

Introduzione

Come si legge un libro di matematica

Una delle principali difficoltà nell'affrontare la preparazione di un esame di matematica (ma quello che si dice vale in gran parte anche per le altre materie scientifiche) è costituita dallo stabilire un rapporto corretto con il testo che si deve studiare. Lo studente in genere ha ascoltato le lezioni, nelle quali ha capito alcune cose, mentre altre, che gli sono sfuggite, si ripromette di vederle più tardi sul libro in modo da chiarirne i punti oscuri; come pure sul testo dovrà lavorare per fissare le parti che a lezione gli erano chiare. In ambedue i casi, il libro di testo è uno strumento indispensabile per l'apprendimento dei contenuti del corso, e non di rado la buona riuscita di un esame dipende più dal modo in cui si è affrontato lo studio che dal tempo passato sul libro.

Ora, una delle caratteristiche dei testi di matematica è di essere ellittici; sono libri densi, nei quali cioè ben poco spazio è riservato ai commenti e alle divagazioni, che semmai trovano posto a lezione; libri in cui quasi ogni parola è essenziale e richiede di essere ricordata. Né d'altra parte potrebbero essere molto diversi, a meno di non pensare a volumi di parecchie centinaia di pagine.

È dunque forte la tentazione di affrontarne lo studio a colpi di memoria bruta; di considerare cioè il libro, e in particolare alcune sue parti quali ad esempio le dimostrazioni dei teoremi, come un testo immutabile e fissato una volta per tutte, che si può solo sapere, vale a dire essere in grado di recitare senza interruzioni, o non sapere, e dunque aver bisogno di ulteriore lavoro.

In realtà le cose non stanno così, o meglio non stanno esattamente così. Se infatti è innegabile che la memoria svolge un ruolo importante nella preparazione (e questo vale per qualsiasi cosa si stia leggendo, da un romanzo all'annuncio pubblicitario), è anche vero che l'apprendimento mnemonico è troppo e troppo poco. Troppo, perché ad affrontare un testo per pura memoria sfugge un dato importante che invece è bene tener sempre presente, e cioè che la maggior parte delle dimostrazioni è una semplice variazione su un numero molto limitato di metodi fondamentali. Troppo poco, perché pur essendo ellittico, un testo di

matematica non è completo, e in genere, specie quando si è abbastanza progrediti nel corso, molti passaggi facili (o almeno giudicati tali da chi scrive) sono solo accennati o addirittura taciuti.

Una lettura proficua di un libro di matematica dovrebbe in primo luogo avere di mira questi due obiettivi: da una parte individuare nella molteplicità delle dimostrazioni il gioco di alcune tecniche che si ripetono costantemente; e dall'altra completare le dimostrazioni esplicitando tutti gli anelli mancanti o solo accennati. Il primo di questi aspetti è più complesso e può essere visto solo a posteriori, quando cioè si sia completata la lettura; il secondo invece non richiede grandi conoscenze e può essere intrapreso fin dall'inizio. Ambedue mirano non tanto a sostituire e a rendere inutile il lavoro mnemonico, che resta sempre una parte importante, quanto piuttosto a sopperire alle lacune di memoria aiutando a ricostruire passaggi eventualmente dimenticati, e a garantire che la mancanza di alcuni dettagli non causi la perdita irreparabile dell'intera struttura.

Per non restare nel vago, vediamo su un esempio come si può operare in una prima lettura del testo. La scelta del secondo paragrafo ("Il valore assoluto") del primo capitolo delle *Lezioni* è dovuta unicamente al fatto che il primo paragrafo non contiene altro che la descrizione del sistema dei numeri reali, un argomento in parte già noto e in parte da imparare a memoria. (Riporteremo in corsivo il testo, e in tondo i "commenti di lettura".)

Sia $a \in \mathbb{R}$. Si definisce valore assoluto (o modulo) di a il massimo tra i due numeri a e $-a$:

$$|a| = \max\{a, -a\} = \begin{cases} a & \text{se } a \geq 0 \\ -a & \text{se } a < 0. \end{cases}$$

In effetti, se $a \geq 0$, il massimo tra a e $-a$ è a , che è positivo, mentre $-a$ è negativo. Se invece $a < 0$, a è negativo e $-a$ è positivo. Allora il massimo tra a e $-a$ è a se $a \geq 0$, e $-a$ se $a < 0$. Vediamo qualche esempio. $|3|$ è 3, perché $3 \geq 0$. $|-1|$ è uguale a $-(-1)$ cioè a 1, perché $-1 < 0$. Quanto fa $|-a|$? Farà $-a$ se $-a \geq 0$, cioè se $a \leq 0$, e a se $-a < 0$, cioè se $a > 0$. Proprio come $|a|$, tranne che c'è > 0 invece di ≥ 0 , e ≤ 0 invece di < 0 . Ma con questo non cambia nulla, perché quando a è zero anche $-a$ è zero, e allora l'uguale può stare sia sopra che sotto. Quanto fa $|1 - x^2|$?

Si deduce immediatamente dalla definizione che $|a| \geq 0$,

Come? Se $a \geq 0$, $|a| = a \geq 0$, mentre se $a < 0$, $|a| = -a > 0$.

e che $|a| = 0$ se e solo se $a = 0$.

Se $a = 0$, $|a| = a = 0$. Viceversa, se $|a| = 0$, a non può essere negativo, perché altrimenti $|a| = -a > 0$. Ma allora $a = |a| = 0$.

Si ha inoltre $|-a| = |a|$ (questo si era già visto), e $a \leq |a|$; $-a \leq |a|$ (per forza, $|a|$ è il massimo tra a e $-a$). Siano ora a e b due numeri reali. Dalla proprietà (AB) e dalla relazione $b \leq |b|$ segue che

$$a + b \leq a + |b| \leq |a| + |b|.$$

La prima disuguaglianza è proprio la (AB), perché nella disuguaglianza $b \leq |b|$ si somma a a destra e a sinistra. La seconda è identica; basta aggiungere $|b|$ ai due membri di $a \leq |a|$. Si doveva però dire a rigore: “dalle relazioni $a \leq |a|$ e $b \leq |b|$ ”, e non: “dalla relazione $b \leq |b|$ ”; meglio ancora: “dalla relazione $b \leq |b|$, che vale per ogni numero reale b ”.

Ricordando che $|a + b| = \max\{(a + b), -(a + b)\}$, si deduce dalle precedenti relazioni la disuguaglianza triangolare $|a + b| \leq |a| + |b|$. [2.1]

Certo, infatti sia $a + b$ che $-(a + b)$, e dunque anche il maggiore tra essi, è minore o uguale a $|a| + |b|$.

Se ora in quest'ultima relazione poniamo $a + b = -c$ (e dunque $b = -c - a$), e ricordiamo che $|-a| = |a|$, otteniamo $|c| \leq |a| + |a + c|$,

Vediamo. Se poniamo $a + b = -c$, abbiamo $|c| = |-c| \leq |a| + |-(a + c)| = |a| + |a + c|$.

e dunque $|a + c| \geq |c| - |a|$. [2.2] Viceversa, da quest'ultima relazione si può risalire alla [2.1] sostituendo a c il suo valore $-a - b$;

Infatti si ottiene $|-b| \geq |-(a + b)| - |a|$, cioè $|b| \geq |a + b| - |a|$.

cosicché le [2.1] e [2.2] sono due relazioni equivalenti. Naturalmente possiamo scambiare tra loro a e c nella [2.2], e ottenere la relazione equivalente $|a + c| \geq |a| - |c|$,

Scambiare tra loro vuol dire chiamare a quello che prima era c e viceversa.

Questo si può fare perché la [2.2] vale per ogni a e c .

da cui, confrontando con la [2.2], si ottiene $|a + c| \geq ||a| - |c||$. [2.3]

Infatti $|a + c|$ è maggiore di $|a| - |c|$ e di $|c| - |a| = -(|a| - |c|)$, dunque anche del maggiore tra essi, cioè di $||a| - |c||$.

Siano ora $x, y \in \mathbb{R}$. Si definisce distanza tra x e y il numero reale $d(x, y) = |x - y|$. [2.4] Si verificano facilmente le seguenti proprietà:

(d₁) $d(x, y) \geq 0$; $d(x, y) = 0$ se e solo se $x = y$.

(d₂) $d(x, y) = d(y, x)$.

(d₃) Per ogni $z \in \mathbb{R}$, $d(x, y) \leq d(x, z) + d(z, y)$ (disuguaglianza triangolare).

Verifichiamo:

(d₁) $d(x, y) \geq 0$; $d(x, y) = 0 \Leftrightarrow x = y$. Questo perché $|x - y|$ è sempre positivo, e si annulla solo se $x - y = 0$, cioè se e solo se $x = y$.

(d₂) $d(x, y) = |x - y|$; $d(y, x) = |y - x| = |-(x - y)| = |x - y|$.

(d₃) Bisogna dimostrare che $|x - y| \leq |x - z| + |z - y|$. Questa disuguaglianza è simile alla [2.1], che si chiamava anch'essa disuguaglianza triangolare e che diceva $|a + b| \leq |a| + |b|$. Qui dovrebbe essere $a = x - z$ e $b = z - y$. Ma allora $a + b = (x - z) + (z - y) = x - y$. Bene, è proprio ciò che si cercava. Infatti $|a + b| = |x - y| \leq |a| + |b| = |x - z| + |z - y|$, che è quello che si voleva dimostrare.

Se $x_0 \in \mathbb{R}$ e $r > 0$, chiameremo intorno sferico di centro x_0 e di raggio r l'insieme di tutti i numeri reali che distano da x_0 meno di r ; in simboli: $I(x_0, r) = \{x \in \mathbb{R} : d(x, x_0) < r\} = \{x \in \mathbb{R} : |x - x_0| < r\}$.

Facciamo qualche esempio. $I(2, 1) = \{x \in \mathbb{R} : |x - 2| < 1\}$. Vediamo di trovare per quali x risulta $|x - 2| < 1$. Se $x < 2$, si ha $|x - 2| = 2 - x$, e dunque dovremo porre $2 - x < 1$, che è verificata per $x > 1$. Se invece $x \geq 2$, risulta $|x - 2| = x - 2$, e dunque si avrà $x - 2 < 1$, cioè $x < 3$. In conclusione, la disuguaglianza $|x - 2| < 1$ sarà verificata per gli $x < 2$ e > 1 , e per quelli ≥ 2 e < 3 , in breve per gli x compresi tra 1 e 3. In definitiva, $I(2, 1) = \{x \in \mathbb{R} : 1 < x < 3\}$. Allo stesso modo si vede che $I(1, 3) = \{x \in \mathbb{R} : |x - 1| < 3\} = \{x \in \mathbb{R} : -2 < x < 4\}$. Più in generale, $I(x_0, r) = \{x \in \mathbb{R} : x_0 - r < x < x_0 + r\}$.

Con questo termina il paragrafo esaminato, ma non la nostra introduzione alla lettura di un testo matematico. Infatti in questo paragrafo manca il punto centrale del discorso matematico: la dimostrazione di un teorema.

Un teorema è costituito da una o più assunzioni (le ipotesi) e da una conclusione (la tesi). La sua dimostrazione consiste in una serie di argomentazioni che conducono a provare come la tesi sia una conseguenza delle ipotesi, nonché di altre proprietà già note, vuoi perché sono parte delle definizioni degli oggetti in gioco, vuoi perché sono state dimostrate in precedenza. La lettura di una dimostrazione mira a individuare i vari elementi che vi intervengono: le ipotesi, le definizioni, i risultati precedenti. Ogni frase della dimostrazione rinvia in genere a uno o più di questi elementi; talvolta i rimandi sono espliciti (per il Teorema 5.2. ..., dalla Definizione 1.4. ...), in alcuni casi vengono sottaciuti, ed è compito di chi legge renderli palesi, in modo da porre in luce la trama completa dell'argomentazione.

Da questo punto di vista, la lettura di un teorema non differisce molto da quella di un qualsiasi testo matematico. La domanda che ci si pone davanti ad ogni affermazione del tipo: da $\mathcal{P}$ segue $\mathcal{Q}$, è sempre la stessa: perché? Ma a differenza di passi quale il paragrafo che abbiamo appena commentato, e proprio a causa della sua struttura compiuta, un teorema (o lemma, o proposizione, ...) richiede anche un altro tipo di indagine, in particolare riguardo al ruolo delle ipotesi nell'economia della dimostrazione. Qui le domande a cui bisogna rispondere si situano a livelli successivi di elaborazione, e non sono sempre le stesse per tutti i teoremi. Vediamo le più semplici.

In primo luogo: dove entrano le ipotesi nella dimostrazione? A questa domanda si sarà già risposto tramite l'analisi precedente, che avrà permesso di individuare ed esplicitare tutti i passaggi della dimostrazione.

Viene ora la seconda questione: le ipotesi sono tutte necessarie? Essa permette una risposta di due tipi. Il primo, formale, consiste nel ripetere il passo precedente, facendo vedere che per ognuna delle assunzioni c'è un passaggio della dimostrazione che non potrebbe essere compiuto senza di essa. Naturalmente, questa risposta non prova altro se non che la particolare dimostrazione non potrebbe funzionare senza l'ipotesi K in questione, ma non dimostra che in assenza di questa il teorema diventerebbe falso. Non si dimentichi che in generale un teorema può essere provato in più modi, cosicché l'ipotesi K , se è essenziale nella dimostrazione riportata nel libro, può divenire superflua in un'altra dimostrazione, magari più complessa della prima; in altre parole, è possibile che K sia un'ipotesi di comodo, fatta allo scopo di semplificare una dimostrazione altrimenti molto complessa.¹ Entra allora in gioco il secondo tipo di risposta, più sostanziale del precedente: far vedere che togliendo l'ipotesi K il teorema non è più vero, individuando una situazione nella quale tutte le ipotesi (ad eccezione di K) siano verificate senza che la tesi sussista. Si tratta, come si dice, di trovare un controesempio. Perlopiù questo esempio si presenta immediatamente, e non ci insegna praticamente nulla; talvolta però esso è più riposto, e dunque contribuisce a chiarire il significato del teorema in esame.

Illustriamo quanto detto con un esempio, questa volta preso un po' più avanti nel testo. Va da sé che il lettore è invitato a leggere, o quanto meno a rileggere, la parte che segue quando avrà studiato il teorema in questione, cioè il *Teorema di Weierstrass* sui massimi e minimi di una funzione continua (vedi *Lezioni*, cap. 3, Teorema 10.4). Esso dice:

Una funzione continua f in un insieme compatto ha massimo e minimo.

Le ipotesi sono dunque due: che la funzione f sia continua in un insieme $Q \subset \mathbb{R}$, e che Q sia compatto. Quest'ultima poi si divide in due, dato che in $\mathbb{R}$ sono compatti tutti e soli gli insiemi chiusi e limitati.

Dove entrano queste ipotesi? Esaminando la dimostrazione si vede che la compattezza di Q si usa quando dalla successione x_n si estrae una sottosuccessione x_n^* che converge a un punto $x_0 \in Q$; la continuità della funzione f , quando si conclude che

$$\lim_{n \rightarrow \infty} y_n^* \equiv \lim_{n \rightarrow \infty} f(x_n^*) = f(x_0).$$

In ambedue i casi le ipotesi sono essenziali per condurre a termine la dimostrazione. Sono però veramente essenziali per la validità del teorema?

Cominciamo dalla compattezza. Ci domandiamo: il teorema sussiste ugualmente se supponiamo soltanto che Q sia chiuso, ma non limitato? o che sia limitato, ma

¹ In genere, quando si fa un'ipotesi di comodo, è buona norma segnalarlo. Non è detto però che le buone norme vengano sempre seguite, neanche da chi scrive. Un'altra possibilità, più rara in un testo elementare, è che l'autore non si sia accorto di aver aggiunto un'ipotesi superflua, o non abbia saputo condurre a termine la dimostrazione senza di essa.

non chiuso? Qui il controesempio è presto trovato: la funzione x non ha massimo né minimo in $\mathbb{R}$ (un insieme chiuso ma non limitato), mentre la funzione $\operatorname{tg} x$ non ha né massimo né minimo nell'intervallo aperto e limitato $(-\pi/2, \pi/2)$. In ambedue i casi l'estremo inferiore della funzione è $-\infty$, e quello superiore è $+\infty$. Ma allora ci possiamo chiedere: e se aggiungessimo l'ipotesi che la funzione $f(x)$ è limitata? Neanche così va bene: la funzione $\operatorname{arctg} x$ è continua e limitata in $\mathbb{R}$, e la funzione... (il lettore completi) è continua e limitata in $(-1, 1)$, ma esse non hanno né massimo né minimo.

Veniamo ora alla continuità: si può trovare una funzione $f(x)$ definita in un insieme Q compatto (ma non continua) e che non abbia né massimo né minimo? Anche qui l'esempio si trova subito: basta prendere una funzione $f(x)$ definita in un intervallo compatto $[a, b]$ e con un unico punto di massimo e uno di minimo, ambedue interni ad $[a, b]$, e cambiarne il valore in questi punti. Un esempio potrà essere allora la funzione $\sin x$, che ha massimo in $\pi/2$ e minimo in $3\pi/2$, ambedue interni a $[0, 2\pi]$. La funzione $f(x)$ definita in $[0, 2\pi]$ e che vale $\sin x$ se $x \neq \pi/2$ e $x \neq 3\pi/2$ e vale 0 in questi due punti, non ha né massimo né minimo.

Con ciò abbiamo terminato l'analisi del teorema di Weierstrass. Ci si potrebbe però spingere più oltre, come abbiamo in parte fatto durante la discussione della compattezza, verso un terzo livello di elaborazione. Potremmo chiederci: si possono indebolire le ipotesi, magari ottenendo una tesi anch'essa più debole, ma ancora significativa?

Cerchiamo di rosicchiare qualcosa alla continuità: supponiamo di voler dimostrare solamente che la funzione $f(x)$ ha minimo in Q ; possiamo fare a meno della continuità di $f(x)$, sostituendola con qualche ipotesi più generale?

La risposta viene in questo caso dall'analisi della dimostrazione. Qui avevamo preso una *successione minimizzante* x_n (tale cioè che $f(x_n) \rightarrow \inf_Q f(x)$) e ne avevamo estratto (grazie alla compattezza di Q) una successione $x_n^* \rightarrow x_0 \in Q$. Vogliamo dimostrare che $f(x_0) = \inf_Q f(x)$. È qui che entra la continuità della funzione $f(x)$: questa implica infatti che $f(x_n^*) \rightarrow f(x_0)$, e dato che per costruzione $f(x_n^*) \rightarrow \inf_Q f(x)$ si può concludere che $f(x_0) = \inf_Q f(x)$, e dunque che x_0 è un punto di minimo.

Ma, a pensarci bene, per dimostrare che $f(x_0) = \inf_Q f(x)$ è sufficiente far vedere che $f(x_0) \leq \inf_Q f(x)$, dato che il segno $<$ è impossibile. Ma allora non sarà necessario supporre che $f(x_n^*) \rightarrow f(x_0)$ (ovvero la continuità della funzione f), ma basterà che sia $f(x_0) \leq \lim_{n \rightarrow \infty} f(x_n^*)$, o meglio, poiché non sappiamo se questo limite esiste (avendo abbandonato l'ipotesi che la funzione $f(x)$ sia continua), che risulti

$$f(x_0) \leq \min \lim_{n \rightarrow \infty} f(x_n).$$

Ciò conduce alla definizione di semicontinuità inferiore (o superiore, se si cerca il massimo) e alla generalizzazione del teorema di Weierstrass a cui abbiamo accennato negli esercizi 10.1 e 10.2 del capitolo 3 delle *Lezioni*.

Lo stesso esame ci dice come si può evitare l'ipotesi di compattezza, sostituendola con opportune ipotesi sul comportamento della funzione $f(x)$ in vicinanza della frontiera di Q .

Limitiamoci per semplicità al caso in cui Q è un intervallo $[a, b)$, con b finito o $+\infty$. Come abbiamo visto, la compattezza di Q serviva per garantire che da una successione minimizzante x_n era possibile estrarre una sottosuccessione convergente a un punto $x_0 \in Q$. Nel nostro caso, potrebbe accadere che $x_n \rightarrow b$, e allora, poiché $b \notin Q$, la dimostrazione crolla.² Dobbiamo dunque escludere questa possibilità, introducendo un'ipotesi sul comportamento della funzione $f(x)$ in un intorno di b . Un'ipotesi sicuramente sufficiente è che risulti

$$\lim_{x \rightarrow b} f(x) = +\infty,$$

dato che in questo caso una successione *minimizzante* x_n non può tendere a b . Più in generale, sarà sufficiente supporre che

$$\inf_Q f(x) < \min_{x \rightarrow b} \lim f(x).$$

Con questo terzo livello di elaborazione, con il quale siamo già vicini alle soglie della ricerca matematica, terminiamo questa introduzione. Nell'esposizione siamo stati volutamente prolissi, sia per mostrare le particolari tecniche che differenziano la lettura di un testo matematico da quella di un'opera letteraria, sia per mettere in evidenza come anche in un paragrafo iniziale, nel quale dunque non si dà nulla per scontato, molti passaggi vengano solo accennati, lasciando al lettore il compito di completarli. Naturalmente un lettore esperto può tralasciare le cose più semplici, concentrandosi sui passaggi più delicati. Non bisogna però mai dimenticare che ciò che sembra evidente quando si ha un'esperienza abbastanza ampia può non esserlo affatto all'inizio, e che l'esperienza si acquista anche esercitandosi su cose che più tardi sembreranno banali.

Anzi, si può dire che lo scopo di un testo matematico è proprio quello di diventare ovvio.

² In effetti, questo è quanto accade in generale: ad esempio la funzione $1/x$ non ha minimo in $[1, +\infty)$, e tutte le successioni minimizzanti x_n tendono a $+\infty$.