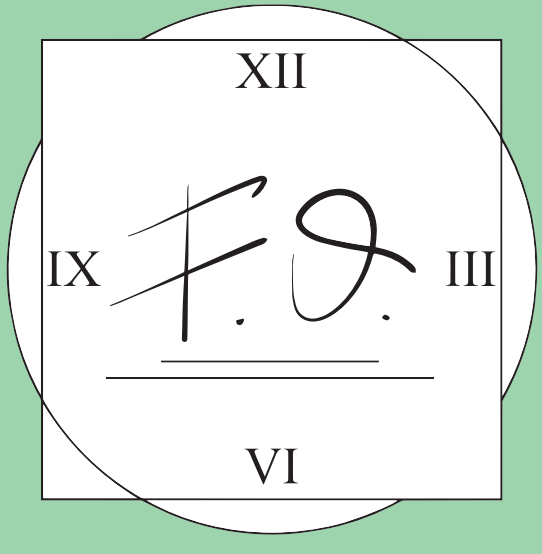
Fuori Orario

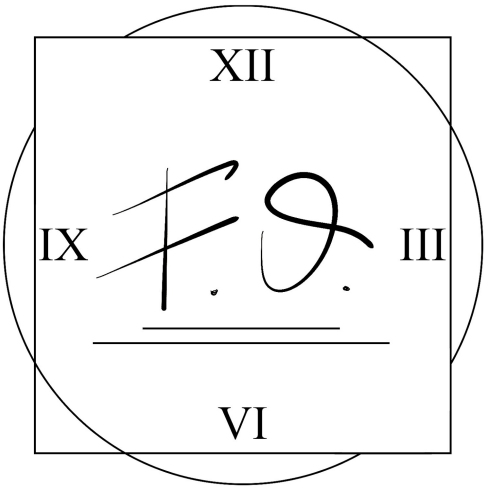
Resoconti di una giornata di seminari
3 Maggio 2023



Dipartimento di Matematica - Università degli Studi di Milano

FUORI ORARIO

Resoconti di una giornata di seminari
3 Maggio 2023



Fuori Orario
Resoconti di una giornata di seminari

3 Maggio 2023
Dipartimento di Matematica
Università degli Studi di Milano

Versione definitiva: Settembre 2023

Impaginazione in LaTeX a cura di Riccardo Formenti
Progetto grafico di Riccardo Formenti

INDICE

INTRODUZIONE

Caro lettore o cara lettrice,
non so se tu sia una matricola a cui è capitato questo libretto per sbaglio in mano, se semplicemente aspettavi di fare la collezione e aggiungere la copertina verde al tuo scaffale o se sei un appassionato di matematica, in qualsiasi caso è bello che tu sia qui a leggere questa introduzione. Vorrei dirti qualcosa di più riguardo a quello che c'è dietro a questo libretto che forse non sai, nella speranza che anche tu prenda a cuore questa iniziativa. Fuori Orario è una giornata di seminari di matematica, interamente organizzata da studenti, con l'idea di poter approfondire qualche aspetto interessante di questa disciplina, mettendosi in gioco e condividendo una passione comune. Ci sono seminari di diverse difficoltà e riguardanti svariati argomenti, per questo il libro che hai in mano non è pensato per essere letto interamente, ma spero che tu possa trovare quel titolo che attiri il tuo interesse.

Mi presento anche io, penso sia bello che tu possa sapere un minimo chi e cosa ci siano dietro a queste pagine. Innanzitutto vorrei ricordare che c'è un gruppo di organizzatori dove ognuno ha dato il suo contributo importante per la realizzazione della giornata. Io sono solo il portavoce di questa organizzazione. Ho iniziato a fare parte di Fuori Orario nel 2020, ero al secondo anno di matematica e non conoscevo bene questa iniziativa. In quel periodo eravamo tutti chiusi in casa per il lockdown e un giorno mi arriva un messaggio da uno studente (il fondatore di Fuori Orario) che mi propone di dare una mano nell'organizzazione. Ho risposto di sì: mi sembrava un bel modo di essere vicina all'università, dove mi sono sempre sentita come in una grande famiglia, in un periodo in cui eravamo tutti distanti. Non sapevo bene a cosa andassi incontro, ma sono rimasta affascinata dalla passione e voglia di poter dare il proprio contributo degli altri organizzatori. Ricordo ancora quando abbiamo aperto la pagina Instagram da zero e ora mi fa effetto realizzare che questa è già la settima edizione di Fuori Orario. Il terzo anno in cui aiutavo nell'organizzazione ho anche partecipato come speaker: per quanto mi riguarda è stata una grande conquista parlare in pubblico. Non avrei mai pensato di partecipare come speaker, ma devo dire che sono stata molto soddisfatta e contenta di aver potuto approfondire un argomento di mio interesse e superare la paura del pubblico. Anche per questo aspetto Fuori Orario offre una grande possibilità di mettersi in gioco. Purtroppo nell'edizione di quest'anno non ho potuto partecipare per sovrapposizioni con esami, anche se avrei parlato volentieri. . . ho quindi solo dato il mio contributo nel portare avanti questa iniziativa, che di anno in anno cerca nuova gente che se ne prenda cura.

Sono molto contenta di come proceda, quest'anno ci sono stati 10 seminari e una buona affluenza di pubblico: si conferma ancora un'iniziativa molto apprezzata. Mi auguro possa continuare così e che ci sia un continuo ricambio di persone che si impegnano per Fuori Orario. Ogni anno è bello vedere che c'è qualcuno di nuovo che si aggiunge, perché ogni aiuto si rivela essenziale per il lavoro che c'è da fare: tra il bando, la gestione dei social e del sito, delle grafiche e delle stampe dei libretti. Ringrazio quindi l'organizzazione di questa settima edizione composta da Francesco Pagliarin, Elisa Severgnini, Alessandro Scapin, Margherita Recchia, Kevin Arana, Lorenzo Paveri, Matteo de Berardinis, Matteo Anoffo, Riccardo Formenti, Sara Donè, Simone Puleio, Francesco Zuccon e Federica Franzi. Magari non sembra, ma ogni piccolo contributo è stato fondamentale. Ringrazio anche i professori che ci sostengono ogni anno in questa iniziativa. Penso di dare ancora il mio contributo l'anno prossimo almeno per i primi mesi, ma ancora per me è tutto da definire, in ogni caso, come è sempre successo, anche noi organizzatori siamo solo di passaggio ed è bello vedere che Fuori Orario va avanti sempre e comunque, anche col cambio generazionale, proprio come io ho preso il testimone da chi mi ha preceduto e sono pronta a cederlo a chi vorrà.

Ora che ci conosci un po' meglio, ti lascio sfogliare questo libretto!
Buona lettura!

Elisa Pedrini

GIOCANDO...CON LE MATRICI

MATTEO ANOFFO

Esistono moltissimi tipi di problemi matematici, dalle classiche equazioni diofantee agli intricatissimi problemi di geometria, passando per i problemi di combinatoria e per quelli in cui si deve trovare la stategia vincente per un gioco. Quest'ultimo sarà il tipo di problema che intendo affrontare, e risolvere, assieme a voi oggi. In particolare, determineremo chi riuscirà a vincere, adottando la giusta strategia, un interessante gioco con le matrici a coefficenti in campi finiti.

1. Il gioco

Due ragazzi, Tommaso e Valerio, si sfidano al seguente gioco. A turno, alternandosi, scelgono un elemento dal gruppo G delle matrici invertibili $n \times n$ a coefficenti in $\mathbb{Z}/p\mathbb{Z}$ con le seguenti regole:

- un giocatore non può scegliere un elemento che è già stato scelto in precedenza
- un giocatore può scegliere solo un elemento che commuta con tutti gli elementi precedenti
- un giocatore che non può scegliere nessun elemento perde la partita

Comincia Valerio. Ora ci chiediamo, in base a n e a p, chi ha una strategia vincente?

2. Prime osservazioni

Notiamo innanzitutto che G è finito. Indichiamo poi con S l'insieme degli elementi scelti dai due giocatori e con $Z(S)$ il centralizzante di s, ovverosia il sottogruppo di G formato da tutti gli elementi che commutano con ogni elemento di S. Chiaramente si ha prima dell'inizio del gioco $S = \varnothing$ e $Z(S) = G$. A ogni turno a S viene aggiunto un elemento e $Z(S)$ viene sostituito da un suo sottogruppo. Dal teorema di Lagrange segue banalmente che se $Z(S)$ ha ordine dispari in un qualche turno allora resterà di ordine dispari per tutti i turni a venirel, garantendo la vittoria di Valerio. Analogamente se S contiene un elemento di ordine pari allora lo conterrà per ogni turno garantendo la parità dell'ordine di $Z(S)$ e la seguente vittoria di Tommaso.

3. Il caso $p > 2$

Iniziamo dal caso più semplice e mostriamo che in questo caso, e solo in questo scopriremo più avanti, Tommaso ha una strategia vincente. Poiché $p > 2$ esiste la matrice -1. Pertanto qualsiasi matrice B scelga Valerio inizialmente a Tommaso è sufficente rispondere con la matrice $-B$. Infatti -1 ha ordine pari e questo assicura la parità dell'ordine di una matrice tra B e $-B$.

4. Il caso $p = 2$

La questione è ora più complessa. Cominciamo ricordando alcuni risultati.

PROPOSIZIONE 1. *Ogni campo finito ha ordine potenza di un primo*

Dimostrazione. Ogni campo finito ha caratteristica un numero primo p e contiene quindi un sottocampo isomorfo a $\mathbb{Z}/p\mathbb{Z}$ su cui è spazio vettoriale □

COROLLARIO 2. *Ogni estensione finita di $\mathbb{Z}/2\mathbb{Z}$ ha ordine una potenza di 2*

COROLLARIO 3. *Il gruppo moltiplicativo di un'estensione di $\mathbb{Z}/2\mathbb{Z}$ ha ordine dispari*

DEFINIZIONE 4. Una estensione di campi K/k si dice $algebrica$ se è ottenuta aggiungendo a k radici di polinomi a coefficenti in k

DEFINIZIONE 5. Sia k campo. k di dice $algebricamente\ chiuso$ se contiene tutte le radici di ogni polinomio a coefficenti in k

DEFINIZIONE 6. Sia k campo. Definiamo $chiusura\ algebrica$ di k la più piccola estensione algebrica K algebricamente chiusa contenente k

TEOREMA 7. *Per ogni k campo esiste K chiusura algebrica di k*

Dimostrazione. Sia $\mathscr{X} = \{\ x_f \mid f$ è un polinomio in $k[X]\ \}$ un insieme di variabili indicizzate su tutti i polinomi a una variabile a coefficenti in k e consideriamo l'anello dei polinomi in più variabili $k[\mathscr{X}]$. Sia ora $I = \{\ f(x_f)\ \}$. Mostriamo che $I \neq 1$. Ragioniamo per assurdo e supponiamo che esistano dei polinomi g per cui si ha $\sum g f(x_f) = 1$. Consideriamo ora la proiezione sul quoziente
$\pi : k[\mathscr{X}] \longrightarrow k[\mathscr{X}]/I$ e notiamo che ogni elemento della somma scritta sopra è mandato in 0 e quindi avremmo $\pi(1) = 0$, assurdo. Dal lemma di Krull sappiamo quindi che esiste m ideale massimale contenente I. Chiamiamo ora $k = k_0$, $k[\mathscr{X}]/m = k_1$ e creiamo una successione di campi procedendo in modo ricorsivo come appena fatto. Notiamo che ogni campo contiene tutte le radici dei polinomi del campo precedente. Sia ora $K = \cup k_n$ e notiamo che K è un campo algebricamente chiuso. Ora per trovare la chiusura algebrica di k è sufficente prendere il sottocampo di K formato da tutti gli elementi algebrici su k, ovvero quelli ottenuti da estensioni algebriche □

NB: la chiusura algebrica di un campo non è unica, ma lo è a meno di isomorfismo

DEFINIZIONE 8. Un polinomio p di grado n nell'anello dei polinomi $k[X]$ si dice $separabile$ se contiene esattamente n radici distinte in una chiusura algebrica K di k

PROPOSIZIONE 9. *Sia p polinomio a coefficenti in un campo k. p ha una radice doppia se e solo se p e il suo polinomio derivato hanno una radice in comune*

Dimostrazione. Segue banalmente dalla regola di Leibniz □

COROLLARIO 10. *Il polinomio $x^n + 1$ è separabile in $(\mathbb{Z}/2\mathbb{Z})[X]$ se n è dispari. Il polinomio $x^n + x + 1$ è separabile in $(\mathbb{Z}/2\mathbb{Z})[X]$ se n è pari*

Dai risultati precedenti sappiamo quindi che per ogni n esiste un polinomio separabile di grado n in $(\mathbb{Z}/2\mathbb{Z})[X]$. Useremo questo fatto per trovare una matrice di ordine dispari in G. Sia infatti $p(x) = x^n + P_{n-1}x^{n-1} + ... + P_1x + P_0$ un polinomio separabile e sia $g = \begin{pmatrix} 0 & 0 & ... & 0 & -P_0 \\ 1 & 0 & ... & 0 & -P_1 \\ . & . & . & . & -P_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & ... & 1 & -Pn-1 \end{pmatrix}$

In particolare $det(g) = (-1)^n P_0 \neq 0$ e il suo polinomio caratteristico è proprio p. Abbiamo quindi che g appartiene a G e poiché il polinomio p è separabile abbiamo che in una chiusura algebrica di $\mathbb{Z}/2\mathbb{Z}$ g è diagonalizzabile. Inoltre l'ordine di g è uguale all'ordine della sua diagonalizzata che è dispari. Infatti l'ordine di ques'ultima è uguale al prodotto dell'ordine dei suoi elementi, ma poiché l'ordine di ognuno di essi è dispari, essendo elementi di un gruppo di ordine dispari, il loro prodotto sarà dispari.
Notiamo infine che ogni matrice che commuta con g deve essere diagonalizzabile, rendendo quindi $Z(S)$ un sottogruppo di ordine dispari e garantendo a Valerio vittoria certa (ammesso che riesca a trovare la strategia giusta).

UNA SFIDA IMPOSSIBILE

LORENZO PAVERI

1. Introduzione

Negli anni novanta del XIX secolo Samuel Loyd annunciò che avrebbe dato un premio di 1000 dollari (30000 attuali) a chi avesse risolto il gioco del 15.

Il gioco del 15 si presenta come una griglia quadrata 4×4 sulla quale sono presenti 15 tessere mobili in questo preciso ordine:

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

Si può muovere una tessera solamente in orizzontale o in verticale spostandola sulla casella vuota e lo scopo del gioco consiste nell'ottenere la configurazione in cui tutti i numeri sono in ordine crescente (nel senso delle righe).
In altre parole l'obiettivo è riuscire a scambiare la casella 14 con la casella 15 per ottenere la seguente configurazione:

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

I soldi che Loyd promise quando lanciò la sfida erano al sicuro perché, come scopriremo, il gioco non può essere risolto!
Per dimostrarlo useremo alcune proprietà dei gruppi di permutazione S_n.

2. Il gioco del 15 e i gruppi di permutazione

Per prima cosa dobbiamo "matematizzare" il gioco. Per farlo numeriamo i pezzi del puzzle da 1 a 16 (dove il sedicesimo pezzo è la casella vuota).
Numeriamo anche le posizioni nel puzzle da 1 a 16 a partire dalla casella in alto a sinistra, muovendosi lungo la riga e poi spostandosi alla riga sotto partendo sempre

dalla casella a sinistra.

Definiamo ora i concetti di configurazione e di mossa:

DEFINIZIONE 1. Una configurazione C e una mossa M sono delle particolare permutazioni in S_{16} definite in questo modo:
$C(i)$:= posizione che occupa il pezzo i;
$M(i)$:= nuova posizione del pezzo che occupava la posizione i.

OSSERVAZIONE 2. Abbiamo definito due insiemi distinti entrambi di cardinalità 16:

$$\{Pezzi\} \qquad \{Posizioni\}$$

In particolare una configurazione è una biiezione dall'insieme dei pezzi nell'insieme delle posizioni, mentre una mossa è una biiezione dall'insieme delle posizioni in se stesso, quindi:

$$\{Pezzi\} \xrightarrow{C} \{Posizioni\}$$

$$\{Posizioni\} \xrightarrow{M} \{Posizioni\}$$

ESEMPIO 3. La configurazione a cui si vuole giungere è descritta dall'identità. Quella di partenza dallo scambio (14 15).

ESEMPIO 4. La configurazione C associata alla seguente disposizione dei pezzi:

8	12	2	5
11	1	6	
7	14	10	15
9	4	3	13

è descritta dall'elemento:

$$(1\ 6\ 7\ 9\ 13\ 16\ 8)(2\ 3\ 15\ 12)(4\ 14\ 10\ 11\ 5) \in S_{16}.$$

ESEMPIO 5. Le mosse utilizzate per scambiare le tessere in questi due casi:

8	12	2	5
11	1	6	
7	14	10	15
9	4	3	13

$\longmapsto$

8	12	5	6
11	1	2	
7	14	10	15
9	4	3	13

6	4	1	9
10	12	8	
2	13	7	5
14	3	15	11

$\longmapsto$

6	4	9	8
10	12	1	
2	13	7	5
14	3	15	11

sono le stesse e sono descritte dalla seguente permutazione:

$$M = (3\ 7\ 4) \in S_{16}.$$

OSSERVAZIONE 6. Se in S_{16} eseguiamo il prodotto tra una configurazione C e una mossa M (in questo ordine: prima C e poi M) si ottiene nuovamente una configurazione. In particolare il prodotto tra C e M verrà indicato con CM.
D'altra parte se si esegue il prodotto tra due mosse si ottiene ancora una mossa.

ESEMPIO 7. Consideriamo nuovamente l'esempio 5.
La configurazione di partenza è stata descritta nell'esempio 4:

$$C = (1\ 6\ 7\ 9\ 13\ 16\ 8)(2\ 3\ 15\ 12)(4\ 14\ 10\ 11\ 5).$$

La mossa che consideriamo è: $M = (3\ 7\ 4)$.
Il prodotto

$$CM = (1\ 6\ 7\ 9\ 13\ 16\ 8)(2\ 3\ 15\ 12)(4\ 14\ 10\ 11\ 5)(3\ 7\ 4)$$

è il 16-ciclo:

$$(1\ 6\ 4\ 14\ 10\ 11\ 5\ 3\ 15\ 12\ 2\ 7\ 9\ 13\ 16\ 8).$$

Richiamiamo ora un risultato della teoria dei gruppi di permutazione che ci sarà utile nella dimostrazione del perché è impossibile risolvere il gioco del 15.

TEOREMA 8. *(1) Ogni elemento di S_n può essere scritto come prodotto finito di scambi.*
(2) Un elemento di S_n non può essere scritto sia come prodotto di un numero pari di scambi che come prodotto di un numero dispari di scambi.

Siamo ora in grado di dimostrare perché la sfida che lanciò Loyd non aveva soluzione.

TEOREMA 9. *Risulta impossibile risolvere il gioco del 15.*

Dimostrazione. Osserviamo che la configurazione corrispondente alla situazione di partenza è lo scambio $C := (14\ 15)$, d'altra parte la configurazione a cui si vorrebbe arrivare è l'identità id.
Supponiamo per assurdo che il gioco si possa risolvere. Esisteranno allora una serie di mosse $M_1, M_2, ..., M_k$ tali che $id = CM_1M_2 \cdots M_k$.

Dal momento che ogni mossa possibile consiste nello scambiare il pezzo 16 (ossia lo spazio vuoto) con un altro pezzo, allora possiamo vedere tutte le mosse M_r come degli scambi (ij) e quindi scrivere: $id = C\tau_1...\tau_s$ con i τ_j scambi.

Osserviamo che la tessera vuota occupa la stessa posizione sia nella situazione di partenza che in quella di arrivo. Quindi ogni volta che la tessera vuota viene spostata in una delle direzioni possibili dovrà successivamente essere spostata nella direzione opposta. Di conseguenza il numero di scambi τ_i che compare nell'uguaglianza appena scritta deve essere pari.

Otteniamo così una contraddizione perché l'elemento che compare a sinistra nell'uguaglianza è una permutazione pari, mentre l'elemento a destra è una permutazione dispari e ogni elemento di S_n può essere o una permutazione pari o dispari e non entrambi. □

UNA REGOLARIZZAZIONE AL GIORNO TOGLIE IL MEDICO DI TORNO

URMILA BOSISIO

La *regolarizzazione* di una soluzione è strettamente legata ai cosiddetti "problemi inversi"; al contrario dei "problemi diretti", in cui a partire da una causa si osserva un effetto, si vuole qui ricostruire una causa *a partire* da un effetto. Matematicamente, supponendo di avere a disposizione un dato $y \in Y$, si vuole ricostruire $x \in X$ dove le due variabili sono legate tra loro da una $f : X \to Y$ tramite $y = f(x)$, con X, Y due spazi metrici (o, più spesso, $y = f(x) + \eta$ con η rumore additivo, che introduce degli errori di misurazione). I due esempi qui discussi sono la ricostruzione di segnali, in cui y è un segnale rumoroso e x quello originale, con f che modellizza la perturbazione, e la diagnostica di immagini, dove y rappresenta delle misurazioni fisiche mentre x una certa quantità che si vuole indagare, ed f è la relazione fisico-matematica che le lega.

Naturalmente, ci si potrebbe chiedere come mai sia necessario impiegare tecniche di regolarizzazione per risolvere un problema inverso. Vi sono, infatti, delle difficoltà nella semplice inversione di f: in primo luogo, non sempre f risulta invertibile; in secondo luogo, spesso il problema non è *ben posto*. Il problema "cerco $x \in X$ tale che $f(x) = y$ dato $y \in Y$" si dice *ben posto* (secondo Hadamard) se la soluzione x esiste, è unica e dipende con continuità dal dato y. Effettivamente, non sempre la soluzione x risulta unica e spesso f non è continua; questo fa sì che, avendo generalmente dei dati non esatti (dovuti alle imprecisioni delle misurazioni fisiche), invertire ingenuamente la funzione risulta in una ricostruzione decisamente poco accurata della soluzione effettiva.

Gli approcci di regolarizzazione quindi sono necessari per ottenere dei risultati soddisfacenti; i metodi più semplici sono quello ai minimi quadrati, che prevede la ricerca di x tale da minimizzare $\|f(x) - y\|_Y^2$ (dove con $\|\cdot\|_Y$ si intende la norma indotta dalla metrica su Y), e quello di minimizzazione della norma, che mira invece alla ricerca di x che soddisfi $f(x) = y$ e che minimizzi la norma $\|x\|_X$ (dove con $\|\cdot\|_X$ si intende la norma indotta dalla metrica su X). All'atto pratico, spesso si utilizza una commistione dei due; per questo, si vedano gli esempi successivi.

Image deblurring

Per prima cosa si ricorda che, data f funzione regolare, a partire da una funzione k detta *nucleo* si può definire l'operatore di convoluzione K come:

$$K(f) = g \text{ con } g(x) = (f * k)(x) = \int_{-\infty}^{+\infty} f(y)k(x-y)dy$$

Ciò risulta utile nella modellizzazione del problema di *image deblurring* (cioè, di ricostruzione di un'immagine a partire da una sfocata): se f dà il valore di grigio di ogni pixel dell'immagine originale (dove 0=nero, 255=bianco), K modellizza la sfocatura e quindi g è l'immagine sfocata che si ha a disposizione. In particolare, intuitivamente si ha che il nucleo k darà la sfocatura rispetto a un *impulso*, che in questo campo può essere rappresentato da un'immagine tutta nera con un solo pixel bianco. Su questa base, quindi, si può definire la **PSF** (o **Point Spread Function**), che contiene l'informazione di come viene modificato un singolo punto dalla sfocatura. Si veda in figura 1 un esempio pratico del suo funzionamento. Matematicamente, si

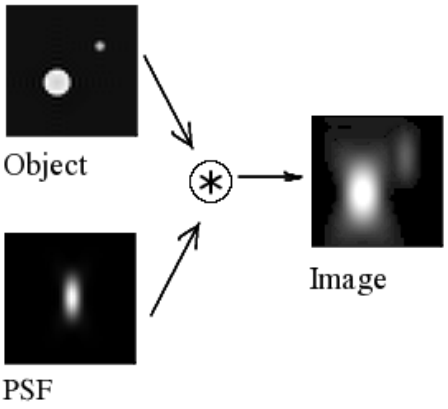


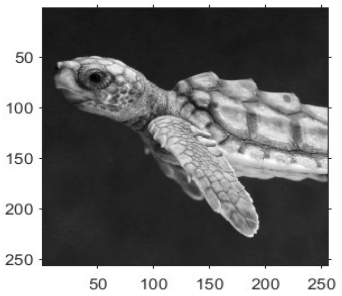
FIGURA 1

hanno F, G, $\eta \in \mathbb{R}^{n\times n}$ l'immagine originale in bianco e nero a n pixel, l'immagine sfocata e il rumore additivo supposto noto rispettivamente.

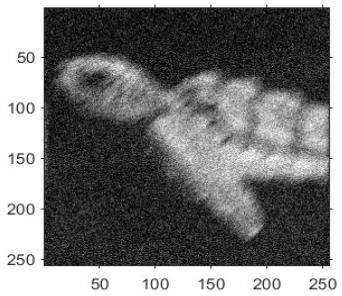
Come esempio, si consideri quindi l'immagine di una tartaruga e la sua controparte sfocata e rumorosa (si veda 2).

Per prima cosa, bisogna costruire un'ipotesi della PSF; guardando le due immagini, è verosimile che si comporti come in figura 3.

A questo punto, è possibile ricostruire in modo "naive", semplicemente deconvolvendo a partire dall'ipotesi di PSF appena fatta (si vedano i risultati in 4). È facile notare come, senza rumore, il metodo ricostruisce l'immagine in modo fedele mentre nel momento in cui si aggiunge rumore il risultato non risulta particolarmente soddisfacente. Ciò era prevedibile; già nell'introduzione infatti si era sottolineato come in generale l'inversione degli operatori senza regolarizzazione comportino, nel

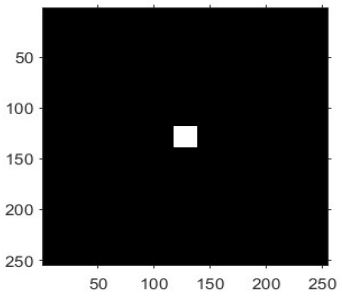


(A) Immagine "vera"

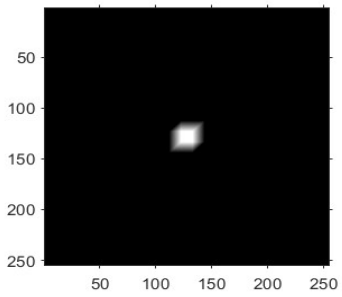


(B) Immagine sfocata

FIGURA 2



(A) Punto



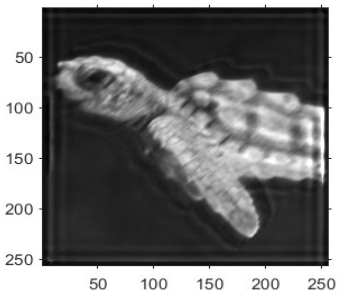
(B) Ipotesi della PSF

FIGURA 3

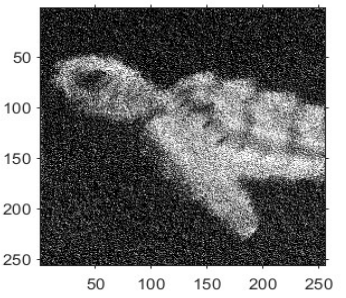
caso di dati poco precisi (e, dunque, presenza di rumore), delle soluzioni poco precise.

Si può quindi utilizzare un approccio di regolarizzazione che è l'algoritmo di Weiner, basato sull'approccio ai minimi quadrati. In particolare, utilizza una stima dell'ordine di grandezza del rumore per tentare di eliminarlo il quanto più possibile; in dipendenza da tale stima, darà evidentemente dei risultati diversi. In figura 5 si possono vedere le ricostruzioni con l'algoritmo di Weiner a seconda del "guess" sulla potenza del rumore.

Si noti che comunque, a causa del rumore, non è possibile avere una ricostruzione "perfetta"; la situazione risulta, comunque, nettamente migliore.

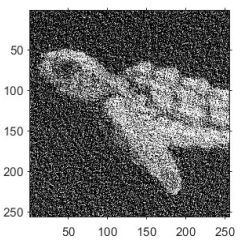


(A) Senza rumore $\eta = 0$

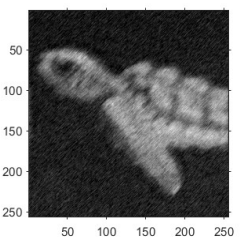


(B) Rumore η di ordine $1e - 2$

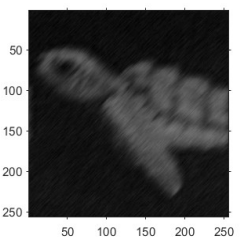
FIGURA 4



(A) Sottostimando potenza



(B) Potenza originale



(C) Sovrastimando potenza

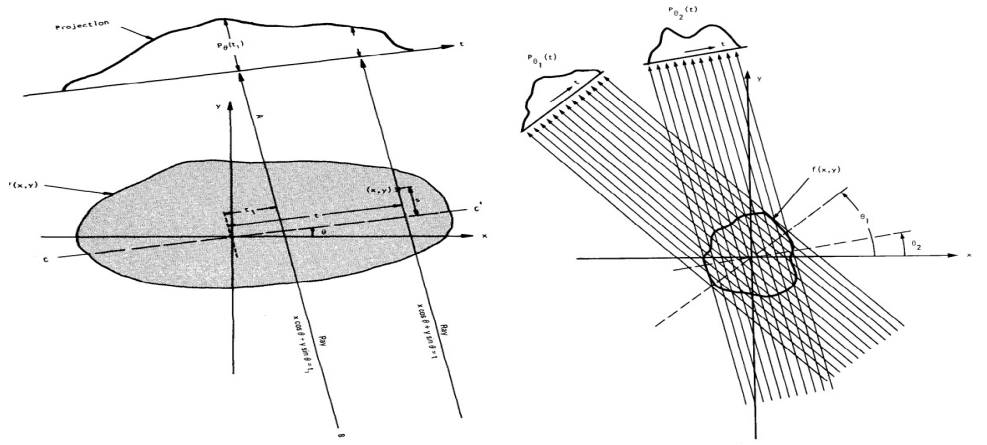
FIGURA 5

Tomografia

Si considera ora il secondo esempio, che è quello della tomografia: a partire da delle proiezioni, si vuole ricostruire l'interno di un corpo. Nella pratica, tale procedura è utilizzata per le TAC in 3D con raggi del tipo *fan-beam*, cioè a ventaglio; per semplicità, si discuterà il caso bidimensionale con raggi paralleli.

Matematicamente, si può modellizzare la situazione in questo modo: si sceglie un angolo θ e si mandano dei raggi paralleli, misurando poi la quantità di materia incontrata da ogni raggio per avere una cosiddetta "proiezione" ad angolo θ (si veda figura 6). Per una più completa spiegazione tramite trasformate di Radon e Fourier e per gli algoritmi di ricostruzione (senza regolarizzazione) si vedano i testi in bibliografia.

In particolare, è possibile discretizzare il tutto riportandosi ad un sistema $Af = g$, dove f è la quantità di interesse, g sarà il cosiddetto *sinogramma* (ricavato dalle proiezioni, quindi il dato) e A una matrice che dipende dal numero di angoli e dal numero di raggi paralleli per ogni angolo.



(A) Proiezione lungo un angolo θ (B) Iterazione del procedimento a diversi angoli

FIGURA 6

Per farsi un'idea di cosa significhi ricostruire tramite proiezioni, si veda figura 7 con l'immagine campione *phantom*, in cui si utilizzano dapprima 180 proiezioni e poi solo 9, dove si ricostruisce invertendo la matrice A in assenza di rumore.

Si nota facilmente che con più proiezioni si ottiene un risultato migliore; d'altronde, è preferibile evitare di utilizzare troppi raggi in quanto, dal punto di vista più concreto, ciò risulta in un'esposizione più lunga da parte del paziente e dell'operatore sanitario ai raggi: cosa che, ovviamente, si vuole evitare. In generale, quindi, si vorranno mettere a punto metodi che funzionino bene anche con un numero limitato di proiezioni.

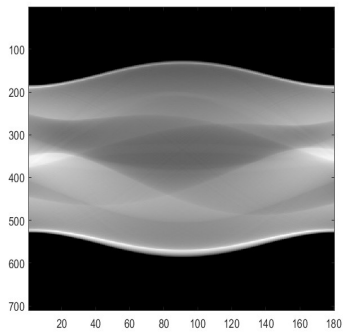
Ora, come ormai si dovrebbe aver capito, aggiungendo rumore è necessario l'utilizzo di metodi di regolarizzazione. In particolare, in quest'ambito è possibile considerare i due metodi più semplici: il metodo della SVD troncata e il metodo di Tikhonov.

Il primo si basa sulla scrittura di una matrice M come $U\Sigma V^T$ con $\Sigma = \text{diag}\{\sigma_1, ..., \sigma_n\}$ quadrata e diagonale, con σ_i i cosiddetti *valori singolari*, ed U, V ortogonali. Si può infatti dimostrare che in tal caso una soluzione per il metodo dei minimi quadrati è data da:

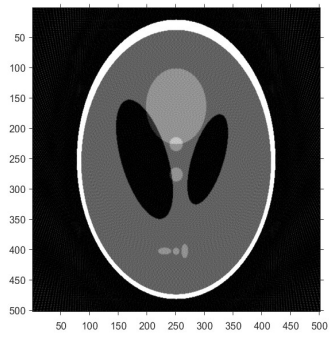
$$x = \sum_{i=1}^{n} \frac{u_i^T y}{\sigma_i} v_i$$

che implica, in particolare, che i valori singolari più piccoli sono responsabili della dipendenza sensibile dai dati: si applica quindi un troncamento della SVD ad un qualche indice $r \in \{1, ..., n\}$ tale per cui $\sigma_1 \approx ... \approx \sigma_r \gg \sigma_{r+1}, ..., \sigma_n$, in modo da avere quindi:

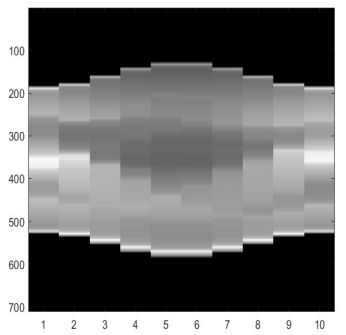
$$x = \sum_{i=1}^{r} \frac{u_i^T y}{\sigma_i} v_i$$



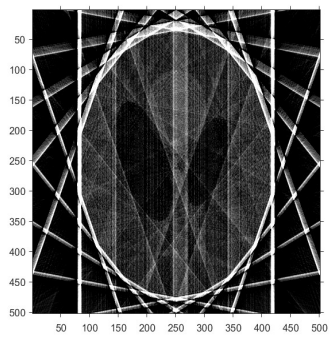
(A) Sinogramma per 180 proiezioni



(B) Relativa ricostruzione



(C) Sinogramma per 9 proiezioni



(D) Relativa ricostruzione

FIGURA 7

cioè, rigorosamente, si applica a σ_i un filtro di troncamento:

$$\omega_\alpha(s^2) = \begin{cases} 1 & s^2 > \alpha \\ 0 & s^2 \leqslant \alpha \end{cases}$$

dove α è un parametro che deve essere scelto empiricamente.

Il secondo metodo, invece, si basa sulla risoluzione del problema ai minimi quadrati modificato, in cui si penalizza una soluzione troppo grande in modulo; si cerca x tale che minimizzi $\|y - Mx\|_Y^2 + \alpha^2\|x\|_X^2$. Si può dimostrare che, in questo caso, ciò equivale ad applicare un filtro di troncamento del tipo:

$$\omega_\alpha(s^2) = \frac{s^2}{s^2 + \alpha}$$

Anche qui α è un parametro che deve essere scelto. In particolare, il buon funzionamento dei due metodi dipende proprio dalla scelta di tale α: per α troppo piccolo, si rischia di non filtrare abbastanza e di ritrovare ancora del rumore amplificato mentre per α troppo grande si rischia di filtrare anche delle componenti della sorgente.

Si può testare quanto detto finora con un esempio suggestivo della misurazione di una noce, utilizzando dati reali provenienti da un esperimento condotto dal dipartimento di Matematica e Statistica e da quello di Fisica dell'Università di Helsinki, in collaborazione con l'*Institut für Mathematik* di Berlino.

In figura 8 si possono vedere i risultati con l'approccio naive di inversione, mentre in figura i risultati utilizzando il filtro di Tikhonov con parametro $\alpha = 10$ fissato.

Si noti che tale parametro è stato scelto a seguito di varie prove numeriche; non è infatti possibile stabilire teoricamente a priori un valore che possa andare bene in ogni caso. Inoltre, esistono anche diversi metodi di regolarizzazione che ad esempio utilizzano diverse norme oltre a quella euclidea qui utilizzata, come ad esempio la norma $\mathscr{L}^1$ o la seminorma data dalla variazione totale.

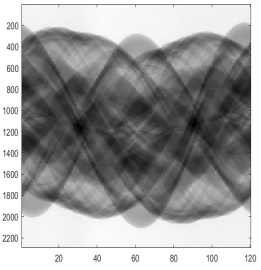
Bibliografia

Curtis R. Vogel (1987), *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics

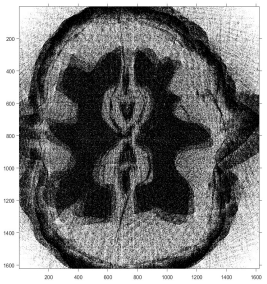
Avinash C. Kak, Malcolm Slaney (1987) *Principles of Computerized Tomographic Imaging*, Classics in Applied Mathematics

Jennifer L. Mueller, Samuli Siltanen (2012), *Linear and Nonlinear Inverse Problems with Practical Applications*, Computational Science and Engineering

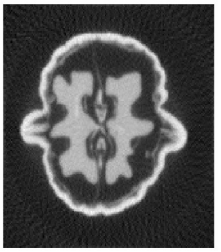
Keijo Hämäläinen, Lauri Harhanen, Aki Kallonen, Antti Kujanpää, Esa Niemi, Samuli Siltanen (2015), *Tomographic X-ray data of a walnut*, arXiv



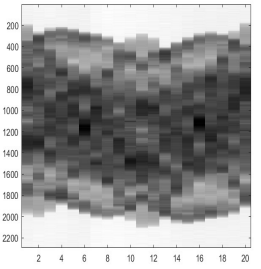
(A) Sinogramma



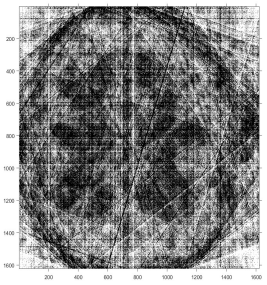
(B) Ricostruzione naive



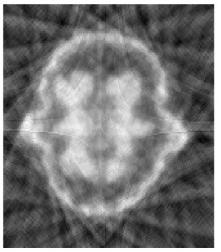
(C) Tikhonov con 120 proiezioni



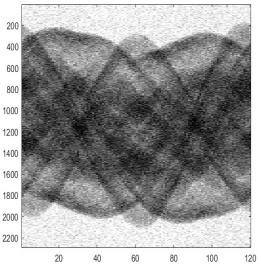
(D) Sinogramma con 20 proiezioni



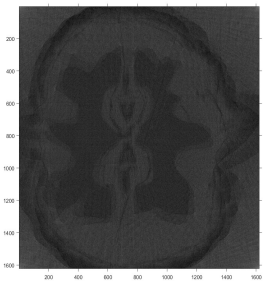
(E) Ricostruzione naive con 20 proiezioni



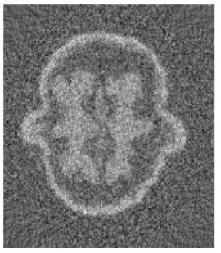
(F) Tikhonov con 20 proiezioni



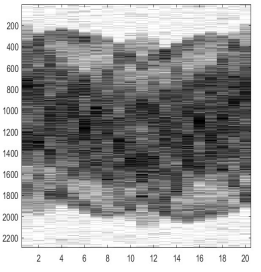
(G) Sinogramma (con rumore)



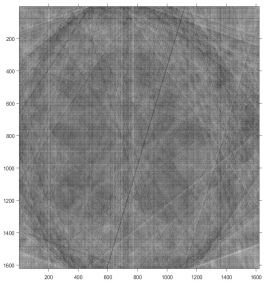
(H) Ricostruzione naive (con rumore)



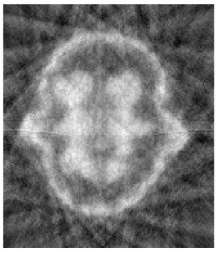
(I) Tikhonov (con rumore) con 120 proiezioni



(J) Sinogramma (con rumore) con 20 proiezioni



(K) Ricostruzione naive (con rumore) con 20 proiezioni



(L) Tikhonov (con rumore) con 20 proiezioni

FIGURA 8

TUTTE LE STRADE PORTANO A ROMA, MA QUALE LA PIU' VELOCE?

SCAPIN ALESSANDRO

1. Introduzione

Immaginiamo di dover prendere un aereo per spostarci da una città ad un'altra. É lecito pensare che alle compagnie aeree, per risparmiare, convenga far seguire ai propri voli le traiettorie più veloci ed efficienti, quindi potremmo aspettarci che, rappresentando la Terra su una cartina, tali traiettorie siano dei segmenti. Infatti, come ben sappiamo, dati due punti nel piano (o nello spazio) il cammino più breve per arrivare dal primo al secondo è il segmento che li congiunge. Osserviamo ora questa immagine che mostra le tratte di alcuni voli di Etihad:

Ciò che osserviamo è che in realtà le traiettorie sono curve; perché questo avviene? In questo seminario cercheremo di capire se la nostra idea intuitiva è corretta oppure se hanno ragione le compagnie aeree. Spoiler: hanno ragione loro. Cercheremo allora di discutere matematicamente quali siano le strade più brevi per unire due punti su una superficie.

2. Superfici e curve

Ricordiamo la definizione di superficie:

DEFINIZIONE 1. Sia $V \subseteq \mathbb{R}^2$ un aperto omeomorfo a $B_1(0)$. Definiamo *superficie elementare* una funzione

$$P : V \to \mathbb{R}^3 \quad (u, v) \mapsto P(u, v) = (x(u, v), y(u, v), z(u, v))$$

che soddisfi le seguenti proprietà:

(1) P è di classe C^∞;
(2) P è iniettiva;
(3) per ogni (u, v) la Jacobiana di P valutata in (u, v) ha rango massimo.

Senza soffermarci sui dettagli tecnici di questa definizione, di fatto una superficie è una varietà bidimensionale (ossia un "oggetto" bidimensionale) in $\mathbb{R}^3$ dotato di una parametrizzazione.

Denotiamo con $\mathscr{M} = P(V)$; nel linguaggio parlato, anche in questo seminario, ci riferiremo quindi alla superficie come al suo sostegno (l'insieme immagine) più che alla sua parametrizzazione.
Generalmente si studiano le *superfici regolari*, ma il nostro studio sarà locale e le cose si complicherebbero un po' considerando diverse parametrizzazioni. Niente di trascendentale, ma esulerebbe dal nostro obiettivo.

Una *curva* in $\mathbb{R}^2$ (in generale in $\mathbb{R}^n$) è una mappa $x : I \subset \mathbb{R} \to \mathbb{R}^2$ della forma

$$x(t) = (x_1(t), x_2(t)),$$

dove I è un intervallo e $x_i : I \to \mathbb{R}$ sono funzioni lisce. A partire da essa definiamo il concetto di *curva* in $\mathscr{M}$ come una mappa $\alpha : I \subset \mathbb{R} \to \mathscr{M}$ della forma

$$\alpha(t) = P(x_1(t), x_2(t)).$$

In altre parole una curva in $\mathscr{M}$ è ottenuta applicando la parametrizzazione della superficie ad una curva in $\mathbb{R}^2$.
Durante il seminario assumeremo che le curve considerate siano *regolari*, ovvero con vettore tangente mai nullo.

Dati $p, q \in \mathscr{M}$ denotiamo con

$$\mathscr{M}_{p,q} = \{\, \alpha : [0, 1] \to \mathscr{M} \ : \ \alpha(t) = P(x_1(t), x_2(t)) \text{ regolari } \ : \ \alpha(0) = p, \ \alpha(1) = q \,\}$$

l'insieme delle curve regolari che uniscono p, q. Mettere come dominio di definizione proprio $[0, 1]$ è una scelta puramente convenzionale e solo per comodità.

Indichiamo con L la funzione

$$L(x, y) = \sum_{i,j=1}^{2} g_{ij}(x) y_i y_j,$$

dove g_{ij} sono i coefficienti della prima forma fondamentale, ovvero il prodotto scalare che si definisce sullo spazio tangente come restrizione del prodotto scalare di $\mathbb{R}^3$. Per chi non avesse familiarità con questo concetto non si preoccupi, basti fidarsi del fatto che su ogni superficie si può costruire una matrice (di entrate le g_{ij} di cui sopra) che permette di "dare senso geometrico" alla definizione di lunghezza che daremo a breve.

In particolare possiamo calcolare L sui punti di una curva, introducendo anche una dipendenza dal tempo

$$L(t, x, x') = \sum_{i,j=1}^{2} g_{ij}\left(x_1(t), x_2(t)\right) x_i'(t) x_j'(t).$$

DEFINIZIONE 2. La l*unghezza di una curva* è definita come

$$\begin{aligned}\mathscr{L}(\alpha) &= \int_0^1 \left(\sum_{i,j=1}^{2} g_{ij}\left(x_1(t), x_2(t)\right) x_i'(t) x_j'(t) \right)^{\frac{1}{2}} dt \\ &= \int_0^1 \sqrt{L(x, x'(t))}\, dt.\end{aligned}$$

Il problema che ci siamo posti inizialmente, riscritto in questo linguaggio, è il seguente: dati $p, q \in \mathscr{M}$, cerchiamo

$$\min_{\alpha \in \mathscr{M}_{p,q}} \mathscr{L}(\alpha),$$

ossia quella curva (se esiste) che unisca i punti p, q e che abbia la lunghezza minore possibile.
In analisi ci hanno sempre insegnato che, sotto le dovute ipotesi, quando vogliamo trovare il minimo o il massimo di una funzione dobbiamo cercare i punti in cui la derivata si annulla. Qui non possiamo fare questa cosa, semplicemente perché il dominio di $\mathscr{L}$ è uno spazio di funzioni (in particolare di curve) e non sappiamo cosa significhi derivare.

Entra allora in gioco il calcolo delle variazioni: non daremo le definizioni nel contesto più generale possibile, ma solo per quello che serve a noi.

DEFINIZIONE 3. Sia $I \subset \mathbb{R}$ un aperto contenente 0; definiamo *variazione liscia* di $\alpha \in \mathscr{M}_{p,q}$ una funzione liscia

$$\Gamma : I \times [0, 1] \to \mathscr{M}$$

tale che

$$\Gamma(0, s) = \Gamma_0(s) = \alpha(s)$$

e che $\forall\, a \in I$

$$\Gamma(a, 0) = p \quad \Gamma(a, 1) = q.$$

In altre parole, una variazione di una curva è una famiglia (dipendente da un parametro) di curve che condividono il punto di partenza e di arrivo e che sono "simili" tra loro.
Possiamo allora ricondurci ad una derivata in senso classico, infatti possiamo considerare la funzione

$$t \mapsto \Gamma_t \mapsto \mathscr{L}(\Gamma_t)$$

che va da $I \subset \mathbb{R}$ a $\mathbb{R}$ e calcolarne la derivata.

DEFINIZIONE 4. Diciamo che α è un *punto critico* per $\mathscr{L}$ se

$$\frac{d}{dt}(\mathscr{L} \circ \Gamma_t) = 0$$

per ogni variazione liscia Γ di α.

PROPOSIZIONE 5. *Se α è un minimo (o un massimo) per $\mathscr{L}$ allora α è un punto critico.*

Prima di andare avanti definiamo un nuovo funzionale, che ci sarà utile a breve:

$$E(\alpha) = \int_0^1 L(t, \alpha, \alpha')\, dt.$$

TEOREMA 1. Una curva α è un punto critico per E oppure per $\mathscr{L}$ se e solo se soddisfa

$$\frac{d}{dt}\left(\frac{\partial L}{\partial x_k'}(t, \alpha, \alpha')\right) - \frac{\partial L}{\partial x_k}(t, \alpha, \alpha') = 0 \tag{1}$$

oppure

$$\frac{d}{dt}\left(\frac{\frac{\partial L}{\partial x_k'}}{2\sqrt{L}}\right) - \frac{\frac{\partial L}{\partial x_k}}{2\sqrt{L}} = 0. \tag{2}$$

L'equazione (2) è detta *equazione di Eulero Lagrange* per il funzionale $\mathscr{L}$.

OSSERVAZIONE 6. Queste equazioni non sono differenziali; conosciamo già la funzione L e stiamo cercando gli zeri di questa combinazione di derivate.

Ad esempio, in fisica matematica 1, le traiettorie di un punto materiale soggetto ad un'energia potenziale V, ad una fissata energia E, sono quelle che minimizzano il funzionale

$$\int_\gamma \sqrt{E - V}.$$

Tali traiettorie vengono appunto trovate risolvendo le equazioni di Eulero-Lagrange associate a questo funzionale.

3. Equazione delle geodetiche

Per trovare le curve che minimizzano il funzionale lunghezza dobbiamo quindi risolvere le equazioni di Eulero-Lagrange (2); risolviamo (1), poi scopriremo che questo basta.

Divertiamoci a fare derivate:

$$L(t, x, x') = \sum_{i,j=1}^{2} g_{ij}(x(t)) x_i'(t) x_j'(t)$$

quindi

$$\frac{\partial L}{\partial x_k} = \sum_{i,j=1}^{2} g_{ij,k} x_i' x_j'.$$

La scrittura $g_{ij,k}$ sta a significare $\frac{\partial g_{ij}}{\partial x^k}$ e la continueremo ad usare.
Invece

$$\begin{aligned}\frac{\partial L}{\partial x'_k} &= \sum_{i,j=1}^{2} g_{ij}\left(\frac{\partial}{\partial x'_k}\left(x'_i\right)x'_j + x'_i\frac{\partial}{\partial x'_k}x'_j\right)\\ &= \sum_{i,j=1}^{2} g_{ij}\left(\delta_{ik}x'_j + x'_i\delta_{jk}\right)\\ &= \sum_{j=1}^{2} g_{kj}x'_j + \sum_{i=1}^{2} g_{ik}x'_i\\ &= 2\sum_{j=1}^{2} g_{kj}x'_j.\end{aligned}$$

Allora

$$\frac{d}{dt}\frac{\partial L}{\partial x'_k} = 2\left(\sum_{j,l=1}^{2} g_{kj,l}x'_l x'_j + \sum_{j=1}^{2} g_{jk}x''_j\right).$$

Dato che a parte k sono tutti indici muti, possiamo cambiargli nomi e invertirli; inoltre ricordando che g è simmetrica la precedente equazione può essere riscritta come

$$\frac{d}{dt}\frac{\partial L}{\partial x'_k} = \sum_{i,j=1}^{2} g_{jk,i}x'_i x'_j + \sum_{i,j=1}^{2} g_{ik,j}x'_i x'_j + 2\sum_{j=1}^{2} g_{jk}x''_j.$$

Mettendo tutto insieme e dividendo per 2, otteniamo

$$\sum_{i=1}^{2} g_{ik}x''_i + \frac{1}{2}\sum_{i,j=1}^{2}\left(g_{jk,i} + g_{ik,j} - g_{ij,k}\right)x'_i x'_j = 0. \tag{3}$$

Denotiamo con g^{ij} la matrice inversa di g_{ij}; moltiplichiamo allora per g^{km} e sommiamo su k.

$$\sum_{i,k=1}^{2} g^{km}g_{ik}x''_i + \frac{1}{2}\sum_{i,j,k=1}^{2} g^{km}\left(g_{jk,i} + g_{ik,j} - g_{ij,k}\right)x'_i x'_j = 0.$$

Però per definizione di matrice inversa, al variare di k il prodotto $g^{mk}g_{ki}$ è il prodotto della m-esima riga e della i-esima colonna, quindi vale δ_{mi}. Possiamo infine riscrivere l'ultima equazione come

$$x''_m + \frac{1}{2}\sum_{i,j,k=1}^{2} g^{km}\left(g_{jk,i} + g_{ik,j} - g_{ij,k}\right)x'_i x'_j = 0. \tag{4}$$

La quantità

$$\Gamma^m_{ij} = \frac{1}{2}\sum_k g^{km}(g_{jk,i} + g_{ik,j} - g_{ij,k})$$

viene detta *simbolo di Christoffel*.

L'equazione (4) è detta *equazione delle geodetiche*; le curve che la soddisfano sono dette *geodetiche*.
Un calcolo diretto mostra che se x è un punto critico per E allora $L(x, x')$ è costante (non dipende da t), x è punto critico per $\mathscr{L}$.

Per quanto dimostrato e fidandoci dell'ultima affermazione, ogni punto critico del funzionale lunghezza è una geodetica. Per essere precisi però non possiamo dire che "le geodetiche sono le curve che minimizzano la distanza tra due punti", infatti l'equazione (4) (e quindi essere una geodetica) è solo una condizione *necessaria* per minimizzare la distanza. Così come esistono funzioni scalari in cui la derivata si annulla in punti che non sono estremanti (il punto 0 per x^3), esistono geodetiche che non sono globalmente minimizzanti. Inoltre, bisognerebbe spendere due parole sull'indipendenza di tutto il discorso dalla scelta del dominio di definizione.
Insomma, per rispondere quindi in modo completo al problema iniziale servirebbe un (ampio) discorso ulteriore, che non è possibile fare in due parole. Per farla breve, come detto *non* è vero che ogni geodetica minimizza la distanza tra due suoi punti, però se i punti sono presi sufficientemente vicini tra loro si può dimostrare che invece la geodetica risulta minimizzante. Per approfondire meglio gli argomenti trattati consigliamo di leggere testi di geometria differenziale.

ALLA RICERCA DELLA COMPATTEZZA PERDUTA

FRANCESCO PAGLIARIN

1. Introduzione

Solitamente, quando vengono introdotti gli spazi L^p si parla e si discute subito della loro completezza.
In questo seminario invece vogliamo occuparci di un'altra proprietà fondamentale, ovvero della compattezza, spesso trascurata, esponendo una caratterizzazione che è un analogo del teorema di Ascoli-Arzelà per gli spazi L^p.
Iniziamo con alcune definizioni e osservazioni preliminari, proseguiamo con la dimostrazione del teorema, poi con un suo miglioramento e infine concludiamo con un curioso controesempio.

Definizione 1 ($L^p(\mathbb{R}^n)$). Lo spazio $L^p(\mathbb{R}^n)$ con $p \in [1, +\infty)$, è lo spazio vettoriale delle funzioni misurabili definite su tutto $\mathbb{R}^n$ a valori in $\mathbb{R}$ aventi la potenza p-esima del modulo integrabile, cioè:

$$L^p(\mathbb{R}^n) := \{f : \mathbb{R}^n \to \mathbb{R} \quad \text{tale che} \quad \|f\|_p := \Big(\int_{\mathbb{R}^n} |f|^p \, dx\Big)^{\frac{1}{p}} < +\infty\}$$

Teorema 2. *Quozientando lo spazio sopra definito tramite la relazione di equivalenza di "uguaglianza quasi ovunque", $\|\cdot\|_p$ risulta essere una norma e lo spazio ottenuto, indicato ancora con $L^p(\mathbb{R}^n)$, è completo.*

Definizione 3 (ε-rete).
Sia (X, d) uno spazio metrico. Siano $E, A \subseteq X$. Diciamo che E è una *ε-rete* per A se per ogni $a \in A$ $\exists e \in E$ tale che $d(a, e) < \varepsilon$.

Definizione 4 (Totale limitatezza).
Sia (X, d) uno spazio metrico. Un sottoinsieme $A \subseteq X$ è detto *totalmente limitato* se $\forall \varepsilon > 0$ esiste una ε-rete finita E per A. Ovvero, A può essere ricoperto con un numero finito di bolle di diametro al più ϵ

Osservazione 5. Per spazi normati finito dimensionali questo è equivalente alla limitatezza (in quanto si ha che la bolla chiusa risulta compatta)

Elenchiamo ora alcuni risultati di topologia generale.

Teorema 6.

Sia (M, d) uno spazio metrico. Allora:

$$M \text{ è compatto} \iff M \text{ è completo e totalmente limitato}$$

Da questo teorema segue facilmente il seguente:

TEOREMA 7.
Sia (X, d) spazio metrico completo e M un suo sottoinsieme. Allora le seguenti affermazioni sono equivalenti:

(i) *M relativamente compatto (i.e. $\overline{M}$ è compatto)*
(ii) *M totalmente limitato*
(iii) *Per ogni successione $\{x_n\}_{n\in\mathbb{N}} \subset M$ esiste una sottosuccessione $\{x_{n_k}\}_{k\in\mathbb{N}}$ convergente (il cui limite non è necessariemente in M).*

Siamo ora pronti per iniziare a dimostrare i fatti principali sulla compattezza degli spazi L^p. Prima però, un lemma:

LEMMA. *Sia X spazio metrico. Assumo che per ogni $\varepsilon > 0$ esiste: una $\delta = \delta(\varepsilon) > 0$, uno spazio metrico $W = W(\varepsilon)$ e una mappa $\Phi = \Phi(\varepsilon)$ tale che $\Phi : X \to W$ e tale per cui $\Phi(X)$ sia totalmente limitato e $\forall x, y \in X$ tale che $d(\Phi(x), \Phi(y)) < \delta$ allora $d(x, y) < \varepsilon$.*
Allora X è totalmente limitato

Dimostrazione. Per ogni ε definisco δ come da ipotesi. Essendo $\Phi(X)$ totalmente limitato, esiste una δ-rete $\{\Phi(x_1), ..., \Phi_(x_n)\}$ per $\Phi(X)$. Allora segue dalle ipotesi che $\{x_1, ..., x_n\}$ è una ε-rete per X. Ecco che quindi X è totalmente limitato. □

OSSERVAZIONE 8. Questo lemma, oltre a giocare un ruolo fondamentale per dimostrare il teorema di Kolmogorov-Riesz-Frechet, può essere usato anche per dimostrare il classico teorema di Ascoli-Arzelà.

2. Il teorema

Siamo ora in grado di dimostrare il principale risultato del seminario:

TEOREMA 9 (Kolmogorov–Riesz–Frechet). *Sia $p \in [1, +\infty)$. Un sottoinsieme F di $L^p(\mathbb{R}^n)$ è totalmente limitato se e solo se:*

(i) *F è limitato*
(ii) *$\forall \varepsilon > 0$ $\exists R$ tale che, $\forall f \in F$:*

$$\int_{|x|>R} |f(x)|^p \, \mathrm{d}x < \varepsilon^p$$

"Uniforme integrabilità" per gli spazi L_p
(iii) *$\forall \varepsilon > 0$ $\exists \rho \in \mathbb{R}$ tale che, $\forall f \in F$ e $y \in \mathbb{R}^n$ con $|y| < \rho$*

$$\int_{\mathbb{R}^n} |f(x + y) - f(x)|^p \, \mathrm{d}x < \varepsilon^p$$

"Uniforme continuità" per gli spazi L_p

Dimostrazione. "⇐"
Sia $\varepsilon > 0$ fissato. Sia Q un cubo aperto centrato nell'origine tale per cui $|y| < \frac{\rho}{2}$ per ogni $y \in Q$. Siano $Q_1, ..., Q_N$ traslati di Q mutuamente disgiunti tali per cui $\overline{\cup_i Q_i} \supset B_R(0)$ (indicando così la bolla di raggio R centrata nell'origine).
Sia P la mappa proiezione da $L^p(\mathbb{R}^n)$ a Y, $P : L^p(\mathbb{R}^n) \to Y$, con Y definito come lo spazio lineare generato dalle funzioni caratteristiche dei cubi Q_i.
Ovvero:

$$Y := span\{\chi_{Q_j} \quad \text{con } j = 1, ..., N\}$$

In particolare P è definita nel modo seguente:

$$Pf(x) = \begin{cases} \frac{1}{|Q_j|}\int_{Q_j} f(y)\,\mathrm{d}y & x \in Q_j \quad j = 1, .., N \\ 0 & \text{altrimenti} \end{cases}$$

Allora, per costruzione otteniamo:

$$\|f - Pf\|_p^p \leqslant \varepsilon^p + \sum_{i=1}^{N}\int_{Q_j} |f(x) - Pf(x)|^p\,\mathrm{d}x = \varepsilon^p + \sum_{i=1}^{N}\int_{Q_j} \Big|\frac{1}{|Q_j|}\int_{Q_j} f(x) - f(z)\,\mathrm{d}z\Big|^p\,\mathrm{d}x$$

Ora, usiamo la diseguaglianza di Jensen, Tonelli, essendo l'integrando positivo, e un cambio di variabile di integrazione, dove osserviamo che $y = x - z \in 2Q$ quando $x, z \in Q_j$, e che quindi $|y| < \rho$.

$$\begin{aligned}
\|f - Pf\|_p^p &\leqslant \varepsilon^p + \sum_{i=1}^{N}\int_{Q_j} \frac{1}{|Q_j|}\int_{Q_j} |f(x) - f(z)|^p\,\mathrm{d}z\mathrm{d}x \\
&\leqslant \varepsilon^p + \sum_{i=1}^{N}\int_{Q_j} \frac{1}{|Q_j|}\int_{2Q} |f(x) - f(x+y)|^p\,\mathrm{d}y\mathrm{d}x \\
&\leqslant \varepsilon^p + \frac{1}{|Q|}\int_{2Q}\int_{\mathbb{R}^n} |f(x) - f(x+y)|^p\,\mathrm{d}x\mathrm{d}y \\
&\leqslant \varepsilon^p + \frac{1}{|Q|}\int_{2Q} \varepsilon^p\,\mathrm{d}y = (2^n + 1)\varepsilon^p
\end{aligned}$$

Usando quindi la disuguaglianza triangolare si ottiene che $\|f\|_p \leqslant (2^n + 1)^{\frac{1}{p}}\varepsilon + \|Pf\|_p$.
Dalla linearità di P, se $f, g \in F$ e $\|Pf - Pg\|_p < \varepsilon$, allora $\|f - g\|_p < ((2^n + 1)^{\frac{1}{p}} + 1)\varepsilon$.
Inoltre, essendo P limitato ($\|P\| = 1$ essendo una proiezione), e essendo F limitato da (i), anche $P(F)$ è limitato. Essendo Y finito dimensionale, $P(F)$ è anche totalmente limitato. Allora F è totalmente limitato, per il lemma.
"⇒"
La limitatezza (i) di F è ovvia, essendo F totalmente limitato per ipotesi.
Sia ora $\varepsilon > 0$ fissato. Sia $\{g_1, ..., g_m\}$ la ε-rete di F associata. Sia R tale per cui valga:

$$\int_{|x|>R} |g_j(x)|^p\,\mathrm{d}x < \varepsilon^p, \quad j = 1, .., m$$

Ciò segue facilmente dal teorema di convergenza dominata, essendo le $g_j \in L^p(\mathbb{R}^n)$
Se $f \in B_\varepsilon(g_j)$ allora $\|f - g_j\|_p < \varepsilon$ e:

$$\Big(\int_{|x|>R} |f(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} \leqslant \Big(\int_{|x|>R} |f(x) - g_j(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} + \Big(\int_{|x|>R} |g_j(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} \leqslant$$
$$\leqslant \|f - g_j\|_p + \Big(\int_{|x|>R} |g_j(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} < 2\varepsilon$$

In cui ho usato la diseguaglianza di Minkovski per gli spazi L^p. Ho dimostrato così (ii).
Ora, uso ancora la ε-rete $\{g_1, ..., g_m\}$Questa volta però, wlog suppongo che $g_j \in C_c^\infty(\mathbb{R}^n)$. Ciò è possibile dalla densità di queste ultime in $L^p(\mathbb{R}^n)$.
Scegliamo ora ρ tale che:

$$\int_{\mathbb{R}^n} |g_j(x+y) - g_j(x)|^p \,\mathrm{d}x < \varepsilon^p, \quad j = 1, .., m \quad \text{per } |y| < \rho$$

Che è possibile essendo le g_j uniformemente continue e a supporto compatto.
Allora, per $f \in B_\varepsilon(g_j)$:

$$\Big(\int_{\mathbb{R}^n} |f(x+y) - f(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} \leqslant \Big(\int_{\mathbb{R}^n} |f(x+y) - g_j(x+y)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} + \Big(\int_{\mathbb{R}^n} |g_j(x+y) - g_j(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} +$$
$$+ \Big(\int_{\mathbb{R}^n} |g_j(x) - f(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} < 3\varepsilon$$

Dimostrando così anche (iii). □

3. Un miglioramento

In realtà, si dimostra che la condizione (i) di limitatezza risulta essere ridondante. Vale infatti questo miglioramento del teorema, dovuto a Sudakov:

Teorema 10 (Kolmogorov-Riesz-Frechet-Sudakov). *Sia $p \in [1, +\infty)$. Un sottoinsieme F di $L^p(\mathbb{R}^n)$ è totalmente limitato se e solo se:*

(ii) $\forall \varepsilon > 0\ \exists R$ *tale che,* $\forall f \in F$*:*

$$\int_{|x|>R} |f(x)|^p \,\mathrm{d}x < \varepsilon^p$$

(iii) $\forall \varepsilon > 0\ \exists \rho \in \mathbb{R}$ *tale che,* $\forall f \in F$ *e* $y \in \mathbb{R}^n$ *con* $|y| < \rho$

$$\int_{\mathbb{R}^n} |f(x+y) - f(x)|^p \,\mathrm{d}x < \varepsilon^p$$

Per quanto appena dimostrato, dobbiamo solo mostrare che le condizioni (ii) e (iii) implicano la (i), cioè la limitatezza. Definisco $T_y f(x) := f(x - y)$. Fisso $\varepsilon = 1$ e siano ρ e R le rispettive quantità definite in (ii) e (iii). Vale allora la seguente

maggiorazione:

$$\begin{aligned}\|f\chi_{B_R(z)}\|_p \leqslant &\|(T_y f - f)\chi_{B_R(z)}\|_p + \|f\chi_{B_R(z-y)}\|_p \leqslant \\ &\leqslant \|T_y f - f\|_p + \|f\chi_{B_R(z-y)}\|_p \leqslant 1 + \|f\chi_{B_R(z-y)}\|_p\end{aligned}$$

In cui $y \in \mathbb{R}^n$ e $|y| < \rho$. Per induzione si verifica che:

$$\|f\chi_{B_R(0)}\|_p \leqslant N + \|f\chi_{B_R(-Ny)}\|_p$$

Scegliendo così N tale che $N|y| > 2R$, si osserva che $B_R(-Ny) \cap B_R(0) = \varnothing$. Ecco che così:

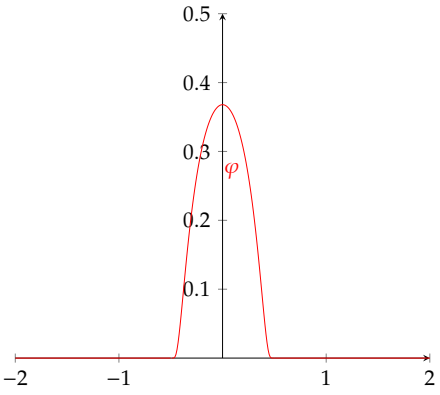
$$\|f\|_p \leqslant \|f\chi_{B_R(0)}\|_p + \|f\chi_{\mathbb{R}^n \setminus B_R(0)}\|_p \leqslant N + 2$$

Uniformemente in f.

4. Un curioso controesempio

Vogliamo ora illustrare un controesempio per il teorema, costruito tramite opportune traslazioni di una bump-function.
Sia $\varphi \in C_c(\mathbb{R})$ così definito:

$$\varphi(x) := \begin{cases} e^{-\frac{1}{1-|2x|^2}} & \text{per } |x| < \frac{1}{2} \\ 0 & \text{altrimenti} \end{cases}$$

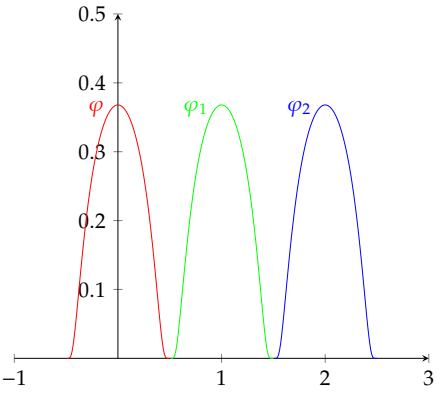


Inoltre, definisco:

$$\mathscr{F} = \bigcup_{n \in \mathbb{N}} \varphi_n$$

Con

$$\varphi_n(x) := \varphi(x - n) = \tau_n \varphi(x)$$



Osservo che:

$$\|\varphi\|_p = \|\varphi_n\|_p = c > 0 \quad \forall n \in \mathbb{N}$$

Quindi (i) del teorema è soddisfatta.
Inoltre:

$$\|\tau_y \varphi_n - \varphi_n\|_p = \|\tau_y \varphi - \varphi\|_p$$

Quindi, per convergenza dominata su φ, per ogni $\varepsilon > 0$ esiste un $\rho \in \mathbb{R}$ per cui per ogni $y \in \mathbb{R}^n$ tali che $|y| < \rho$ vale:

$$\|\tau_y \varphi_n - \varphi_n\|_p = \|\tau_y \varphi - \varphi\|_p < \varepsilon$$

Quindi la (iii) del teorema è soddisfatta.
Purtroppo però la (ii) non vale, quindi $\mathscr{F}$ non è relativamente compatto in L^p con $p \in [1, +\infty)$.
Inoltre, è un insieme chiuso poichè tutte e sole le successioni che convergono sono quelle definitivamente uguali ad un certo elemento dell'insieme (essendo tutti elementi $2c$ separati), quindi la chiusura è lui stesso. E' poi limitato per quanto osservato prima. Ecco che quindi ho appena trovato un sottoinsieme di $L^p(\mathbb{R}^n)$ chiuso, limitato ma non compatto. Questo mostra come negli spazi infinito dimensionali, il teorema di Heine-Borel sia generalmente falso.

Bibliografia

[H1] Hanche-Olsen, Holden, *The Kolmogorov-Riesz compactness theorem*
[H2] Hanche-Olsen, Holden, *An improvement of the Kolmogorov–Riesz compactness theorem*
[B] Brezis, Haim, *Functional analysis, Sobolev spaces and partial differential equations*

QUANTO SONO COMPLICATI I LINGUAGGI UMANI?

ALBERTO PANERAI

Per rispondere al quesito proviamo a formalizzare il "linguaggio" in ambito matematico/informatico e vediamo se e come i linguaggi umani (qualsiasi dall'Inglese all'Italiano passando dall'Esperanto e dal Latino) possono essere "catturati" da queste costruzioni.
Premessa: Ci concentreremo sul lato sintattico e non semantico quindi la celebre frase di Chomsky *"colorless green ideas sleep furiously"*[8] andrà benissimo.
Nonostante la vaghezza ("fuzziness") dei linguaggi umani qualcosa si può dire! Esistono dei risultati che riguardano l'inglese e lo svizzero-tedesco che daranno degli indizi per indagare questa tematica ancora abbastanza aperta.

Introduzione ai Linguaggi Formali e Automi

Prendiamo un insieme finito (non vuoto!) di simboli Σ altrimenti detto "Alfabeto". Ora considero una concatenazione finita di elementi di Σ e la chiamo "parola". Le parole su Σ le indico con Σ^* [1]. Un Linguaggio $\mathscr{L}$ è un insieme di parole definite sull'Alfabeto Σ.

Osservazione 1. $\mathscr{L} \subseteq \Sigma^*$

Dati due linguaggi A e B oltre alle classiche operazioni insiemistiche $\cup$ e $\cap$ abbiamo

Definizione 2. $A^c := \Sigma^* \setminus A$,
$A \cdot B := \{x \in \Sigma^* | x = yw, y \in A, w \in B\}$,
$A^0 := \{\epsilon\}$, $A^n := A \cdot A^{n-1}$,
$A^* := \bigcup_{n=0}^{+\infty} A^n$, $A^+ := \bigcup_{n=1}^{+\infty} A^n$.

Una prima distinzione tra i linguaggi è data dai Linguaggi Ricorsivi e i Linguaggi Ricorsivamente Enumerabili o R.E.
Per i nostri scopi daremo una definizione "intuitiva":

Definizione 3. Il Linguaggio $\mathscr{L}$ è Ricorsivamente Enumerabile se esiste una *procedura* che stampa tutti gli elementi di $\mathscr{L}$.

[1] con la concatenazione e la parola vuota ϵ elemento neutro Σ^* è un monoide "libero su Σ".

Con *procedura* si pensi a un programma che una volta eseguito fa scorrere sul monitor del computer tutti gli elementi di $\mathscr{L}$ [2]. Per Linguaggio $\mathscr{L}$ Ricorsivo chiederemo invece:

DEFINIZIONE 4. Il Linguaggio $\mathscr{L}$ è Ricorsivo se sia $\mathscr{L}$ sia $\mathscr{L}^c$ sono Ricorsivamente Enumerabili.

Equivalentemente si può definire $\mathscr{L}$ ricorsivo se il *problema di appartenenza* a $\mathscr{L}$ è *decidibile*, ovvero se esiste un algoritmo che su ogni input $x \in \Sigma^\star$ restituisce 1 se $x \in \mathscr{L}$ e 0 altrimenti.
Un esempio comune di linguaggio ricorsivamente enumerabile ma non ricorsivo è $HALT := \{(i, x) \mid \text{programma } i \text{ che si arresta su input } x\}$.

OSSERVAZIONE 5. I Linguaggi Naturali sono Ricorsivamente Enumerabili.

Noi assumeremo che ciò sia vero[3] e cercheremo di capire meglio in che "zona" possano trovarsi all'interno dei linguaggi R.E.
Ci chiediamo innanzitutto se i linguaggi in questione siano finiti. La risposta è *No* perché esistono frasi lunghe quanto voglio.. ad esempio:

ESEMPIO 6. I THINK I THINK I THINK...[4]
Io penso che io penso che io penso...

Ora ci serve una *Grammatica Formale* per generare/individuare un determinato linguaggio.
Supponiamo che per i linguaggi in esame ci siano chiare evidenze di[5] frasi o enunciati ben formati come *The man took the book* e mal formati come *The took man the book*.

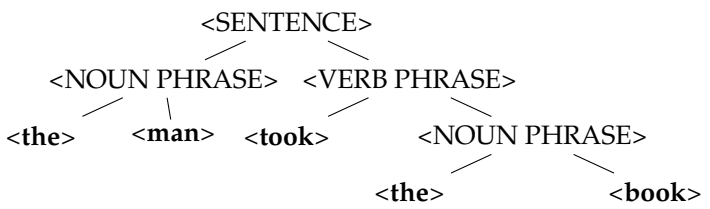


FIGURA 1. esempio di *Parse Tree*.

Diamo un'occhiata all'Albero Sintattico o *Parse Tree* in figura. La radice e i nodi che non sono foglie sono etichettati da *categorie sintattiche* o *meta-variabili* del linguaggio

[2] notare che non si richiede che il programma si arresti!

[3] l'intuizione meccanicista di esseri umani equiparabili a 'computer biologici' con risorse e tempo in quantità finita rende valida questa assunzione. In *"The Vastness of Language"*[5] Paul Postal e Terry Langendoen argomentano invece che i Linguaggi Naturali *Non* siano R.E.

[4] corrisponde a (I THINK)* il cui significato sarà chiarito meglio in seguito.

[5] o meglio: *Convergenza di opinioni su...*

come "Sentence", "Noun Phrase" e "Verb Phrase".[6] Le parole costituenti la frase sono tutte sulle foglie. [7]

OSSERVAZIONE 7. Purtroppo non abbiamo un insieme ben definito di regole per riconoscere o *parserizzare* tutti gli enunciati in un determinato linguaggio umano come l'Inglese o l'Italiano..

Abbiamo usato le seguenti regole per il *parsing*:

- <SENTENCE> → <NOUN PHRASE><VERB PHRASE>
- <NOUN PHRASE> → <**the**><**man**>
- <VERB PHRASE> → <**took**><NOUN PHRASE>
- <NOUN PHRASE> → <**the**><**book**>

Nulla ci vieta di pensare a queste regole come *produzioni* o regole di *riscrittura*.[8] Quello che vogliamo fare ora è andare *a ritroso*. Cioè non limitarci a testare la correttezza grammaticale degli enunciati bensì *generare* enunciati.
Esploriamo un altro esempio *Logico-Linguistico*:

ESEMPIO 8. I Teoremi del Sistema "MIU": Th_{MIU}.[9]

Il *Sistema Formale "MIU"* è costituito da un alfabeto di tre lettere M,I e U da cui posso costruire stringhe o parole finite come "MU","UIM","MUUMUU",.. Inoltre è presente un unico *assioma* **MI** e le seguenti regole di *inferenza*:

(1) Se una parola finisce in "I" posso aggiungere una "U" alla fine.
(2) Se ho una parola della forma "Mx" posso ottenere "Mxx" dove x rappresenta una qualsiasi stringa di M,I e U.
(3) Se esiste un'occorrenza di "III" in una parola, posso sostituirla con "U".
(4) Se è presente "UU" all'interno di una parola, esso può essere cancellato.

A partire quindi dall'assioma **MI** posso *generare teoremi* applicando le regole assegnate come in figura.[10]

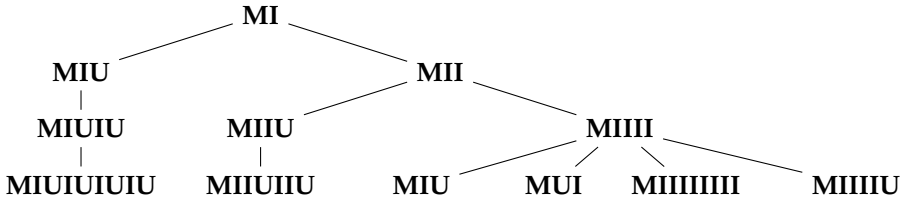


FIGURA 2. Albero di derivazione dei teoremi di MIU.

[6]pensate in italiano per esempio alle proposizioni oggettive, soggettive, etc..

[7]Di solito per l'analisi "grammaticale" di una frase "inizio" dalle foglie e "risalgo" verso la radice. Qui è l'esatto contrario.

[8]nel caso in esame posso "produrre" pochissime frasi come "the book took the book". Provate invece con "the little boy ran quickly"..

[9]ispirato al "MU Puzzle" in *Gödel,Escher,Bach*[4]..

[10]provate a ricavare "MU"..

Se ora penso a un linguaggio $\mathscr{L}$ costituito dai Teoremi di MIU, posso individuare delle regole di *produzione* che mi *generano* tale linguaggio.[11]

- S $\rightarrow$ **MI**
- α**I** $\rightarrow$ α**IU**
- **M**α $\rightarrow$ **M**$\alpha\alpha$
- α**III**β $\rightarrow$ α**U**β
- α**UU**β $\rightarrow$ $\alpha\beta$

Uso meta-variabili S come simbolo di *start* o inizio, α e β invece per indicare stringhe arbitrarie di M,I e U. Inizio producendo l'assioma **MI** e a questo punto scelgo di riscrivere **MI** in **MIU** oppure in **MII** e così via..

Definiamo ora una *Grammatica Formale* $\mathscr{G}$:

Definizione 9. $\mathscr{G} := < V, \Sigma, S, P >$.

V è l'insieme delle meta-variabili (saranno indicati con lettere maiuscole A,B,..) o dei simboli *non-terminali*, Σ è l'alfabeto o l'insieme dei simboli *terminali* (saranno indicati con lettere minuscole a,b,..), $S \in V$ è il simbolo *iniziale* o di *start* mentre P è un insieme di *produzioni* $(\alpha \rightarrow \beta)$ con $\alpha \in (V \cup \Sigma)^+$ e $\beta \in (V \cup \Sigma)^*$.
Nota Bene: Assumeremo che V, Σ e P siano *finiti* e $V \cap \Sigma = \varnothing$.
Introduciamo la relazione di *derivazione* o *riscrittura* in un passo:
Presi α,β,γ,$\delta \in (V \cup \Sigma)^*$..

Definizione 10. $x \equiv \gamma\alpha\delta \Rrightarrow_{\mathscr{G}} y \equiv \gamma\beta\delta$
se $(\alpha \rightarrow \beta) \in P$.

Mentre la derivazione in più passi:

Definizione 11. $x \Rrightarrow^*_{\mathscr{G}} y$
se $x \equiv x_0 \Rrightarrow_{\mathscr{G}} x_1 \Rrightarrow_{\mathscr{G}} ...x_{n-1} \Rrightarrow_{\mathscr{G}} x_n \equiv y$.

Allora il Linguaggio *generato* da $\mathscr{G}$ lo posso definire come tutte e sole le parole ottenute da una riscrittura in uno o più passi a partire dal simbolo iniziale:

Definizione 12. $\mathscr{L}(\mathscr{G}) := \{x \in \Sigma^* | S \Rrightarrow^*_{\mathscr{G}} x\}$.

[11] notare che le ultime due regole a differenza delle altre accorciano la lunghezza della parola..

Noam Chomsky[8] ha sviluppato la seguente "gerarchia" o "cipolla" di linguaggi generati da grammatiche formali a seconda delle tipologie di regole di produzione:

Tipo - nome	**Regole di Produzione**	**Esempi di linguaggi**
Tipo 0 - senza vincoli / recursively enumerable	$\alpha(\neq \epsilon) \rightarrow \beta$	**HALT** $\mathbf{Th}_{MIU}$
Tipo 1 - dipendente da contesto / context sensitive	$\alpha A \beta \rightarrow \alpha\ \gamma(\neq \epsilon)\ \beta$	$\{a^n b^n c^n\}_{n \geqslant 0}$ $\{a^m b^n c^m d^n\}_{m,n \geqslant 0}$
Tipo 2 - libero da contesto / context free	$A \rightarrow \gamma(\neq \epsilon)$	$\{0^n 1^n\}_{n \geqslant 0}$ $\{, (), (()), ((())), ..\}$
Tipo 3 - regolare / regular	$A \rightarrow aB, A \rightarrow a$	$\{a^n\}_{n \geqslant 0}$ $\{baa!, baaa!, baaaa!..\}$ linguaggi finiti

FIGURA 3. "Gerarchia" di Chomsky.

Nota Bene: Nei linguaggi *context sensitive* una richiesta equivalente per le regole di produzione è $\alpha \rightarrow \beta$ dove $|\alpha| \leqslant |\beta|$ cioè "ogni riscrittura dev'essere non decrescente per quanto riguarda la lunghezza delle parole".
Risulta chiaro dalle definizioni che ogni linguaggio regolare è *context free* e così via..
Ho una "cipolla" che raffigura le inclusioni tra i linguaggi.[12]

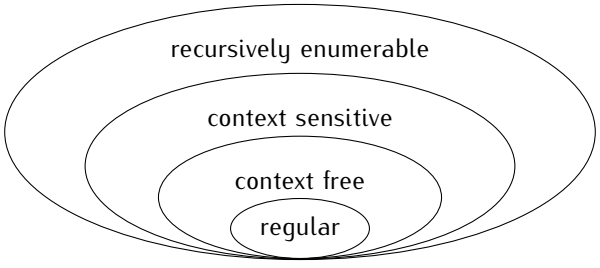


FIGURA 4. "cipolla" di linguaggi.

Concentriamo l'attenzione sui linguaggi regolari...
Essi sono particolarmente "maneggevoli" perché si può vedere[13] che formano la più piccola famiglia di linguaggi contenenti i linguaggi finiti e chiusa rispetto alle operazioni razionali (unione $\cup$,prodotto $\cdot$ e "star di Kleene" *).

[12]queste inclusioni sono proprie.. gli esempi sono stati scelti in modo tale che non siano contenuti nei livelli successivi.

[13]grazie al Teorema di Kleene.[3]

Riprendendo un esempio precedente di linguaggio infinito dato dalla ripetizione di una semplice frase: $\{(\textbf{I THINK})^n\}_{n\geqslant 0}$ pensandoci bene mi accorgo che è regolare. Inoltre si presta a essere "descritto" da un' espressione regolare o *regexp*[14] della forma /(**I THINK**)*/.
Le regexp sono importanti perché sono "riconosciute" da *Automi a stati finiti*, un modello di macchina a memoria finita che introduciamo con un esempio:

ESEMPIO 13. Il linguaggio delle pecore o *sheep-talk* := { baa!, baaa!, baaaa!...}

Noto subito che la regexp corrispondente è data da / **baaa***! /[15].
L'Automa che la riconosce è modellato dal seguente grafo diretto: Il grafo diretto[16] è

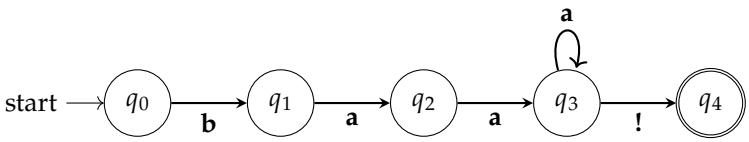


FIGURA 5. Automa a stati finiti che riconosce lo *sheep-talk*

costituito da un numero finito di vertici o nodi che rappresentano gli *stati* dell'automa collegati a sua volta da un set finito di archi orientati che danno le *transizioni* tra gli stati.

a	b	a	!	b		

FIGURA 6. INPUT : **"aba!b"** ,OUTPUT : **REJECT**

b	a	a	a	a	a	!

FIGURA 7. INPUT : **"baaaaa!"** ,OUTPUT : **ACCEPT**

La macchina funziona così: Dato in input una parola essa viene rappresentata da una stringa di caratteri su un nastro come in figura.
L'automa "legge"[17] la prima lettera e entra nello stato *iniziale*. Se la lettera coincide con l'etichetta di un arco uscente da q_0 (nel nostro caso abbiamo un unico arco **b**) passa allo stato indicato dalla transizione (in questo caso q_1) altrimenti non fa nulla. Nel caso in cui ci sia stata una transizione di stati (siamo passati da q_0 a q_1) viene letta la lettera successiva nella parola e il processo si ripete.

[14] usate spesso in UNIX..

[15] sarebbe più precisamente /b̂aaa*!$/ che individua tutte e sole le stringhe 'baa' seguite da zero o più occorrenze di 'a' e infine '!'..

[16] per chi non avesse mai visto questo costrutto pensate alla Metropolitana dove i nodi sono le stazioni e "diretto" vuol dire che ho un unico senso di percorrenza per ogni tratto..

[17] si pensi a un puntatore sulle lettere..

Se la parola viene letta tutta e mi ritrovo alla fine del processo in uno stato *finale* (il nostro q_4) la parola viene *accettata* dall'Automa altrimenti viene *rifiutata*. Negli esempi considerati per lo *sheep-talk* "aba!b" viene rifiutato perché l'automa si ferma subito allo stato inziale senza proseguire mentre "baaaaa!" viene accettata (passo da tutti gli stati in ordine e itero per tre volte lo stato q_3). Se la parola viene accettata la sequenza di stati $\{q_0, .., q_n\}$ in cui l'automa è "passato" (per "baaaaa!" avrò $\{q_0, q_1, q_2, q_3, q_4\}$) viene chiamato *percorso accettante* dell'Automa per la parola data.

Formalmente un automa a stati finiti $\mathscr{A}$ è dato da:

DEFINIZIONE 14. $\mathscr{A} := \{Q, \Sigma, q_0, F, \delta\}$.

Q è un set finito di n stati $\{q_0,...,q_{n-1}\}$, Σ è un alfabeto finito, $q_0 \in Q$ è lo stato iniziale, $F \subseteq Q$ è la collezione di stati finali mentre δ è la funzione di transizione per cui $\delta(q, i) = q'$ significa: "se mi trovo nello stato q e leggo la lettera **i** allora mi sposto nello stato q'".
Il risultato notevole di Kleene riguardo ai linguaggi regolari è il seguente:

OSSERVAZIONE 15. Un linguaggio $\mathscr{L}$ è regolare se e solo se è riconosciuto da un automa a stati finiti.

Prima di procedere al primo risultato sui linguaggi umani abbiamo bisogno di un paio di osservazioni e di un lemma molto importante:

OSSERVAZIONE 16. Se $\mathscr{L}$ è regolare non è detto che $\mathscr{L}' \subset \mathscr{L}$ lo sia.[18]

OSSERVAZIONE 17. I linguaggi regolari sono chiusi per intersezione, differenza e complementazione.

LEMMA 18 (Pumping Lemma). *Sia $\mathscr{L}$ un linguaggio regolare infinito sull'alfabeto Σ. Allora $\exists N > 0$ tale che per ogni $\underline{w} \in \mathscr{L}$ con $|\underline{w}| = n > N$, $\exists x, y, z \in \Sigma^*$ t.c. $|xy| \leqslant N$, $|y| \geqslant 1$ con $\underline{w} = xyz$ e $xy^kz \in \mathscr{L}$ $\forall k \geqslant 1$.*

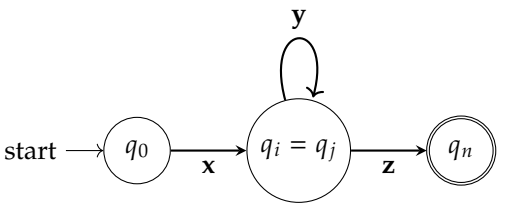


FIGURA 8. Automa a stati finiti che riconosce xy^kz.

Dimostrazione. Preso $N = |Q|$ il numero totale di stati dell'automa,
$\underline{w} \in \mathscr{L}$ di lunghezza $|\underline{w}| = n > N$,l'automa individuerà un *percorso accettante* $\{q_0, q_1, .., q_i, .., q_j, .., q_n\}$ in cui ,poiché $n > N$, ci sono almeno due stati coincidenti.
Considero allora i due indici $i < j$ minimi tali che $q_i = q_j$. Il percorso da q_0 a q_i sarà

[18] si consideri $\{a^h b^k\}_{h,k \geqslant 0}$..

individuato da una sottoparola **x**, da q_i a q_j da una sottoparola **y** e infine da q_j a q_n da una sottoparola **z**.
L'automa accetta $\underline{w} = xyz \in \mathscr{L}$ ma dato che è presente un ciclo o *loop* da q_i a q_j accetterà anche ogni $\underline{w'} = xy^kz$ con $k \geqslant 1$ perché il percorso accettante è lo stesso di $\underline{w}$. □

Primo Risultato

Abbiamo finalmente tutto il necessario per il primo risultato:

TEOREMA 19. *L'Inglese non è un linguaggio regolare.*[19]

Dimostrazione. Consideriamo le seguenti frasi in Inglese **[ENG]**:

- The cat likes tuna fish.
- The cat *the dog chased* likes tuna fish.
- The cat *the dog* **the rat bit** *chased* likes tuna fish.
- The cat *the dog* **the rat** the elphant liked **bit** *chased* likes tuna fish.[20]

e così via ad libitum.. Possono essere racchiuse nello schema:

DEF. $\mathscr{S}$:= (the+noun)^n (transitive verb)$^{n-1}$ likes tuna fish.

Indico con A = {the cat, the dog, the elephant,..} e B = {likes,chased,bit,ate,..}. Introduco la seguente *regexp*:

DEF. **R** := /**A*B*likes tuna fish**/

Ottengo quindi:

OSS. **[ENG]** ∩ **R** ≡ $\{x^n y^{n-1}$likes tuna fish$\}_{x\in A, y\in B} \equiv \mathscr{S}$.

Ora, se per assurdo l'Inglese fosse regolare dovrei avere per quanto visto in precedenza che $\mathscr{S}$ è regolare. Usando il Pumping Lemma si dimostra che $\mathscr{S}$ non può essere regolare.
Se lo fosse infatti dovrei avere una parola sufficientemente lunga con una sotto-parola che posso *pompare* o ripetere quante volte voglio senza "uscire" dal linguaggio.
Un breve studio per casi[21] mi fornisce però sempre parole "pompate" sgrammaticate che sono tutte ∉ [ENG].[22] Concludo quindi che l'Inglese non può essere regolare. □

Nota Bene: Lo schema $\mathscr{S}$ è stato ottenuto tramite quello che viene chiamato *center-embedding*[23] che corrisponde a una riscrittura della forma $A \Rrightarrow^* \alpha A \beta$..

---

[19] dimostrazione dovuta a Partee (1990)[1] et al..

[20] queste frasi innestate nel caso dell'Italiano hanno un analogo in "Alla fiera dell'est" di Branduardi.. "..E venne il cane che morse il gatto che si mangiò il topo che al mercato mio padre comprò.."

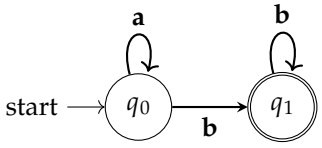
[21] "soli x","soli y" e "misto".. provare per credere..

[22] "(the+noun)^n (transitive verb)^m likes tuna fish" ∉ [ENG] per $m \neq n-1$.

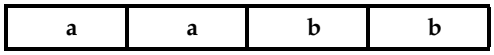
[23] si può addirittura dimostrare che un linguaggio è generato da una macchina a stati finiti se e soltanto se la sua grammatica non contiene "ricorsioni" "center-embedded"..

Secondo Risultato

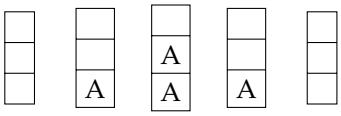
Abbiamo mostrato che i Linguaggi Naturali come l'Inglese non sono regolari o riconoscibili da automi a stati finiti. Sono liberi da contesto o *context-free*? Per rispondere al quesito oltre all'Inglese iniziamo a esplorare altre lingue come per esempio lo Svizzero Tedesco[24].
L'ultimo linguaggio considerato era della forma $\{a^n b^n\}_{n\geqslant 1}$ che usando il Pumping Lemma si dimostra non essere regolare. L'intuizione dietro a questo risultato sta nel fatto che un automa per riconoscerlo dovrebbe in linea di principio "contare" un numero arbitrario di occorrenze di 'a' e confrontarlo col numero di occorrenze di 'b'. Ma avendo una "memoria" limitata (il numero di stati è finito) l'automa non riesce in questo compito.
$\{a^n b^n\}_{n\geqslant 1}$ però è *context-free* e per riconoscerlo devo "potenziare" gli automi a stati finiti che ho a disposizione. Serve una memoria potenzialmente infinita. Ecco allora gli automi a stati finiti *a pila (stack)* noti anche come *Pushdown Automata*. Vediamo operativamente come questo nuovo automa riconosce per esempio **"aabb"**.



(A) Automa a stati finiti che riconosce "aabb".



(B) "aabb"



(C) la pila/stack durante la computazione

L'automa inizia come nel caso regolare dallo stato iniziale e legge la prima lettera della parola. Abbiamo però una pila o stack vuota all'inizio. Leggendo un'occorrenza di 'a' posiziono la meta-variabile 'A' in cima alla pila[25]. Rimango nello stato q_0 ,leggo un'altra 'a' e aggiungo un'altra 'A' in cima alla pila. Ora quando leggo la lettera 'b' elimino una 'A' dalla cima della pila[26]. Cosicché una volta letta tutta la parola mi ritrovo nello stato finale q_1 e la pila è stata svuotata.
Una parola viene accettata se è letta tutta, mi ritrovo in uno stato finale e la pila è vuota.

OSSERVAZIONE 20. $\mathscr{L}$ è *context-free* se e solo se è riconosciuto da un automa (a stati finiti) *a pila*.

Abbiamo anche i seguenti risultati analoghi al caso regolare:

OSSERVAZIONE 21. Se $\mathscr{L}$ è context-free e $\mathscr{L}'$ è regolare allora $\mathscr{L} \cap \mathscr{L}'$ è context-free.

[24] fa parte del gruppo dei dialetti Alemannici, è noto come "Schwiizerdütsch"..
[25] l'azione si chiama "PUSH".
[26] l'azione si chiama "POP".

LEMMA 22. *Pumping Lemma (versione context-free): se $\mathscr{L}$ è context-free (e infinito) $\exists K \geqslant 0$ tale che $\forall \underline{z} \in \mathscr{L}$ con $|\underline{z}| \geqslant K$, $\underline{z} = uvwxy$, $|vx| \geqslant 1$, $|vwx| \leqslant K$ e $uv^i wx^i y \in \mathscr{L}$ $\forall i \geqslant 1$.*

Tornando al nostro problema, lo Svizzero-Tedesco ha come il Latino forme dative e accusative che danno luogo a fenomeni chiamati "dipendenze incrociate" o *cross-serial dependencies*.

ESEMPIO 23. Jan sait das mer *em Hans* **es huus** *halfed* **aastriiche**.
(Jan says that we *helped Hans* **paint the house**.)[27]

In questo esempio abbiamo una forma in *dativo* "*em Hans..halfed*" e l'altra in **accusativo** "**es huus..aastriiche**" che danno luogo a una prima "dipendenza incrociata". Finalmente possiamo affermare:

TEOREMA 24. *lo Svizzero Tedesco non è context-free.*[28]

Dimostrazione. riprendiamo la frase dell'esempio recente e complichiamola ulteriormente con un altro livello di accusativo..

ES. Jan sait
das mer d'chind *em Hans* **es huus** haend wele laa *halfe* **aastriiche**.
(Jan says that we have wanted to let the children *help Hans* **paint the house**.)[29]

Viene naturale ora considerare una ripetizione potenzialmente infinita di d'chind..laa e *em Hans..halfe*.[30] Usiamo una opprtuna regexp:

DEF. $\mathscr{R}$:= / Jan sait das mer (*d'chind*)* (*em Hans*)* **es huus** haend wele (*laa*)* (*halfe*)* **aastriiche**. /

Suppongo lo Svizzero-Tedesco [ScD] sia context-free e lo interseco con $\mathscr{R}$ e ottengo:

OSS. $\mathscr{R} \cap$ [ScD] = {Jan sait das mer $(d'chind)^n (em\ Hans)^m$
es huus haend wele $(laa)^n (halfe)^m$ **aastriiche**.}$_{n,m\geqslant 1}$

$\mathscr{R} \cap$ [ScD] però è della forma $\{wa^n b^m x c^n d^m y\}_{n,m\geqslant 1}$ e questo linguaggio per il "nuovo" Pumping Lemma non è context-free..[31] Quindi lo Svizzero-Tedesco non può essere context-free. □

Conclusione

Riassumiamo quello che abbiamo provato:

(1) L'Inglese non è regolare.
(2) Lo Svizzero-Tedesco non è context-free.

[27] "Jan dice che abbiamo aiutato Hans a dipingere la casa"
[28] dimostrazione dovuta a Schieber[6] (1985).
[29] "Jan dice che abbiamo voluto lasciare che i bambini aiutino Hans a dipingere la casa"
[30] corrisponde nel primo caso a "let the children let the children let the children"/"lasciare che i bambini lascino che i bambini lascino che.." e nel secondo a "help Hans help Hans"/"aiutare ad aiutare Hans.."
[31] come nel caso regolare procedo a un analisi per casi e trovo che le sotto-parole "pompate" fanno perdere correttezza grammaticale..

Abbiamo usato nel primo caso frasi con *center-embedding* e nel secondo *cross-serial dependencies*. Ci si potrebbe domandare se questi fatti valgano solo per un certo tipo di linguaggi umani, d'altronde sono entrambe lingue di origine germanica.
Culy[7] nel 1985 ha fatto vedere però che la morfologia del "Bambara" ,un dialetto del Mali che quindi è quasi totalmente scollegato dalle lingue europee, presenta anch'esso delle dipendenze incrociate.
Ma allora in fin dei conti dove "stanno" i linguaggi umani? Questi indizi mi dicono che non sono nè regolari (di tipo III) nè context-free (di tipo II).. forse sono context-sensitive (di tipo I)? La direzione di ricerca[32] sembra propendere per un livello intermedio tra context-free e context-sensitive.
La questione è ancora aperta..

Bibliografia

[1] "Speech and Language Processing" di Daniel Jurafsky, James H. Martin.
[2] "Mathematical Methods in Linguistics" di Barbara H. Partee, Alice Meulen, Robert E. Wall.
[3] "Introduction to Automata Theory, Languages, and Computation" di E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman.
[4] "Gödel, Escher, Bach: an Eternal Golden Braid" di Douglas Hofstadter.
[5] "The Vastness of Natural Languages" di D. Terence Langendoen e Paul M. Postal.
[6] "Evidence against the context-freeness of natural languages" di Stuart M. Shieber.
[7] "The complexity of the vocabulary of Bambara" di Christopher Culy.
[8] "Three models for the description of language" di Noam Chomsky.

[32] A.Joshi propone le "mildly context sensitive languages"..

A PROPOSITO DELLA ζ DI RIEMANN

MARCO DELLA PENNA

1. Il problema

Nonostante il nome del seminario possa trarre in inganno, tratteremo del seguente problema di probabilità. Prima di tutto qualche definizione preliminare: diremo che un naturale n è squarefree se non è divisibile per nessun quadrato perfetto (1 non è un quadrato perfetto) o, analogamente, se $n = \prod_{i=1}^{k} {p_i}^{\alpha_i}$ è la fattorizzazione in primi di n allora $\forall i\ \alpha_i = 1$. Adesso passiamo a una versione poco matematica dell'enunciato del problema: immaginiamo di scegliere "a caso" due numeri naturali e di farne il g.c.d. (massimo comune divisore) e chiamiamolo M, adesso contiamo tutti i divisori di M che sono squarefree...quanti ce ne aspettiamo? A priori potrebbero essere tantissimi (non esiste un limite superiore) mentre invece ne abbiamo sempre almeno uno: 1 infatti è squarefree e divide tutti i numeri. Passiamo ora a una versione rigorosa dell'enunciato del problema.

Sia $A_n = \{1, .., n\}$ lo spazio discreto con la densità uniforme e sia Ω_n lo spazio prodotto $A_n \times A_n$. Sia ora X_n la variabile aleatoria che associa a un evento (a,b) di Ω_n il numero di divisori squarefree di M=g.c.d.(a,b) e sia E_n il valore atteso di X_n. Calcolare, se esiste, $E = \lim_{n\to\infty} E_n$. Per chi non fosse familiare con il linguaggio della probabilità il problema si traduce come di seguito: ogni numero tra 1 e n viene scelto con probabilità $1/n$ e a ogni coppia di numeri scelta viene associato un numero x_i (ovvero il numero degli sqf divisori di M) che avrà dunque una certa probabilità $p(x_i)$ =#{coppie a cui viene associato x_i}$/n^2$. Dopodichè si calcola $E_n = \sum x_i p(x_i)$, si noti che fissato n la somma è finita.

2. Osservazioni preliminari

Dato un certo numero M con fattorizzazione in primi $n = \prod_{i=1}^{k} {p_i}^{\alpha_i}$ si nota immediatamente che il numero di sqf divisori di M è 2^k: infatti per ogni primo che divide M posso scegliere se prenderlo con esponente 1 o con esponente 0 (2 scelte per ognuno dei k primi che dividono M). Abbiamo dunque notato che gli x_i del problema precedente sono in realtà delle potenze di 2, chiamiamo dunque $x_i = 2^i$. Inoltre in un certo senso abbiamo ridotto il problema al calcolo della probabilità che un certo primo divida M. E' evidente che, fissati due primi p e q, la probabilità che p divida a (il primo dei due numeri scelti) e la probabilità che q divida a sono eventi indipendenti quando

n tende a ∞ (mentre questo non è vero per n fissato). Notiamo inoltre che, affinchè p divida M, condizione NS è che p divida sia a che b e ciò avviene con probabilità $1/p^2$ quando n tende a ∞. Ma allora possiamo claimmare che

$$E = \sum_{i=0}^{\infty} 2^i p(2^i) = \sum_{i=0}^{\infty} 2^i \sum_{p_1,..,p_i primi} \prod_{j=1}^{i} 1/{p_j}^2 \prod_{p \neq p_j \forall j} (1 - 1/p^2) =$$

$$\sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} \prod_{j=1}^{i} 2/{p_j}^2 \prod_{p \neq p_j \forall j} (1 - 1/p^2) = \sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} (\sum_{k=0}^{i} 2^k (-1)^{i-k}) \frac{1}{(p_1..p_i)^2} =$$

$$\sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} \frac{1}{(p_1..p_i)^2} = (*)$$

Spieghiamo un po' più nel dettaglio i passaggi fatti:

La prima uguaglianza segue dal calcolo esplicito di $p(2^i)$ che è eseguito come di seguito: fissato un certo i dobbiamo valutare la probabilità che M abbia esattamente 2^i sqf divisori e dunque la probabilità che M sia diviso da esattamente i numeri primi. Quindi basta fissare $p_1, .., p_i$ primi e chiedere che M sia diviso da questi ma non da tutti gli altri (da qui il termine che contiene $1 - 1/p^2$); dopodichè la formula segue dall'indipendenza degli eventi e facendo variare i primi scelti. Le altre uguaglianze sono semplici calcoli di combinatoria, segnaliamo in particolare che nell'ultimo passaggio si è usato il binomio di Newton.

3. Traccia di una dimostrazione rigorosa

In realtà le osservazioni preliminari contengono già una dimostrazione quasi soddisfacente, l'unico passo che non è stato giustificato è il passaggio al limite con conseguente indipendenza degli eventi "p divide a" al variare di p primo. In questa sezione vogliamo fornirne una giustificazione anche se non entreremo nello specifico dei calcoli di combinatoria, perchè analoghi a quelli già svolti nella sezione precedente.

Sia n fissato. Supponiamo di sapere che $a = \prod_{i=1}^{k} {p_i}^{\alpha_i}$ e chiamiamo $p_a(2^i)$ la probabilità di x_i sapendo già che è uscito "a" come primo numero e sia inoltre E_{n_a} il valore atteso relativo (una probabilità condizionata nel gergo probabilistico). Allora si verifica che, se k=0, $p_a(1) = E_{n_a} = 1$, se k=1 $p_a(1) = 1 - \frac{[\frac{n}{p_1}]}{n}$, $p_a(2) = \frac{[\frac{n}{p_1}]}{n}$, $E_{n_a} = 1 + \frac{[\frac{n}{p_1}]}{n}$ e in generale

$$E_{n_a} = \sum_{m|a,sqf} \frac{[\frac{n}{m}]}{n}$$

dove [x] indica la parte intera di x. Ma allora

$$E_n = \sum_{a \leq n} \sum_{m|a,sqf} \frac{[\frac{n}{m}]}{n^2} = \sum_{m \leq n,sqf} (\frac{[\frac{n}{m}]}{n})^2 = (**)$$

Dalla disuguaglianza $\frac{\frac{n}{m}-1}{n} \leq \frac{[\frac{n}{m}]}{n} \leq \frac{\frac{n}{m}}{n}$ e notando che $(\frac{\frac{n}{m}-1}{n})^2 \geq \frac{1}{m^2} - \frac{2}{mn}$ si ha che

$$(*) = \sum_{m,sqf} \frac{1}{m^2} = \lim_{n\to\infty}\Big(\sum_{m\leq n,sqf} \frac{1}{m^2} - 2\frac{\log(n)+1}{n}\Big) \leq \lim_{n\to\infty}\Big(\sum_{m\leq n,sqf} \frac{1}{m^2} - 2\frac{\sum_{i=1}^{n} 1/i}{n}\Big) \leq$$

$$\lim_{n\to\infty} \sum_{m\leq n,sqf} \frac{1}{m^2} - \sum_{m,sqf} \frac{2}{mn} \leq \lim_{n\to\infty} (**) \leq \sum_{m,sqf} \frac{1}{m^2} = (*)$$

Quindi

$$\lim_{n\to\infty} (**) = (*)$$

che era proprio ciò che volevamo dimostrare.

4. Ma cosa c'entra la Zeta di Riemann in tutto questo?

Iniziamo con la sua definizione:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

che è ben definita per ogni numero complesso s che soddisfa $Re(s) > 1$ ma che per noi oggi sarà definita sui naturali maggiori di 1 (per semplicità).

4.1. Il prodotto di Eulero

Si verifica facilmente che, siccome ogni intero n ha un'unica fattorizzazione in primi distinti, vale la seguente

$$\zeta(s) = \prod_{p,primo} \sum_{i=0}^{\infty} \frac{1}{p^{is}} = \prod_{p,primo} \frac{1}{1-\frac{1}{p^s}}$$

dove nell'ultima uguaglianza si è usata la somma della serie geometrica.

Ma allora si ha che

$$(*) = \prod_{p,primo} (1 + \frac{1}{p^2}) = \prod_{p,primo} \frac{(1-\frac{1}{p^4})}{(1-\frac{1}{p^2})} = \frac{\zeta(2)}{\zeta(4)}$$

4.2. Si, tutto bellissimo, ma quanto vale il risultato?

Beh intanto sappiamo che il risultato è finito e ciò non era assolutamente scontato a priori, però ci piacerebbe fornirne una stima migliore. Per farlo ci occorre la seguente formula di cui daremo una traccia di dimostrazione rigorosa nell'appendice.

4.2.1. *Fattorizzazione di Hadamard del seno.*

$$\sum_{j=0}^{\infty} \frac{(\pi x)^{2j}(-1)^j}{(2j+1)!} = \frac{\sin(\pi x)}{\pi x} = \prod_{k=1}^{\infty}(1 - \frac{x^2}{k^2})$$

La prima uguaglianza è semplicemente lo sviluppo di Taylor nell'origine della funzione, ma, siccome la serie ha raggio di convergenza ∞, l'uguaglianza vale ovunque. La seconda uguaglianza passa attraverso l'analisi complessa; per coloro i quali non volessero addentrarsi nei meandri di questa materia, forniamo la seguente (insoddisfacente) giustificazione:

Nel caso in cui la serie e il prodotto non fossero infiniti la tesi sarebbe ovvia per il teorema di Ruffini, perchè avremmo due polinomi che hanno gli stessi zeri (ovvero tutti gli interi relativi non nulli) e che coincidono anche nel punto x=0 (siccome $lim_{x\to 0}\frac{\sin(\pi x)}{\pi x} = 1$), tuttavia il fatto che la serie e il prodotto siano infiniti complica non poco le cose. Diamo per buona la formula e calcoliamo i primi termini dello sviluppo di Taylor del prodotto di Hadamard, chiamiamo per semplicità $t = x^2$:

$$p'(t) = (\prod_{k=1}^{\infty}(1 - \frac{t}{k^2}))' = p(t)(\ln(\prod_{k=1}^{\infty}(1 - \frac{t}{k^2})))' = p(t)(\sum_{k=1}^{\infty}(ln(1 - \frac{t}{k^2}))' = p(t)\sum_{k=1}^{\infty}\frac{1}{t-k^2}$$

$$p''(t) = p'(t)\sum_{k=1}^{\infty}\frac{1}{t-k^2} - p(t)\sum_{k=1}^{\infty}\frac{1}{(t-k^2)^2}$$

e quindi in particolare $p'(0) = -\zeta(2)$ mentre $p''(0) = (\zeta(2))^2 - \zeta(4)$. Tuttavia dallo sviluppo di Taylor segue che $p'(0) = -\frac{\pi^2}{3!}$ e $p''(0) = 2\frac{\pi^4}{5!}$ e in definitiva si ha che

$$(*) = \frac{\zeta(2)}{\zeta(4)} = \frac{\frac{\pi^2}{6}}{\frac{\pi^4}{90}} = \frac{15}{\pi^2}$$

5. Osservazioni finali e generalizzazioni

Si noti che $\frac{15}{\pi^2} \approx 1,52$ e la nostra primissima stima (banale) sul risultato era che fosse maggiore di 1. Dunque il risultato ottenuto è bassissimo e ciò potrebbe essere un po' controintuitivo pensando che quando n tende all'infinito il numero massimo di sqf divisori di una coppia può tendere all'infinito. Erroneamente si potrebbe pensare quindi che, in media, gli sqf divisori di un numero non siano così tanti ma questo è assolutamente falso: si pensi che già ogni numero maggiore di 1 ha almeno 2 sqf divisori. Da dove salta fuori allora 1,52? Esso è conseguenza del fatto che due numeri scelti a caso hanno, in generale, un g.c.d. molto basso. A riprova di ciò lascio il seguente esercizio sulla falsa riga di quello proposto in questo seminario.

ES. Provare che scelti due numeri naturali a caso (nello stesso senso di questo esercizio) la probabilità che essi siano coprimi vale $\frac{1}{\zeta(2)} = \frac{6}{\pi^2} \approx 60\%$

Chiarito dunque questo aspetto passiamo a una generalizzazione: se invece di due numeri a caso ne avessimo scelti $m \geq 2$ cosa sarebbe cambiato? Sostanzialmente

nulla infatti analogamente a quanto visto nel capitolo 2 troveremmo che

$$E = \sum_{i=0}^{\infty} 2^i p(2^i) = \sum_{i=0}^{\infty} 2^i \sum_{p_1,..,p_i primi} \prod_{j=1}^{i} 1/{p_j}^m \prod_{p \neq p_j \forall j} (1 - 1/p^m) =$$

$$\sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} \prod_{j=1}^{i} 2/{p_j}^m \prod_{p \neq p_j \forall j} (1 - 1/p^m) = \sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} (\sum_{k=0}^{i} 2^k (-1)^{i-k}) \frac{1}{(p_1..p_i)^m} =$$

$$\sum_{i=0}^{\infty} \sum_{p_1,..,p_i primi} \frac{1}{(p_1..p_i)^m} = \frac{\zeta(m)}{\zeta(2m)}$$

Questo è infatti il risultato in generale; notiamo in particolare che esso è esplicitabile (come fatto per il caso m=2) soltanto per valori pari di m, poichè non sono noti i valori dispari della zeta. Notiamo anche che $\lim_{m\to\infty} \zeta(m) = 1 \Rightarrow \lim_{m\to\infty} \frac{\zeta(m)}{\zeta(2m)} = 1$ come era ovviamente prevedibile.

6. Appendice

In questa sezione ci proponiamo di dire qualcosa in più sui prodotti infiniti e di giustificare la fattorizzazione di Hadamard del seno. Per ovvi motivi di spazio non scenderemo nel dettaglio, ma elencheremo una carrellata di fatti che saranno un ottimo esercizio per il lettore. Per una trattazione più approfondita rimandiamo al corso di analisi complessa.

FATTO 1: $\pi \cot(\pi z) = \frac{1}{z} + \sum_{n\neq 0}(\frac{1}{z-n} + \frac{1}{n})$

DIM(cenni): Entrambe le funzioni sono meromorfe con stessi poli e stessi residui, dunque la differenza h è intera. $h'(z) = \frac{\pi^2}{\sin^2(\pi z)} - \sum_{n\in Z} \frac{1}{(z-n)^2}$ è periodica di periodo 1 dunque per mostrare che h' è limitata basta mostrarlo su $0 \leq Re(z) \leq 1$ e questo è abbastanza facile. Dunque, per Liouville, h' è costante e facilmente si arriva allora a dire che h è identicamente nulla.

FATTO 2:(Teo. Weierstrass) Sia $\{z_j\}$ una successione di numeri complessi e sia $\{p_j\}$ una successione di interi tale che $\sum_{j=1^\infty}(\frac{r}{|z_j|})^{p_j+1}$ converge per ogni $r > 0$; sia inoltre $E(z,n) = (1-z)e^{z+\frac{z^2}{2}+...+\frac{z^n}{n}}$. Allora $\prod_{j=1^\infty} E(z/z_j, p_j)$ è una funzione intera che si annulla esattamente negli z_j.

DIM(cenni): Anzitutto ricordiamo che, in analogia con le serie, si dice che un prodotto converge se esiste finito e DIVERSO DA 0 il limite dei prodotti parziali. Dopodichè si verifica facilmente che il prodotto degli z_j converge solo se converge la somma dei $\log(z_j)$. Inoltre si deve chiaramente avere che z_j tende a 1 quindi scriviamo $z_j = 1 + a_j$ con a_j che tende a 0. Diciamo che il prodotto degli z_j converge assolutamente se converge la somma dei moduli dei $\log(z_j)$ e verifichiamo che ciò avviene solo se converge assolutamente la serie degli a_j. A questo punto abbiamo, in un certo senso, scaricato la convergenza di un prodotto su quella di una serie. Un esercizio non banalissimo è mostrare che $|1 - E(z,n)| \leq |z|^{n+1}$ ma allora, sfruttando l'ipotesi sui p_j si verifica immediatamente la convergenza assoluta su ogni compatto.

L'olomorfia della funzione prodotto viene garantita da un altro teorema (sempre di Weierstrass).

Consideriamo ora la funzione $f(z) = \pi z \prod_{n\neq 0} E(z/n, 1)$: dal teorema di Weierstrass essa è intera e si annulla in tutti e soli gli interi proprio come $\sin(\pi z)$. In ogni disco che non contiene interi vale inoltre che

$$\frac{d}{dz}\log(f(z)) = \frac{f'(z)}{f(z)} = \frac{1}{z} + \sum_{n\neq 0}(\frac{1}{z-n} + \frac{1}{n}) = \pi\cot(\pi z) = \frac{d}{dz}\log\sin(\pi z)$$

quindi $\log(f(z)) = \log\sin(\pi z) + C$ da cui è immediato verificare che C=0 e vale la formula cercata:

$$f(z) = \sin(\pi z)$$

7. Sitografia

https://math.stackexchange.com/questions/3995219/prime-zeta-function-and-squarefree-divisors

https://www.robertobigoni.it/Matematica/Trascendenti/f10/f10.htm

Per l'appendice sono state sfruttate le note del corso di analisi complessa redatte dal prof Peloso. Vi consiglio inoltre il seguente link per vedere molti altri modi (attenzione che non tutti sono corretti) di calcolare la $\zeta(2)$, celebre risultato noto come "problema di Basilea".

https://math.stackexchange.com/questions/8337/the-basel-problem

IL TEOREMA FONDAMENTALE DELL'ALGEBRA

FRANCESCO ALESSIO ZUCCON

Introduzione

In questo breve seminario si vanno ad esplorare alcune delle numerose prove di un risultato di cruciale importanza per l'evoluzione storica della matematica: il *Teorema fondamentale dell'algebra*.
Oltre alla rivoluzione che portarono i vari tentativi antecedenti al successo di Gauss nel 1799, di notevole interesse è anche il ricco spettro di ambiti della matematica che abbraccia.
Se all'interno di molteplici corsi o libri di matematica è possibile verificare questo fatto osservandone applicazioni di diversa natura, in questa sede si predilige evidenziare questo aspetto attraverso la diversificazione delle prove successivamente presentate.
Denotando con $R_p(\mathbb{C})$ l'insieme delle radici in $\mathbb{C}$ del polinomio $p \in \mathbb{C}[z]$, si delinea, infine, un enunciato di riferimento per il teorema protagonista del seminario:

TEOREMA 1. *Sia* $p \in \mathbb{C}[z]$, $n = deg(p) > 0 \implies R_p(\mathbb{C}) \neq \varnothing$

1. Dimostrazione con Analisi Complessa

La dimostrazione più classica presentata in epoca moderna è quella oggetto di questa sezione.
Essa si basa su un risultato sufficientemente elementare nell'ambito dell'Analisi Complessa, che pare il luogo forse più naturale in cui il teorema possa emergere.
Il risultato succitato, dimostrato da Cauchy nel 1844, ma noto come *Teorema di Liouville*, è il seguente:

TEOREMA 2. *Sia* $f : \mathbb{C} \to \mathbb{C}$ *mappa intera e limitata* $\implies$ f *costante.*

Dimostrazione. Si consulti [1] (Corollario 4.5). □

Senza dilungarsi sugli aspetti di analisi complessa da cui ottenere questo utile strumento, è presto dimostrato il teorema fondamentale dell'algebra:

Dimostrazione. Sia $p \in \mathbb{C}[z]$ tale che $R_p(\mathbb{C}) = \varnothing$, allora la mappa $f(z) := \frac{1}{p(z)}$ è una mappa intera. Se per assurdo $deg(p) > 0$, allora $\lim_{|z|\to\infty} |p(z)| = \infty$, pertanto $\lim_{|z|\to\infty} |f(z)| = 0$, e quindi f è limitata. Da Teorema 2 allora f è costante, dunque anche p, da cui la contraddizione con il grado. □

Osservazione 3. Lo studio del teorema fondamentale dell'algebra tramite l'analisi complessa ha rivelato la presenza di un minimo per $|p(z)|$, il che sarà in seguito sfruttato, pur ottenendolo in modo indipendente.

2. Dimostrazione Analitica

Introduzione

La prova protagonista della seconda sezione è stata sviluppata da Cauchy e risulta peculiare in quanto potrebbe essere considerata la più elementare conosciuta, che si appoggia unicamente ad elementi comuni ad un primo corso di Analisi Matematica. Come già accennato, l'idea è di sfruttare l'intuizione che proviene dall'Analisi Complessa della presenza del minimo per $|p(z)|$, combinandola con lo sviluppo di Taylor, la completezza dei numeri reali e la proprietà di $\mathbb{C}$ di essere un campo algebricamente chiuso.

Dimostrazione. Sia $p \in \mathbb{C}[z]$ tale che $deg(p) > 0$, allora $\lim_{|z|\to\infty}|p(z)| = \infty$, pertanto $\exists R \in \mathbb{R}$ con $R > 0$ e $|p(z)| > |p(0)|, \forall z \in \mathbb{C} : |z| > R$.
Si definisca $D_R := \{z \in \mathbb{C} \text{ t.c. } |z| \leqslant R\}$, dal Teorema di Heine-Borel esso è compatto (rispetto alla topologia euclidea su $\mathbb{C}$), pertanto dal Teorema di Weierstrass $\exists z_0 \in D_R$ t.c. $|p(z_0)| = min\{|p(z)| : z \in D_R\}$, e dalla definizione di R si ha banalmente $|p(z_0)| = min\{|p(z)| : z \in \mathbb{C}\}$.
Sia per assurdo $|p(z_0)| \neq 0$, allora dallo sviluppo di Taylor centrato in z_0 si ha: $p(z) = \sum_{i=0}^{n} a_i(z - z_0)^i$, con $a_i \in \mathbb{C}$ e $a_0 = p(z_0)$.
Sia poi $k := min\{m \in \mathbb{N} - \{0\} : a_m \neq 0\}$, si ha $p(z) = a_0 + a_k z^k + o((z - z_0)^k)$, e, poiché $\mathbb{C}$ è campo algebricamente chiuso, $\forall \epsilon > 0 \exists z_\epsilon \in \mathbb{C}$ t.c. $a_k(z_\epsilon - z_0)^k = -\epsilon a_0$, ottenendo $p(z) = (1 - \epsilon)a_0 + R_k(z_\epsilon)$, ove $R_k(z_\epsilon) = o((z_\epsilon - z_0)^k)$.
Tuttavia, per continuità si ha $\lim_{\epsilon\to 0}|z_\epsilon - z_0| = 0$, e poiché $k \geqslant 1$, per completezza del campo dei numeri reali segue che $\exists \epsilon > 0$ t.c. $|R_k(z_\epsilon)| \leqslant |o(z_\epsilon - z_0)| < \frac{a_0\epsilon}{2}$, pertanto per la disuguaglianza triangolare:

$$|p(z_\epsilon)| \leqslant |(1 - \epsilon)a_0| + |R_k(z_\epsilon)| < (1 - \frac{1}{2}\epsilon)|a_0| < |a_0|$$

il che è contro la minimalità di z_0.
Dunque, si ha $z_0 \in R_p(\mathbb{C}) \neq \varnothing$, la tesi. □

3. Dimostrazione Topologica

Introduzione

La dimostrazione geometrica presentata in seguito sfrutta un'invariante topologico molto versatile ed importante: il gruppo fondamentale. In particolare, si sfrutta il fatto che lo spazio S^1 abbia come gruppo fondamentale $\mathbb{Z}$ (per maggiori dettagli si consulti [2], capitolo 1).

Dimostrazione. Sia $p \in \mathbb{C}[z]$ tale che $R_p(\mathbb{C}) = \emptyset$, allora, per $r \in [0, +\infty)$, la mappa $f_r : [0,1] \to S^r$, ove

$$f_r(s) := \frac{p(re^{2\pi is})/p(r)}{|p(re^{2\pi is})/p(r)|}$$

è ben definita, risultando un loop di S^1 puntato in 1.
Inoltre, f_r definisce implicitamente in funzione del parametro r un'omotopia, rendendo tutte tali funzioni omotopicamente equivalenti tra loro, in particolare equivalenti ad $f_0 = 1$ omotopicamente banale.
Posto $p(z) := z^n + a_1 z^{n-1} + ... + a_n$, si scelga $r \in \mathbb{R}$ t.c. $r > max\{1, \sum_{i=1}^n |a_i|\}$, se poi vale $|z| = r$, si ha:

$$|z^n| > (\sum_{i=1}^n |a_i|)|z^{n-1}| \geqslant \sum_{i=1}^n |a_i z^{n-i}| \geqslant |\sum_{i=1}^n a_i z^{n-i}| = |p(z) - z^n|$$

Al variare del parametro $t \in [0,1]$ e a partire dal polinomio p, si definisce polinomio $p_t(z) := z^n + t(a_1 z^{n-1} + ... + a_n) = z^n + t(p(z) - z^n)$, il quale, in virtù della disuguaglianza (3), non può ammettere radici in S^r.
Infatti, se per assurdo $z_t \in R_{p_t}(S^r)$, allora, poiché $p_t(z_t) = 0$, $|z_t| = r$ e $|t| \leqslant 1$, si ha una contraddizione:

$$|{z_t}^n| > (\sum_{i=1}^n |a_i|)|{z_t}^{n-1}| \geqslant |\sum_{i=1}^n t a_i {z_t}^{n-i}| = |p(z_t) - {z_t}^n| = |{z_t}^n|$$

Allora viene indotta un'omotopia $H : [0,1]x[0,1] \to S^1$, con $H_r(t,s) := f_{r,t}(s)$, ove analogamente a quanto fatto in precedenza si ha:

$$f_{r,t}(s) := \frac{p_t(re^{2\pi is})/p_t(r)}{|p_t(re^{2\pi is})/p_t(r)|}$$

Tuttavia, $H(0,s) = w_n(s) := e^{2\pi ins}$ per $n = deg(p)$ e $H(1,s) = f_r(s)$, dunque w_n è omotopicamente banale in quanto si è mostrato tale f_r.
Inoltre, essendo w_n un loop di S^1, dato un generatore $\pi_1(S^1, 1) =< \alpha > \simeq \mathbb{Z}$, si ha $w_n \sim_{hom} \alpha^n$, del resto per banale omotopia segue $n = 0$, da cui la tesi. □

4. Dimostrazione Algebrica

Introduzione

L'ultima prova presentata mira ad dimostrare 1 nel modo il più possibile indipendente da strumenti analitici, con il solo utilizzo di due risultati elementari: il *Teorema di Bolzano* e la presenza delle radici quadrate in $\mathbb{C}$.
Vengono, tuttavia, a servire ulteriori risultati di tipo pressoché combinatorio presentati qui di seguito. Il primo è noto come *Formula di Viète*:

TEOREMA 4. *Sia A dominio di integrità, $p(x) = \sum_{i=0}^n a_i \cdot x^i$, con $n \neq 0$, $a_n \neq 0$ e $R_p(\overline{Frac(A)}) = \{\lambda_i\}_{i=1}^n$, allora $\forall k \in \{1, ..., n\}$:*

$$\sum_{1 \leqslant i_1 \leqslant ... \leqslant i_k \leqslant n} \lambda_{i_1} \cdot \lambda_{i_2} \cdots \lambda_{i_k} = (-1)^k \cdot \frac{a_{n-k}}{a_n}$$

Dimostrazione. Ovvia. □

Per introdurre al secondo risultato, è necessaria la seguente definizione:

DEFINIZIONE 5. Sia A anello, $p \in A[x_1, ..., x_n]$ è detto *polinomio simmetrico* se $\forall \sigma \in Sym(\{1, ..., n\})$ vale $p(x_1, ..., x_n) = p(x_{\sigma(1)}, ..., x_{\sigma(n)})$
Si usa la notazione $S_n := Sym(\{1, ..., n\})$ e $S_{A[x_1,...,x_n]} = A[x_1, ..., x_n]^{S_n}$ insieme dei polinomi simmetrici a coefficienti in A.
Si definiscono polinomi simmetrici elementari:

$$e_0(x_1, ..., x_n) := 1$$
$$e_{j\geqslant 1}(x_1, ..., x_n) := \sum_{1\leqslant i_1\leqslant ...\leqslant i_j\leqslant n} x_{i_1} \cdot x_{i_2} \cdot \cdots x_{i_j}$$

Il secondo risultato è il *Teorema fondamentale dei polinomi simmetrici*:

TEOREMA 6. *Sia A anello, $p \in S_{A[x_1,...,x_n]} \implies \exists q \in A[x_1, ..., x_n]$ t.c.:*

$$p(x_1, ..., x_n) = q(e_1(x_1, ..., x_n), ..., e_n(x_1, ..., x_n))$$

Dimostrazione. Si veda [3] (Teorema 5.36). □

Si hanno ora tutti gli strumenti per procedere con l'ultima prova:

Dimostrazione. Innanzitutto, si può supporre $p \in \mathbb{R}[x]$: qualora $p \in \mathbb{C}[x]\backslash\mathbb{R}[x]$, posto $q(x) := p(x) \cdot \overline{p(x)} \in \mathbb{R}[x]$ segue $R_p(\mathbb{C}) \neq \varnothing \iff R_q(\mathbb{C}) \neq \varnothing$.
Sia $n := deg(p) \geqslant 0$, si procede per induzione su $k := max\{m \in \mathbb{N} : 2^m | n\}$.
Wlog $n = 2^k \cdot t$, ove $t \in \mathbb{N}$ dispari. Se $k = 0$ il risultato segue dal teorema di Bolzano, da cui la base induttiva.
Sia, poi, $k > 0$, e sia $K := \mathbb{R}(p)$ campo di spezzamento di $p(x) = \sum_{i=1}^n a_i \cdot x^i$, e siano $R_p(K) = \{\lambda_i\}_{i=1}^n$, allora si mostra $R_p(K) \subset \mathbb{C}$.
In dipendenza del parametro reale t, si definisca:

$$q_t(x) := \prod_{1\leqslant i<j\leqslant n} (x - \lambda_i - \lambda_j - t \cdot \lambda_i\lambda_j) \in K[x]$$

Innanzitutto, si osserva che $q_t \in \mathbb{R}[x]$. Infatti, si consideri

$$h_{t,x}(x_1, ..., x_n) := \prod_{1\leqslant i<j\leqslant n} (x - \lambda_i - \lambda_j - t \cdot \lambda_i\lambda_j) \in \mathbb{R}[x_1, ..., x_n]$$

come polinomio in più variabili, e risulta chiaro $h_{t,x} \in S_{\mathbb{R}[x_1,...,x_n]}$.
Dunque da Teorema 6 si ha che $\exists s \in \mathbb{R}[x_1, ..., x_n]$, t.c.

$$h_{t,x}(x_1, ..., x_n) = s(e_1(x_1, ..., x_n), ..., e_n(x_1, ..., x_n))$$

Pertanto, vale $q_t(x) = h_{t,x}(\lambda_1, ..., \lambda_n) = s(e_1(\lambda_1, ..., \lambda_n), ..., e_n(\lambda_1, ..., \lambda_n))$, ma ora sfruttando Teorema 4 si può osservare che $\forall i \in \{0, 1, ..., n\}$:

$$e_i(\lambda_1, ..., \lambda_n) = (-1)^i \cdot \frac{a_{n-i}}{a_n} \in \mathbb{R}$$

Dunque, $q_t \in \mathbb{R}[x]$ come volevasi.
Inoltre, $deg(q_t) = n \cdot (n-1)/2 = 2^{k-1} \cdot t \cdot (n-1)$, ma $t \cdot (n-1)$ è dispari, dunque si

può applicare l'ipotesi induttiva al polinomio q_t, ossia $R_{q_t}(\mathbb{C}) \neq \emptyset$.
Ma $R_{q_t}(K) = \{\lambda_i + \lambda_j + t \cdot \lambda_i \cdot \lambda_j\}_{1 \leqslant i < j \leqslant n}$, quindi $\exists i, j \in \{1, ..., n\}$ dipendenti da t tali che $\lambda_i + \lambda_j + t \cdot \lambda_i \cdot \lambda_j \in \mathbb{C}$.
Tuttavia, t varia in $\mathbb{R}$, dunque per finitezza delle radici esiste almeno una coppia di radici per cui $\exists t, s \in \mathbb{R}$ t.c. $t \neq s$ e $\lambda_i + \lambda_j + t \cdot \lambda_i \cdot \lambda_j = \lambda_i + \lambda_j + s \cdot \lambda_i \cdot \lambda_j \in \mathbb{C}$.
Dunque, $\lambda_i \cdot \lambda_j \in \mathbb{C}$, da cui anche $\lambda_i + \lambda_j \in \mathbb{C}$, pertanto $\lambda_i, \lambda_j \in \mathbb{C}$ come radici del polinomio di secondo grado $r(z) = (z - \lambda_i) \cdot (z - \lambda_j) \in \mathbb{C}[z]$.
Da ciò, $R_p(\mathbb{C}) \neq \emptyset$, la tesi. □

Bibliografia

(1) E. M. Stein, R. Shakarchi, *Complex Analysis*.
(2) A. Hatcher, *Algebraic Topology*.
(3) J.S. Milne, *Fields and Galois Theory*.

TEOREMA DI SEIFART VAN KAMPEN

EDWARD KEVIN ARANA MEDINA

1. Introduzione

In matematica, e più precisamente in topologia algebrica, il teorema di Seifert-Van Kampen è uno dei principali strumenti per il calcolo del gruppo fondamentale di uno spazio topologico. Venne dimostrato indipendentemente da Herbert Seifert ed Egbert Van Kampen agli inizi del 1930. Il teorema afferma che se uno spazio topologico X è unione di due aperti A e B che verifichino certe proprietà di connessione allora la struttura del suo gruppo fondamentale è esprimibile in termini dei gruppi fondamentali di A, B e dell'intersezione $A \cap B$. In tal modo il teorema permette di calcolare il gruppo fondamentale di uno spazio complicato partendo da gruppi fondamentali di spazi più semplici.

2. Enunciato del teorema

Sia X uno spazio topologico. E supponiamo che U,$V \subset X$ tali che :

- U,V siano aperti, non vuoti e connessi per archi
- $U \cup V = X$
- $U \cap V$ sia connesso per archi

E sia $F : \pi_1(U \cap V, q) \to \pi_1(U, q) * \pi_1(V, q)$ una mappa definita nel seguente modo: $F(\gamma) = (i_*\gamma)^{-1}(j_*\gamma)$. E sia $\overline{F(\pi_1(U \cap V, q))}$ la chiusura normale dell'immagine di F in $\pi_1(U, q) * \pi_1(V, q)$. Allora $\forall q \in U \cap V$ vale il seguente isomorfismo:

$$\pi_1(X, q) \cong \frac{\pi_1(U,q)*\pi_1(V,q)}{\overline{F(\pi_1(U\cap V,q))}}$$

2.1. dimostrazione

Innanzitutto consideriamo quattro mappe di inclusione: $i : U \cap V \to U$, $j : U \cap V \to V$, $k : U \to X$ e infine $l : V \to X$. Queste mappe di inclusione inducono i seguenti omomorfismi di gruppi: $i_* : \pi_1(U \cap V) \to \pi_1(U)$, $j_* : \pi_1(U \cap V) \to \pi_1(V)$, $k_* : \pi_1(U) \to \pi_1(X)$ e infine $l_* : \pi_1(V) \to \pi 1(X)$. Ora considerando anche il prodotto libero $\pi_1(U, q) * \pi_1(V, q)$, consideriamo le mappe canoniche $T_U : \pi_1(U, q) \hookrightarrow \pi_1(U, q) * \pi_1(V, q)$ e $T_V : \pi_1(V, q) \hookrightarrow \pi_1(U, q) * \pi_1(V, q)$. Sappiamo che per la proprietà del prodotto libero, k_* e l_* inducono un unico omomorfismo di gruppi

$\Phi : \pi_1(U,q) * \pi_1(V,q) \to \pi_1(X,q)$ tale che $k_* = \Phi \circ T_U$ e $l_* = \Phi \circ T_V$. Ora possiamo dividere la dimostrazione in due parti: nella prima parte dimostriamo che Φ è suriettiva e nella seconda parte dimostriamo che il nucleo di Φ è esattamente $\overline{F(\pi_1(U \cap V, q))}$.

2.2. Notazioni

Siccome dovremmo considerare cammini e le loro classe di omotopia in vari spazi topologici, se a e b sono cammini in X che vivono in uno dei seguenti sottoinsiemi U,V,$U \cap V$ e X allora userò le seguenti notazioni: $a \sim_U b$, $a \sim_V b$, $a \sim_{U\cap V} b$, $a \sim_X b$ per indicare che a è un cammino omotopico a b in U,V,$U \cap V$, o in X rispettivamente. Inoltre scriverò $[a]_U$ per indicare una classe di a in $\pi_1(U,q)$ e similmente per gli altri insiemi. Inoltre dobbiamo considerare due tipi di prodotti in questa dimostrazione: ovvero la moltiplicazione delle classi di cammini come elementi del gruppo fondamentale ed la moltiplicazione intesa come concatenazione di parole nel prodotto libero di gruppi. Dunque per indicare la moltiplicazione di classi di cammini utilizzerò la seguente notazione: $[a]_U \cdot [b]_U = [a \cdot b]_U$. Per enfatizzare la distinzione tra i due prodotti, denoterò la moltiplicazione nel prodotto libero con un asterisco: $[a]_U * [b]_U * [c]_V = [a \cdot b]_U * [c]_V \in \pi_1(U,q) * \pi_1(V,q)$. Dunque la mappa Φ può essere definita nel seguente modo: $\Phi([a_1]_U * [a_2]_V * \cdots * [a_{m-1}]_U * [a_m]_V) = k_*[a_1]_U \cdot l_*[a_2]_V \cdots k_*[a_{m-1}]_U \cdot l_*[a_m]_V = [a_1]_X \cdot [a_2]_X \cdots [a_{m-1}]_X \cdot [a_m]_X = [a_1 \cdot a_2] \cdots [a_{m-1} \cdot a_m]_X$ Definisco $N = \overline{F(\pi_1(U \cap V, q))}$ e dunque devo provare tre cose: (1) Φ è suriettivo; (2) $N \subset Ker(\Phi)$; e (3) $Ker(\Phi) \subset N$.

3. Step 1

3.1. Φ è suriettiva

Dimostrazione. Sia $a : I \to X$ un qualsiasi laccio in X con punto base in q. Grazie al lemma di Lebsegue, possiamo scegliere una numero naturale n sufficientemente grande tale che a mappi ogni sottointervallo del tipo $[\frac{(i-1)}{n}, \frac{i}{n}]$ tutto in U oppure in V. (Ecco perché è importante l'ipotesi che U e V siano aperti). Indicando con a_i la restrizione di a ad $[\frac{(i-1)}{n}, \frac{i}{n}]$. La classe di cammini di a in X si fattorizza nel seguente modo: $[a]_X = [a_1 \cdot \cdots \cdot a_n]_X$. Pero abbiamo un problema perché questi cammini a_i non sono dei lacci in generale! Per porre rimedio a questo inconveniente scegliamo per ogni $i = 1, ..., n-1$ un cammini h_i da q ad $a(\frac{i}{n})$. Se $a(\frac{i}{n}) \in U \cap V$, scegliamo h_i come cammino che vive in $U \cap V$; altrimenti, scegliamo dentro nell'insieme che contiene $a(\frac{i}{n})$. (E qui capiamo perché è importante l'ipotesi che $U, V, U \cap V$ siano connessi per archi). Dunque definiamo $\tilde{a_i} = h_{i-1} \cdot a_i \cdot h_i^{-1}$. (Dove h_0 e h_n sono i lacci costanti C_q). Dunque segue facilmente che a si possa fattorizzare anche in questo modo: $[a]_X = [\tilde{a_1} \cdot \cdots \cdot \tilde{a_n}]_X$. E dunque possiamo concludere immediatamente che Φ sia suriettiva. □

4. Step 2

4.1. $N \subset Ker(\Phi)$

Dimostrazione. Se riusciamo a far vedere che $F(\pi_1(U \cap V, q))$ è contenuto nel $Ker(\Phi)$, allora anche la chiusura normale N sarà contenuta nel nucleo perché $Ker(\Phi)$ è normale. Sia $[a]_{U\cap V} \in \pi_1(U \cap V, q)$ un elemento arbitrario. Allora $\Phi \circ F([a]_{U\cap V}) = \Phi((i_*[a]^-_{U\cap V}1 * (j_*[a]_{U\cap V}))) = \Phi([a^-1]_U * [a]_V) = [a^-1 \cdot a]_X = 1$. □

5. Step 3

5.1. $Ker(\Phi) \subset N$

Dimostrazione. Questa è la parte più lunga della dimostrazione. Sia $\gamma = [a_1]_U * [a_2]_V * \cdots * [a_{m-1}]_U * [a_m]_V \in \pi_1(U, q) * \pi_1(V, q)$ un elemento arbitrario del prodotto libero, e supponiamo che $\Phi(\gamma) = 1$. Allora sappiamo che $[a_1 \cdot \cdots \cdot a_k]_X = 1$, o che equivalentemente $a_1 \cdot \cdots \cdot a_k \sim_X C_q$. Dobbiamo mostrare che $\gamma \in N$. Sia $H : I \times I \to X$ l'omotopia di cammini da $a_1 \cdot \cdots \cdot a_k$ ad C_q in X. Nuovamente per il lemma di Lebesgue possiamo suddividere $I \times I$ in quadrati di lunghezza $1/n$ in maniera tale che H mappi ogni quadrato $S_{ij} = [\frac{(i-1)}{n}, \frac{i}{n}] \times [\frac{(j-1)}{n}, \frac{j}{n}]$ tutto dentro U o V. Sia v_{ij} l'immagine del vertice $(i/n, j/n)$ tramite H; e sia a_{ij} la restrizione di H al segmento orizzontale $[\frac{(i-1)}{n}, \frac{i}{n}] \times \frac{j}{n}$, e b_{ij} la restrizione di H al segmento orizzontale $[\frac{i}{n} \times \frac{(j-1)}{n}, \frac{j}{n}]$. La restrizione di H al lato orizzontale inferiore del quadrato $I \times I$, dove $t = 0$, è uguale al prodotto di cammini $a_1 \cdot \cdots \cdot a_k$. Prendendo come n una potenza di due sufficientemente grande, siamo sicuri che gli estremi del cammino a_i in questo prodotto siano della forma i/n, quindi il cammino ottenuto restringendo H al lato inferiore del quadrato $I \times I$ può essere riscritto nel seguente modo: $H_0 \sim a_1 \cdot \cdots \cdot a_k \sim (a_{10} \cdots a_{p0}) \cdots (a_{r0} \cdots a_{n0})$. Nel prodotto libero questo significa che: $\gamma = [a_{10} \cdots a_{p0}]_U * \cdots * [a_{r0} \cdots a_{n0}]_V$. Ora per abbiamo lo stesso problema che avevamo nello step 1 ossia che vorremmo fattorizzare γ nel seguente prodotto $[a_{10}]_U * [a_{20}]_U *$ ... e cosi via. Pero questi sono di nuovo cammini e non lacci in q. Per ovviare a questo problema per ogni i e j scegliamo un cammino h_{ij} da q a v_{ij} che abita in $U \cap V$ se $v_{ij} \in U \cap V$, altrimenti in U oppure in V. Se per caso v_{ij} risulta essere il cammino costante in q allora poniamo h_{ij} uguale al laccio costante in q. Dunque definiamo i seguenti lacci: $\tilde{a_{ij}} = h_{i-1,j} \cdot a_{ij} \cdot h^-_{ij}1$, $\tilde{b_{ij}} = h_{i,j-1} \cdot b_{ij} \cdot h^-_{ij}1$ ognuno dei quali sarà contenuto interamente in U o in V. E quindi γ può essere fattorizzato nel seguente modo: $\gamma = [\tilde{a_{10}}]_U * [\tilde{a_{20}}]U * \cdots [\tilde{a_{n0}}]_V$. L'idea principale della dimostrazione è questo: Vogliamo mostrare che $\gamma \equiv [\tilde{a_{11}}]_U * \cdots * [\tilde{a_{n1}}]_V$ mod N. Ripetendo questo ragionamento possiamo salire all'altra riga, e cosi via per induzione, fino ad ottenere che $\gamma \equiv [\tilde{a_{1n}}]_U * \cdots * [\tilde{a_{nn}}]_V$ mod N. Ma l'intero lato superiore orrizontale del quadrato $I \times I$ è mappato da H nel punto q e cosi quindi $\tilde{a_{in}}$ è uguale al laccio costante C_q, e quest'ultimo prodotto è uguale all'identità. Questo mostra che $\gamma \in N$, e questo conclude la dimostrazione. □

6. Bibliografia

(1) John M. Lee. *Introduction to Topological Manifolds*. Springer, 2019.
(2) Manetti.*Topologia*. Springer, 2016.

Fuori Orario è una giornata di seminari *per studenti da studenti*, giunta quest'anno alla sua settima edizione, organizzata dai giovani del Dipartimento di Matematica dell'Università degli Studi di Milano. Nata per condividere specifici argomenti di notevole bellezza che non trovano spazio all'interno delle lezioni ordinarie, questa iniziativa permette a noi studenti di via Saldini di metterci in gioco. Siamo così stati in grado di condividere le nostre idee e confrontarci dal vivo con grande passione.

Stampato con il contributo dell'Università derivante dai fondi previsti per le attività culturali e sociali.