Richiami sulla cardinalità di insiemi

Marco Bramanti
Politecnico di Milano

March 9, 2018

Richiamiamo alcuni fatti riguardanti la cardinalità (numerosità) degli insiemi infiniti[1].

L'idea, che sta alla radice stessa del concetto di numero, è la seguente. Quando osserviamo due insiemi *finiti* di oggetti, siamo soliti dire che, ad esempio, "l'insieme A è più numeroso di B perché A ha 10 elementi e B ne ha 7", oppure che "gli insiemi A e B sono ugualmente numerosi perché entrambi hanno 5 elementi". In altre parole, è naturale contare gli elementi di ciascun insieme per fare un confronto tra le due numerosità. Questo però non è l'unico modo di procedere: posso capire che le sedie in una stanza sono tante quante le persone non perché ho contato le sedie e le persone, ma semplicemente perché osservo che ogni persona è seduta su una sedia, nessuno è in piedi e nessuna sedia è vuota. In altre parole, ho percepito che esiste una corrispondenza biunivoca tra l'insieme delle sedie e l'insieme delle persone, e questo è sufficiente a dire che i due insiemi sono ugualmente numerosi. Se poi notassi che tutte le persone sono sedute e rimane qualche sedia libera, concluderei che le sedie sono più numerose delle persone. In questo caso ho percepito che esiste una corrispondenza biunivoca tra l'insieme delle persone e un sottoinsieme proprio dell'insieme delle sedie.

Quando si passa dagli insiemi finiti agli insiemi infiniti, è ancora possibile confrontare le numerosità di due insiemi? Certamente non è possibile confrontando il *numero* di elementi dei due insiemi: sono infiniti! Ma il concetto di corrispondenza biunivoca aiuta ancora.

Definizione 0.1 *Due insiemi X, Y (qualsiasi) si dicono avere uguale cardinalità (o uguale potenza), e si scrive* $\operatorname{card} X = \operatorname{card} Y$, *se esiste una corrispondenza biunivoca tra X e Y, cioè una funzione*

$$f : X \to Y$$

[1] In queste note c'è molto di più di ciò che ci servirà in senso stretto, che è poco più della distinzione tra insieme numerabile e non numerabile. Sono argomenti delicati, però, in cui se si dice troppo poco è facile generare fraintendimenti. Qualche informazione in più dovrebbe dare un quadro concettuale più chiaro. E comunque si tratta di argomenti molto interessanti per chi è curioso.

iniettiva e suriettiva. (Esplicitamente: a ogni $x \in X$ corrisponde uno e un solo $f(x) \in Y$ e viceversa ad ogni $y \in Y$ corrisponde uno e un solo $x \in X$ tale che $f(x) = y$).

Questa definizione, così come molti importanti risultati di base sulla cardinalità di insiemi, sono dovuti a George Cantor (1845-1918), considerato il padre della teoria degli insiemi. Il suo primo scritto sulla teoria degli insiemi è del 1874.

Esempio 0.2 *L'insieme $\mathbb{N} = \{0, 1, 2, 3, ...\}$ è in corrispondenza biunivoca con $\mathbb{N}_1 = \{1, 2, 3, ...\}$, dove la corrispondenza biunivoca è quella che associa ad ogni numero naturale il suo successivo.*

Esempio 0.3 *L'insieme $\mathbb{N}$ è in corrispondenza biunivoca con l'insieme $\mathbb{Z}$ degli interi relativi, dove la corrispondenza è:*

$$\begin{aligned} f : \mathbb{N} &\to \mathbb{Z} \\ 0 &\mapsto 0 \\ 1 &\mapsto 1 \\ 2 &\mapsto -1 \\ 3 &\mapsto 2 \\ 4 &\mapsto -2 \\ 5 &\mapsto 3 \\ 6 &\mapsto -3 \\ &... \end{aligned}$$

Formalmente: per $n = 0, 1, 2, ...$ si pone

$$\begin{aligned} f(2n) &= -n \\ f(2n-1) &= n. \end{aligned}$$

Quindi $\mathbb{N}$ e $\mathbb{Z}$ hanno la stessa cardinalità.

Gli esempi precedenti hanno un aspetto curioso, perché in entrambi i casi abbiamo trovato una corrispondenza biunivoca tra un certo insieme e un suo sottoinsieme proprio: $\mathbb{N}_1 \subsetneq \mathbb{N}$ e $\mathbb{N} \subsetneq \mathbb{Z}$. Questo ha un aspetto controintuitivo, perché per gli insiemi finiti non capita mai. Difatti questa proprietà caratterizza esattamente gli insiemi infiniti:

Definizione 0.4 *Un insieme X si dice infinito quando esiste un suo sottoinsieme proprio $X_0 \subsetneq X$ che ha la stessa cardinalità di X.*

Ad esempio, i due esempi fatti dimostrano formalmente che $\mathbb{N}$ e $\mathbb{Z}$ sono insiemi *infiniti*. Se A è un insieme *finito*, può essere in corrispondenza biunivoca

con tanti altri insiemi *ma non con suoi sottoinsiemi propri*[2].

Il primo a rendersi conto che tra gli insiemi infiniti poteva accadere questo "strano fenomeno" (che un insieme e un suo sottoinsieme proprio siano in corrispondenza biunivoca) fu Galileo Galilei (1564-1642), che notò il "paradosso dei quadrati" (i numeri naturali sono in corrispondenza biunivoca con i quadrati dei numeri naturali). Bernard Bolzano (1781-1848), nella sua opera "I paradossi dell'infinito" (pubblicata postuma nel 1851) fu il primo a capire che questa proprietà è *caratteristica* degli insiemi infiniti, cioè può essere presa come definizione stessa di insieme infinito (anche se a quel tempo il concetto di insieme non era stato ancora formalizzato).

Ci interessa ora descrivere e confrontare le cardinalità degli insiemi infiniti. Cominciamo dalla seguente

Definizione 0.5 *Un insieme E si dice numerabile se ha la stessa cardinalità di $\mathbb{N}$. Quindi E si può rappresentare come successione: $E = \{x_n\}_{n=0}^{\infty}$.*

E' importante capire l'ultima osservazione: se E è numerabile e $f : \mathbb{N} \to E$ una corrispondenza biunivoca tra i due, ogni elemento di E è del tipo $f(n)$ per uno e un solo $n \in \mathbb{N}$, quindi posso disporre gli elementi di E in un elenco ordinato che ha un inizio (e non ha una fine)

$$x_0, x_1, x_2, ...$$

L'Esempio 0.3 dimostra che $\mathbb{Z}$ è numerabile, mentre $\mathbb{N}$ è numerabile per definizione (è in corrispondenza biunivoca con se stesso!). Vale anche il seguente:

Teorema 0.6 (Cantor) *L'insieme $\mathbb{Q}$ dei numeri razionali è numerabile.*

Dimostrazione. In base all'ultima osservazione fatta, per provare che $\mathbb{Q}$ è numerabile è sufficiente dare un criterio per disporre in successione $x_0, x_1, x_2, ...$ tutti i numeri razionali, in modo che si possa stabilire univocamente a quale naturale n corrisponde un assegnato numero razionale.

Cominciamo a disporre in una successione $x_1, x_2, x_3, ...$ i razionali positivi. Facciamo così: poiché i razionali positivi non sono altro che le frazioni $\frac{n}{m}$ con $n, m = 1, 2, 3...$, possiamo disporre in successione queste frazioni con questo criterio:

prima tutte quelle in cui $n + m = 2$ (ce n'è una sola):

$$\frac{1}{1};$$

poi tutte quelle in cui $n + m = 3$, disposte in ordine crescente rispetto al numeratore n:

$$\frac{1}{2}; \frac{2}{1};$$

[2] Una delle possibili formalizzazioni del concetto di numero si fonda esattamente in queste riflessioni. In un certo senso si può dire che il numero n è la classe di tutti gli insiemi che sono in corrispondenza biunivoca con un certo insieme fissato (che, a posteriori, diremo avere n elementi). Rendere rigorosa quest'idea però non è facile, e non proseguiremo su questa linea di discorso: qui interessa parlare degli insiemi infiniti.

poi tutte quelle in cui $n + m = 4$, disposte in ordine crescente rispetto al numeratore n:

$$\frac{1}{3}; \frac{2}{2}; \frac{3}{1};$$

e così via. Ad ogni passo c'è un numero finito di frazioni che soddisfa il requisito, perciò ogni passo dell'algoritmo termina e consente di passare al passo successivo. Si costruisce così la tabella di corrispondenza biunivoca tra $\mathbb{N}_1$ e $\mathbb{Q}^+$:

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & \dots \\ \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{2}{2} & \frac{3}{1} & \dots \end{array}$$

Per essere precisi, questa corrispondenza non è ancora biunivoca, perché alcune frazioni sono uguali tra loro: ad esempio $\frac{2}{2} = \frac{1}{1}$. Appena, nel costruire la tabella, si trova una frazione che indica un razionale già elencato, si butta via l'ultimo doppione e si riempie il suo posto con chi viene dopo in successione, così:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & \dots \\ \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{3}{1} & \dots \end{array}$$

Questo costruisce una corrispondenza biunivoca tra $\mathbb{N}_1$ e $\mathbb{Q}^+$. Consideriamo ora gli elementi di $\mathbb{Q}^+$ elencati in successione:

$$\mathbb{Q}^+ = \{x_n\}_{n=1}^{\infty}$$

e costruiamo la corrispondenza biunivoca tra $\mathbb{N}$ e $\mathbb{Q}$ a questo modo:

$$\begin{aligned} f &: \mathbb{N} \to \mathbb{Q} \\ f &: 0 \longmapsto 0 \\ f &: 2n \longmapsto x_n \\ f &: 2n-1 \longmapsto -x_n \end{aligned}$$

per $n = 1, 2, 3...$ Con ciò la dimostrazione è completa. ■

A questo punto può sorgere il sospetto che, pur di essere sufficientemente abili, si riesce sempre a costruire una corrispondenza biunivoca tra due insiemi infiniti. Se fosse così, tutti gli insiemi infiniti sarebbero numerabili. Non è così, però:

Teorema 0.7 (Cantor) *L'intervallo $[0,1] \subset \mathbb{R}$ non è numerabile.*

Dimostrazione. Per assurdo, sia $[0,1]$ numerabile, e sia $[0,1] = \{x_n\}_{n=0}^{\infty}$. Ogni numero reale compreso nell'intervallo $[0,1]$ si può rappresentare in forma decimale

$$\begin{aligned} x_n &= 0, a_1^{(n)} a_2^{(n)} a_3^{(n)} .... \text{ dove:} \\ a_i^{(n)} &\in \{0,1,2,3,4,5,6,7,8,9\}. \end{aligned}$$

Si noti in particolare che

$$0,0000000... = 0$$
$$0,99999999... = 1.$$

L'idea è costruire un numero $\overline{x} \in [0,1]$ che non appartenga all'elenco $\{x_n\}_{n=0}^{\infty}$ (che stiamo supponendo, per assurdo, essere l'elenco completo di tutti i reali dell'intervallo $[0,1]$): questo darà l'assurdo. Poniamo

$$\overline{x} = 0, b_1 b_2 b_3 ....$$

dove, per ogni $i = 1, 2, 3, ...$ abbiamo definito

$$b_i = \begin{cases} 5 & \text{se } a_i^{(i)} \in \{0,1,2,3,4\} \\ 4 & \text{se } a_i^{(i)} \in \{5,6,7,8,9\} \end{cases}$$

(la regola è scelta in modo tale che sia sempre $b_i \neq a_i^{(i)}$). Per capire meglio l'algoritmo, facciamo un esempio: se la tabella comincia con le righe

$$x_1 = 0,30954...$$
$$x_2 = 0,20167...$$
$$x_3 = 0,89712...$$

si ha

$$a_1^{(1)} = 3$$
$$a_2^{(2)} = 0$$
$$a_3^{(3)} = 7$$

e le prime cifre di $\overline{x}$ saranno

$$\overline{x} = 0,554...$$

A questo punto, per costruzione

- $\overline{x} \neq x_1$ perché differiscono (almeno) per la prima cifra: $b_1 \neq a_1^{(1)}$
- $\overline{x} \neq x_2$ perché differiscono (almeno) per la seconda cifra: $b_2 \neq a_2^{(2)}$
- $\overline{x} \neq x_3$ perché differiscono (almeno) per la terza cifra: $b_3 \neq a_3^{(3)}$
- ...

e quindi $\overline{x}$ è un numero reale appartenente all'intervallo $[0,1]$ ma non è nessuno dei numeri in elenco, assurdo perché l'elenco doveva essere completo. ■

Abbiamo quindi scoperto che gli insiemi infiniti non sono tutti ugualmente numerosi. Si può dimostrare (non lo facciamo) il seguente:

Teorema 0.8 *Ogni intervallo $I \subset \mathbb{R}$, limitato o illimitato, chiuso o aperto, purché non ridotto a un punto, ha la stessa cardinalità di $\mathbb{R}$.*

Definizione 0.9 *La cardinalità di $\mathbb{R}$ viene anche detta cardinalità del continuo, o potenza del continuo.*

Per quanto visto si ha quindi:

Teorema 0.10 *L'insieme $\mathbb{R}$ non è numerabile. La cardinalità del continuo è diversa dalla cardinalità numerabile.*

Ma come si fa a decidere quale tra due insiemi infiniti di diversa cardinalità ha cardinalità *maggiore*?

Definizione 0.11 *Si dice che $\operatorname{card} X < \operatorname{card} Y$ se X non ha la stessa cardinalità di Y e inoltre X ha la stessa cardinalità di un sottoinsieme proprio di Y.*

Si presti attenzione alla precedente definizione, solo apparentemente ovvia: ricordiamo che ogni insieme infinito ha (per definizione) la stessa cardinalità di *qualche* suo sottoinsieme proprio; perciò il solo fatto che X abbia la stessa cardinalità di un sottoinsieme proprio di Y non implica ancora che X abbia cardinalità minore: potrebbe averla uguale. La prima parte della definizione quindi (che X e Y non abbiano la stessa cardinalità) non può essere soppressa. Il modo più semplice in cui si realizza il fatto che "X ha la stessa cardinalità di un sottoinsieme proprio di Y" è quando $X \subsetneqq Y$. (In questo caso X è in corrispondenza biunivoca con se stesso, che è sottoinsieme proprio di Y).

Alla luce di queste osservazioni si può quindi riformulare l'ultimo teorema visto così:

Teorema 0.12 *L'insieme $\mathbb{R}$ non è numerabile. La cardinalità del continuo è maggiore dalla cardinalità numerabile.*

Dimostrazione. Difatti, $\mathbb{N} \subset \mathbb{R}$ (quindi $\mathbb{N}$ è in corrispondenza biunivoca con un sottoinsieme proprio di $\mathbb{R}$) e d'altro canto sapevamo già che i due insiemi non hanno la stessa cardinalità. ∎

Presentiamo, senza dimostrazione, qualche altro risultato che arricchisce un po' il quadro concettuale sulla cardinalità di insiemi:

Teorema 0.13 *L'insieme $\mathbb{R}$ ha la stessa cardinalità dell'insieme delle parti $\mathcal{P}(\mathbb{N})$, cioè l'insieme di tutti i sottoinsiemi di $\mathbb{N}$:*

$$\operatorname{card} \mathbb{R} = \operatorname{card} \mathcal{P}(\mathbb{N}) .$$

Teorema 0.14 *Se $\{E_n\}_{n=1}^{\infty}$ è una successione di insiemi tali che ogni E_n è numerabile, anche $\bigcup_{n=1}^{\infty} E_n$ è numerabile.*

Se $E_1, E_2, ..., E_n$ sono n insiemi numerabili, allora anche $E_1 \times E_2 \times ... \times E_n$ è numerabile.

Se $E_1, E_2, ..., E_n$ sono n insiemi aventi ciascuno la cardinalità del continuo, allora anche $E_1 \times E_2 \times ... \times E_n$ ha la cardinalità del continuo.

Ad esempio, $\mathbb{C}$ e $\mathbb{R}^n$ hanno la cardinalità del continuo, mentre $\mathbb{Q}^n$ è numerabile.

La cardinalità del numerabile è la più piccola cardinalità infinita:

Teorema 0.15 *Ogni insieme infinito contiene al suo interno un insieme numerabile. Di conseguenza, non esiste un insieme infinito con cardinalità minore di quella del numerabile.*

Finora conosciamo solo due cardinalità infinite, ci si può chiedere se ce ne siano altre. La risposta è affermativa:

Teorema 0.16 *Per ogni insieme X si ha* $\operatorname{card} X < \operatorname{card} \mathcal{P}(X)$.

Questo implica che la scala delle cardinalità non ha un limite superiore: di ogni insieme esiste sempre un insieme più numeroso.

Riflettiamo ora un momento sulla cardinalità degli spazi di funzioni (di variabile reale, per semplicità). L'insieme $\mathcal{P}(\mathbb{R})$ dei sottoinsiemi di $\mathbb{R}$ è in corrispondenza biunivoca con l'insieme delle funzioni caratteristiche di tali sottoinsiemi ($E \leftrightarrow \chi_E$), ossia con l'insieme delle funzioni

$$f : \mathbb{R} \to \{0, 1\}.$$

Naturalmente le funzioni

$$f : \mathbb{R} \to \mathbb{R}$$

ci aspettiamo che siano molte di più delle funzioni $f : \mathbb{R} \to \{0, 1\}$ (per ogni $x \in \mathbb{R}$, per assegnare il valore $f(x)$ invece di avere solo due scelte (0 o 1) ne ho un'infinità non numerabile). Quindi senz'altro si ha:

$$\operatorname{card}\{f : \mathbb{R} \to \mathbb{R}\} \geq \operatorname{card} \mathcal{P}(\mathbb{R}) > \operatorname{card} \mathbb{R}.$$

Infine, abbiamo detto che esiste la più piccola cardinalità infinita (quella numerabile) e abbiamo dato un criterio con cui, data una cardinalità infinita, costruirne una maggiore: passare da X a $\mathcal{P}(X)$. Così facendo, al primo passo da $\mathbb{N}$ si passa a $\mathcal{P}(\mathbb{N})$ cioè alla cardinalità del continuo. Ci si può chiedere se questo passo sia il più piccolo possibile o vi siano gradi intermedi. In altre parole:

esiste un insieme X per cui si abbia

$$\operatorname{card} \mathbb{N} < \operatorname{card} X < \operatorname{card} \mathbb{R} \ ?$$

Cantor era convinto di no, e l'affermazione della *non esistenza* di una cardinalità intermedia tra $\operatorname{card} \mathbb{N}$ e $\operatorname{card} \mathbb{R}$ prende il nome di "ipotesi del continuo". Cantor non riuscì però a dimostrare questa sua convinzione, che rimase a lungo un problema aperto. In effetti la teoria "informale" degli insiemi come era sviluppata nell'800 non consentiva di studiare a fondo la questione. Agli inizi del 20° secolo la teoria degli insiemi divenne una teoria assiomatica formale, a cominciare con Ernst Zermelo (1871-1953), nel 1908. Nell'ambito della teoria assiomatica degli insiemi, Kurt Godel (1906-1978) nel 1940 dimostrò che l'ipotesi

del continuo non può essere dimostrata falsa usando il sistema di assiomi di Zermelo-Fraenkel. D'altra parte, nel 1963 Paul Cohen (1934-2007) dimostrò che l'ipotesi del continuo non può essere neppure dimostrata vera a partire da quegli assiomi. L'ipotesi del continuo è quindi una "proposizione formalmente indecidibile" all'interno della teoria degli insiemi standard. Questo significa che si può scegliere di assumere la sua validità come ulteriore assioma della teoria degli insiemi, oppure fare una scelta diversa. Così come si può decidere di studiare la geometria euclidea o una geometria in cui l'assioma delle parallele è rimosso o sostituito con uno che afferma qualcosa di diverso, analogamente si può sviluppare la teoria degli insiemi aggiungendo alla lista degli assiomi della teoria standard l'ipotesi del continuo, o la sua negazione, o non aggiungendo niente: si ottengono teorie diverse ma ugualmente lecite. E questo è molto poco intuitivo...

Tornando agli spazi di funzioni, diamo la seguente

Definizione 0.17 (Spazio separabile) *Sia (X, d) uno spazio metrico. Si dice che X è separabile se esiste un suo sottoinsieme denso numerabile, ossia un insieme $X_0 \subset X$ numerabile, tale che ogni elemento di X è limite di una successione di elementi di X_0.*

Quindi in uno spazio separabile esistono "relativamente pochi" elementi (un sottoinsieme numerabile) che permettono di approssimare tutti gli altri. Ad esempio $\mathbb{R}$ o $\mathbb{R}^n$ sono insiemi separabili, perché $\mathbb{Q}$ è denso in $\mathbb{R}$ e $\mathbb{Q}^n$ è denso in $\mathbb{R}^n$. Ci interessa però discutere questo concetto per gli spazi di funzioni. Come già osservato, gli spazi di funzioni hanno solitamente una cardinalità non numerabile, anzi maggiore della potenza del continuo. E' particolarmente significativo, quindi, che, non ostante questo, talvolta si trovi un sottospazio X_0 avente solo un'infinità numerabile di elementi, che però sono sufficienti ad approssimare bene quanto si vuole qualsiasi elemento di X.