L'Analisi Numerica

Dario A. Bini, Università di Pisa

18 agosto 2019

Sommario

In questo articolo viene fatta un'introduzione molto discorsiva dell'Analisi Numerica descrivendone le caratteristiche specifiche e inquadrando questa disciplina nell'ambito della matematica e del contesto applicativo.

1 Matematica e Mondo Reale

Esistono molti luoghi comuni (e tante barzellette) sulla matematica e i matematici. L'immagine più benevola che si incontra nella mentalità comune è quella di una disciplina fine a sé stessa e per questo abbastanza inutile. Il matematico è visto come un personaggio strano che vive nel suo mondo fantastico e si occupa di cose strane e di importanza trascurabile per la vita di ogni giorno.

Alcune delle idee più diffuse e clamorosamente errate sulla matematica sono:

- la rigidità del pensiero matematico capace di esprimere solo meccanismi ripetitivi e automatici;
- la totale mancanza di fantasia;
- l'aridità e l'incapacità di esprimere qualcosa di nuovo, per cui è difficile per molti immaginare cosa possa esserci di più oltre alle formule e ai concetti visti nelle scuole superiori;
- la completa inutilità (a parte i conti della spesa e gli strumenti geometrici elementari) della matematica.

Invece, alcuni degli ingredienti essenziali nella ricetta del buon matematico sono la fantasia, l'immaginazione, la creatività, la libertà di pensiero non vincolato a rigidi schemi mentali, l'attrazione per l'eleganza, l'attrazione per la singolarità, il rigore logico formale. Queste caratteristiche rendono la matematica più vicina ad una forma di arte che non a una fredda scienza esatta. In matematica si generano idee continuamente nuove e il mondo matematico è per certi versi molto più ricco del mondo reale.

Di fatto la matematica è un bellissimo giocattolo che

– non stanca mai e non si rompe mai,

– sviluppa il pensiero libero la creatività e la fantasia,

– permette di creare quasi senza limiti strutture eleganti col massimo rigore logico,

– non ha colorazioni politiche né religiose,

– è proprietà di tutti,

– è uno strumento indispensabile per risolvere una gran parte di problemi del mondo reale.

Quest'ultimo punto è un aspetto che viene spesso trascurato. Infatti la matematica con i suoi modelli ha la capacità di descrivere ed analizzare i problemi del mondo reale in modo estremamente efficace. I modelli matematici vengono comunemente utilizzati in molti campi del mondo scientifico, industriale e tecnologico a tal punto che la matematica, a insaputa di molti, ha assunto un ruolo importante e pervasivo nella nostra vita quotidiana.

Ogni giorno ciascuno di noi fa uso inconsapevole di matematica. Un esempio significativo è il mondo di internet. Le ricerche di informazioni sul Web con i motori di ricerca tipo Google vengono fatte con tecnologie matematiche. L'ordinamento delle informazioni in base alla loro importanza viene calcolato usando la teoria di Perron-Frobenius delle matrici non negative, la scelta del documento più vicino alla richiesta fatta, così come la ricerca di sinonimi dipendenti dal contesto, si basa sulla decomposizione ai valori singolari (SVD) di una matrice.

La codifica delle informazioni e la loro trasmissione fatta minimizzando la possibilità di errore viene realizzata con tecniche algebriche nel campo della teoria dei codici. Stesse tecniche sono usate per codificare l'informazione musicale sui CD. Proprietà dei numeri primi e risultati di teoria dei numeri sono alla base dei codici di crittografia per proteggere le informazioni e le transazioni finanziarie su internet. Le stesse metodologie sono usate nei bancomat e nelle carte di credito.

La telefonia mobile e la telefonia sul Web sono altre utilità quotidiane che esistono grazie alla matematica. Infatti le modalità con cui i pacchetti di informazione vocale vengono costruiti e instradati in rete utilizzano protocolli basati su modelli matematici di teoria delle code in cui intervengono catene di Markov. Queste vengono risolte attraverso lo studio di complesse equazioni in cui l'incognita è una matrice. Lo stesso tipo di problema si incontra nelle reti wireless che ormai sono presenti in molte abitazioni.

Catene di Markov trovano applicazione in molti altri campi. In particolare nella filogenetica e nello studio nelle catene di amminoacidi del DNA.

Tutta la tecnologia digitale, inclusa la fotografia i filmati la musica, la registrazione e l'analisi di suoni, il riconoscimento vocale, il riconoscimento di immagini, è essenzialmente una tecnologia matematica. La compressione jpeg delle immagini richiede il calcolo di trasformate discrete dei coseni, il filtraggio di immagini e di suoni richiede il calcolo di trasformate discrete di Fourier. Trasformate di Fourier e trasformate di Radon assieme a complesse equazioni integrali intervengono nei modelli di tomografia mediante raggi X la cui risoluzione permette di svolgere efficienti indagini cliniche. Le previsioni del tempo vengono realizzate risolvendo complesse equazioni differenziali che regolano l'evoluzione temporale di pressioni e temperature nell'atmosfera terrestre. Equazioni differenziali intervengono nello studio dell'inquinamento, nei modelli matematici dell'attività cardiaca, nello studio del comportamento del flusso sanguigno nel caso di patologie. Equazioni differenziali della fluidodinamica, risolte con i metodi della fluidodinamica computazionale, sono alla base della progettazione degli scafi e della geometria delle vele nelle imbarcazioni da regata. La vittoria della barca svizzera Alinghi alle edizioni della Coppa America del 2003 e del 2007 è avvenuta grazie anche al lavoro del team diretto da Alfio Quarteroni che ha sviluppato modelli matematici per una migliore progettazione delle vele e dello scafo, in particolare il bulbo, dell'imbarcazione.

Strumenti di statistica vengono diffusamente usati in ambito politico economico e commerciale attraverso i sondaggi di opinione, attraverso proiezioni per prevedere sviluppi di differenti possibili scenari, e nelle exit poll così come nelle indagini epidemiologiche e nell'analisi di efficacia di farmaci. Modelli matematici sono sviluppati per lo studio degli andamenti della borsa, nell'analisi dei rischi dalle compagnie di assicurazione.

Metodologie di ottimizzazione sono applicate per progettare in modo più efficiente processi lavorativi, gli orari nei trasporti, per migliorare la viabilità del traffico.

La teoria dei controlli e dei sistemi dinamici è largamente usata per gestire in modo automatico sistemi complessi che evolvono nel tempo, in particolare per il decollo, il volo e l'atterraggio dei velivoli. Il cosiddetto *pilota automatico* è di fatto un *pilota matematico* poiché è realizzato attraverso l'uso del filtro di Kalman.

La progettazione dei robot industriali per l'assemblaggio delle auto e per la lavorazione nelle catene di montaggio richiede la risoluzione di complessi sistemi di equazioni non lineari algebrici e non per i quali sono state sviluppate tecniche risolutive avanzate basate sulle basi di Groebner.

Oltre agli utenti tradizionali della matematica come l'ingegneria e la fisica con i loro modelli matematici, altre aree di ricerca apparentemente lontane come la biologia, la psicologia e la sociologia, hanno sviluppato al loro interno i settori di Biologia matematica, Psicologia matematica e Sociologia matematica.

2 La Matematica Computazionale e l'Analisi Numerica.

Nella risoluzione dei problemi matematici che governano il mondo reale da un lato è importante dimostrare che il modello è coerente, la soluzione esiste, magari è anche unica, ha proprietà qualitative (tipo regolarità e positività) utili per la validità del modello. Dall'altra è importante poter calcolare la soluzione esplicitamente o, nel caso ciò non fosse possibile, è importante poterla approssimare con sufficiente precisione. Occorre quindi disporre di efficienti metodi di calcolo e di strumenti per la loro analisi di efficienza.

La parte della matematica che si occupa degli aspetti computazionali e del calcolo effettivo della soluzione di problemi è la matematica computazionale. Essa è trasversale a tutte le discipline dall'algebra alla geometria dall'analisi alla probabilità e alla statistica e riguarda parte dell'informatica in particolare per le metodologie implementative. Accanto alle classiche discipline quali l'algebra, la geometria e l'analisi hanno avuto grande sviluppo l'algebra computazionale, e la geometria computazionale che studiano metodi di risoluzione adatti ad una implementazione a computer. Lo sviluppo delle discipline computazionali ha generato ulteriori problemi teorici di grande interesse che alimentano la ricerca astratta.

L'analisi numerica interseca ampiamente la matematica computazionale trattando gli aspetti computazionali di alcuni settori della matematica che hanno un ruolo preminente nelle applicazioni. La caratteristica fondamentale che distingue un metodo di risoluzione dell'analisi numerica da altri metodi di risoluzione è che in un metodo numerico i numeri reali utilizzabili oltre allo zero sono tutti e soli quelli che si possono scrivere in una base assegnata $\mathtt{B}$ (ad esempio 10) con un numero finito $\mathtt{n}$ di cifre e con un esponente, più precisamente come

$$\pm\mathtt{0.c_1c_2c_3\cdots c_n}\times\mathtt{B^e},\quad \mathtt{c_1}\neq\mathtt{0}$$

col significato di $\pm B^e\cdot\sum_{i=1}^{n}c_i\cdot B^{-i}$, dove $\mathtt{B}$ è la base della rappresentazione, $\mathtt{c_1},\ldots,\mathtt{c_n}$ sono le cifre, cioè degli interi compresi tra 0 e $\mathtt{B-1}$, $\mathtt{e}$ è l'esponente, ed $\mathtt{n}$ è il numero di cifre consentite.

Ad esempio 0.1234500000×10^5 rappresenta il numero intero 12345 con $\mathtt{B}=\mathtt{10}$, $\mathtt{n}=\mathtt{10}$. Questa rappresentazione è chiamata *floating point*. Si osservi che l'insieme dei numeri floating point non ha alcuna struttura algebrica, cioè non è chiuso rispetto alle operazioni aritmetiche, ad esempio il prodotto di due numeri di 10 cifre ha generalmente un numero di cifre superiore a 10. Questo crea particolari difficoltà quali ad esempio il fatto che per poter utilizzare sempre numeri floating point, il risultato di ogni operazione aritmetica va riportato dentro l'insieme dei numeri floating point mediante un procedimento di arrotondamento o di troncamento. Questo introduce inevitabilmente un errore con cui occorre convivere.

La rappresentazione floating point e l'aritmetica conseguente ha il vantaggio che le operazioni aritmetiche hanno un costo computazionale uniforme. Infatti indipendentemente dal numero di operazioni svolte il numero di cifre rimane sempre lo stesso. Mentre nelle aritmetiche usate ad esempio in algebra computazionale in cui le operazioni tra interi o tra razionali vengono implementate ed eseguite senza errori, il numero di cifre degli interi a numeratore e a denominatore delle frazioni ottenute nei calcoli crescono considerevolmente col numero di operazioni svolte. Per cui il costo computazionale di ciascuna operazione aritmetica non è limitabile a priori e cresce man mano che il metodo di calcolo procede nell'esecuzione. Naturalmente il vantaggio di avere un costo uniforme è pagato dalla presenza degli errori che possono essere generati nello svolgimento di ogni operazione aritmetica.

Questi problemi legati agli errori rischiano di rimanere nascosti all'utente e vengono aggravati dal fatto che il computer usa la base 2 mentre noi utilizziamo la base 10. Per cui certi numeri come 1/10 che sono rappresentabili senza errore in base 10, rappresentati in base 2 hanno uno sviluppo periodico e il loro troncamento necessario ad una rappresentazione con un numero finito di cifre introduce necessariamente degli errori. Ad esempio, se in un sistema di calcolo dotato di aritmetica floating point assegnamo ad una variabile `x` il valore di 1/10 scrivendo ad esempio `x=0.1` e se poi chiediamo al sistema di calcolo di mostrare sullo schermo il valore di `x`, ritroviamo sullo schermo il valore `0.1` e questo non è una sorpresa. è però un imbroglio poiché nella memoria del computer la variabile `x` non contiene di fatto 1/10 ma il numero floating point che in base 2 rappresenta al meglio 1/10 con le cifre (di solito 52) usate nella rappresentazione floating point. Il sistema di calcolo ha infatti convertito dalla base 2 alla base 10 il risultato presente in memoria e nell'arrotondamento effettuato nel riportare il numero sullo schermo l'errore presente in memoria è rimasto nascosto.

Definiamo *metodo di risoluzione numerica* o più semplicemente *metodo numerico* un procedimento costituito da un numero *finito* di operazioni aritmetiche o logiche applicate ad un insieme *finito* di dati numerici rappresentati in floating point che fornisce in uscita un numero *finito* di valori numerici, anch'essi rappresentati in floating point, che in qualche forma descrivono la soluzione del problema che dobbiamo risolvere.

Allora possiamo definire l'analisi numerica come *la disciplina che sviluppa e tratta strumenti matematici utili per costruire e analizzare metodi di risoluzione* numerica *di problemi matematici in generale.*

Tipici problemi classicamente parte dell'analisi numerica sono i problemi di algebra lineare, dalla risoluzione dei sistemi lineari al calcolo degli autovalori e al problema dei minimi quadrati; i problemi della risoluzione di equazioni e sistemi di equazione non lineari, algebrici e non; i problemi di integrazione di funzioni; i problemi di interpolazione e di approssimazione di funzioni; la risoluzione di equazioni differenziali ordinarie ai valori iniziali e ai valori ai limiti; la risoluzione di equazioni differenziali alle derivate parziali.

I vincoli sulla *finitezza* posti sul numero di cifre consentite e sul numero di operazioni consentite è legato alla *finitezza delle risorse* quali il tempo di calcolo e lo spazio di memoria, e comporta il considerare problematiche computazionali specifiche usualmente non rilevanti per le altre discipline. In particolare la finitezza delle risorse di tempo e spazio rende fondamentale tenere in considerazione la **complessità computazionale** di un metodo di risoluzione. La finitezza delle rappresentazioni numeriche conduce alle problematiche relative agli **errori**. Il trattare oggetti continui o processi asintotici con un numero finito di dati e di operazioni aritmetiche porta alle problematiche della **approssimazione** e della **discretizzazione.**

Riportiamo di seguito una descrizione di queste problematiche:

3 La complessità computazionale

Nella risoluzione di un problema è importante che il metodo di risoluzione non impieghi un tempo eccessivo quando viene eseguito mediante un computer. Si pensi ad esempio al problema delle previsioni meteorologiche dove occorre risolvere complesse equazioni differenziali. Se il metodo di risoluzione dovesse impiegare una settimana di tempo di calcolo per determinare le previsioni del tempo nell'arco dei due giorni successivi allora sarebbe completamente inutile. La situazione sarebbe ancora più drammatica nel caso della TAC in cui il metodo di calcolo non può impiegare dei mesi o degli anni per fare l'analisi di un paziente ricoverato d'urgenza. Diventa quindi importante valutare i metodi di risoluzione anche in base al loro costo computazionale e scegliere quello dotato delle caratteristiche migliori.

Un esempio significativo a questo proposito è quello della risoluzione di un sistema lineare di n equazioni ed n incognite.

Si consideri il metodo di Cramer che esprime le soluzioni come quozienti di due determinanti che vengono calcolati con la regola di Laplace dello sviluppo per righe. Un analisi del suo costo mostra che il numero di operazioni aritmetiche richieste è almeno $n!$ (n fattoriale). Il metodo di eliminazione Gaussiana richiede invece circa $\frac{2}{3}n^3$ operazioni aritmetiche. Se ad esempio $n = 50$, il metodo di Cramer richiede almeno 3×10^{64} operazioni mentre il metodo di Gauss circa 83333. Con un personal computer come quelli largamente diffusi che esegue in un secondo circa cento milioni di operazioni aritmetiche il primo metodo impiegherebbe 9.6×10^{48} anni, il secondo poco meno di un millesimo di secondo. Anche usando un super computer quale il Blue Gene dell'IBM che esegue circa 500 tera operazioni al secondo (1 tera= 10^{12}), il metodo di Cramer non fornirebbe il risultato in meno di 10^{42} anni.

Diventa quindi fondamentale nell'analisi dei metodi di risoluzione tenere conto del costo computazionale e cercare di ridurre il più possibile il numero di operazioni richieste per la risoluzione di un problema. Viene inoltre naturale porre la questione di valutare il numero minimo di operazioni che sono sufficienti a risolvere un problema assegnato. Tale numero viene chiamato *complessità aritmetica del problema*. Ad esempio, per il problema del calcolo del valore di un polinomio $p(x) = \sum_{i=0}^{n} a_i x^i$ di grado n in un punto x assegnato, si possono utilizzare vari metodi. Una possibilità è quella di calcolare prima le potenze x^i per $i = 2, \ldots, n$ eseguendo $n - 1$ moltiplicazioni, e poi calcolare la sommatoria con n addizioni e n moltiplicazioni per un totale di $3n-1$ operazioni aritmetiche. Un'altra possibilità consiste nell'utilizzare lo schema di Horner che è espresso dalla formula

$$p(x) = a_0 + x(a_1 + x(a_2 + x(\cdots + x(a_{n-1} + xa_n)\cdots))),$$

che richiede $2n - 1$ operazioni aritmetiche. è stato dimostrato da Alexander Ostrowski che $2n - 1$ è il minimo numero possibile di operazioni aritmetiche. Quindi la complessità del problema del calcolo dei valori di un polinomio di grado n in un punto x dati i coefficienti e il valore di x è di $2n - 1$ operazioni aritmetiche.

4 Gli errori

Già nel momento stesso in cui si rappresenta un dato con la notazione in virgola mobile si commette potenzialmente un errore, visto che il numero può avere una rappresentazione periodica o è irrazionale. Quindi già prima di iniziare i calcoli abbiamo a che fare con degli errori presenti nei dati. Non solo, ma anche quando eseguiamo operazioni aritmetiche si introducono degli errori visto che dobbiamo mantenere un numero finito di cifre nella rappresentazione in base. Ad esempio se la nostra rappresentazione è in base 10 con 5 cifre e consideriamo i numeri $x = 0.12345 \times 10^1$ e $y = 0.12345 \times 10^{-1}$ risulta $x + y = 0.1246845 \times 10^1$ che riportato ad una rappresentazione con 5 cifre diventa 0.12468×10^1. È quindi affetto da errore.

Questi errori presenti nei dati e generati dalle operazioni aritmetiche possono amplificarsi nel corso dell'esecuzione del metodo di risoluzione e rovinare completamente il risultato. Un esempio significativo a riguardo è dato dal calcolo della funzione esponenziale e^x. Per questo possiamo usare la formula

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

dove la somma viene arrestata quando tutte le cifre del risultato calcolato non cambiano più. Un'altra possibilità è considerare il fatto che $e^x = 1/e^{-x}$ per cui

$$e^x = 1/(1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \cdots)$$

Le formule sono matematicamente equivalenti. Ma se le applichiamo in aritmetica floating point ad esempio con $x = 15$, la prima ci fornisce il valore

$$\mathtt{3.26901737247211} \times \mathtt{10}^{\mathtt{6}}$$

la seconda

$$\mathtt{3.24842918574941} \times \mathtt{10}^{\mathtt{6}}$$

I due numeri hanno solo due cifre in comune. Quale dei due risultati è più preciso?

5 Approssimazione e discretizzazione

È certo che con un numero finito di operazioni aritmetiche a partire da un insieme di numeri razionali si possono ottenere solamente risultati razionali. Nel caso il problema da risolvere avesse soluzioni irrazionali non possiamo sperare di calcolare esattamente la soluzione ma dobbiamo accontentarci di approssimarla con sufficiente precisione. Un esempio lo abbiamo visto nel calcolo dell'esponenziale attraverso il troncamento di una serie di potenze. Altri esempi sono il calcolo di radici o di altre funzioni trascendenti, o il calcolo delle soluzioni di equazioni algebriche. In particolare sappiamo che dalla teoria di Evariste Galois le equazioni di grado maggiore o uguale a 5 non ammettono formule risolutive espresse mediante radicali e frazioni algebriche. Per cui la loro risoluzione non può essere fatta se non *approssimando* le soluzioni con processi iterativi.

In questo contesto nascono nuove problematiche computazionali quali dare condizioni di convergenza per le successioni che approssimano le soluzioni, valutare la velocità di convergenza, progettare metodi che abbiano la massima velocità di convergenza possibile per minimizzare i costi computazionali, studiare il comportamento asintotico in aritmetica floating point.

In certe situazioni, come nel caso del motore di ricerca Google e del metodo di PageRank, anche se la soluzione in principio può essere calcolata esattamente, essendo razionale, risulta molto più conveniente dal punto di vista computazionale approssimarla con una successione convergente. Per cui l'approccio iterativo viene utilizzato anche quando apparentemente non sarebbe necessario.

L'approssimazione spesso interviene nel momento in cui l'oggetto da calcolare non è un numero bensì una funzione $f(x)$. Ecco allora che cerchiamo di approssimare $f(x)$ con funzioni più facili da manipolare quali sono i polinomi. Infatti i polinomi dipendono da un *numero finito di parametri* e quindi soddisfano alle nostre condizioni di finitezza. Questa forma di approssimazione può essere ottenuta mediante interpolazione o minimizzando una media del discostamento delle due funzioni ottenuta mediante una norma. A seconda della norma utilizzata si hanno metodologie e problematiche nettamente diverse quali le tecniche di approssimazione uniforme col metodo di Remez, e l'approssimazione dei minimi quadrati. Tecniche di questo tipo possono essere usate per la risoluzione numerica di equazioni differenziali.

Strettamente correlata all'approssimazione è il procedimento di *discretizzazione* che interviene allorché un oggetto di natura continua quale una funzione, deve essere sostituito da un oggetto di natura discreta, anzi di natura finita, rappresentabile con un numero finito di parametri. Il processo di interpolazione, in cui si sostituisce alla funzione un polinomio, è una forma di discretizzazione. La risoluzione numerica di equazioni differenziali avviene con processi di discretizzazione in cui la soluzione viene sostituita da una combinazione lineare *finita* di funzioni assegnate in cui i coefficienti vanno determinati imponendo le condizioni differenziali.

6 L'algebra lineare numerica

Un trattamento speciale lo richiede la risoluzione di problemi che appartengono al campo dell'algebra lineare quali la risoluzione di sistemi lineari, il calcolo di autovalori e autovettori, la risoluzione di problemi dei minimi quadrati. Infatti, ogni problema di natura lineare, una volta discretizzato si riduce ad un problema di algebra lineare. Inoltre, nella fase di discretizzazione, per avere una maggior precisione occorre utilizzare un numero elevato di punti di discretizzazione con un conseguente innalzamento delle dimensioni dei problemi di algebra lineare trattati. Possiamo dire che il trattamento numerico della totalità dei problemi di natura lineare conduce ad una formulazione di algebra lineare e che una gran parte di problemi non lineari, trattati con tecniche di linearizzazione, comportano la risoluzione di problemi di algebra lineare. Questo fatto ha portato a un grande sviluppo dei metodi numerici per l'algebra lineare configurando una disciplina autonoma e ben caratterizzata che va sotto il nome di *Numerical Linear Algebra* che viene inquadrata come parte essenziale dell'analisi numerica.