

Seminario Didattico per studenti di 3^a e 4^a liceo

Matematica e Fisica in Dialogo: Il Potere del Metodo di Gauss

Franco Dassi
franco.dassi@unimib.it

1 Introduzione

Tutti voi conoscete le equazioni. Per esempio tutti voi conoscete che cosa significa la seguente scrittura

$$3x + 2 = 1. \quad (1)$$

In questa equazione riconosciamo una variabile x . Ci poniamo quindi il problema di trovare il valore di questa variabile incognita tale per cui abbiamo un'identità. Per raggiungere questo obiettivo vengono utilizzati i cosiddetti *principi di equivalenza*. Questi principi li conoscete di già e corrispondono alle operazioni che solitamente eseguite per risolvere un'equazione, ossia trovare il valore di x .

In particolare

“portare a dx cambiando il segno”,

corrisponde a questo principio

“sommare a dx e sx quantità uguali *non* altera l'identità”.

In Equazione (1) possiamo sommare a destra e sinistra il valore “-2” in modo da isolare la variabile x . Se procediamo in questo modo si vede che a sinistra dell'equazione il “+2” scompare e appare il “-2” dall'altra parte.

$$\begin{aligned} 3x + 2 &= 1 \\ 3x + 2 - 2 &= 1 - 2 \\ 3x &= 1 - \mathbf{2} \\ 3x &= -1 \end{aligned}$$

L'effetto finale è +2 si è spostato dall'altra parte dell'equazione cambiando il segno. Qualcuno qui potrebbe contestare il fatto che qui si stia sottraendo a destra e sinistra la stessa quantità e quindi non si stia utilizzando il principio di equivalenza sopra citato. Ebbene in matematica la differenza viene vista come la somma per il valore opposto. Di conseguenza stiamo effettivamente utilizzando il principio di equivalenza citato in precedenza. Nello stesso modo il

“portare sotto”

nasconde il seguente principio

“dividere a dx e sx quantità uguali *non* altera l'identità”.

In particolare per Equazione (1) abbiamo

$$\begin{aligned} 3x &= -1 \\ \frac{3x}{3} &= \frac{-1}{3} \\ x &= \frac{-1}{3} \end{aligned}$$

Anche in questo caso l'effetto di dividere entrambi i membri non si vede direttamente e il risultato finale sembra proprio quello di “aver portato sotto il 3”, ma questo passaggio si basa sui principi di equivalenza.

Con quest'ultimo passaggio abbiamo *risolto* l'equazione, ossia abbiamo trovato il valore di x che sostituito in Equazione (1) ci permette di avere un'identità. Infatti,

$$3x + 2 = 1 \quad \implies \quad 3\left(-\frac{1}{3}\right) + 2 = 1 \quad \implies \quad 1 = 1.$$

Un'osservazione molto importante sui principi di equivalenza è che essi *non* alterano la soluzione dell'equazione stessa se applicati in modo corretto e non sbagliando i conti. L'idea è applicarli in cascata fino a quando l'equazione “dice” la soluzione. Infatti se provassimo a sostituire la soluzione finale $x = -1/3$ ai vari passaggi intermedi si dovrebbe sempre ottenere un'identità.

L'esempio che vi ho fatto vedere prima è una classica equazione lineare. Tutti voi sapete che questa è la forma più semplice di equazione. Infatti, Equazione (1) può essere ulteriormente complicata “aumentando il grado”, ossia facendo apparire potenze maggiori di 1 della variabile incognita x . Espressioni di questo tipo

$$x^2 + 2x - 1 = 6 \quad \text{oppure} \quad x^3 + 4x^2 - x - 6 = 7, \quad (2)$$

pongono il medesimo problema di Equazione (1):

“trovare il valore di x in modo tale da avere un'identità.”

Purtroppo, risolvere questo tipo di equazioni è più complicato per questi motivi fondamentali

- non è sicuro che una soluzione x esista;
- c'è ne possono essere più di una
- non esiste un metodo semplice e sempre applicabile come i principi di equivalenza per risolverle.

Dato che le equazioni lineari sono quelle che si studiano per prime e che sono semplici da risolvere, si potrebbe pensare erroneamente che non siano utili né che possano essere sfruttati per risolvere problemi complicati. Ebbene questo è sbagliato. Vedremo che le equazioni lineari e in particolare i sistemi di equazioni lineari sono alla base della simulazione numerica e che anche il famoso motore di ricerca di Google si basa sulla risoluzione di un sistema lineare.

2 Calore di una stanza

Consideriamo un semplice modello per determinare la distribuzione di temperatura in una stanza conoscendo quella delle pareti. Vedremo che per giungere a questo obiettivo ci ridurremo alla risoluzione di un sistema lineare di grande dimensione.

Modellizziamo la stanza un quadrato e scegliamo un numero finito di punti dove calcoleremo un'approssimazione della vera temperatura, vedi Figura 1.

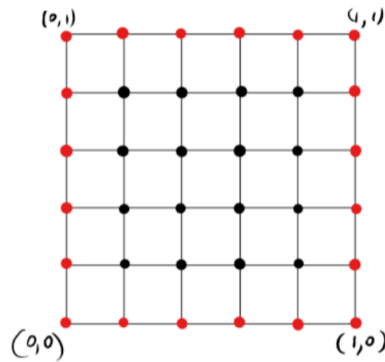


Figura 1: Discretizzazione di una stanza.

Supponiamo di conoscere il valore della temperatura nei pallini rossi che rappresentano le pareti della stanza. Il nostro obiettivo è quello di calcolare la temperatura nei punti all'interno della stanza, i.e., i pallini neri di Figura 1.

Intuitivamente maggiore è il numero di pallini, migliore sarà la descrizione della temperatura della stanza di questo modello. Ma vedremo che questo comporterà un maggiore sforzo perché il sistema lineare che dovremmo risolvere diventa molto grande.

Nella discretizzazione rappresentata in Figura 1 abbiamo 36 pallini equidistribuiti che rappresentano la temperatura in punto specifico della stanza. In particolare distinguiamo i seguenti casi:

- 20 pallini rossi di cui supponiamo di conoscere la temperatura;
- 16 pallini neri all'interno della stanza di cui vogliamo conoscere la temperatura.

Per calcolare la temperatura in questi punti neri facciamo la seguente supposizione:

La temperatura di un punto è la media di quelli vicini.

Sfruttiamo la connessione dei quadrati per stabilire quali sono i vicini a un dato punto. Di conseguenza la temperatura di un punto nero è data dalla media dei quattro punti a lui connessi, in formule avremmo che

$$T_C = \frac{1}{4}(T_N + T_S + T_E + T_O), \quad (3)$$

dove T_C è la temperatura del punto centrale che vogliamo calcolare mentre T_N , T_S , T_E e T_O sono le temperature dei vicini a nord, sud, est e ovest.

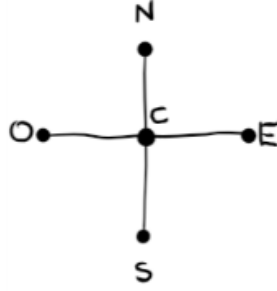


Figura 2: Punti connessi a uno dato.

Equazione (3) descrive matematicamente la supposizione che avevamo fatto sul valore della temperatura in un punto e può essere riscritta nel seguente modo

$$T_C - \frac{1}{4}T_N - \frac{1}{4}T_S - \frac{1}{4}T_E - \frac{1}{4}T_O = 0, \quad (4)$$

che ricorda meglio l'equazione di un sistema lineare le cui incognite sono T_C , T_N , T_S , T_E e T_O . Tuttavia a seconda del punto nero che stiamo considerando in Equazione (4) alcune temperature saranno note. Di conseguenza queste temperature non sono più incognite e “possono essere portate a destra”. Per capire meglio questo fatto consideriamo la numerazione dei punti interni descritta in Figura 3.

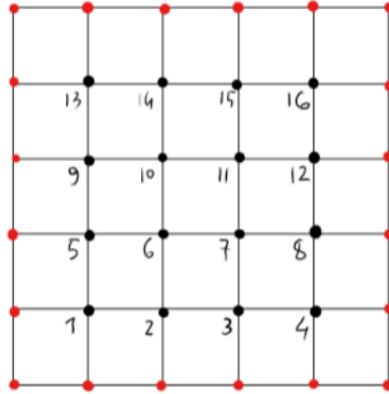


Figura 3: Punti connessi a uno dato.

Distinguiamo i seguenti casi

- per il punto 6 *tutti* i connessi sono pallini neri quindi avremmo esattamente Equazione (4), i.e.,

$$T_6 - \frac{1}{4}T_{10} - \frac{1}{4}T_2 - \frac{1}{4}T_7 - \frac{1}{4}T_5 = 0,$$

dove tutte le temperature dei punti adiacenti non sono note.

- per il punto 2 c'è un punto rosso, ossia un punto sul bordo per cui è nota la temperatura, quindi Equazione (4) avrà valori di temperatura noti che possono essere portati a destra dell'equazione, i.e.,

$$T_2 - \frac{1}{4}T_6 - \frac{1}{4}T_3 - \frac{1}{4}T_1 = \frac{1}{4}T_S,$$

dove T_S è nota ed è la temperatura del punto rosso che è connesso a sud del punto 2, mentre T_2 , T_6 , T_3 e T_1 non sono note.

- per il punto 1 ci sono 2 punti rossi vicini, uno a sud e uno a ovest, quindi Equazione (4) avrà valori di temperatura noti che possono essere portati a destra dell'equazione, i.e.,

$$T_1 - \frac{1}{4}T_5 - \frac{1}{4}T_2 = \frac{1}{4}T_S + \frac{1}{4}T_E,$$

dove T_S e T_E sono note e sono la temperatura dei punti rosso connessi a sud ed est del punto 1, mentre T_1 , T_5 e T_2 non sono note.

In questo modo si vede che avremmo un sistema lineare composto da 16 equazioni in 16 incognite. Tale sistema è abbastanza complicato da risolvere a mano. Risulta quindi necessario l'utilizzo di un computer per ottenere il valore di tutte e sedici le incognite.

3 Sistemi lineari e computer

Partiamo da un esempio. Supponete di voler risolvere il seguente sistema lineare

$$\begin{cases} x + y + z = 1 \\ x + y = 4 \\ x + 2y - z = 3 \end{cases} . \quad (5)$$

Ognuno di voi non è un automa quindi risolverà questo sistema lineare in diversi modi in base alle proprie conoscenze e osservazioni. Un computer non ragiona in questo modo: lui esegue delle istruzioni senza spaventarsi nè dei conti nè del tempo che gli occorrerebbe a eseguirle. Tuttavia ha bisogno di avere una procedura precisa che può essere *sempre* applicata e che tiene conto di tutti i possibili casi che possono succedere. Oltre al problema di trovare questa procedura c'è anche il problema di capire come “passare” nel computer un sistema lineare.

In questa sezione ci occuperemo di dare una risposta a entrambe queste domande.

Occupiamoci della prima domanda: come il computer rappresenta un sistema lineare. Il sistema lineare è rappresentato da una tabella. In particolare il sistema precedente è rappresentato nel seguente modo

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 4 \\ 1 & 2 & -1 & 2 \end{array} \right] . \quad (6)$$

In questa tabella ogni riga rappresenta una specifica equazione, mentre ogni colonna rappresenta il coefficiente di fronte alla variabile. In particolare la prima colonna contiene tutti i coefficienti di fronte alla x , la seconda tutti quelli

di fronte alla y e, infine, la terza tutti quelli di fronte alla z . Da notare il fatto che ci sia esplicitamente scritto “0” nella seconda riga dato che la variabile z non appare nell’equazione corrispondente. L’ultima colonna rappresenta tutti i termini a destra dell’uguale che non dipendono da una variabile.

Osservazione 3.1. *Graficamente inseriamo una linea per dividere la parte dei coefficienti da quella dei termini a destra. Nel computer non è così: non viene rappresentata la riga. Tutti questi numeri vengono messi in aree di memoria adiacenti una in seguito all’altra, ma lui sa quante sono le gli elementi che rappresentano ogni equazione.*

1	1	1	1	1	1	0	4	1	2	-1	2
---	---	---	---	---	---	---	---	---	---	----	---

Vedremo che giocando con questi coefficienti è possibile risolvere un sistema lineare. Inoltre per comodità utilizzeremo la rappresentazione di Equazione (6) invece di quella descritta nella Osservazione 3.1.

Passiamo ora al modo in cui un computer risolve un sistema lineare. Il metodo che descriveremo si basa sulla *riduzione* di un sistema lineare. La procedura di riduzione consiste nel sostituire un’equazione del sistema lineare con una combinazione dell’equazione sostituita con un’altra.

È chiaro che la riduzione non cambia la soluzione del sistema, ossia non cambia il valore delle variabili che soddisfano tutte le equazioni. Infatti se supponiamo che

$$x + y + z = 1 \quad \text{e} \quad x + y = 4,$$

allora necessariamente vale

$$(x + y + z) + (x + y) = (1) + (4) \implies 2x + 2y + z = 5.$$

Inoltre tale relazione vale anche quando mettiamo un coefficiente di fronte a una delle due equazioni o a entrambe. Per esempio

$$(x + y + z) - 3(x + y) = (1) - 3(4) \implies -2x - 2y + z = -11.$$

Osservazione 3.2. *Con combinazione si intende la somma di due equazioni moltiplicate per un opportuno coefficiente, i.e.,*

$$\text{Eqn1} + \alpha \text{Eqn2},$$

dove α è un generico numero reale.

Nel sistema di Equazione (5) possiamo per esempio sostituire la terza equazione,

$$x + 2y - z = 3,$$

con lei stessa meno la prima

$$(x + 2y - z) - (x + y + z) = 3 - 1 \implies y - 2z = 2. \quad (7)$$

Questa operazione *non* altera la soluzione del sistema lineare, ossia la soluzione del sistema di Equazione (5) è la stessa del sistema

$$\begin{cases} x + y + z = 1 \\ x + y = 4 \\ y - 2z = 2 \end{cases}. \quad (8)$$

Possiamo pensare di applicare la riduzione anche al sistema di Equazione (8): sostituiamo la seconda equazione con la differenza fra la prima e la seconda

$$(x + y + z) - (x + y) = 1 - 4 \quad \implies \quad z = -3.$$

In questo modo il sistema diventa:

$$\begin{cases} x + y + z &= 1 \\ z &= -3 \\ y - 2z &= 2 \end{cases}.$$

Per semplicità possiamo ordinare le equazioni in base al numero di variabili che appaiono invertendo semplicemente l'ordine delle equazioni

$$\begin{cases} x + y + z &= 1 \\ y - 2z &= 2 \\ z &= -3 \end{cases}. \quad (9)$$

La risoluzione del sistema di Equazione (9) è estremamente semplice. Partendo dall'ultima equazione e sostituendo le variabili trovate "all'indietro" è possibile trovare i valori di x , y e z . Quest'ultima procedura risolutiva è molto semplice da descrivere:

step 1) porta tutte le variabili che hai trovato a destra;

step 2) sostituisci i valori;

step 3) fai somme e prodotti;

step 4) dividi per il coefficiente di fronte alla variabile da trovare;

step 5) ripeti i passi 1-4 per la variabile più in alto.

Inoltre è anche la più conveniente per risolvere sistemi della forma di quello di Equazione (9). Non esiste un modo migliore e penso che tutti concordano che sia *la* procedura che utilizzerebbero senza pensarci due volte. Di conseguenza risulterebbe anche facile implementare in un computer questa procedura di sostituzione all'indietro per risolvere un sistema lineare. Rimane solamente un piccolo problema: non tutti i sistemi hanno questa forma. Però pensandoci bene è possibile sfruttare il processo di riduzione per ottenere un sistema lineare con questa forma. Vedremo inoltre che tale procedura sarà anche facilmente schematizzabile e quindi facilmente traducibile in istruzioni macchina da passare al computer.

Cerchiamo ora di capire come la riduzione ci può aiutare a ottenere un sistema della forma di Equazione (9). Guardando da un'altra prospettiva la riduzione notiamo che può essere sfruttata per eliminare una variabile in tutte le equazioni che abbiamo tranne una. Infatti, in Equazione (7) notiamo che abbiamo eliminato la variabile x dalla terza equazione. Possiamo applicare la stessa procedura per eliminare la variabile x dalla seconda e dalla terza equazione. Una volta eliminata la x da tutte le equazioni tranne una, si passa a eliminare la y da tutte tranne una sempre grazie alla riduzione e così via. Partendo dalla rappresentazione a tabella del sistema, Equazione (6), la riduzione è molto semplice perché si riconduce a fare somme di numeri nelle colonne corrispondenti.

Possiamo allora schematizzare la procedura come segue

step 1) parti dalla colonna più a sinistra

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 4 \\ 1 & 2 & -1 & 2 \end{array} \right];$$

step 2) se in testa c'è un valore diverso da zero procedi, altrimenti scambia le righe in modo tale che ci sia un valore diverso da zero. Tale valore si chiamerà *pivot*

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 4 \\ 1 & 2 & -1 & 2 \end{array} \right];$$

step 3) fai in modo che sotto il pivot ci siano solamente zeri, per raggiungere tale obiettivo utilizza la riduzione

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & -3 \\ 0 & 1 & -2 & -1 \end{array} \right];$$

step 4) ripeti gli step 1-3 sulla tabella più piccola

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & -3 \\ 0 & 1 & -2 & -1 \end{array} \right].$$

Una volta applicata la procedura precedente otteniamo la seguente tabella

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & -1 & -3 \end{array} \right],$$

dalla struttura a scalini della tabella è chiaro che la soluzione del sistema a essa associato si ottiene tramite una sostituzione all'indietro.

Un aspetto interessante è lo **step 2**. È sempre possibile trovare un valore diverso da zero? Che cosa succede se non lo troviamo? Dato un sistema lineare, non è detto a priori che sia possibile trovare una soluzione. Ora se nella procedura descritta in precedenza *non* è possibile trovare un pivot, questo significa che il sistema lineare non ha soluzione. Di conseguenza possiamo cambiare lo **step 2** nel seguente modo:

step 2) se in testa c'è un valore diverso da zero procedi, altrimenti scambia le righe in modo tale che ci sia un valore diverso da zero. Tale valore si chiamerà *pivot*. Se non si riesce a trovare questo valore il sistema non ammette soluzione.

Un sistema lineare potrebbe avere anche infinite soluzioni. Anche questo caso può essere gestito, ma vedremo sempre casi in cui la soluzione esiste ed è unica quindi non ne parleremo.

3.1 Problemi

Partendo da questa soluzione si potrebbe pensare che risolvere un sistema lineare sia una questione chiusa con una soluzione efficace. In realtà non è così. Ci sono una serie di problemi e questioni aperte legate a questo aspetto. Uno dei problemi più evidenti è il cosiddetto *fill-in*.

Solitamente i sistemi lineari che si risolvono al computer sono molto grandi e *sparsi*, ossia hanno molti coefficienti uguali a zero. Tale struttura è particolarmente interessante, infatti si può salvare molto spazio in termini di memoria. Infatti anziché salvare tutte le entrate è più efficace in termini di memoria utilizzata salvare solamente i coefficienti non nulli. Tuttavia se salvassimo i meri coefficienti non sarebbe più possibile ricostruire la matrice, quindi si decide di salvare le terne

$$(a_{i,j}, i, j).$$

Si può capire bene quanto sia notevole il guadagno. Se prima cresceva al crescere della dimensione della matrice, ora dipende solamente del numero di coefficienti diversi da zero: più precisamente sarà tre volte il numero di coefficienti diversi da zero.

Quando applichiamo la procedura precedente per risolvere un sistema lineare non è garantito che la struttura sparsa viene mantenuta. Anzi in generale osserviamo che tale caratteristica viene persa. Di conseguenza la nostra procedura fallisce solamente perché il computer non è in grado di gestire la matrice. Notare che “gestire” significa avere memoria sufficiente per usare la matrice, non vuol dire che la procedura precedente fallisce.

Per evitare questo problema sono nati i cosiddetti *metodi iterativi*. Tali metodi evitano il fenomeno del fill-in ma sono molto più complicati rispetto al metodo che vi ho precedentemente descritto.

4 Page Rank di google

Ogni volta che voi effettuate una ricerca tramite il motore di ricerca di google utilizzate il risultato di un sistema lineare. Supponete di voler cercare informazioni riguardanti la serie televisiva Obi-Wan Kenobi. I motori di ricerca agiscono nel seguente modo

1. cercando nel web i siti che contengono la parola chiave;
2. mettono in ordine di importanza i siti.

Il secondo passaggio è molto importante perché permette di trovare il sito più pertinente e che contiene le maggiori informazioni. Pensate a quanto è grande il mondo del web e soprattutto pensate a quanti siti possono contenere la parola “Obi-Wan Kenobi”, trovare il migliore potrebbe essere un vero problema. Google si occupa di quest’ultimo aspetto e ha trovato il modo al momento migliore di ordinare i siti salvo da “attacchi”.

I primi rudimentali motori di ricerca effettuavano questo ordine in base alla ricorrenza della parola cercata nella pagina. Si basavano su questo principio:

il sito più importante è quello in cui appare la parola cercata più volte.

Purtroppo questo modo di classificare è facilmente attaccabile. Nel lontano 1997 quando è uscito il celebre film “Titanic” con Leonardo di Caprio, prima della nascita di google, quando si cercava la parola chiave “Leonardo di Caprio” il risultato della ricerca dava una serie di siti a luci rosse. Questo perché questi siti per acquisire importanza cavalcavano l’onda del successo del film e riempivano le loro pagine con la parola “Leonardo di Caprio” ripetuta molte volte.

Di conseguenza è necessario utilizzare un altro criterio. Uno potrebbe pensare che un buon modo sia quello di contare il numero di siti che puntano al proprio, ma anche questo è facilmente attaccabile. Si possono creare infatti una serie di siti fittizi che puntano al proprio e quindi accrescere la sua importanza.

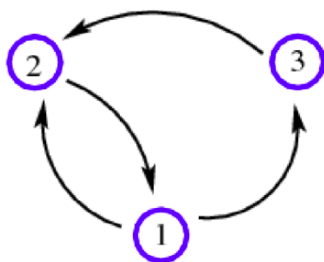


Figura 4: Grafo di un mini Word Wide Web.

I fondatori di Google trovarono il modo migliore per classificare i siti che si basa sul concetto di grafo e sulla risoluzione di un sistema lineare. Il criterio è il seguente

Una pagina che punta a altre distribuisce la sua importanza in parti uguali a tutte quelle che punta.

Questo principio richiede la conoscenza dei grafi. Infatti la connessione fra i siti del web è rappresentabile da un grafo, vedi Figura 4. Tale grafo può essere riassunto tramite una matrice

$$H := \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

dove ogni colonna e riga rappresenta ogni sito internet. Alla riga i colonna j viene messo un valore 1 se il sito j è puntato da quello i , altrimenti viene messo 0. Di conseguenza l’idea dei fondatori di Google si traduce nel sistema lineare

$$\begin{cases} w_1 &= w_2 \\ w_2 &= 0.5w_1 + w_3 \\ w_3 &= 0.5w_2 \end{cases}.$$

Partendo da questa distribuzione delle connessioni via web, possiamo affermare che il sito più importante è w_1 .