1}F_{n+1}F_{n+2} = 1.$$

Ulteriore conseguenza dell'identità (1) è la seguente:

$$\gcd(F_m, F_k) = \gcd(F_{k-1}\cdot F_{m-k}, F_k) = \gcd(F_{m-k}, F_k) = F_{\gcd(m,k)}.$$

Dal binomio di Newton e dall'identità di Binet si ha:

$$F_n = 2^{1-n}\sum_{k=0}^{\lfloor\frac{n-1}{2}\rfloor}\binom{n}{2k+1}\cdot 5^k,$$

$$F_{2n} = \sum_{k=0}^{n}\binom{n}{k}\cdot F_k,$$

mentre è semplice provare, unicamente per induzione, l'identità[8]:

$$F_{n+1} = \sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n-k}{k}.$$

La successione di Fibonacci, solo in quanto autonoma e a termini interi, è periodica modulo m per ogni numero naturale $m \geq 2$; in particolare, detto $\pi(m)$ il periodo della successione di Fibonacci modulo m, si ha:

- $\frac{\log m}{\log \varphi} < \pi(m) \leq m^2 - 1$;
- $k \mid \pi(F_k)$;
- se p è un primo congruo a 2 o 3 modulo 5, $\pi(p) \mid p^2 - 1$;
- se p è un primo congruo a 1 o 4 modulo 5, $\pi(p) \mid p - 1$;
- se, al solito, $\nu_p(m) = \max\{h \in \mathbb{N} : p^h | m\}$, vale $\nu_5(F_k) = \nu_5(k)$, in quanto:

$$F_{5k} = 5F_k\left(25F_k^4 + 25(-1)^k F_k^2 + 1\right).$$

[8] *Suggerimento:* ambo i membri soddisfano la medesima equazione di ricorrenza.

Si noti inoltre che la successione $\left\{\frac{F_{n+1}}{F_n}\right\}_{n=1}^{\infty}$ converge linearmente verso φ, in quanto:

- il fatto che $\left|\frac{F_{n+1}}{F_n}-\frac{F_{n+2}}{F_{n+1}}\right| = \frac{1}{F_n\,F_{n+1}} < \frac{5}{\varphi^{2n+1}-1}$ garantisce che $\left\{\frac{F_{n+1}}{F_n}\right\}_{n=1}^{\infty}$ sia una successione di Cauchy, dunque ammetta limite $L\in\mathbb{R}$;
- posto $L=\lim_{n\to+\infty}\frac{F_{n+1}}{F_n}$, si ha $L\geq 1$ e $L=1+\frac{1}{L}$, dunque $L=\varphi$;
- $|F_{n+1}-\varphi\,F_n| = \frac{1}{\sqrt{5}}\left|\bar{\varphi}^{n-1}+\bar{\varphi}^{n+1}\right| \leq \frac{2}{\sqrt{5}\,\varphi^n}$ garantisce che, $\forall\epsilon>0$, si abbia definitivamente $\left|\frac{F_{n+1}}{F_n}-\varphi\right| \leq \frac{2}{5\,F_n^{2-(1+\epsilon)}}$.

Più semplicemente si consideri che, con la notazione propria delle frazioni continue, si ha:

$$\frac{F_{n+1}}{F_n} = [(1,)^n].$$

In generale, sussiste un legame molto stretto tra successioni per ricorrenza autonome e lineari, problemi di Cauchy per equazioni differenziali ordinarie a coefficienti costanti, orbite nel gruppo moltiplicativo $\mathrm{GL}_k(\mathbb{Q})$. Si considerino, ad esempio, le identità:

$$\begin{pmatrix} F_{n+2}\\ F_{n+1}\end{pmatrix} = \begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix} F_{n+1}\\ F_{n}\end{pmatrix},\quad \begin{pmatrix} L_{n+2}\\ L_{n+1}\end{pmatrix} = \begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix} L_{n+1}\\ L_{n}\end{pmatrix},$$

dove figura la matrice compagna (*companion*, o *matrice di Frobenius*) del polinomio x^2-x-1, polinomio caratteristico della ricorrenza. Tale matrice è diagonalizzabile in quanto ha autovalori distinti (il suo polinomio caratteristico è un elemento irriducibile di $\mathbb{Q}[x]$). Abbiamo poi che la traccia di una matrice (ossia la somma degli elementi che figurano sulla diagonale) è invariante per cambiamenti di base, ed è dunque pari alla somma degli autovalori. Da queste semplici considerazioni seguono facilmente le identità:

$$L_k = \mathrm{Tr}\left(\begin{pmatrix}1&1\\1&0\end{pmatrix}^k\right),\qquad \begin{pmatrix}1&1\\1&0\end{pmatrix}^k = \begin{pmatrix}F_{k+1}&F_k\\F_k&F_{k-1}\end{pmatrix}.$$

4 Disuguaglianze

Una funzione $f: D\subseteq\mathbb{R}^n\to\mathbb{R}$ è detta *convessa* se

$$\forall\lambda\in[0,1],\,\forall x,y\in D,\quad f(\lambda x+(1-\lambda)y)\leq\lambda\,f(x)+(1-\lambda)\,f(y),$$

o, equivalentemente,

$$\forall\lambda_1,\ldots,\lambda_n,\ \lambda_i\in[0,1],\ \sum_{j=1}^{n}\lambda_j=1\ \text{ si ha }\ f\left(\sum_{j=1}^{n}\lambda_j\,x_j\right)\leq\sum_{j=1}^{n}\lambda_j\,f(x_j),$$

Una funzione è detta *concava* se è l'opposto di una funzione convessa.
Condizioni sufficienti affinché una funzione $f:\mathbb{R}\to\mathbb{R}$ risulti convessa sull'intervallo $[a,b]$ sono:

- $f\in C^0([a,b]),\ \forall(c,d)\in[a,b]^2,\ f\left(\frac{c+d}{2}\right)\leq\frac{f(c)+f(d)}{2}$ (continuità e convessità per punti medi);
- $f\in C^1([a,b]), f'(x)$ è una funzione debolmente crescente su $[a,b]$;
- $f\in C^2([a,b]), f''\geq 0$.

In tali ipotesi f soddisfa la *disuguaglianza di Jensen*[9]:

$$\forall(c_1,\ldots,c_k)\in[a,b]^k,\ \forall(\lambda_1,\ldots,\lambda_k)\in[0,1]^k\cap\left\{\sum_{j=1}^{k}\lambda_j=1\right\},\ f\left(\sum_{j=1}^{k}\lambda_j\,c_j\right)\leq\sum_{j=1}^{k}\lambda_j\,f(c_j),$$

[9] L'uguaglianza può verificarsi solo nel caso in cui i vari c_i coincidano.

inoltre il suo sopragrafico è convesso e giace al di sopra di ogni retta tangente; se dunque $f \in C^1([a,b])$ si ha:

$$\forall (x,y) \in [a,b]^2, f(x) \geq f'(y)(x-y) + f(y),$$

mentre per le rette secanti si ha:

$$\forall (c,d) \in [a,b]^2 : a \leq c < d \leq b, \begin{cases} \forall y \in [c,d], & f(y) \leq \dfrac{f(d)-f(c)}{d-c}\, y + \dfrac{d\, f(c) - c\, f(d)}{d-c}, \\ \forall y \in [a,b] \setminus [c,d], & f(y) > \dfrac{f(d)-f(c)}{d-c}\, y + \dfrac{d\, f(c) - c\, f(d)}{d-c}. \end{cases}$$

Supponiamo ora che $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ sia una funzione convessa e D sia un dominio semplicemente connesso. I risultati che seguono, pur semplici nella loro formulazione e dimostrazione, sono centrali nella teoria delle funzioni convesse:

- f è una funzione continua sulla parte interna di D; la cardinalità dei punti della parte interna di D per cui f non è differenziabile è al più numerabile;
- se f ammette massimo, lo assume sulla frontiera del dominio D.

Siano ora $(a_1, \ldots, a_k)$ e $(b_1, \ldots, b_k)$ due sequenze di numeri reali non negativi con le seguenti proprietà:

- $\forall i < j,\ a_i \geq a_j$ e $b_i \geq b_j$ (ordinate debolmente descrescenti);
- $\forall i \in [1,k],\ A_i \doteq \sum_{j=1}^{i} a_j \geq \sum_{j=1}^{i} b_j \doteq B_i$ (la prima sequenza maggiora la seconda);
- $\sum_{j=1}^{k}(a_j - b_j) = 0$ (medesima somma).

La *disuguaglianza di Karamata*, anche nota come *disuguaglianza di Hardy-Littlewood*, asserisce che in tali ipotesi una qualunque funzione reale convessa realizza:

$$\sum_{i=1}^{k} f(a_i) \geq \sum_{i=1}^{k} f(b_i).$$

Dimostrazione. Se f è convessa, la funzione

$$\delta_f(a,b) = \frac{f(b) - f(a)}{b-a}$$

è simmetrica nei suoi argomenti e crescente al crescere del secondo argomento.
Se nelle ipotesi del teorema poniamo dunque

$$c_i = \delta_f(a_i, b_i),$$

abbiamo:

$$\sum_{i=1}^{k}(f(a_i) - f(b_i)) = \sum_{i=1}^{k} c_i(a_i - b_i) = \sum_{i=1}^{k} c_i(A_i - A_{i-1} - B_i + B_{i-1}) = \sum_{i=1}^{k-1}(c_i - c_{i+1})(A_i - B_i),$$

ma

$$c_i = \delta_f(a_i, b_i) \geq \delta_f(b_i, a_{i+1}) \geq \delta_f(a_{i+1}, b_{i+1}) = c_{i+1}$$

e la tesi è provata. □

Se ora le sequenze $(a_1, \ldots, a_k)$ e $(b_1, \ldots, b_k)$ soddisfano le ipotesi del teorema di Karamata, per ogni funzione convessa f e per ogni sequenza di numeri reali non negativi $\{\lambda_i\}_{i=1}^{k}$ tale per cui la somma $\sum_{i=1}^{k} \lambda_i(a_i - b_i)$ sia nulla, si ha:

Teorema 4.1 (Weighted Karamata)**.**

$$\sum_{j=1}^{k} \lambda_j\, f(a_j) \geq \sum_{j=1}^{k} \lambda_j\, f(b_j).$$

Dimostrazione. E' sufficiente provare il risultato nel caso $\lambda_i \in \mathbb{N}$, quindi nel caso $\lambda_i \in \mathbb{Q}$, quindi concludere per continuità. □

Siano ora $\{a_i\}_{i=1}^k$ e $\{b_i\}_{i=1}^k$ due sequenze di numeri reali. Il polinomio di secondo grado

$$p(x) = \sum_{j=1}^{k} (a_j\, x + b_j)^2,$$

in quanto somma di quadrati, assume valori non negativi per ogni $x \in \mathbb{R}$. L'importante risultato:

Teorema 4.2 (Cauchy-Schwarz)**.**

$$\left(\sum_{j=1}^{k} a_j\, b_j\right)^2 \leq \left(\sum_{j=1}^{k} a_j^2\right) \cdot \left(\sum_{j=1}^{k} b_j^2\right)$$

segue dunque dalla constatazione che $p(x)$ ha discriminante non positivo; l'uguaglianza ha luogo unicamente nel caso in cui esista una costante $\lambda \in \mathbb{R}$ tale da realizzare:

$$\forall j \in [1, k],\ \ a_j = \lambda\, b_j.$$

A partire da una disuguaglianza banale, è in realtà possibile ottenere la disuguaglianza di Cauchy-Schwarz per *amplificazione* (o *interpolazione*). Siano $v, w \in \mathbb{R}^n \setminus \{0\}$. Allora:

$$\|v - w\| \geq 0,$$

da cui, per l'identità di polarizzazione (anche nota come *Teorema del coseno* o *Teorema di Carnot*):

$$< v, w > \ \leq\ \frac{1}{2}(\|v\|^2 + \|w\|^2).$$

Il membro sinistro è un prodotto scalare, in particolare una forma bilineare: segue che rimpiazzando v con λv e w con $\frac{1}{\lambda}\, w$ il suo valore non muta. Possiamo perciò asserire:

$$\forall v, w \in \mathbb{R}^n \setminus \{0\},\ \ \forall \lambda > 0, \qquad < v, w > \ \leq\ \frac{1}{2}(\lambda^2 \|v\|^2 + \frac{1}{\lambda^2}\, \|w\|^2).$$

Operiamo ora la seguente accortezza: scegliamo λ^2 in modo da minimizzare il membro destro. Poiché

$$\min_{\substack{x,y \geq 0 \\ xy = k}} (x + y) = 2\sqrt{k},$$

(volendo, per AM-GM) la migliore scelta per λ^2 risulta essere $\frac{\|w\|}{\|v\|}$, e da tale scelta segue:

$$< v, w > \ \leq \|v\| \|w\|,$$

che è la disuguaglianza di Cauchy-Schwarz in forma vettoriale. Addizionalmente, notiamo che lo scarto tra il quadrato del membro destro e il quadrato del membro sinistro può essere precisamente quantificato:

Teorema 4.3 (Sharpened Cauchy-Schwarz Inequality)**.**

$$\left(\sum_{j=1}^{n} a_j b_j\right)^2 - \left(\sum_{j=1}^{n} a_j^2\right) \cdot \left(\sum_{j=1}^{n} b_j^2\right) = \sum_{1 \leq i < j \leq n} (a_i b_j - a_j b_i)^2 \ \ \geq 0.$$

Dimostrazione. È sufficiente raffrontare i monomi di grado 4 che figurano nel termine sinistro con quelli che figurano nel termine centrale. □

Siano ora $\{a_i\}_{i=1}^k$ e $\{b_i\}_{i=1}^k$ due sequenze debolmente crescenti di numeri reali non negativi. Si ha:

Teorema 4.4 (Disuguaglianza di riarrangiamento)**.**

$$\forall \sigma \in S_k, \quad \sum_{j=1}^{k} a_j b_j \geq \sum_{j=1}^{k} a_j b_{\sigma(j)} \geq \sum_{j=1}^{k} a_j b_{k+1-j}.$$

Dimostrazione. Supponiamo che per $\sigma_1, \sigma_2 \in S_k$ si abbia

$$\sigma_1 = (n_1\, n_2)\, \sigma_2,$$

con $n_1 < n_2$ e $\sigma_1^{-1}(n_1) < \sigma_2^{-1}(n_2)$.
In tal caso diciamo che σ_1 può essere ottenuta da σ_2 attraverso una *inversione*, e vale:

$$\sum_{j=1}^{k}(a_j\, b_{\sigma_2(j)} - a_j\, b_{\sigma_1(j)}) = (a_{\sigma_1^{-1}(n_1)} b_{n_2} + a_{\sigma_1^{-1}(n_2)} b_{n_1}) - (a_{\sigma_1^{-1}(n_1)} b_{n_1} + a_{\sigma_1^{-1}(n_2)} b_{n_2}),$$

$$\sum_{j=1}^{k}(a_j\, b_{\sigma_2(j)} - a_j\, b_{\sigma_1(j)}) = (b_{n_2} - b_{n_1})(a_{\sigma_1^{-1}(n_1)} - a_{\sigma_1^{-1}(n_2)}) \leq 0, \tag{2}$$

ed avremmo ottenuto la disuguaglianza opposta nel caso in cui fosse valsa $\sigma_1^{-1}(n_1) > \sigma_2^{-1}(n_2)$.
Posto dunque:

$$S(\sigma) = \sum_{j=1}^{k} a_j b_{\sigma(j)},$$

abbiamo:

$$\operatorname{sign}(S(\sigma) - S((n_1\, n_2)\, \sigma) = \operatorname{sign}(\sigma^{-1}(n_2) - \sigma^{-1}(n_1)),$$

tuttavia ogni permutazione di S_k può essere scritta come prodotto di al più $k-1$ trasposizioni:

$$\sigma = (1\, \eta_1)(2\, \eta_2) \ldots (k-1\, \eta_{k-1}),$$

con il vincolo addizionale $\eta_i \geq i$ (assumiamo che $(i\, i)$ rappresenti la permutazione identica e), segue:

$$\forall \sigma \in S_k : \sigma \neq e, \;\; S(\sigma) \leq S(e).$$

Se ora $\{a_i\}_{i=1}^k$ e $\{c_i\}_{i=1}^k$ sono due sequenze di numeri reali non negativi, l'una crescente e l'altra decrescente, possiamo ricalcare le nostre orme semplicemente cambiando di segno la (2) e ottenere:

$$\sum_{j=1}^{k} a_i\, c_i \leq \sum_{j=1}^{k} a_i\, c_{\sigma_i}.$$

La dimostrazione è perciò conclusa prendendo $c_i = b_{k+1-i}$.
Notiamo che la chiusura transitiva della relazione antisimmetrica

$$\sigma_1 < \sigma_2 \quad \Longleftrightarrow \quad \sigma_1 \text{ può essere ottenuta da } \sigma_2 \text{ tramite un'inversione}$$

munisce il gruppo S_k di un ordinamento parziale; in tale contesto, si può asserire con certezza

$$S(\sigma_i) \geq S(\sigma_j) \qquad \text{oppure} \qquad S(\sigma_i) \leq S(\sigma_j)$$

solo se σ_i e σ_j sono in relazione, ossia appartengono ad una stessa catena. Incondizionatamente, la permutazione identica maggiora ogni altro elemento di S_k e la permutazione $\sigma(i) = k+1-i$ è maggiorata da ogni altro elemento di S_k. □

Sommando ora le k disuguaglianze di riarrangiamento date dalle potenze del k-ciclo $\sigma = (1, 2, \ldots, k)$ in S_k abbiamo[10]:

Teorema 4.5 (Chebyshev)**.** *Se $\{a_i\}_{i=1}^k$ e $\{b_i\}_{i=1}^k$ sono due sequenze non decrescenti di numeri reali non negativi,*

$$k \cdot \sum_{j=1}^{k} a_k b_k \geq \left(\sum_{j=1}^{k} a_k \right) \cdot \left(\sum_{j=1}^{k} b_k \right).$$

La disuguaglianza cambia verso se le due sequenze sono ordinate in modo opposto, l'uguaglianza ha luogo se e solo se tutti gli elementi di una o dell'altra sequenze coincidono.

Presi ora k numeri reali non negativi $x_1, \ldots, x_k$ ed una sequenza di numeri reali non negativi $a_1, \ldots, a_k$, denotiamo con $T[a_1, \ldots, a_k]$ la quantità:

$$T[a_1, \ldots, a_k] = \sum_{\sigma \in S_k} x_{\sigma(1)}^{a_1} x_{\sigma(2)}^{a_2} \cdot \ldots \cdot x_{\sigma(k)}^{a_1} \doteq \sum_{sym} \prod_{j=1}^{k} x_j^{a_j}.$$

Vale il seguente:

Teorema 4.6 (Schur)**.**

$$\forall a, b \in \mathbb{R}^+, \quad T[a + 2b, 0, 0] + T[a, b, b] \geq 2 \cdot T[a + b, b, 0].$$

Dimostrazione. E' sufficiente verificare che per ogni terna di numeri reali non negativi (x, y, z) si ha:

$$x^a(x^b - y^b)(x^b - z^b) + y^a(y^b - z^b)(y^b - x^b) + z^a(z^b - x^b)(z^b - y^b) \geq 0.$$

Senza perdere di generalità possiamo assumere che valga $x \geq y \geq z$, alché il terzo addendo risulta non negativo ed è sufficiente provare:

$$x^a(x^b - z^b) - y^a(y^b - z^b) \geq 0,$$

ossia:

$$x^{a+b} - y^{a+b} - z^b(x^a - y^a) \geq 0.$$

Ciò è semplice, in quanto:

$$x^{a+b} - y^{a+b} - z^b(x^a - y^a) \geq x^{a+b} - y^{a+b} - y^b(x^a - y^a) = x^a(x^b - y^b) \geq 0.$$

□

Inoltre, se $\{a_i\}_{i=1}^k$ e $\{b_i\}_{i=1}^k$ sono due sequenze che soddisfano le ipotesi della disuguaglianza di Karamata, vale il seguente risultato, anche noto come *disuguaglianza di riarrangiamento generalizzata* o *disuguaglianza di bunching*:

Teorema 4.7 (Muirhead)**.**

$$T[a_1, \ldots, a_k] \geq T[b_1, \ldots, b_k].$$

L'uguaglianza si verifica unicamente nei casi in cui coincidano tutte le variabili x_i,
o coincidano le sequenze $\{a_i\}_{i=1}^k$ e $\{b_i\}_{i=1}^k$.

Dimostrazione. Si consideri che per ogni sequenza $\{a_i\}_{i=1}^k$, per ogni coppia di numeri naturali (n_1, n_2) che realizza $1 \leq n_1 < n_2 \leq k$ e per ogni $\rho \in \left[0, \frac{a_{n_1} - a_{n_2}}{2}\right]$, la sequenza

$$\{b_i\}_{i=1}^k \doteq (a_1, a_2, \ldots, a_{n_1 - 1}, a_{n_1} - \rho, a_{n_1 + 1}, \ldots, a_{n_2 - 1}, a_{n_2} + \rho, \ldots, a_k) \tag{3}$$

[10] Incidentalmente, notiamo come la disuguaglianza di Chebyshev proceda in direzione opposta a quella della disuguaglianza di Cauchy-Schwarz.

è maggiorata dalla sequenza $\{a_i\}_{i=1}^k$; in particolare:

$$\sum_{sym}\prod_{j=1}^{k} x_j^{a_j} \;\geq\; \sum_{sym}\prod_{j=1}^{k} x_j^{b_j}, \tag{4}$$

in quanto vale:

$$\sum_{sym}\left(x_{n_1}^{a_{n_1}}x_{n_2}^{a_{n_2}}+x_{n_1}^{a_{n_2}}x_{n_2}^{a_{n_1}}\right) \;\geq\; \sum_{sym}\left(x_{n_1}^{a_{n_1}-\rho}x_{n_2}^{a_{n_2}+\rho}+x_{n_1}^{a_{n_2}+\rho}x_{n_2}^{a_{n_1}-\rho}\right),$$

a sua volta implicata da:

$$\left(x_{n_1}^{a_{n_1}-a_{n_2}-\rho}-x_{n_2}^{a_{n_1}-a_{n_2}-\rho}\right)\cdot\left(x_{n_1}^{\rho}-x_{n_2}^{\rho}\right)\geq 0.$$

D'altro canto, se $\{a_i\}_{i=1}^k$ maggiora $\{c_i\}_{i=1}^k$, è possibile mandare la prima sequenza nella seconda con una successione di trasformazioni del tipo esposto in (3), e la tesi è provata. □

Proviamo ora che il logaritmo è una funzione concava sui reali positivi:

$$\forall a\neq b\in\mathbb{R}^+,\ (a-b)^2>0 \;\longrightarrow\; (a+b)^2>4ab \;\longrightarrow\; \frac{a+b}{2}>\sqrt{ab},$$

$$\log\left(\frac{c+d}{2}\right)\geq\frac{\log(c)+\log(d)}{2} \;\longleftarrow\; \frac{c+d}{2}\geq\sqrt{cd}.$$

Dalla concavità del logaritmo segue che per ogni $(x_1,\ldots,x_k)\in\mathbb{R}^{+k}$ si ha:

$$\log\left(\frac{1}{k}\sum_{j=1}^{k}x_j\right)\geq\frac{1}{k}\sum_{j=1}^{k}\log(x_j),$$

ossia:

Teorema 4.8 (Disuguaglianza Aritmo-Geometrica, AM-GM)**.**

$$\frac{x_1+\ldots+x_k}{k}\geq\sqrt[k]{x_1\cdot\ldots\cdot x_k};$$

l'uguaglianza si verifica solo nel caso in cui gli x_i coincidano.

Come nel caso della disuguaglianza di Karamata, è semplice produrre una versione pesata, per $\lambda_i>0$:

Teorema 4.9 (Weighted AM-GM)**.**

$$\sum_{j=1}^{k}\lambda_j\,x_j\geq\left(\sum_{j=1}^{k}\lambda_j\right)\cdot\left(\prod_{j=1}^{k}x_j^{\lambda_j}\right)^{\frac{1}{\sum_{j=1}^{k}\lambda_j}};$$

l'uguaglianza si verifica solo nel caso in cui gli x_i coincidano.

Si noti che la sostituzione $x_i=\frac{1}{y_i}$ conduce alla:

Teorema 4.10 (Disuguaglianza Geometrico-Armonica, GM-HM)**.**

$$\sqrt[k]{x_1\cdot\ldots\cdot x_k}\geq\frac{k}{x_1+\ldots+x_k};$$

l'uguaglianza si verifica solo nel caso in cui gli x_i coincidano.

Teorema 4.11 (Disuguaglianza delle medie)**.** *Se $m_1 > m_2$,*

$$\left(\frac{1}{k}\sum_{j=1}^{k} x_j^{m_1}\right)^{\frac{1}{m_1}} \geq \left(\frac{1}{k}\sum_{j=1}^{k} x_j^{m_2}\right)^{\frac{1}{m_2}},$$

dove l'uguaglianza è verificata nel solo caso in cui tutti gli x_i coincidano.

Dimostrazione. Studiamo dapprima il caso $m_1 > m_2 > 0$. Ponendo $y_j = x_j^{\frac{1}{m_2}}$ abbiamo che è sufficiente provare:

$$\forall t > 1, \quad \left(\frac{1}{k}\sum_{j=1}^{k} y_j^t\right)^{\frac{1}{t}} \geq \frac{1}{k}\sum_{j=1}^{k} y_j,$$

dove, per omogeneità, non è restrittivo supporre $\sum_{j=1}^{k} x_j = k$.
Posto dunque $z_i = y_i - 1$, è sufficiente provare:

$$\sum_{j=1}^{k} (1 + z_j)^t \geq k.$$

Tuttavia per $t > 1$ la funzione $f(x) = (1+x)^t$ risulta convessa sull'intervallo $[-1, +\infty)$: il suo grafico giace dunque al di sopra del grafico di una qualunque tangente. In particolare, presa la tangente nell'origine, si ha:

$$(1 + x)^t \geq 1 + t\,x,$$

nota anche come *disuguaglianza di Bernoulli*. L'applicazione di tale risultato permette di asserire:

$$\sum_{j=1}^{k} (1 + z_j)^t \geq k + t\sum_{j=1}^{k} z_j = k,$$

come desiderato. Il caso $0 < m_1 < m_2$ può essere ricondotto al caso appena analizzato attraverso la sostituzione $y_i = \frac{1}{x_i}$; in ultima istanza si ha, per $r > 0$:

$$\left(\sum_{j=1} x_j^{-r}\right)^{-\frac{1}{r}} \leq \sqrt[k]{x_1 \cdot \ldots \cdot x_k} \leq \left(\sum_{j=1} x_j^{r}\right)^{\frac{1}{r}}$$

attraverso la sostituzione $x_i = y_i^r$ nelle disuguaglianze geometrico-armonica e aritmo-geometrica. □

Teorema 4.12 (Disuguaglianza di Newton)**.** *Se $(x_1, \ldots, x_n)$ è una n-upla di numeri reali positivi e*

$$d_k \doteq \binom{n}{k}^{-1} [t^{n-k}] \prod_{j=1}^{n} (t + x_j),$$

allora $\forall k \in [1, n-1]$ si ha:

$$d_{k-1} d_{k+1} \leq d_k^2.$$

Dimostrazione. Preso il polinomio omogeneo bivariato

$$p(x, y) = \prod_{j=1}^{n} (x + y\,x_j),$$

i valori del rapporto $\frac{x}{y}$ per cui p si annulla sono tutti reali positivi, e tale proprietà, per il teorema di Rolle, si conserva per derivazione. Segue, in particolare, che il polinomio omogeneo bivariato

$$q_k(x, y) \doteq \frac{\partial^{n-2} p}{(\partial y)^{k-1} (\partial x)^{n-k-1}} = n! \left(d_{k-1}\frac{x^2}{2} + d_k\,x\,y + d_{k+1}\frac{y^2}{2}\right)$$

ha discriminante non negativo, da cui la tesi. □

Teorema 4.13 (Disuguaglianza di MacLaurin)**.** *Nelle ipotesi del precedente teorema si ha:*

$$\sqrt[k]{d_k} \geq \sqrt[k+1]{d_{k+1}}.$$

Dimostrazione. Posto $d_0 = 1$, per la disuguaglianza di Newton si ha

$$(d_0\, d_2) \cdot (d_1\, d_3)^2 \cdot (d_2\, d_4)^3 \cdot \ldots \cdot (d_{k-1}\, d_{k+1})^k \leq d_1^2 \cdot d_2^4 \cdot d_3^6 \cdot \ldots \cdot d_k^{2k},$$

da cui:

$$d_{k+1}^k \leq d_k^{k+1},$$

equivalente alla tesi. □

Teorema 4.14 (Disuguaglianza di Young)**.** *Se $p > 1$ e $\frac{1}{p} + \frac{1}{q} = 1$,*

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}.$$

Dimostrazione. Si consideri la funzione $f(x) = x^{p-1}$, con inversa $g(x) = x^{\frac{1}{p-1}} = x^{q-1}$.
Sia f che g sono funzioni crescenti su $\mathbb{R}^+$, dunque:

$$\int_0^{|a|} x^{p-1}\, dx + \int_0^{|b|} x^{q-1}\, dx \geq |ab|.$$

□

Nelle stesse ipotesi della disuguaglianza di Young si ha inoltre:

Teorema 4.15 (Disuguaglianza di Hölder)**.**

$$\sum_{j=1}^{k} |x_j\, y_j| \leq \left(\sum_{j=1}^{k} |x_j|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{k} |y_j|^q \right)^{\frac{1}{q}}.$$

Dimostrazione. Utilizzando la seguente notazione:

$$\|x\|_p = \left(\sum_{j=1}^{k} |x_j|^p \right)^{\frac{1}{p}},$$

abbiamo che, per la disuguaglianza di Young, vale:

$$\frac{\sum_k |x_k||y_k|}{\|x\|_p \cdot \|y\|_q} = \sum_k \frac{|x_k|}{\|x\|_p} \cdot \frac{|y_k|}{\|y\|_q} \leq \frac{1}{p} \sum_k \frac{|x_k|^p}{\|x\|_p^p} + \frac{1}{q} \sum_k \frac{|y_k|^q}{\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

□

Teorema 4.16 (Disuguaglianza di Minkowski)**.** *Per ogni numero reale $p > 1$ si ha:*

$$\left(\sum_{j=1}^{k} |x_j + y_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{k} |x_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{k} |y_j|^p \right)^{\frac{1}{p}}.$$

Dimostrazione. Utilizzando la medesima notazione della precedente dimostrazione, abbiamo che:

$$\|x + y\|_p^p \leq \sum_{j=1}^{k} |x_j| \cdot |x_j + y_j|^{p-1} + \sum_{j=1}^{k} |y_j| \cdot |x_j + y_j|^{p-1},$$

dunque applicando la disuguaglianza di Hölder al membro destro si ha:

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p) \|x + y\|_p^{p-1},$$

da cui la tesi. □

5 Esercizi

1. Si provi che l'esponenziale di una funzione convessa è convesso, mentre l'esponenziale di una funzione concava non necessariamente è concavo;

2. Si provi che la somma di due funzioni convesse è una funzione convessa, mentre il prodotto di due funzioni convesse può non essere convesso;

3. Si provi che l'inversa di una $f : \mathbb{R} \to \mathbb{R}$ convessa e crescente, ove definita, è concava e crescente;

4. (*Disuguaglianza di Nesbitt*) Si provi che per qualunque terna (a, b, c) di numeri reali positivi si ha:
$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} \geq \frac{3}{2};$$

5. (*Disuguaglianza di Mahler*) Si provi che:
$$\prod_{j=1}^{n} (|x_j| + |y_j|)^{\frac{1}{n}} \geq \prod_{j=1}^{n} |x_j|^{\frac{1}{n}} + \prod_{j=1}^{n} |y_j|^{\frac{1}{n}};$$

6. (*IMO 1984*) Si provi che, per ogni terna (x, y, z) di numeri reali non negativi tale da verificare $x + y + z = 1$, si ha:
$$0 \leq xy + xz + yz - 2xyz \leq \frac{7}{27};$$

7. Si provi che, se a, b, c sono tre numeri reali non negativi e $a + b + c = 3$, allora:
$$\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \geq a^2 + b^2 + c^2;$$

8. Diciamo che un elemento di S_k è una *dismutazione* se non ammette punti fissi, ossia se $\forall j \in [1, k],\ \sigma(j) \neq j$. Si determini, in funzione del parametro k, qual è la differenza tra il numero di dismutazioni pari (ossia di segno $+1$) e il numero di dismutazioni dispari (ossia di segno -1) in S_k;

9. (*Polinomi di Chebyshev*) Si provi che per ogni numero naturale k esiste un polinomio T_k a coefficienti interi, di grado k, che realizza $T_k(\cos\theta) = \cos(k\theta)$ per ogni $\theta \in \mathbb{R}$; se ne determinino inoltre i coefficienti e si indaghi sulle relazioni che intercorrono tra T_{k+1}, T_k e T_{k-1};

10. $p(x)$ e $q(x)$ sono due polinomi a coefficienti interi e positivi che realizzano $p(1) = q(1) = k$ e $p(k + 1) = q(k + 1)$. Si provi che $p(x)$ coincide con $q(x)$;

11. Si provi che un polinomio a coefficienti interi che verifica $p(2) = p(3) = p(5) = p(7) = 11$ non può assumere il valore 8 su alcun intero;

12. Dato un numero primo p, si provi che il polinomio $q(x) = x^{p-1} + x^{p-2} + \ldots + 1$ è irriducibile su $\mathbb{Q}$ e se ne calcoli il discriminante;

13. Si determini il polinomio minimo su $\mathbb{Q}$ di $\cos\frac{2\pi}{7}$;

14. Si provi che il polinomio $x^4 + x + 1$ è irriducibile su $\mathbb{Q}$;

15. (*Lemma di Artin*) Si provi che, per ogni numero primo p, il polinomio $x^p + x + 1$ risulta irriducibile su $\mathbb{F}_p$ e dunque irriducibile su $\mathbb{Q}$;

16. Si provi che, per ogni numero naturale positivo $m \leq 2^n$, è possibile fissare $\varepsilon_m \in \{-1, +1\}$ in modo che si abbia:
$$\sum_{m=1}^{2^n} \varepsilon_m \cdot m^k = 0 \qquad \forall k \in [1, n-1];$$

17. (*Problema 6 Cesenatico 2007*) Data la successione

$$\begin{cases} x_1 &= 2 \\ x_{n+1} &= 2\,x_n^2 - 1 \end{cases}$$

si provi per ogni numero naturale $n > 1$ si ha $\gcd(n, x_n) = 1$;

18. Sia $x_0 > 0$ e $x_{k+1} = x_k^2 + x_k$. Si provi che, per ogni numero naturale n, si ha:

$$\sum_{j=1}^{n} \frac{1}{x_j + 1} \leq \frac{1}{x_0};$$

19. (*Gauss Digamma serie*) Si provi che:

$$\sum_{n=0}^{+\infty} \frac{1}{n^2+1} = \frac{1}{2}\left(1 + \pi\,\frac{e^{2\pi}+1}{e^{2\pi}-1}\right);$$

20. (*Teorema di Lucas*) Sia $p(z) \in \mathbb{C}[z]$ un polinomio privo di radici multiple.
Detto *K l'inviluppo convesso* delle sue radici,

$$K = \left\{ z \in \mathbb{C} :\ z = \alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_n\xi_n,\ \ \alpha_i \in [0,1],\ \ \sum_{i=1}^{n} \alpha_i = 1 \right\},$$

si provi che ogni radice di $p'(z)$ giace all'interno di K;

21. Si provi che, se $p(x) \in \mathbb{Z}[x]$ è privo di radici multiple, realizza $|p(0)| = 1$ ed ha un'unica radice ξ di modulo minore di 1, allora è irriducibile su $\mathbb{Q}$;

6 Testi e link consigliati

1. Cut The Knot,
`http://www.cut-the-knot.org`;
2. MathWorld,
`http://mathworld.wolfram.com`;
3. MathLinks,
`http://www.mathlinks.ro`;
4. American Mathematical Association,
`http://www.maa.org`;
5. Massimo Gobbino, Training Olimpico,
`http://www2.ing.unipi.it/~d9199/Home_Page/OT_Index.html`;
6. Generatingfunctionology, Wilf,
`http://www.math.upenn.edu/~wilf/DownldGF.html`;
7. Analytic Combinatorics, Flajolet,
`http://algo.inria.fr/flajolet/Publications/AnaCombi/anacombi.html`
8. Problems from the european mathematycal olympiad, AGA, 10€;
9. Problem Solving Through Problems, Loren C.Larson, 50$;
10. Problem-Solving strategies, Arthur Engel, 50$;
11. William Lowell Putnam Mathematical Competition 1985-2000: Problems, Solutions, Kiran Kedlaya, 50$;
12. Che cos'è la matematica, Courant-Robbins, Bollati-Boringhieri, 23€;
13. Advanced Euclidean Geometry, Roger A.Johnson, 15$;
14. 102 Combinatorial Problems, Titu Andreescu, 40$.