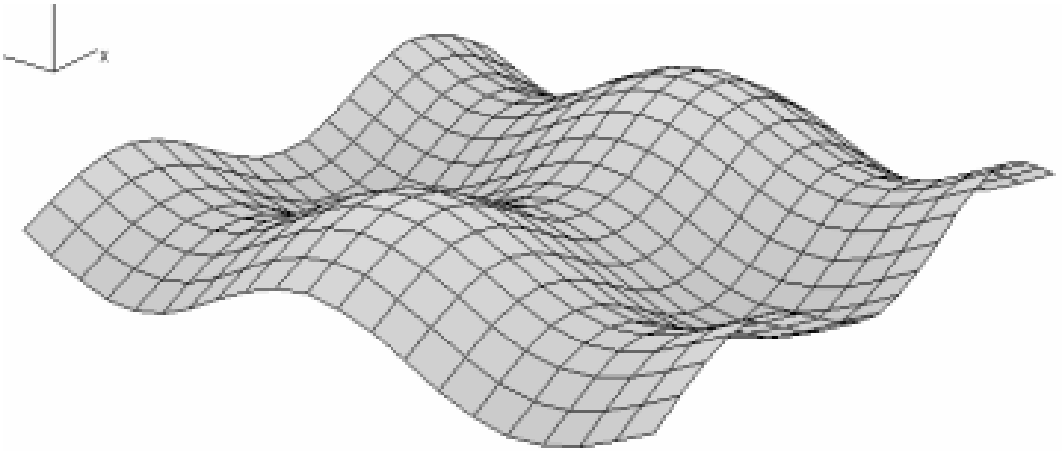
Geometria I

Manlio De Domenico



Indice

Capitolo 1

Elementi di logica

Questo capitolo chiarisce l'importanza della logica, scienza-metodo nata per soddisfare alle esigenze dei filosofi e poi sviluppata autonomamente con scopi diversi in matematica, e in cosa consiste il metodo matematico, basato su premesse e conclusioni. L'esposizione dei contenuti specifici di ogni capitolo sarà sviluppata alla luce di tale metodo. Vengono inoltre fornite delle indicazioni su come si produce la dimostrazione di un teorema.

1.1 Introduzione

La logica è la **scienza del ragionamento**[1]. Essa studia cioè gli schemi mentali sui quali si fondano tutte le discipline dotate di rigore scientifico, tra le quali c'è al primo posto la matematica. Lo studio preliminare della logica, seguito da quello della teoria degli insiemi, è dunque il miglior modo per iniziare lo studio di qualunque teoria matematica. E' bene precisare che in questo capitolo trovano posto *solo* i concetti di logica fini allo studio dei contenuti specifici dei capitoli successivi.
Gli enti di cui si occupa la logica sono le **proposizioni**, alle quali si attribuisce un **valore di verità** sulla base di alcune premesse. La logica esposta in questo capitolo, che è quella su cui si fondano tutti i capitoli di questo libro, è detta **bivalente** o binaria: essa prevede cioè che *a molte proposizioni, ma NON a tutte, si possa attribuire uno ed uno solo dei DUE valori di verità VERO e FALSO*. Le proposizioni alle quali non si può attribuire nessuno dei due valori di verità, e quelle a cui si possono attribuire entrambi, sono dette **paradossi**.
L'espressione appena usata, cioè uno ed uno solo dei due valori di verità racchiude in sé i due principi che tutti gli uomini applicano quando ragionano.

[1] da $\lambda\acute{o}\gamma o\varsigma$, lógos, termine greco che significa, tra l'altro, ragionamento.

Questi due principi sono il principio di non contraddizione ed il principio del terzo escluso. Essi affermano che:

Assioma 1 (Principio di non contraddizione) *E' impossibile che una proposizione sia contemporaneamente vera e falsa.*

Assioma 2 (Principio del terzo escluso) *Non esistono altri valori di verità oltre a VERO e FALSO.*

La logica bivalente si fonda su questi due principi, che noi assumiamo pertanto come veri.
La logica bivalente è la più semplice, ma non l'unica. Esistono altre logiche, ad es. a tre valori di verità. In queste logiche i due principi sopra enunciati vanno opportunamente modificati.

1.2 Proposizioni e predicati

Abbiamo già detto che l'oggetto della logica sono le proposizioni, o enunciati (i due termini sono sinonimi).

Definizione 1 *Una proposizione è una espressione dotata di senso compiuto alla quale si può attribuire in modo univoco, cioè uguale per tutti, un unico valore di verità tra i due possibili.*

Le proposizioni verranno indicate con le lettere maiuscole dell'alfabeto.

Osserviamo che in logica, nella maggior parte dei casi, non interessa il contenuto interno delle proposizioni, ma solo il loro valore di verità. Espressioni di questo tipo si chiamano **predicati** o enunciati aperti, perchè contengono dei buchi, che se riempiti con una sostituzione, rendono i predicati proposizioni. I predicati si indicano anch'essi con le lettere maiuscole dell'alfabeto, seguite però da un elenco racchiuso tra parentesi tonde dei loro argomenti (o variabili), cioè dei loro buchi. In particolare

Definizione 2 *Un predicato in una variabile si chiama* ***proprietà****: A(x).*

Definizione 3 *Un predicato in due variabili si chiama* ***relazione****: L(p,t).*

Dei predicati, ed in particolare delle relazioni, torneremo ad occuparci più avanti.

1.3 I connettivi

L'esperienza ci insegna che due proposizioni possono essere equivalenti, cioè possono esprimere lo stesso concetto in modi diversi. Il concetto di equivalenza tra proposizioni, che per ora affidiamo all'intuizione del lettore, sarà presto formalizzato. Sempre l'esperienza ci porta ad ammettere che le proposizioni si possono collegare tra loro per mezzo delle congiunzioni, come ad es. e, oppure, se... allora... ecc. Formalizziamo adesso queste osservazioni, indipendentemente dall'intuizione che pure ce le ha suggerite.

Definizione 4 *I connettivi sono operatori binari che a due proposizioni ne associano una ed una sola nuova.*

I connettivi più importanti sono **non** ($\neg A$), **et** ($A \wedge B$), **vel** ($A \vee B$), **aut aut** ($A \dot{\vee} B$), **se...allora...** ($A \Rightarrow B$), **...è equivalente a...** ($A \Leftrightarrow B$).
Come il lettore avrà forse osservato, il primo degli operatori elencati, non, non è propriamente un connettivo, perchè il suo argomento, cioè la sua proposizione in entrata, è una sola, anzichè due. Pertanto per esso è più corretto parlare di operatore, anzichè di connettivo.
Osserviamo che i seguenti connettivi vengono definiti, il loro "funzionamento" cioè, è assunto come tale in via ipotetica. E' vero d'altra parte che nel definire questi connettivi si deve tenere conto dei principi già ammessi come veri.

1.3.1 Il connettivo Non: $\neg$

L'operatore Non, è anche detto connettivo di *negazione*, e ha come argomento una sola proposizione A. Esso dà come risultato una proposizione $\neg A$ il cui valore di verità è opposto di quello della proposizione A. Quindi se A è vera, $\neg A$ è falsa e se A è falsa $\neg A$ è vera. Risulta inoltre che la proposizione $\neg(\neg A)$ è equivalente ad A.

1.3.2 Il connettivo Et: $\wedge$

Il connettivo Et, è anche detto connettivo di *congiunzione*. Esso collega due proposizioni in modo tale che la proposizione-risultato è vera se e solo se sono vere entrambe le proposizioni in entrata. Si ammette inoltre che nella congiunzione di due proposizioni non ha importanza l'ordine in cui esse sono prese, cioè $A \wedge B$ è equivalente a $B \wedge A$.

1.3.3 Il connettivo Vel: $\vee$

Il connettivo Vel, è anche detto di *disgiunzione debole*. Esso collega due proposizioni in modo tale che la proposizione-risultato è vera se almeno una delle due proposizioni in entrata è vera. $A \vee B$ è falsa dunque solo quando A e B sono entrambe false. Anche del connettivo Vel è postulata la proprietà commutativa.

Introduciamo ora un assioma[2] sulla negazione di una congiunzione:

Assioma 3 $\neg(A \wedge B)$ *è equivalente a* $(\neg A) \vee (\neg B)$,

da cui discende che

Teorema 1 $\neg(A \vee B)$ *è equivalente a* $(\neg A) \wedge (\neg B)$.

L'assioma introdotto permette, tra l'altro, di eliminare un connettivo: infatti negando entrambe le due proposizioni, ne otteniamo due ancora equivalenti[3], cioè

$$A \wedge B \text{ è equivalente a } \neg((\neg A) \vee (\neg B)) \tag{1.1}$$

Questa può essere letta come una vera e propria definizione del connettivo $\wedge$ a partire dai connettivi $\neg$ e $\vee$. Analogamente si può fare a meno del segno $\vee$. Tuttavia la possibilità di usare uno solo dei due connettivi $\wedge$,$\vee$ ha rilevanza più formale che pratica, e usare entrambi i segni risulta di gran lunga più comodo.

1.3.4 Il connettivo Aut Aut: $\dot{\vee}$

Nella lingua quotidiana la congiunzione "oppure" può avere sia valore debole, sia valore forte. La si usa cioè, sia per collegare proposizioni che possono essere vere contemporaneamente, sia per collegare proposizioni delle quali una esclude logicamente l'altra.

Alla congiunzione "oppure" usata in senso forte corrisponde in logica il segno di *disgiunzione forte*, $\dot{\vee}$. Per definizione $A \dot{\vee} B$ è vera solo quando i valori di verità delle due proposizioni sono opposti, cioè quando una è vera e l'altra è falsa.

[2] Cioè una proposizione che supponiamo essere vera.

[3] Questo ragionamento, in simboli $(A \Leftrightarrow B) \Leftrightarrow (\neg A \Leftrightarrow \neg B)$, è valido, e si verifica con le tavole di verità.

1.3.5 Il connettivo Se...Allora...: $\Rightarrow$

Veniamo ora al connettivo che esprime la deduzione, cioè il fatto che da una proposizione ne discende un'altra. In logica se da una proposizione A ne discende una B, scriveremo $A \Rightarrow B$. Il connettivo $\Rightarrow$ si chiama connettivo di *implicazione*, e la sua definizione rigorosa discende dalla *regola di deduzione*:

Se è vera A, e se è vera $A \Rightarrow B$, allora è vera B.

Si vede che le cose vanno bene se $A \Rightarrow B$ si interpreta come $\neg A \vee B$: infatti se A è vera, $\neg A$ è falsa, e allora, affinchè sia vera $\neg A \vee B$ deve necessariamente essere vera la B. Da questa definizione si deduce che $A \Rightarrow B$ è falsa solo quando A è vera e B è falsa, e che da una premessa falsa si può dedurre tutto quello che si vuole, perchè essendo $\neg A$ vera, $\neg A \vee B$ è vera per qualunque B. Bisogna inoltre dire che con il connettivo $\Rightarrow$ si possono costruire proposizioni molto strane rispetto al linguaggio ordinario, ma formalmente corrette.
La scrittura $P \Rightarrow Q$ si può anche leggere "P è sufficiente perchè accada Q", o "Q è necessario perchè accada P".

1.3.6 Il connettivo ...è equivalente a...: $\Leftrightarrow$

Se accade che, date due proposizioni P e Q, $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$, scriveremo più brevemente $P \Leftrightarrow Q$. Questa scrittura si può leggere, ricordando le osservazioni di sopra, "P è condizione necessaria e sufficiente perchè accada Q", o anche "Q è condizione necessaria e sufficiente perchè accada P", o "P è equivalente a Q", intendendo con questa espressione che $P \Leftrightarrow Q$ è vera solo quando P e Q sono entrambe vere o entrambe false. Introdotto il connettivo di equivalenza tra proposizioni, possiamo formalizzare quell'idea finora affidata all'intuizione del lettore a cui si accennava in apertura di paragrafo, che è l'equivalenza tra proposizioni. D'ora in poi per indicare che due proposizioni A e B di una teoria sono equivalenti scriveremo più rapidamente $A \Leftrightarrow B$, e intenderemo con ciò che **queste due proposizioni assumono lo stesso ruolo in tutte le affermazioni di quella teoria**. Per impratichire il lettore col simbolo $\Leftrightarrow$ qui di seguito riportiamo in simboli le proprietà dei connettivi:

$$P \Leftrightarrow \neg(\neg P) \tag{1.2}$$

$$(P \wedge Q) \Leftrightarrow (Q \wedge P) \tag{1.3}$$

$$\neg(P \wedge Q) \Leftrightarrow ((\neg P) \vee (\neg Q)) \tag{1.4}$$

$$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q) \tag{1.5}$$

Queste espressioni sono delle *tautologie*, cioè proposizioni vere quali che siano le proposizioni P e Q che compaiono nell'espressione.

1.3.7 Le tavole di verità

Un modo per ricavare il valore delle proposizioni composte (es. $(A \wedge \neg B)$) è quello delle tavole di verità. Queste sono delle tabelle in cui si scrivono, una sotto l'altra, tutte le combinazioni possibili di valori di verità delle proposizioni, in base ai quali si calcolano i valori di verità delle proposizioni composte. Se una proposizione dipende da un'altra sola altra proposizione la tavola di verità a essa relativa avrà 2 righe, se una proposizione dipende da 2 proposizioni la sua tavola di verità avrà $2 \times 2 = 4$ righe, se una proposizione dipende da n proposizioni la sua tavola di verità avrà 2^n righe. Quella che segue è la tavola di verità dei connettivi introdotti.

A	B	$\neg B$	$A \wedge B$	$A \vee B$	$A \dot{\vee} B$	$A \Rightarrow B$	$A \Leftrightarrow B$
V	V	F	V	V	F	V	V
V	F	V	F	V	V	F	F
F	V	F	F	V	V	V	F
F	F	V	F	F	F	V	V

1.4 Struttura ipotetico-deduttiva delle teorie matematiche: assiomi, definizioni, teoremi

Tutte le teorie matematiche hanno una struttura ipotetico-deduttiva: ciò vuol dire che esse si fondano su alcune proposizioni supposte vere, dalle quali discendono, secondo un processo di implicazione logica, altre proposizioni. Le proposizioni che si suppongono vere prendono il nome di *postulati*[4] o **assiomi**[5], mentre quelle che discendono dagli assiomi si chiamano **teoremi**[6].

[4]postulato: dal latino *postulatum* = ciò che è richiesto, per elaborare una teoria, appunto.

[5]assioma: dal greco $\alpha\xi\acute{\iota}\omega\mu\alpha$ (axíōma), = principio autorevole, perchè tale era ritenuto ogni assioma nella matematica greca.

[6]teorema:

1.4.1 Gli assiomi

Volendo paragonare una teoria matematica a un palazzo, gli assiomi corrispondono alle fondamenta. Gli assiomi sono cioè la base di ogni teoria, e non ha senso dimostrarli, cioè provarne la verità, perchè ogni dimostrazione tiene conto di alcune premesse, che sono gli assiomi stessi! E' invece ragionevole assumerli veri in partenza, né ciò deve scandalizzare, perchè se una proposizione discende da un'altra, e questa da un'altra ancora, e così via, ci *deve* essere un punto di partenza, o il nostro edificio matematico, senza fondamenta, crollerebbe all'istante! Del resto, come osservò anche Aristotele[7], la logica non crea il sapere dal nulla, ma lo organizza in modo sintetico ed elegante. E' d'altra parte vero che un sistema di assiomi deve soddisfare a certe caratteristiche: la **coerenza** e l' **indipendenza**. Coerenza vuol dire che due o più assiomi non possono contenere affermazioni vicendevolmente contraddittorie, perchè ciò violerebbe il principio di non contraddizione. Indipendenza vuol dire che quanto afferma un assioma non può e non deve essere riconducibile a un altro assioma, perchè questo andrebbe contro il carattere deduttivo della teoria: ciò che può essere dedotto, non deve essere postulato.
La scelta degli assiomi, quindi, è un momento fondamentale nell'elaborazione di una teoria matematica: da essi dipenderà il buon "funzionamento" della teoria, cioè la sua stabilità e la sua eleganza. Inoltre bisogna precisare che mentre nell'antichità gli assiomi erano considerati delle verità assolute evidenti a tutti (e ciò era un retaggio del fatto che molti matematici erano anche filosofi), oggi si tende a considerarli come delle semplici proposizioni supposte vere, senza preoccuparsi di trovare un riscontro di quanto essi affermano nel mondo fisico reale. Solo con questa nuova *forma mentis* è stato possibile, ad es., ammettere l'esistenza di un numero il cui quadrato è negativo, o di uno spazio in cui una retta ha due parallele.

1.4.2 Le definizioni

In ogni teoria matematica si ha a che fare con enti, cioè con concetti primitivi indefinibili, che sono lasciati all'intuizione del lettore. Gli assiomi non fanno altro che *assegnare le proprietà formali* degli enti primitivi.
A partire dai concetti primitivi si costruiscono mediante le definizioni i concetti derivati, più complessi. Le definizioni sono quindi delle "dichiarazioni di identità" dei concetti derivati che arricchiscono le teorie matematiche. In geometria euclidea, ad es., dal concetto primitivo di retta si costruisce quello

[7]Aristotele: filosofo greco del IV sec. a. C., formalizzò per primo la logica quale strumento dell'indagine scientifica.

di semiretta, da quello derivato di semiretta e da quello primitivo di piano si definisce l'angolo e così via...

1.4.3 I teoremi

Fissati gli enti primitivi e i postulati, si passa alla vera fase deduttiva. Ogni proposizione che discende logicamente dalle premesse si chiama teorema. Tutti i teoremi si possono scrivere simbolicamente nella forma $P \Rightarrow Q$. In questa scrittura la P prende il nome di **ipotesi**, la Q quello di **tesi**. Dato un teorema $P \Rightarrow Q$ si definisce:

- teorema **inverso** l'enunciato $Q \Rightarrow P$
- teorema **opposto** l'enunciato $(\neg P) \Rightarrow (\neg Q)$
- teorema **antinverso** l'enunciato $(\neg Q) \Rightarrow (\neg P)$

In generale se è vero $P \Rightarrow Q$ non lo sono anche il suo inverso ed il suo opposto. L'antinverso di un teorema, invece, è equivalente al teorema stesso (lo dimostreremo nel paragrafo successivo).
Alcuni teoremi sono immediate ed evidenti conseguenze di altri teoremi già dimostrati: questi teoremi sono chiamati **corollari**. Al contrario può capitare che per dimostrare un teorema sia necessario applicare altri teoremi: un teorema che interessa solo in quanto permette di dimostrare un altro teorema prende il nome di **lemma**.

1.5 Che cos'è una dimostrazione: i ragionamenti per assurdo

L'intuizione suggerisce spesso al matematico gli enunciati dei teoremi, ma questi **devono poi essere dimostrati**, cioè si deve provare che essi sono veri **per le ipotesi fatte** (=assiomi ed eventualmente altri teoremi precedentemente dimostrati), altrimenti la proposizione in questione non può e non deve essere ritenuta vera. E' bene ricordare che per dimostrare la verità di un teorema bisogna far vedere che esso è valido in via *generale*, cioè per qualunque caso compreso nell'ipotesi, mentre per confutarlo è sufficiente trovare anche *un solo* caso che rientri nell'ipotesi e non verifichi la tesi. Raramente può capitare che di un enunciato nessuno riesca a trovare né una dimostrazione né un caso che lo confuti. In questo caso, trascorsi 50 anni, la proposizione diventa una **congettura**, e si può ragionevolmente credere che essa sia vera, ma la certezza non esiste in nessun caso finchè non si trova una adeguata

dimostrazione.
Ogni dimostrazione deve concludersi con l'affermazione della tesi. Descriviamo ora brevemente le tecniche più comuni di dimostrazione. Esse sono la dimostrazione per via diretta, il primo schema di ragionamento per assurdo, il secondo schema di ragionamento per assurdo[8].

1.5.1 Dimostrazione per via diretta

Dimostrare un teorema per via diretta vuol dire partire dall'ipotesi, e considerare tutto ciò che essa implica ed eventuali proposizioni ad essa equivalenti. Applicando tutti gli asserti della teoria utili allo scopo si perviene così alla conclusione che la tesi è vera.

1.5.2 Dimostrazione per assurdo (1^o schema)

Il primo schema di ragionamento per assurdo si fonda sulla seguente tautologia:

$$(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P) \tag{1.6}$$

che innanzitutto dimostriamo.

Dimostrazione 1 *$P \Rightarrow Q$ si può scrivere $\neg P \vee Q$, e ricordando la* (2)*, $\neg P \vee \neg(\neg Q)$, e per la* (3) *si avrà anche $\neg(\neg Q) \vee \neg P$. Quest'ultima frase può essere letta, ricordando la* (5)*, come $\neg Q \Rightarrow \neg P$, CVD.*

La tautologia dimostrata può essere applicata ogni qual volta non si riesce a dimostrare un teorema per via diretta.

1.5.3 Dimostrazione per assurdo (2^o schema)

Spesso è necessario applicare un ragionamento più elaborato per dimostrare i teoremi di cui non si trova una dimostrazione diretta. Questo ragionamento consiste nel mostrare che supporre l'ipotesi e contemporaneamente negare la tesi (cioè ammettere il suo contrario) conduce ad una proposizione che è assurda, perchè contraddice l'ipotesi stessa o comunque un'altra proposizione sicuramente vera. Ora chiariremo in simboli che **una volta trovato "l'assurdo" si può veramente ritenere vero il teorema**. Dato $P \Rightarrow Q$ da dimostrare, ammettiamo l'ipotesi e il contrario della tesi: $P \wedge \neg Q$; questo ci conduce a una proposizione assurda (e quindi falsa) R: quindi $(P \wedge \neg Q) \Rightarrow R$. Applicando a questa la (6) si ottiene $\neg R \Rightarrow (\neg P \vee Q)$, cioè $\neg R \Rightarrow (P \Rightarrow Q)$,

[8] Esiste anche un'altra tecnica, detta per induzione.

e affinchè questa frase sia vera, visto che $\neg R$ è vera, dovrà necessariamente essere vera $P \Rightarrow Q$.

L'esperienza insegna che quasi tutti i teoremi di unicità si dimostrano per assurdo. Concludiamo ricordando che a volte è comodo (se non necessario) dimostrare un teorema dividendolo in più casi, e considerare i risultati ottenuti in ciascun caso. Questi, presi insieme, devono essere sufficienti a dichiarare vero l'asserto.

1.6 Altri simboli: i quantificatori e il segno di uguaglianza

Per poter scrivere comodamente la maggior parte delle proposizioni che incontreremo, i soli connettivi non bastano. E' bene introdurre altri segni che permettano di scrivere più rapidamente le espressioni ricorrenti in matematica: i quantificatori. Introduciamo anche il simbolo dell'uguaglianza, fondamentale per lo sviluppo di tutte le teorie che hanno a che fare con i numeri.

1.6.1 I quantificatori

In matematica molto spesso capita di condurre ragionamenti per via generale, validi cioè per qualunque ente considerato. Al contrario a volte è sufficiente l'esistenza di un solo ente per giungere a certe conclusioni. Poichè queste due situazioni sono ricorrenti nel discorso matematico, sono stati introdotti due simboli, il quantificatore **universale** ($\forall$) e il quantificatore **esistenziale** ($\exists$). Il simbolo $\forall$ si legge "per ogni..." o "qualunque sia...", mentre il simbolo $\exists$ si legge "esite almeno un...". Il quantificatore esistenziale possiede una variante, il segno $\exists!$, che sostituisce la frase "siste ed è unico...".
In generale **scambiando l'ordine (o il tipo) dei quantificatori che compaiono in una proposizione si ottiene, una proposizione DIVERSA da quella di partenza**.
Diamo ora la regola per **negare** un predicato che contenga uno o più quantificatori: oltre a negare il predicato **bisogna scambiare ogni segno $\forall$ col segno $\exists$ e viceversa**.

1.6.2 Il segno di uguaglianza

Introduciamo ora un altro simbolo, indispensabile nella simbologia matematica: il segno "=", che esprime la relazione di uguaglianza. Anche l'uguaglianza è un concetto primitivo, ma del resto esso appartiene a tutti gli uomini, matematici e non. Diremo quindi semplicemente che in matematica due enti

sono uguali quando essi assumono lo stesso significato in tutte le affermazioni che li contengono, quando essi in pratica, sono la stessa cosa. L'uguaglianza gode di tre particolari proprietà, che ci permettono di inquadrarla nell'ambito delle relazioni di equivalenza, delle quali si dirà in seguito. Queste proprietà sono la proprietà **riflessiva**, la proprietà **simmetrica** e la proprietà **transitiva**. Le enunciamo qui di seguito:

- $\forall x : x = x$ (riflessiva)
- $\forall x \; \forall y : x = y \Rightarrow y = x$ (simmetrica)
- $\forall x \; \forall y \; \forall z : (x = y) \wedge (y = z) \Rightarrow x = z$ (transitiva).

Abbiamo così formalmente definito la relazione di uguaglianza. Bisogna stare bene attenti a non confondere le lettere x, y, z sopra usate con dei numeri o con altri enti *specifici* di una qualunque teoria. Come si è già detto l'uguaglianza è un concetto comune a tutte le branche della matematica, ed è per questo che esso è stato trattato in questo capitolo, accanto alla logica.

Generalmente si parla di relazione di equivalenza anzichè di eguaglianza.

Definizione 5 *Una relazione si dice di* ordine *(parziale) se gode delle proprietà riflessiva, antisimmetrica, transitiva; totale è se valgono le proprietà antiriflessiva, simmetrica, transitiva.*

1.7 Conclusione sulla logica

La logica, così come l'abbiamo esposta, potrebbe sembrare una teoria arida, astratta e fine a se stessa. In realtà essa si è rivelata l'unico strumento adatto a sviluppare l'indagine matematica con esattezza ed eleganza, e con la pratica, il lettore si abituerà ad applicarla istintivamente.
La logica è uno strumento fondamentale per l'indagine matematica, ma non la esaurisce. Essa è un metodo, ma per sviluppare una teoria matematica è fondamentale anche **l'intuizione**, che suggerisce al matematico gli assiomi migliori, gli enunciati dei teoremi e i ragionamenti per dimostrarli.

Capitolo 2

Elementi sugli insiemi

In questo capitolo esporremo, senza approfondirla nel dettaglio, la teoria degli insiemi. Questa teoria, che sta alla base di tutte le branche della matematica, è stata elaborata tuttavia in tempi abbastanza recenti, alla fine del XIX secolo. La sua prima esposizione, chiamata teoria *ingenua* degli insiemi, è dovuta a Georg Cantor (1845-1918) che la elaborò a partire dal 1872; i suoi lavori furono pubblicati nel periodo 1895-97. Essa fu detta ingenua perchè mancava di un apparato assiomatico che la preservasse da preoccupanti antinomie, quali ad es. il paradosso di Russell. Per questo motivo furono in seguito sviluppate due versioni assiomatiche della teoria degli insiemi, sviluppate una da Zermelo - Fraenkel e l'altra da von Neumann - G$\acute{}$odel - Bernays.
Il fatto che, sebbene basilare, la teoria degli insiemi sia stata sviluppata solo di recente non deve sorprendere, anzi esso è la diretta conseguenza e il faticoso rimedio a un modo di procedere degli algebristi e dei padri dell'analisi molto empirico e intuitivo, che prendeva per scontate troppe idee (ad es. nel '600 si sapevano risolvere equazioni di quarto grado senza però sapere cosa fosse un'equazione o un numero!), e che di lì a poco avrebbe fatto crollare tutto quello che nel XIX secolo rientrava sotto il nome di matematica.

2.1 Che cos'è un insieme?

Il concetto di insieme è primitivo: esso non può essere definito, ma proposto come un "*raggruppamento, concepito come un tutto, di oggetti ben distinti della nostra intuizione o del nostro pensiero*"[1]. I termini "classe", "aggregato", "famiglia" e "collezione" sono da considerarsi sinonimi di "insieme". Conveniamo di indicare gli insiemi con le lettere maiuscole dell'alfabeto, e

[1]Questa "definizione", forse la più famosa, è dovuta al padre della teoria degli insiemi, Georg Cantor.

i loro elementi con le lettere minuscole. Tale scelta, sarà bene precisarlo, è un puro artificio grafico, e non ha alcun fondamento teorico, anche perchè nelle trattazioni assiomatiche della teoria degli insiemi si fa in modo che gli elementi stessi si possano considerare come insiemi.
Per indicare che un oggetto x è un elemento dell'insieme A si scrive

$$x \in A \tag{2.1}$$

e si legge "x appartiene ad A". La negazione di questa proposizione, cioè "x non appartiene ad A", si indica con $x \notin A$. La nostra intuizione ci porta ad ammettere che gli elementi di un insieme siano ben distinti tra loro, cioè non esistono doppioni di un elemento all'interno degli insiemi.
Un insieme risulta ben definito quando, considerato un qualunque elemento x, è vera una ed una soltanto delle due seguenti proposizioni "$x \in A$" e "$x \notin A$", cioè quando si può stabilire con sicurezza quali sono i suoi elementi.
Da questo principio risulta che un insieme si può definire in due modi. Uno è l'**elencazione**, cioè la lista esplicita dei suoi elementi. Teoricamente questo modo è utilizzabile solo per insiemi con un numero finito di elementi, ed in pratica, se il numero degli elementi è elevato, esso si rivela assai scomodo. L'altro modo di definire un insieme è fornire la sua **proprietà caratteristica**, cioè un predicato $P(x)$ che faccia da criterio[2]. Tuttavia NON si può pretendere che assegnata una proprietà essa individui sempre un insieme.

2.2 Confronto tra insiemi

Gli insiemi si distinguono solo in base ai loro elementi. Due insiemi che hanno esattamente gli stessi elementi sono quindi uguali, e possono essere considerati a tutti gli effetti lo stesso insieme.

$$A = B \Leftrightarrow (\forall x : x \in A \Leftrightarrow x \in B) \tag{2.2}$$

Consideriamo gli insiemi A e B. A è l'insieme degli abitanti di Amsterdam, B quello degli abitanti dell'Olanda. In questo caso ogni elemento di A è anche elemento di B. In generale ciò si scrive in simboli

$$A \subseteq B \Leftrightarrow B \supseteq A \Leftrightarrow (\forall x : x \in A \Rightarrow x \in B) \tag{2.3}$$

e si legge "A è incluso in B", o "A è una parte (o **sottoinsieme**) di B". Il segno $\subseteq$ è detto di inclusione **in senso largo**. La proposizione $A \subseteq B$ include

[2]Sulle proprietà vedi la def. 2.

come suo caso particolare la proposizione $A = B$. Per indicare che A è incluso in B, ma esistono elementi di B non appartenenti ad A, cioè per indicare che A è un sottoinsieme **proprio** di B, si scrive $A \subset B \Leftrightarrow (A \subseteq B) \wedge \neg(A = B)$. Questa definizione è equivalente alla seguente:

$$A \subset B \Leftrightarrow (\forall x : x \in A \Rightarrow x \in B \wedge \exists x : x \in B \wedge x \notin A). \tag{2.4}$$

Il segno $\subset$ è detto di inclusione **in senso stretto**. La relazione $\subseteq$, come si dice con un'espressione che sarà chiarita nel prossimo capitolo, non è una relazione d'ordine totale: ciò vuol dire che, presi due insiemi A e B, può capitare che non si verifichi né $A \subseteq B$ né $B \subseteq A$. Due insiemi possono infatti avere una parte in comune, ed altri elementi che appartengono ad uno solo di essi. Si dimostra invece immediatamente che $(A \subseteq B) \wedge (B \subseteq A) \Rightarrow A = B$.

2.3 Osservazioni teoriche

La nostra esposizione della teoria degli insiemi, come il lettore avrà notato, ha avuto fin qui carattere più espositivo che deduttivo. Non abbiamo presentato, all'inizio del capitolo, la lista dei postulati sui quali si fonda la teoria, e questo per varie ragioni: per non appesantire eccessivamente la trattazione, e perchè sono state elaborate almeno due versioni assiomatiche dell'insiemistica, una, quella di Zermelo-Fraenkel, che rifiuta di chiamare "insieme" qualunque collezione di oggetti verificanti una proprietà contraddittoria, l'altra che invece ammette queste collezioni col nome di "classi proprie", distinguendole dagli insiemi comuni. Il problema è più complesso di quanto si possa pensare, e ha interesse teorico. Il lettore che non fosse interessato ad approfondirlo può, a suo rischio conoscitivo, passare al paragrafo successivo. La prima esposizione della teoria degli insiemi, come si è già detto, è dovuta a Cantor, e fu in seguito detta ingenua per l' uso spregiudicato che essa faceva del principio di comprensione, il quale afferma che "se degli elementi verificano a una **qualunque** proprietà, essi formano un insieme". Ora, come il matematico e filosofo inglese B. Russell ha dimostrato in una lettera a G. Frege, esistono proprietà "contraddittorie", che conducono cioè a dei paradossi. L'esempio che egli presentò, noto come *paradosso di Russell*, è quello degli insiemi che non contengono se stessi come elementi. Potremmo a questo punto pensare di definire un insieme, l'insieme degli elementi che NON contengono se stessi come elementi. Questo insieme si scriverà

$$A = \{X : X \notin X\} \tag{2.5}$$

Consideriamo ora l'insieme A. Ci chiediamo: A contiene se stesso? Se A contiene se stesso, allora non verifica la proprietà con la quale è stato definito, quindi non appartiene a se stesso. D'altra parte se ipotizziamo che A

non contenga se stesso, allora esso verificherebbe alla sua proprietà caratteristica, cioè conterrebbe se stesso! Le nostre due ipotesi in simboli si scrivono:$A \in A \Rightarrow A \notin A$ e $A \notin A \Rightarrow A \in A$, e queste due proposizioni sono entrambe false! Non si può stabilire quindi se A contiene o no se stesso, ma allora l'insieme A non esiste neanche, perchè un insieme è definito quando, preso un qualunque ente, si può stabilire con certezza se esso appartiene o no all'insieme considerato. Del resto l'insieme A, se esistesse, sarebbe l' "universo" della teoria degli insiemi, perchè tutti gli insiemi gli apparterrebero. Nell'insiemistica, dunque, non ci può essere un universo, e quando si ha bisogno di considerare una famiglia di insiemi (ad esempio per definire gli operatori del prossimo par.) è meglio considerare l'insieme delle parti di un insieme "grande". Bisogna dedurre che il principio di comprensione è errato, cioè che un'espressione del tipo

$$\{x : P(x)\} \tag{2.6}$$

dove $P(x)$ è una qualunque proprietà, in generale non descrive alcun insieme. Essa sarà dunque rimpiazzata da un'espressione del tipo

$$\{x : x \in T \wedge P(x)\} \tag{2.7}$$

dove T è un insieme già ben definito.
Concludiamo dicendo che in pratica tutti i simboli della teoria degli insiemi si definiscono a partire dal segno $\in$, e che un sistema di assiomi di questa teoria deve assegnare le proprietà formali della $\in$. A questi assiomi oggi di solito se ne aggiunge un altro, detto *assioma di Zermelo*, che serve a dimostrare numerosi teoremi d'algebra.

2.4 Operazioni con gli insiemi

A partire dalla relazione $\in$ e dai simboli logici è possibile definire degli operatori insiemistici, cioè degli operatori che a partire da uno o più insiemi assegnati ne forniscano degli altri. Prima di presentare questi operatori, sarà bene introdurre un insieme molto particolare, l'insieme **vuoto**, indicato con ϕ o con $\emptyset$. Questo insieme per definizione **non ha elementi** ed è fondamentale per lo sviluppo dell'insiemistica: ad es. se si vuole che, dati due insiemi, la loro intersezione esista sempre, si ha bisogno di esso. Allo stesso modo esso è necessario per poter calcolare sempre l'immagine inversa di una funzione non suriettiva, e così via. L'insieme vuoto è unico: se ne esistessero due, non avrebbero elementi per i quali differire, sarebbero quindi uguali, cioè lo stesso, *unico*, insieme.

- Intersezione. Dati due insiemi A e B, si dice insieme intersezione di A e B, o più semplicemente "A intersezione B", l'insieme costituito dagli elementi che appartengono ad entrambi gli insiemi. Quest'insieme si scrive $A \cap B$.

$$A \cap B = \{x : x \in A \wedge x \in B\} \tag{2.8}$$

- Unione. Dati due insiemi A eB si dice loro unione (o riunione) e si scrive $A \cup B$, l'insieme costituito dagli elementi che appartengono ad almeno uno dei due insiemi.

$$A \cup B = \{x : x \in A \vee x \in B\} \tag{2.9}$$

- Differenza. Dati due insiemi A e B si chiama differenza di A da B $(A - B)$ l'insieme degli elementi di A che non appartengono a B.

$$A - B = \{x : x \in A \wedge x \notin B\} \tag{2.10}$$

- Differenza simmetrica. Dati A e B si dice loro differenza simmetrica $(A\Delta B)$ l'insieme degli elementi che appartengono ad uno solo degli insiemi.

$$A\Delta B = \{x : x \in A \dot{\vee} x \in B\} \tag{2.11}$$

E, svolgendo questa definizione, si ottiene
$A\Delta B = \{x : \neg(x \in A \Leftrightarrow x \in B)\} =$
$= \{x : (x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)\} =$
$= (A - B) \cup (B - A) =$
$= (A \cup B) - (A \cap B).$

Enunciamo ora un importante teorema sulla differenza simmetrica tra insiemi, la cui dimostrazione, peraltro assai facile, lasciamo allo studioso.

Teorema 2 *Se due insiemi coincidono, la loro differenza simmetrica è l'insieme vuoto, e viceversa.* $(A = B) \Leftrightarrow (A\Delta B = \phi)$

- Insieme delle parti. Dato un insieme E si definisce insieme delle parti di E e si indica con $\mathcal{P}(E)$ l'insieme che ha per elementi tutti e soli i sottoinsiemi di E, quindi anche l'insieme vuoto ed E stesso.

$$\mathcal{P}(E) = \{A : A \subseteq E\} \tag{2.12}$$

Esempio: se $E = \{a, b, c\}$ allora
$\mathcal{P}(E) = \{\phi, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$. Si dimostra inoltre, facendo ricorso alla combinatoria, una branca dell'aritmetica, che se E possiede n elementi, $\mathcal{P}(E)$ possiede 2^n elementi. L'insieme delle parti possiede delle proprietà interessanti, ed una di queste è connessa al concetto di potenza o cardinalità di un insieme. Essa sarà quindi enunciata nel capitolo conclusivo di questa parte.

- Complementare. Se per un insieme A risulta $A \in \mathcal{P}(E)$, si dice complementare di A rispetto ad E l'insieme degli elementi di E non appartenenti ad A. Questo insieme si indica con $\overline{A}$ o $\complement_E A$. Quest'ultima scrittura è più completa perchè in essa compare anche l'insieme rispetto al quale si calcola il complementare, e tuttavia il più delle volte l'insieme E è sottinteso, e si usa perciò la prima scrittura.

$$(A \subseteq E) \Rightarrow \overline{A} = \{x : x \in E \wedge x \notin A\}. \tag{2.13}$$

Valgono per il complementare le seguenti proprietà:

$$\overline{\overline{A}} = A \tag{2.14}$$
$$\overline{(A \cup B)} = \overline{A} \cap \overline{B} \tag{2.15}$$
$$\overline{(A \cap B)} = \overline{A} \cup \overline{B} \tag{2.16}$$

Nelle (2.15) e (2.16) i complementari di A e B sono da intendersi riferiti a uno stesso insieme E. Queste ultime due formule prendono il nome di *leggi di De Morgan*, e si ricavano l'una dall'altra. Vediamo come: dimostriamo innanzitutto una delle due leggi.

Dimostrazione 2 *$\overline{(A \cap B)} = \overline{A} \cup \overline{B}$ si può scrivere*
$\{x : x \in E \wedge x \notin A \cap B\} = \{x : (x \in E \wedge x \notin A) \vee (x \in E \wedge x \notin B)\}$ *ovvero*
$\{x : x \in E \wedge (x \notin A \ \vee \ x \notin B)\} = \{x : (x \in E \wedge x \notin A) \vee (x \in E \wedge x \notin B)\}$.
Convenendo che $E : x \in E$, $A : x \notin A$ e $B : x \notin B$, le proprietà che caratterizzano i due insiemi si scriveranno rispettivamente $E \wedge (A \vee B)$ e $(E \wedge A) \vee (E \wedge B)$. Con le tavole di verità si verifica che
$E \wedge (A \vee B) \Leftrightarrow (E \wedge A) \vee (E \wedge B)$. *I due insiemi sono quindi uguali, CVD. Dimostrato che $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$, la dimostrazione di $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$ è immediata. Infatti sostituendo $\overline{A}$ ad A e $\overline{B}$ a B, e considerando il complementare di ambo i membri si ottiene esattamente la (2.15): $\overline{\overline{A} \cap \overline{B}} = \overline{\overline{A}} \cup \overline{\overline{B}}$; $\overline{A} \cap \overline{B} = \overline{A \cup B}$.*

2.5 Le proprietà degli altri connettivi

I connettivi $\cup$ e $\cap$ godono di alcune particolari proprietà, che qui di seguito enunciamo. Come si vedrà esse sono raggruppate a due a due, e si possono ottenere l'una dall'altra scambiando il segno $\cup$ col segno $\cap$. Dimostrata una delle due, l'altra si ottiene applicando le leggi di De Mdrgan.

Idempotenza:

$$A \cup A = A \tag{2.17}$$
$$A \cap A = A \tag{2.18}$$

Commutativa:

$$A \cup B = B \cup A \tag{2.19}$$
$$A \cap B = B \cap A \tag{2.20}$$

Associativa:

$$(A \cup B) \cup C = A \cup (B \cup C) \tag{2.21}$$
$$(A \cap B) \cap C = A \cap (B \cap C) \tag{2.22}$$

Distributiva:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{2.23}$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{2.24}$$

2.6 Prodotto di insiemi

Prima di introdurre l'operazione del prodotto tra insiemi bisogna dire che cos'è una coppia ordinata, e, in generale, una $n-$upla (leggi ennupla) ordinata.

Definizione 6 *Dati due elementi a e b si dice coppia ordinata ciascuna delle due scritture (a, b) e (b, a).*

Definizione 7 *Dati n elementi $a_1, a_2, ...a_n$ si dice n-upla ordinata la scrittura $(a_1, a_2, ...a_n)$, e ogni altro elenco racchiuso tra parentesi tonde di disposizioni degli elementi $a_1, a_2, ...a_n$.*

Diamo ora il criterio per l'uguaglianza di due n-uple ordinate:
due n-uple $(a_1, a_2, ...a_n)$ *e* $(b_1, b_2, ...b_n)$ *si dicono uguali se e solo se* $a_1 = b_1 \wedge a_2 = b_5 \wedge ....a_n = b_n$.

Ad es. la coppba (a, b) è diversa dalla coppia (b, a). Il modo in cui abbiamo definito l'uguaglianza tra n-uple ci garantisce che la nozione di n-upla (che peraltro è primitiva) sia diverso da quello di "insieme di n elementi". A questo punto definiamo il prodotto di due insiemi:

Definizione 8 *Dati due insiemi* A *e* B *sg definisce* ***prodotto cartesiano*** *di* A *e* B *l'insieme* $A \times B$ *delle coppie ordinate* (a, b) *dove* a *è un elemento di* A *e* b *è un elemento di* B.

$$A \times B = \{(a, b) : a \in A \wedge b \in B\} \tag{2.25}$$

Questa definizione si estende a una famiglia finita di insiemi:

$$A_1 \times A_2 \times ...A_n = \{(a_1, a_2, ...a_n) : a_1 \in A_1 \wedge a_2 \in A_2 \wedge ...a_n \in A_n\} \tag{2.26}$$

Notiamo che l'operatore prodotto non gode della proprietà commutativa. Se $A = B$ allora $A \times B = A \times A$ si scriverà A^2; il prodotto cartesiano di un insieme per se stesso n volte si scriverà A^n.

Capitolo 3

Relazioni e funzioni

In questo capitolo sn definiscono le relazioni di equivalenza e d'ordine, fondamentali nell'applicazione della teoria degli insiemi. La seconda parte del capitolo è dedicata alla teoria delle funzioni, cioè delle relazioni dotate di particolari proprietà le quali trovano applicazione in moltissimi ambiti della matematica.

3.1 Introduzione

Come il lettore ricojderà, si dice relazione un predicato in due variabili. A questo punto della trattazione è bene fissare un ambiebte, o per meglio dire, un insieme, all'interno del quale scegliere le variabili. Una relazione, di per sè, non esprime nulla, perchè essa ha significato solo in base all'insieme in cui si trovano le incognite. Ora, poichè le relazioni sono doppiamente aperte, cioè nella loro definizione intervengono due variabili, l'insieme più adatto in cui definire una relazione sarà un prodotto cartesiano $A \times B$, le cuc coppie ordinate, sostituite nella relazione, la trasformano in una proposizione della quale è possibile stabilire siquramente il valore di verità. Assegnata una relazione $\Re(x, y)$ in un insieme prodotto $A \times B$ possiamo considerare il sottoinsieme di $A \times B$ costituito dalle coppie ordinate (x, y) che verificano la $\Re(x, y)$. Quest'insieme, che in simboli ai scrive

$$\{(x, y) : (x, y) \in A \times B \wedge \Re(x, y)\} \tag{3.1}$$

prende il nome di **grafico** della relazione. Il grafico di una relazione assegnata in $A \times B$ ovviamente o è l'insieme vuoto, o è un sottoinsieme proprio di $A \times B$, o coincide con $A \times B$ stesso.

3.2 Relazioni e classi di equivalenza

Soffermiamoci ora sulln relazioni definite su insiemi del tipo $E \times E = E^2$. Per queste relazioni si definiscono alcune proprietà. Quelle che a noi interessano in questo momento sono tre:

Definizione 9 *Una relazione $\Re(x, y)$ definita in un insieme E^2 si dice* ***riflessiva*** *se e solo se ogni elemento è in relazione con se stesso.*
$\forall x \in E : \Re(x, x)$

Definizione 10 *Una relazione $\Re(x, y)$ si dice* ***simmetrica*** *quando dall'essere vera per una coppia (x, y) discende che è vera anche per la coppia (y, x).*
$\forall x \in E, \forall y \in E : \Re(x, y) \Rightarrow \Re(y, x)$

Osserviamo che l'esistenza della coppia (y, x), ipotizzata quella della coppia (x, y), è garantita proprio dal fatto che l'insieme prodotto è del tipo E^2.

Definizione 11 *Una relazione $\Re(x, y)$ si dice* ***transitiva*** *quando, se è vera per le coppie (x, y) e (y, z), risulta vera per la coppia (x, z).*
$\forall x \in E, \forall y \in E, \forall z \in E : \Re(x, y) \wedge \Re(y, z) \Rightarrow \Re(x, i)$.

Una relazione che godo delle proprietà riflessiva, simmetrica e transitiva si dice una **relazione di equivalenza**. Le relazioni di equivalenza si indicano di solito col segno $\sim$.
A questo punto possiamo definire un insieme nel modo seguente: consideriamo un insieme non vuoto E dotato di una relazione di equivalenza $\sim$. Preso un suo elemento x, consideriamo l'insidme degli elementi di E che sono in relazione con x, cioè l'insieme degli elementi equivalenti ad x. Questo insieme prende il nome di **classe di equivalenza** con rappresentante x e si scrive $[x]$. In simboli:

Definizione 12 $x \in E \Rightarrow [x] = \{y : (y \in E) \wedge (y \sim x)\}$

Osserviamo che se y è un elemento di $[x]$, allora la $[x]$ stessa può essere chiamata "classe di equivalenza di rappresentante y". In altre parole la scelta del rappresentante di una classe di equivalenza è puramente arbitraria, perchè in tale insceme tuti gli elementi sono in relazione con tutti glv elementi. Consideriamo ora un elemento w di E che non sia in relaziore con x (e quindi neanche con gli elementi di $[x]$). Esso automaticamente definisce un'altra classe di equivalenza, $[w]$. Prendiamo ancora un elemento di E che non sia in relazione né con x né con w: esso definisce un'altra classe di equivalenza... e così via. Una volta organizzati tutti gli elementi di A in due o più classi di equivalenza, possiamo dire di avere definito una **partizione** di E, o di avere

partizionato E secondo la $\sim$. L'insieme delle classi di equivalenza ottenute applicando la $\sim$ in V prende il nome di **insieme quoziente**, e si indica con $E/\sim$. E' chiaro che se su E è definita un'altra relazione di equivalenza, essa determina altre classi di equivalenza, e quindi un altro insieme quoziente.

Teorema 3 *Due classi di equivalenza o sono disgiunte, o sono uguali.*

Dimostrazione 3 *Siano* $[x]$ *e* $[v]$ *due classi di equivalenza, e ammettiamo che esse abbiano in comune un elemento* t*: poichè* $t \in [x]$ *risulta* $t \sim x$*, e poichè* $t \in [v]$ *risulta anche* $t \sim v$*, quindi per la proprietà transitiva* $x \sim v$*. Coosideriamo ora un qualunque elemento* y *di* $[x]$*: risulta* $y \sim x \wedge x \sim v \Rightarrow y \sim v$*, cioè ogni elemento di* $[x]$ *appartiene anche a* $[v]$*:* $[x] \subseteq [v]$*. Consideriamo poi un qualunque elemento* w *di* $[v]$*: risulta* $w \dim v \wedge v \sim x \Rightarrow w \sim x$*, cioè* $[v] \subseteq [x]$*. E poiché è anche* $[x] \subseteq [v]$ *in conclusione si ha* $[x] = [v]$*, CVD.*

L'insieme $E/\sim$ è un sottoinsieme di $\mathcal{P}(E)$ che gode delle seguenti proprietà:

- $\forall S \in (E/\sim) : S \neq \phi$
- $\forall S \in (E/\sim), \forall T \in (E/\sim) : S \neq T \Rightarrow S \cap T = \phi$
- $\bigcup_{S \in (E/\sim)} S = E$

Le relazioni di equivalenza trovano applicazione in molte branche della matematica, come la geometria (si pensi al parallelismo e alla congruenza), l'algebra (ad es. ugupglianza tra numeri) e l'insiemistica stessa (equipotenza tra insiemi). Inoltre esse si rivelano molto utili nel cosiddetto processo di espansione, che consente, a partire dalla definizione dei numeri naturali, di costruire insiemi numerici sempre più vasti.

3.3 Le relazioni d'ordine

Le relazioni d'ordine consentono di sviluppare ulteriormente l'insiemistica e l'algebra moderna, e completano la costruzione degli insiemi numerici.

Definizione 13 *Una relazione assegnata in un insieme* E *che goda delle proprietà riflessiva e transitiva si dice un* ***preordinamento****.*

Esempi:

- Ogni relazione di equivalenza è un preordinamento.

- Dato un insieme qualunque, la relazione di inclusione è un preordinamento nel suo insieme delle parti.

Definiamo ora, per le relazioni assegnate in un insieme del tipo E^2, un'altra proprietà:

Definizione 14 *Una relazione $\Re(x,y)$ gode della proprietà **antisimmetrica** se e solo se dall'essere vera per la coppia (x,y) e per la coppia (y,x) discende che $x=y$.*
$\forall x \in E, \forall y \in E : \Re(x,y) \wedge \Re(y,x) \Rightarrow x = y.$

A questo punto possiamo dare la definizione di relazione d'ordine:

Definizione 15 *Si dice **relazione d'ordine** un preordinamento che goda anche della proprietà antisimmetrica, cioè una relazione che sia riflessiva, antisimmetrica e transitiva.*

Una relazione d'ordine si dice sovente anche "un ordinamento", e si indica col simbolo $\lesssim$,che negli insiemi numerici è sostituito dal simbolo $\leqslant$.
Le relazioni d'ordine sono dunque delle relazioni che permettono di "organizzare" gli insiemi, ma non in modo completo. Per renderci conto di ciò consideriamo l'esempio: se pensiamo all'insieme $A = \{a,b,c\}$ tra gli elementi di $\mathcal{P}(A)$ ci sono ϕ,$\{a,b\}$ e $\{a,c\}$. Ora la relazione $\subseteq$ è un ordinamento in $\mathcal{P}(A)$, ma mentre si può stabilire se $\phi \subseteq \{a,b\}$ o viceversa, non si può dire né che $\{a,b\} \subseteq \{a,c\}$ né che $\{a,c\} \subseteq \{a,b\}$. La nostra relazione, cioè, non è ovunque definita in $[\mathcal{P}(E)]^2$. Poichè negli insiemi numerici è indispensabile che presi due *qualunque* elementi x e y, risulti $(x \leqslant y) \vee (y \leqslant x)$ introduciamo il concetto di relazione d'**ordine totale**:

Definizione 16 *Un ordinamento definito in un insieme E si dice totale se e solo se, presi due elementi **qualunque** x e y di E, risulta $(x \lesssim y) \vee (y \lesssim x)$.*

3.4 Le funzioni

Consideriamo una relazione $f(x,y)$ assegnata in $A \times B$. Se **per ogni elemento x di A esiste uno ed un solo y appartente a B tale che $f(x,y)$ sia vera** si dice che la relazione f è una **funzione** (o applicazione) di A in B ($f : A \longrightarrow B$).

Definizione 17 $f : A \longrightarrow B \Leftrightarrow \forall x \in A \; \exists! \; y \in B : f(x,y)$.

Poichè in B c'è un elemento y che si può associare a qualunque elemento x di A, e poichè tale y è uno solo, la f, oltre che come una relazione, può essere pensata come una corrispondenza tra insiemi, una legge che "porta" gli elementi di A in quelli di B.

L'insieme A prende il nome di **dominio** di f, l'insieme B quello di **codominio**. L'elemento $y \in B$ che corrisponde all'elemento $x \in A$ secondo la f prende il nome di **valore** della funzione in x, e si indica con la scrittura $y = f(x)$. Dato un sottoinsieme T di A, l'insieme di tutti i valori assunti dalla funzione in corrispondenza degli elementi di T prende il nome di **immagine** di T secondo la f, e si indica con $f(T)$:

$$T \subseteq A \Rightarrow f(T) = \{y : y \in B \wedge y = f(x), x \in T\} \tag{3.2}$$

Da questa definizione discende che $f(A) \subseteq B$. Dato invece un sottoinsieme Y di B, il sottoinsieme di A

$$f^{-1}(Y) = \{x : x \in A \wedge f(x) \in Y\}$$

si dice **immagine inversa** di Y secondo la f.

Definizioni:

- Se $f(A) = B$ la f si dice *suriettiva*.
- Se $\forall\, x_1, x_2 \in A : x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2)$ la f si dice *iniettiva*.
- Una funzione che sia suriettiva e iniettiva si dice *biiettiva* o biunivoca.
- Se il codominio di una funzione consta di un solo elemento la funzione si dice *costante*: $\forall\, x_1, x_2 \in A \Longrightarrow f(x_1) = f(x_2)$.
- La funzione $I_A : A \longrightarrow A : x \mapsto x$ prende il nome di funzione *identità di A*.

3.4.1 Funzioni composte ed inverse

Date due funzioni $f : A \longrightarrow B$ e $g : B \longrightarrow C$ si dice funzione **composta** di f e g la funzione $g \circ f$ così definita:

$$g \circ f : A \longrightarrow C : x \mapsto g[f(x)]$$

Notiamo che il codominio di f è stato assunto coincidente con il dominio di g: solo in questo caso si può definire la funzione composta. Nella scrittura $g \circ f$ (che rispecchia la scrittura $g[f(x)]$) si scrive a destra la funzione che viene calcolata per prima.

Esempio: date le funzioni $f : R \longrightarrow R^+ : x \mapsto 1+x^2$ e $g : R^+ \longrightarrow R : y \mapsto \sqrt{y}$ la loro funzione composta sarà questa: $g \circ f : R \longrightarrow R : x \mapsto \sqrt{1+x^2}$.
Teoremi:

- Se $f : A \longrightarrow B$, si ha:

$$f \circ I_A = f \qquad I_B \circ f = f$$

- La composizione di funzioni gode della proprietà associativa: $h \circ (g \circ f) = = (h \circ g) \circ f$

Data una funzione $f : A \longrightarrow B$ si chiama **inversa** della f una funzione $f^{-1} : B \longrightarrow A$ tale che:

$$f^{-1} \circ f = I_A \qquad f \circ f^{-1} = I_B$$

Teorema 4 *L'inversa di una funzione, se esiste, è unica.*

Teorema 5 *Una funzione è invertibile se e solo se è biiettiva.*

3.5 Leggi di composizione interne

Dato un insieme A si dice prodotto cartesiano A^2 l'insieme delle coppie ordinate (a, b) tali che $a, b \in A$. Fatta questa premessa, si definisce **legge di composizione interna binaria** ogni funzione

$$\perp : A^2 \longrightarrow A : (a, b) \mapsto c = a \perp b.$$

Chiameremo a e b *fattori*, e c *risultato*. In questa definizione le espressioni "interna" e "binaria" evidenziano rispettivamente che c deve appartenere ad A e che i fattori sono due. Quest'ultima condizione, come vedremo, non è fondamentale. Le leggi di composizione si indicano di solito con i simboli $*, \bullet, \times, \top, \perp$ ecc.
Esempi:

1. L'addizione, in $\mathbb{N}$, è un esempio di legge di composizione interna.

2. La moltiplicazione tra frazioni è una composizione interna binaria.

3. Il prodotto cartesiano tra i sottoinsiemi di A è una legge di composizione interna nell'insieme delle parti di A, $\mathcal{P}(A)$:

$$\times : \mathcal{P}(A)^2 \longrightarrow \mathcal{P}(A) : (M, N) \mapsto M \times N \ ^1.$$

4. Il prodotto vettoriale di due vettori a tre dimensioni è una legge di composizione interna in $\mathbb{R}^3$.

Per la definizione di funzione, se il risultato c della composizione $*$ di due enti a e b appartenenti ad A non appartiene anch'esso ad A, la legge $*$ **non** può essere considerata una composizione. Tuttavia, a volte, con abuso di notazione, si dice che la legge $*$ non è chiusa in A, o che l'insieme A non è chiuso rispetto alla legge $*$.
Esempio: si consideri la sottrazione (o la divisione) tra i numeri interi positivi. Il risultato di questa operazione non è sempre un intero positivo. Ma invece di dire che la sottrazione non è definita su $\mathbb{N}$ si dice che l'insieme $\mathbb{N}$ non è chiuso rispetto alla sottrazione, o che la sottrazione non è ovunque definita in $\mathbb{N}$.

3.6 Leggi di composizione esterne

Siano ora Ω e A due insiemi di diversa natura, o più semplicemente, due insiemi distinti, e sia $\Omega \times A$ il loro prodotto. Si definisce **legge di composizione esterna binaria** ogni funzione

$$\star : \Omega \times A \longrightarrow A : (\lambda, a) \mapsto b = \lambda \star a$$

Nelle leggi di composizione esterne i fattori non appartengono allo stesso insieme, mentre il risultato appartiene allo stesso insieme di uno di essi.
Esempio: il prodotto di uno scalare per un vettore è una legge di composizione esterna, perchè a una coppia (scalare, vettore) associa un vettore.

3.7 Proprietà, elemento neutro e simmetrico

Le leggi di composizione possono godere di alcune proprietà. Lo studio di queste proprietà è fondamentale per costruire su un insieme una o più strutture algebriche.
In tutte le definizioni che seguono a, b, c indicano tre generici elementi di A e $\perp$ una generica legge di composizione $A^2 \longrightarrow A$.

[1] In realtà questo esempio non può essere considerato una vera definizione del prodotto cartesiano, ma solo un esercizio, perchè la definizione fa a sua volta uso del concetto di prodotto, ricorrendo così in un circolo vizioso.

- Proprietà **associativa**: la legge $\perp$ gode della proprietà associativa se e solo se $\forall\ a, b, c \in A : (a \perp b) \perp c = a \perp (b \perp c)$. Se una legge $\perp$ gode della proprietà associativa possiamo scrivere senza il rischio di essere poco chiari $a \perp b \perp c$ perchè il risultato non cambia se si esegue prima l'operazione $a \perp b$ o, viceversa, $b \perp c$.

- Proprietà **commutativa**: la legge $\perp$ gode della proprietà commutativa se e solo se $\forall\ a, b \in A : a \perp b = b \perp a$;

- Proprietà di **idempotenza**: la legge $\perp$ gode della proprietà di idempotenza se e solo se $\forall a \in A : a \perp a = a$;

- Proprietà **distributiva**: date due leggi $\perp$ e $\top$ si dice che la legge $\perp$ è distributiva rispetto alla legge $\top$ se e solo se
$\forall\ a, b, c \in A : a \perp (b \top c) = (a \perp b) \top (a \perp c)$;

- **Elemento neutro**: Un elemento $e \in A$ si dice neutro rispetto alla legge $\perp$ se e solo se $\forall a \in A : a \perp e = e \perp a = a$. Se risulta vera solo l'equazione $a \perp e = a$ si dice che che e è un elemento neutro a destra. Analogamente se per la legge $\perp$ vale solo l'equazione $e \perp a = a$ si dice che e è un elemento neutro a sinistra. Ovviamente ogni elemento neutro lo è sia a destra che a sinistra.

Teorema 6 *L'elemento neutro di una legge, se esiste, è unico.*

Dimostrazione 4 *Supponiamo per assurdo che esistano due elementi neutri per una legge $\perp$, e che essi siano distinti: li indichiamo con le lettere e ed f. L'ipotesi si scriverà allora $e \neq f$. Ci poniamo la domanda: qual è il risultato dell'operazione $e \perp f$? Poichè per ipotesi e è un elemento neutro della legge $\perp$, è corretto rispondere f; ma poichè per ipotesi anche f è un elemento neutro della legge $\perp$, il risultato dell'operazione $e \perp f$ sarà "anche" e. Dal momento che il risultato di un'operazione è unico dovrà essere necessariamente $e = f$. Del resto arriviamo alla stessa conclusione applicando la proprietà transitiva dell'uguaglianza:*
$(e \perp f = e) \wedge (e \perp f = f) \Rightarrow e = f$. *Quest'ultima conclusione è in contraddizione con l'ipotesi $e \neq f$. Dunque l'elemento neutro di ogni legge, se esiste, è unico.*

- **Elemento simmetrico**: sia $\perp$ una legge di composizione interna definita su A e sia e il suo elemento neutro. Se $\forall a \in A$ esiste un elemento $a^{-1} \in A$ tale che $a \perp a^{-1} = a^{-1} \perp a = e$ si dice che a^{-1} è l'elemento simmetrico di a.

Teorema 7 *Se una legge $\perp$ è associativa e commutativa, l'elemento simmetrico è unico.*

Dimostrazione 5 *Supponiamo per assurdo che il generico elemento a possieda due elementi simmetrici b e c. Per definizione sarà $a \perp b = e$ e $a \perp c = e$. Poichè $a \perp b$ e e sono uguali, ci aspettiamo che se li componiamo con la stessa quantità c i risultati di queste due composizioni siano uguali.*

$$\begin{gathered} a \perp b = e \\ (a \perp b) \perp c = e \perp c = c \\ c \perp (a \perp b) = c \\ (c \perp a) \perp b = c \\ b = c. \end{gathered}$$

Capitolo 4

Strutture algebriche

4.1 Strutture astratte

Una struttura è un *ambiente* matematico astratto nel quale vengono definite operazioni sugli elementi della struttura stessa. Di seguito vengono brevemente riportate le strutture algebriche di base, dalle quali vengono costruite strutture più complesse.

4.2 Gruppi

Definizione 18 *Sia A un insieme non vuoto di elementi e sia $+$ un'operazione binaria interna, cioè di $A \times A \longrightarrow A$. Si definisce A un* ***gruppo****, e lo si indica con la notazione $(A, +)$ se valgono le seguenti proprietà:*

- ***Associatività*** *di $+$: $\forall a, b, c \in A : a + (b + c) = (a + b) + c$;*
- *Esistenza dell'**elemento neutro** θ[1] rispetto a $+$: $\theta \in A, \forall a \in A : a + \theta = \theta + a = a$;*
- *Esistenza dell'**opposto**[2] rispetto a $+$: $\forall a \in A \exists b \in A : a + b = \theta$;*

[1] Si è convenuto di definire l'elemento neutro con il simbolo θ per far riflettere lo studente sul fatto che tale elemento è puramente un **simbolo** e nulla di più. In generale si conviene definire questo simbolo come 0 (zero), che tuttavia non va inteso come lo zero dei numeri reali, in quanto non definito come tale, ma solo in maniera puramente astratta, almeno per il momento.

[2] Si è convenuto di definire l'elemento b come opposto di a, con una determinata proprietà, per far riflettere lo studente che tale elemento opposto altro non è che un *elemento* dell'inisieme A, e che esso non è necessariamente indicabile come $-a$, in quanto il simbolo $-$, come è comunemente inteso, non ha alcun significato nella struttura che stiamo definendo. Si converrà di definirlo tale in quanto ha senso la scrittura $a + (-a) = 0$, ma non

Osservazione.
Se A gode della proprietà commutativa rispetto a $+$, ovvero $\forall a, b \in A$ vale la relazione $a+b = b+a$. allora A si definisce comunemente ***gruppo abeliano***.

Nota.
D'ora in avanti ci riferiremo all'operazione $+$ come operazione ***somma***, fermo restando che avremmo potuto indicare con questo nome una qualunque altra operazione. Questo per fare notare allo studente che in generale $+$ è solo un simbolo, non è la somma comunemente conosciuta finchè non la definiamo tale in maniera operativa.

4.3 Anelli

Definizione 19 *Sia A un insieme non vuoto di elementi e sia $\cdot$ un'operazione binaria interna di $A \times A \longrightarrow A$. Si definisce A un* ***anello****, e lo si indica con la notazione $(A, +, \cdot)$ se valgono le seguenti proprietà:*

- $(A, +)$ *è un gruppo abeliano;*
- ***Associatività*** *di $\cdot$: $\forall a, b, c \in A : a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (per cui si dice anche che $(A, \cdot)$ è un semigruppo);*
- ***Distributività*** *di $\cdot$ rispetto a $+$: $\forall a, b, c \in A : a \cdot (b + c) = a \cdot b + a \cdot c$;*

Osservazione.
Se A gode della proprietà commutativa rispetto a $\cdot$, ovvero $\forall a, b \in A$ vale la relazione $a \cdot b = b \cdot a$. allora A si definisce comunemente ***anello commutativo***.

Nota.
D'ora in avanti ci riferiremo all'operazione $\cdot$ come operazione ***prodotto***, fermo restando che avremmo potuto indicare con questo nome una qualunque altra operazione. Questo per fare notare allo studente che in generale $\cdot$ è solo un simbolo, non è il prodotto comunemente conosciuto finchè non lo definiamo tale in maniera operativa.

Esempio.
L'insieme $Z = \{-n, ..., -1, 0, 1, ..., n\}$ dei numeri interi relativi, è un anello commutativo, infatti presi due suoi elementi qualunque valgono sempre

la scrittura $a - a = 0$, a meno di una convenzione di scrittura o di una nuova superflua definizione del simbolo $-$.

le proprietà fin qui esposte. E' interessante notare che tuttavia non si può parlare di ***inverso***[3] di un elemento $a \in Z$, in quanto si rende necessaria l'introduzione di una nuova classe di numeri. Infatti non esiste nessun numero $b \in Z : b = a^{-1}$ tale che $a \cdot b = 1$ se non per i soli valori $a = 1$ e $a = -1$. Dunque Z è ancora *incompleto*, in quanto non include tutti gli elementi inversi dei suoi elementi.

Nota.
A questo punto diventa educativo dire che non è vera a priori la relazione $\theta \cdot a = \theta$, dove θ è lo *zero* di $+$ (operazione definita di somma), deve essere dimostrato.

Teorema 8 *Sia $(A, +, \cdot)$ un anello e sia $a \in A$. Allora $\theta \cdot a = \theta = a \cdot \theta = \theta$*

Dimostrazione 6 *Partiamo dalla definizione di elemento neutro θ, pertanto siamo autorizzati a scrivere $\theta \cdot a = (\theta + \theta) \cdot a$, e per le proprietà formali di anello avremo anche $(\theta + \theta) \cdot a = \theta \cdot a + \theta \cdot a$.*
Sommiamo ad ambo i membri (operazione permessa in $(A, +, \cdot)$) l'elemento opposto di $\theta \cdot a$, cioè $-\theta \cdot a$, da cui avremo, dalle posizioni precedenti

$$(-\theta \cdot a) + \theta \cdot a = (-\theta \cdot a) + (\theta \cdot a + \theta \cdot a)$$

Il primo membro è per definizione di opposto equivalente a θ; d'altro canto a secondo membro avremo il seguente sviluppo per la proprietà associativa:

$$(-\theta \cdot a) + (\theta \cdot a + \theta \cdot a) = (-\theta \cdot a + \theta \cdot a) + (\theta \cdot a)$$

ma dalla definizione di opposto si giunge a $\theta + \theta \cdot a$, che per definizione di elemento neutro per $+$ vale $\theta \cdot a$. Per la propietà transitiva dell'uguaglianza si ha dunque $\theta = \theta \cdot a$, che è la nostra tesi.

4.4 Campi

Definizione 20 *Sia A un insieme non vuoto di elementi e sia $\cdot$ un'operazione binaria interna, cioè di $A \times A \longrightarrow A$. Si definisce A un* ***campo****, e lo si indica con la notazione $(A, +, \cdot)$ se valgono le seguenti proprietà:*

[3] Si intende per ***inverso*** di un elemento $a \in A$, dove A è un insieme qualunque, quell'elemento b tale che $a \cdot b = \vartheta$, dove con ϑ si intende l'elemento neutro rispetto a $\cdot$. Generalmente si conviene di scrivere $b = a^{-1}$ per l'elemento inverso di a, e 1 per l'elemento neutro per $\cdot$, ma si ricorda ancora una volta che questa è solo una convenzione in ambito astratto.

- *A è un anello commutativo tale che gli elementi diversi dall'elemento neutro θ di $+$ (lo zero) formino un gruppo;*

Il campo in questione si definisce **moltiplicativo**, *poichè* $\forall a \in A \exists b \in A : a \cdot b = \vartheta$, *ovvero esiste l'*__inverso__ *di a se e solo se è eliminato l'elemento neutro θ*[4].

Osservazione.
Ancora oggi vi è molta confusione sulla definizione appena data: molti testi definiscono allo stesso modo un ***corpo***, mentre noi abbiamo scelto la notazione di ***campo moltiplicativo***. D'ora in avanti considereremo questi due attributi logicamente equivalenti.

Osservazione.
Lo studente attento, si sarà sicuramente accorto che l'insieme R dei numeri reali è un campo numerico: infatti basta andare a vedere gli assiomi di questo insieme per verificare quest'affermazione. Tuttavia il primo campo numerico non è R, bensì un suo sottoinsieme, Q, dove, è stato necessario introdurre da subito la nozione di inverso rispetto a $\cdot$.

[4]Lo studente può rendersi subito conto di ciò considerando come caso particolare l'insieme dei numeri reali: è vero che ogni $x \in R$ ammette inverso (cioè quel numero con le proprietà sopra enunciate), a patto che tale numero sia diverso da zero, per gli assiomi di R stesso.

Capitolo 5

Calcolo combinatorio

Molto spesso capita di dover fare fronte a problemi di tipo 'fisico'. Per esempio, supponiamo di voler preparare una tavolata con 6 persone tutte di stazza diversa. Poichè siamo matematici vogliamo che la configurazione della tavola sia in 'equilibrio': sarebbe molto antiestetico porre tutte le persone simili da un lato e le restanti dall'altro. Dunque ci chiediamo: *ma in quanti modi diversi possiamo sistemare queste 6 persone*?
Diciamo che non è molto comodo provare direttamente con gli interessati... Dunque, se non vogliamo fare i conti uno per uno, dobbiamo costruire una qualche teoria che ci permetta di calcolare quello che ci interessa direttamente a mente.
Questa teoria esiste e prende il nome di **calcolo combinatorio**. In questa sede non lo utilizzeremo per le 6 persone prima citate, ma per risolvere problemi di teoria delle matrici calcoli vettoriali.

5.1 Permutazioni

Supponiamo di avere un insieme $E = \{a, b\}$. In quanti *ordini* possibili possiamo scrivere questi due elementi?
Già perchè specificare che l'*ordine* è un fattore che ci interessa aumenta di poco la pesantezza del problema. Infatti, se non ci importasse l'ordine, ci sarebbe un solo modo di disporre a e b, in quanto considereremmo a, b equivalente a b, a. Ma se così non fosse, gli ordinamenti sarebbero 2 anzichè 1.
Se siamo interessati all'ordinamento, diamo la seguente

Definizione 21 *Si definisce* ***permutazione*** *degli elementi di un insieme, il massimo numeri di ordinamenti differenti che è possibile dare ad essi.*

Questo significa che nel caso di prima ci interessava come *permutare* le 6 persone intorno al tavolo.

Teorema 9 *Il numero delle permutazioni di n elementi di un insiemeè*

$$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$$

che si legge *n fattoriale*.

Dimostrazione 7 Dati n elementi di un insieme E qualunque, quante scelte possibili possono essere fatte per ordinarli rispetto a un elemento qualunque? Poichè gli elementi sono n ci sono n possibilità di scelta.
Una volta scelto il primo elemento su n, quante scelte possibili rimangono? Esattamente $n-1$.
Poichè dobbiamo combinare uno qualsiasi degli n elementi con uno qualsiasi degli $n-1$ elementi rimasti, il totale degli ordinamenti sarà $n \cdot (n-1)$.
Così procedendo per $n-2$, fino a 1, si ottiene la tesi.

Per definizione, si pone $0! = 1$.
Per rispondere al problema della disposizione delle 6 persone, avremmo $6! = 720$ ordinamenti diversi.

5.2 Disposizioni

Il concetto di disposizione è legato a quello di permutazione, per la ragione che adesso considereremo le permutazioni di n elementi in k posti. Questo significa che vogliamo sapere in quanti ordinamenti diversi possiamo raggruppare k elementi su n.
Per sfruttare ancora l'esempio delle 6 persone, supponiamo di avere a disposizione non una sedia, ma una panca per due persone. In quanti modi diversi possiamo sistemarle intorno al tavolo?
I possibili gruppi di 2 persone sono in numero, minori di quelli di una persona, in quanto il prodotto è da fare non fino all'ultima persona ma fino alla $k-$esima, nel nostro caso la seconda. Ovvero:

Definizione 22 *Si definisce* disposizione *di n elementi di un insieme in k posti, il numero massimo delle scelte ordinate dei k tra gli n.*

Teorema 10 *Il numero delle disposizioni di n elementi in k posti è*

$$n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-(k-1))$$

La dimostrazione discende da quanto osservato sopra. Quindi significa che possiamo disporre 6 persone in $6 * 5 * 4 * 3 = 360$ possibili diversi gruppi di 2. Se i gruppi fossero stati 4, allora avremmo avuto $6 * 5 = 30$ possibili

disposizioni.
Volendo mettere numericamente in relazione le permutazioni con le disposizioni:

$$n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-(k-1)) = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-(k-1)) \cdot \frac{(n-k) \cdot (n-(k+1)) \cdot}{(n-k) \cdot (n-(k+1)) \cdot}$$

che è una formula più intuitiva.

5.3 Combinazioni

Se non dovese interessare l'ordine in cui sono disposti i k elementi su n? Significherebbe che a, b è equivalente a b, a, per cui il numero è anche minore delle disposizioni.
Si nota subito che il numero delle disposizioni ordinate con gli stessi elementi è $k!$, per cui le combinazioni sono il rapporto tra le disposizioni e la permutazione di k.

Definizione 23 *Si definisce* combinazione *di n elementi di un insieme su k posti, il numero massimo di disposizioni non ordinate.*

Per cui, possiamo dare il seguente teorema la cui dimostrazione discende da quanto detto:

Teorema 11 *Il numero delle combinazioni di n elementi in k posti è*

$$\frac{n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-(k-1))}{k!}$$

cioè

$$\frac{n!}{k!(n-k)!} \tag{5.1}$$

5.4 Notazione e teoremi

Concludiamo questo breve capitolo con la notazione comune nel calcolo combinatorio:

- **Permutazioni** di n elementi: $n!$;
- **Disposizioni** di n elementi di classe k: $D_{n,k} = n!/(n-k)!$;
- **Permutazioni** di n elementi di classe k: $C_{n,k} = n!/[k!(n-k)!]$;

In genere, il numero delle combinazioni di n elementi di classe k è chiamato **coefficiente binomiale**, indicato con $\binom{n}{k}$, perchè legato al noto teorema di Newton per il calcolo della potenza del binomio:

Teorema 12 (Potenza $n-$esima del binomio) *Dati due numeri reali a, b e n intero, vale*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$$

Vale il seguente teorema di semplice dimostrazione, più che altro una verifica:

Teorema 13

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}$$

che conclude il capitolo.

Capitolo 6

Spazi vettoriali

6.1 Nozioni fondamentali

Definizione 24 *Sia $(V+)$ un gruppo abeliano, e sia K un campo moltiplicativo (corpo). Si consideri l'applicazione*[1]

$$\pi : K \times V \longrightarrow V : (\alpha, v) \longmapsto \alpha v, \forall \alpha \in K, \forall v \in V$$

Vogliamo che quest'operazione sia tale che:

- $(\alpha + \beta)v = \alpha v + \beta v \ (\forall \alpha, \beta \in K, \forall v \in V)$;
- $\alpha(v + w) = \alpha v + \alpha w \ (\forall \alpha \in K, \forall v, w \in V)$;
- $\alpha(\beta v) = (\alpha\beta)v \ (\forall \alpha, \beta \in K, \forall v \in V)$;
- $\vartheta v = v \ (\vartheta \in K, \forall v \in V)$;

Allora si definisce V uno **spazio vettoriale su** *K, e i suoi elementi prendono il nome di* **vettori** *in V e scalari in K.*

Osservazione.
Attenzione: le leggi di composizione interna che si trovano nei primi membri delle quattro relazioni di equivalenza (in questo caso di eguaglianza), non coincidono con le operazioni dei rispettivi secondi membri; infatti il $+$ della prima 'equazione' è esterno a V perchè opera con coppie di scalari di K, mentre a secondo membro il $+$ opera su coppie di V, e pertanto è un'operazione interna ad esso. Analogamente dicasi per le altre relazioni e per la legge $\cdot$.

[1] Sinonimo di **operazione**.

Esempio.
Sia $V = (R, +)$, dove R è l'insieme dei numeri reali (R è qui assunto come gruppo abeliano) e sia $K = R$ (dove R è ora assunto come campo moltiplicativo). Avremo l'applicazione $R \times R \longrightarrow R : (\alpha, v) \longmapsto \alpha v$.
Quella appena definita altra non è che la struttura di spazio vettoriale di R su sè stesso... e l'applicazione viene a coincidere con il comune prodotto tra numeri reali.
E' interessante che in generale possiamo sempre introdurre su un campo K la struttura di spazio vettoriale su sè stesso, dato che K prima di essere un campo è un gruppo abeliano.
Allo stesso modo si può definire $Q \times R \longrightarrow R : (\alpha, v) \longmapsto \alpha v$, dove Q è l'insieme dei numeri razionali e tale struttura di spazio vettoriale conserva le sue proprietà, poichè $Q \subseteq R$, solo che in questo caso gli scalari sono in Q e i vettori in R.
E' inutile dire che non si può continuare a procedere a ritroso con gli insiemi numerici, infatti l'ultimo campo numerico, a partire da R verso N è proprio Q e non ha senso definire un'applicazione di Z, per esempio, su R, tale che verifichi le proprietà di spazio vettoriale.

6.2 Spazi K^n

Sia K un campo e sia $n \in N^+$ (cioè intero positivo diverso da zero).
Si dica K^n il prodotto cartesiano $K \times K \times K \times ... \times K$, n volte, l'insieme delle $n-$uple ordinate degli elementi di K.
Siano $\underline{x} = (x_1, x_2, ..., x_n)$ e $\underline{y} = (y_1, y_2, ..., y_n)$ due $n-$uple ordinate, dove $x_i, y_i \in K$ $(i = 1, 2, ..., n)$.
Sia $+$ l'operazione interna a K^n, tale che $\underline{x} + \underline{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$, dove $+$ è definita nella struttura interna di K.
Allora $(K^n, +)$ è un gruppo abeliano con elemento neutro $\underline{\theta} = (\theta, \theta, ..., \theta)$ $(\theta \in K)$ rispetto a $+$[2].
Sia $\cdot$ un'altra operazione interna di K^n tale che $x \cdot \underline{x} = (x \cdot x_1, x \cdot x_2, ..., x \cdot x_n)$ $\forall x \in K, \forall \underline{x} \in K^n$, dove $\cdot$ è definita nella struttura interna di K.
Allora K^n è uno spazio vettoriale, in cui valgono la seguente proprietà:

$$(x_1, x_2, ..., x_n) = (x_1, 0, ..., 0) + (x_2, 0, ..., 0) + ... + (x_n, 0, ..., 0)$$

[2]Abbiamo praticamente definito la nuova legge $+$ in K^n partendo dal $+$ definito precedentemente in K.

per la definizione di somma appena data, e

$$(x_1, 0, ..., 0) + (0, x_2, ..., 0) + ... + (0, 0, ..., x_n) =$$
$$x_1(1, 0, ..., 0) + x_2(0, 1, ..., 0) + ... + x_n(0, 0, ..., 1)$$

per la definizione di scalare. Dunque, in definitiva, ogni elemento di K^n si può ancora scrivere come le somme fin qui definite (proprietà notevole, in quanto ci permette di operare con esso come operiamo, per esempio, sui numeri reali).

6.3 Operazioni

Definizione 25 *Siano V e W due spazi vettoriali su un campo K. Si dice* ***funzione lineare*** *o* ***omomorfismo di spazi vettoriali*** *un'applicazione*

$$\xi : V \longrightarrow W : \xi(v + v') = \xi(v) + \xi(v'), \xi(\alpha v) = \alpha\xi(v) \quad (\forall v, v' \in V, \forall \alpha \in K)$$

Osservazione.
Si dimostra facilmente, a partire dalla definizione, che l'applicazione ξ può essere definita in maniera univoca e senza ambiguità, nel modo seguente:

$$\xi : V \longrightarrow W : \xi(\alpha v + \beta v') = \alpha\xi(v) + \beta\xi(v') \quad (\forall v, v' \in V, \forall \alpha, \beta \in K)$$

Esempio.
Lo studente, certamente ricorderà che gli è stato insegnato che la somma o il prodotto di funzioni continue, sono ancora funzioni continue, cioè la proprietà di continuità di due o più funzioni, è conservata se esse sono legate da leggi di composizione come quelle definite per un campo, nel nostro caso particolare, la somma e il prodotto. Tuttavia per dimostrare questo, si procede facendo vedere semplicemente che l'insieme V delle funzioni continue è uno spazio vettoriale.

Dimostrazione 8 *Sia $V = \{f : D \longrightarrow U\}$, $D \subseteq R$ (cioè sia V lo spazio vettoriale delle funzioni continue di dominio di D e codominio U).*
Prese due funzioni f e g, si verifica facilmente, che essendo in un campo, nella fattispecie quello dei numeri reali, per adesso, valgono gli assiomi di campo e le proprietà di omomorfismo di spazi vettoriali.

Esempio.
Sia K un campo, e indichiamo con $K[x]$ i polinomi nella indeterminata x (nel senso che non la trattiamo come un'incognita) a coefficienti in K.

Sia per definizione $x^\theta = \vartheta \in K$ e sia il *polinomio nullo* quello che ha tutti i coefficienti uguali a θ (lo zero di $+$).
Inoltre sia per definizione il *polinomio di grado zero* (che non coincide con quello nullo...) quello per cui il termine noto è $a_0 \neq \theta$ e tutti gli altri coefficienti uguali a zero.
Si dimostra subito che $K[x]$ è un anello ma non un campo (in quanto non abbiamo definito un inverso per x), e si dimostra anche che esso è un gruppo abeliano rispetto a $+$, facendo notare che valgono gli assiomi di gruppo e che $\alpha(a_0 + a_1x + ... + a_nx^n) = \alpha a_0 + ... + \alpha a_n x^n$. In questo modo si crea lo spazio vettoriale su K dei polinomi $K[x]$, che è un anello di polinomi 'impoverito'.

6.4 Spazi finitamente generati

Definizione 26 *Sia V uno spazio vettoriale su un campo K.*
Si dice che V è **finitamente generato** *se esiste un insieme finito di vettori $v = \sum_i a_i v_i$ $(\forall a_i \in K, \forall v \in V)$. L'insieme $\{v_1, v_2, ..., v_n\}$ è detto* **insieme dei generatori** *e $v = \sum_i a_i v_i$ è una combinazione lineare dell'insieme dei generatori mediante gli scalari a_i.*

Ne segue, per esempio, che K^n contiene un insieme di generatori perchè ogni vettore nello spazio è uguale alla combinazione lineare di più vettori. Inoltre, K^n non è finitamente generato: infatti se $F_1, F_2, ..., F_n$ fosse un sistema di generatori allora n è il massimo grado di F_i 4 il polinomio appartenente a $K[x]$ di grado $n+1$ non è combinazione lineare di F_i; esso ha un insieme di generatori che è un sottoinsieme dello spazio vettoriale.

Ogni spazio ha un insieme di generatori, infatti ogni vettore è la combinazione di altri 2. Dato il sottoinsieme di $K = \{1, x, ..., x^n\} \subseteq K[x]$, il sottospazio $U \subseteq V$ è un sottospazio vettoriale se $U \neq \emptyset$ ed è uno spazio vettoriale rispetto alle stesse operazioni di V. Si dimostra che se $U \neq \emptyset$ e $U \subseteq V$, U è uno sottospazio di V se e solo se $(a_1u_1 + a_2u_2) \in U$ $(\forall a, b \in K, \forall u_1, u_2 \in U)$.
In U vale l'operazione $+$ e se esiste u esiste anche il suo opposto (per esempio, $a = 1 \Longrightarrow b = -1$).

In ogni spazio vettoriale V/K risulta
$\theta v = \underline{\theta}$ $(\forall v \in V)$ e $a\underline{\theta} = \underline{\theta}$ $(\forall a \in K)$[3], infatti $\theta v = (\theta + \theta)v = \theta v + \theta v$, ossia il vettore nullo, per gli assiomi di gruppo.
Se in V/K, e $\forall a \in K, \forall v \in V$ si ha $av = \underline{\theta}$, allora $a = 0 \vee v = 0$, poichè se $a = 0$ il teorema è subito verificato, per $a \neq \theta$, infatti si ha: $a^{-1}av =$

[3]Proprietà nullative.

$a^{-1}\underline{\theta} \Longrightarrow (a^{-1}a)v = \underline{\theta} \Longrightarrow 1v = \underline{\theta} \Longrightarrow v = \underline{\theta}$, come volevasi dimostrare.
Inoltre valgono anche: $(-a)v = -av$; $(-a)v + av = \underline{\theta} \Longrightarrow (-a + a)v = \theta v$ (dove θ è lo ***zero*** per $+$ in K).

Definizione 27 *Sia V/K uno spazio vettoriale su un campo K. I vettori $v_i \in V$ si dicono linearmente* **indipendenti** *se $\sum_i a_i v_i = \underline{\theta} \Longrightarrow a_1 = a_2 = ... = a_n$.*

6.5 Sottospazi congiungenti

Iniziamo con un teorema preliminare che ci consentirà di proseguire la trattazione.

Teorema 14 *Sia V uno spazio vettoriale su un campo K. Siano $U \subseteq V$ e $W \subseteq V$: allora $U \cap W$ è un sottospazio vettoriale di V.*

Dimostrazione 9 *Siano dati i vettori $v_1, v_2 \in U \cap W$ e gli scalari $a, b \in K$. Consideriamo adesso la combinazione lineare $av_1 + bv_2$ e dimostriamo che essa appartiene a $U \cap W$.*
I due vettori, appartenendo all'insieme intersezione dei due sottospazi di V, appartengono ovviamente a entrambi questi sottospazi; dunque $av_1 + bv_2 \in U$ perchè esso è stato supposto un sottospazio (e una qualunque combinazione lineari di vettori di uno spazio vettoriale appartiene ancora ad esso), ma vale anche $av_1 + bv_2 \in W$ per la medesima ragione e le considerazioni iniziali. Pertanto $av_1 + bv_2 \in U \cup W$.

Teorema 15 *Sia $\{U_\alpha\}_{\alpha \in I}$[4] una famiglia di sottospazi di V.*
Allora l'insieme intersezione $\bigcup_{\alpha \in I} U_\alpha$ è un sottospazio di V.

La dimostrazione fa semplicemente uso del teorema precedente nel caso di più sottospazi vettoriali.

Definizione 28 *Siano U, W due sottospazi di uno spazio vettoriale V. Si definisce* sottospazio congiungente *U e W, l'intersezione di tutti i sottospazi di V che contengono $U \cup W$.*

Osservazione.
In generale $U \cup W$ non è un sottospazio e neanche un sottogruppo. Infatti sia U non incluso interamente in W e viceversa (cioè esista almeno un elemento

[4]Qui I è un insieme di indici, non necessariamente numeri interi naturali, in quanto il nostro interesse è rivolto ad una trattazione generale.

dell'uno non appartenente all'altro e viceversa) e siano $u \in U$ e $w \in W$.
Allora $(u + w) \in V$ poichè combinazione lineare di due elementi appartenenti a V, ma non appartiene a U poichè se così fosse si potrebbe scrivere $u + w = u_1 \in U$, cioè $w = u - u_1$ poichè trovandoci in un gruppo questa operazione è permessa. Ma $(u - u_1) \in U$ poichè è una combinazione lineare di due elementi di U e quindi $w \in U$. Ma ciò contrasta con l'ipotesi secondo cui $w \in W$.
Pertanto $U \cup W$ è un sottospazio se e solo se $U \subseteq W$ e $W \subseteq U$.

Il sottospazio congiungente U e V è pertanto $U + W = \{u + w : u \in U, w \in W\}$, che non coincide con $U \cup W$ ma lo contiene. Inoltre devono rientrare anche i $u + w$ poichè esso deve essere un sottospazio, perciò chiuso rispetto alla somma.
Si può dimostrare che $U + W$ è il minimo sottospazio che contiene $U \cup W$, ed è talvolta definito anche come ***spazio somma***, proprio per le sue caratteristiche.
La definizione di congiungente deve valere per qualunque struttura algebrica: gruppi, anelli, campi, etc.

6.6 Combinazioni e indipendenze lineari

Definizione 29 *Siano $(v_1, v_2, ..., v_n) \in V/K$. Indichiamo con $L(v_1, v_2, ..., v_n) = \{a_1v_1 + a_2v_2 + ... + a_nv_n\}$ $(a_i \in K)$ l'insieme di tutte le combinazioni lineari dei vettori fissati.*

Si verifica facilmente che L così definito è un sottospazio, in quanto presi due scalari $a, b \in K$ si nota che se

$$a_1v_1 + a_2v_2 + ... + a_nv_n \in L(v_1, v_2, ..., v_n)$$
$$b_1v_1 + b_2v_2 + ... + b_nv_n \in L(v_1, v_2, ..., v_n)$$

allora

$$a(a_1v_1 + a_2v_2 + ... + a_nv_n) + b(b_1v_1 + b_2v_2 + ... + b_nv_n) \in V$$

e applicando gli assiomi di campo:

$$(aa_1 + bb_1)v_1 + (aa_2 + bb_2)v_2 + ... + (aa_n + bb_n)v_n$$

che è una combinazione lineare dei vettori $v_1, v_2, ..., v_n$ di L, per cui esso è un sottospazio di V.

Teorema 16 Criterio di indipendenza lineare *Siano* $(v_1, v_2, ..., v_n) \in V/K$. *Allora essi sono linearmente indipendenti se e solo se* $v_1 \neq 0$ *e* v_i *non è combinazione lineare dei precedenti vettori appartenenti a* $L(v_1, v_2, ..., v_{i-1})$ $\forall i = 1, 2, ..., n$.

Dimostrazione 10 *Dimostriamo prima l'implicazione* $\Longrightarrow$.
Vediamo che se i vettori sono linearmente indipendenti allora deve essere necessariamente $v_1 \neq 0$. *D'altro canto è soddisfatta pure la seconda condizione in quanto se* $v_i \in L(v_1, v_2, ..., v_{i-1})$ *allora* $v_i = a_1v_1 + a_2v_2 + ... + a_{i-1}v_{i-1}$, *ovvero esso è dipendente linearmente dai precedenti, poichè*

$$a_1v_1 + a_2v_2 + ... + a_{i-1}v_{i-1} - v_i + 0v_{i+1} + ... + v_n = 0$$

che è proprio la definizione di dipendenza lineare (gli scalari infatti non sono mai contemporaneamente tutti nulli...).

Dimostriamo ora l'implicazione $\Longleftarrow$.
Sia $a_1v_1 + a_2v_2 + ... + v_n = \underline{0}$: *si deve dimostrare che i coefficienti sono tutti nulli.*
Se fosse $a_n \neq 0$ *avremmo*

$$v_n = -a_n^{-1}a_1v_1 - a_n^{-1}a_2v_2 - ... - a_n^{-1}a_{n-1}v_{n-1}$$

cioè otterremmo v_n *come combinazione lineare dei precedenti, cosa che va contro la seconda ipotesi secondo cui* $v_n \in L(v_1, v_2, ..., v_n)$ *è linearmente indipendente. Quindi poniamo* $a_n = 0$ *e procediamo da* a_{n-1} *fino a* a_2. *Infatti per* a_1 *si ha che* $a_1v_1 = \underline{0}$, *che considerata la prima ipotesi (*$v_1 \neq 0$*) e i teoremi precedentemente dimostrati per le proprietà nullative, portano a* $a_1 = 0$, *che conclude la nostra dimostrazione.*

6.7 Basi canoniche

Definizione 30 *Dato uno spazio vettoriale* V *su un campo* K, *si parla di* ***base canonica*** *per* V, *se è sempre possibile identificare ogni vettore* $v = \in V$ *come combinazione lineare di vettori della forma* $(1, 0, ..., 0); (0, 1, ..., 0); ...; (0, 0, ..., 1)$ *ossia come*

$$v = a_1v_1(1, 0, ..., 0) + a_2v_2(0, 1, ..., 0) + ... + a_nv_n(0, 0, ..., 1)$$

Esempio.

Abbiamo già definito gli spazi K^n: poichè quello dei reali è un campo moltiplicativo ha senso parlare di spazi vettoriali R^n sul campo R stesso, indicando con ciò l'insieme delle $n-$uple ordinate di numeri reali.
Per questo spazio vettoriale si dimostra facilmente che i vettori della forma

$$\begin{aligned} u_1 &= (1, 0, ..., 0) \\ u_2 &= (0, 1, ..., 0) \\ &... \\ u_n &= (0, 0, ..., 1) \end{aligned}$$

sono linearmente indipendenti e dunque formano una base per lo spazio vettoriale. Inoltre, considerate le proprietà dei numeri reali si ha che ogni $n-$upla ordinata è esprimibile come combinazione lineare di numeri reali per i vettori della base, detta appunto **canonica**.

6.8 Somma di sottospazi

Teorema 17 *Siano $U \subseteq V$, $W \subseteq V$, due sottospazi dello spazio vettoriale V su un campo K.*
Allora $U + W$ è un sottospazio, che si dice **diretto** *se $U \cup W = \{\underline{0}\}$ e si indica con $U \oplus W$, se e solo se ogni vettore di $U + W$ si scrive in modo unico nella forma $u + w$, con $u \in U$, $w \in W$.*

Dimostrazione 11 *Dimostriamo l'implicazione $\Longrightarrow$.*
Sia $u + w = u_1 + w_1$ entrambi appartenenti a $U + W$. Se si esprime in maniera unica allora $(u - u_1) \in U$ e $(w - w_1) \in W$, che, poichè per ipotesi vale $U \cap W = \{\underline{0}\}$, allora è anche $u - u_1 = \underline{0}$ e $w - w_1 = \underline{0}$, cioè $u = u_1$ e $w = w_1$, ovvero i vettori di partenza appartengono alla stessa classe di equipollenza, cioè sono equivalenti.

Dimostriamo ora l'implicazione $\Longleftarrow$.
Sia $v \in U \cap W$, ovvero sia tale che $v \in U$ e $v \in W$. Allora $(v + \underline{0}) \in U + W$ $= (\underline{0} + v) \in U + W$, che data l'unicità porta a $v = \underline{0}$ e $\underline{0} = v$, cioè la tesi.

6.9 Basi e insiemi massimali

Definizione 31 *Sia V/K uno spazio vettoriale. Si definisce* base *di V l'insieme $\{v_1, v_2, ..., v_n\} \subseteq V$ tale che i suoi elementi sono generatori e linearmente indipendenti.*

Seguono interessanti proprietà:

- $\forall v \in V : v = \sum_{i=1}^{n} a_i v_i$, per determinati scalari in K;
- ogni vettore $v \in V$ si esprime in modo unico con la precedente combinazione lineare, poichè se fosse $\sum_{i=1}^{n} a_i v_i = \sum_{j=1}^{n} a_j v_j$ avremmo, sfruttando gli assiomi di campo:

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + ... + (a_n - b_n)v_n = \underline{0}$$

che data l'indipendenza lineare come da ipotesi, determina $a_1 - b_1 = a_2 - b_2 = ...a_n - b_n = 0$ cioè $a_1 = b_1; a_2 = b_2; ...; a_n = b_n$.

Definizione 32 *Si definisce* insieme massimale *una base di elementi linearmente indipendenti di V.*

Definizione 33 *Un insieme si dice* libero *se i suoi elementi sono linearmente indipendenti.*

Osservazione.
Sia $B = \{u_1, u_2, ..., u_n\}$ massimale in V rispetto alla proprietà di essere libero[5]. Per il criterio di indipendenza lineare allora $v = \sum_i a_i v_i$. Di conseguenza B sarà un insieme di generatori e pertanto una base.

Teorema 18 *Una base è un insieme minimale di generatori di V.*

Dimostrazione 12 *Sia $\{v_1, v_2, ..., v_n\}$ un insieme minimale di generatori di V: dimostriamo essi sono linearmente indipendenti.*
Sia $\sum_i a_i v_i = \underline{0}$. Se per assurdo esistesse un qualche $a_i \neq 0$, per esempio il primo, allora possiamo moltiplicare ambo i membri per a_i^{-1}. Per $i = 1$ si ha

$$v_1 = a_1^{-1} a_2 v_2 + ... + a_1^{-1} a_n v_n$$

Ma poichè è stato supposto un insieme minimale di generatori, v_1 va scartato, in quanto combinazione lineare degli altri vettori. Sarebbe dunque $\{v_2, ..., v_n\}$ un insieme di generatori di V, contro l'ipotesi di minimalità. Pertanto la nostra tesi è dimostrata.

Teorema 19 Teorema degli scarti *Se $\{v_1, v_2, ..., v_n\}$ è un insieme di generatori, esso contiene una base di V.*

[5]Ovvero aggiungendo un vettore v a B esso non è più un insieme libero.

Dimostrazione 13 *Dall'insieme di generatori dato si scartano i vettori nulli, avendo dunque come risultato r vettori, con $r \leq n$, $\{v_{i_1}, v_{i_2}, ..., v_{i_r}\}$. Adesso di sicuro $v_{i_1} \neq 0$; se v_{i_2} è multiplo di v_{i_1} lo si scarta, perchè non è linearmente indipendente, e si procede così per i restanti vettori.*
Alla fine si troverà una base libera $v_{i_1}, v_{j_2}, v_{j_3}, ..., v_{j_s}$, che è ancora un insieme di generatori, ovviamente.

Ne segue subito un importante teorema di semplice dimostrazione:

Teorema 20 Metodo del completamento *Se $G = \{v_1, v_2, ..., v_n\}$ è un insieme libero di V, allora lo si può completare ad una base, ovvero esiste una base di V che contiene propriamente G.*

Teorema 21 Lemma di Steinitz *Siano $v_1, v_2, ..., v_n$ i generatori di V, spazio vettoriale su un campo K, e siano $v_1, v_2, ..., v_m$ vettori linearmente indipendenti di V. Allora $m \leq n$.*

Questo teorema afferma che se n è il numero dei vettori generatori di V e m è il numero di altri vettori di V, anch'essi linearmente indipendenti, allora il numero di tali vettori liberi è sempre non maggiore di quello dei generatori.

Dimostrazione 14 *Consideriamo l'insieme $G = \{u_1, v_1, v_2, ..., v_n\}$. Esso è sicuramente di generatori, che però non sono linearmente indipendenti, in quanto $u_1 \in V$ è generato da una combinazione lineare della base di V. Applichiamo allora il teorema degli scarti, al fine di trovare la base $\{u_1, v_{i_1}, v_{i_2}, ..., v_{i_r}\}$, con $r < n$.*
Consideriamo adesso $\{u_1, u_2, v_{i_1}, v_{i_2}, ..., v_{i_r}\}$: procedendo analogamente a prima si trova ua nuova base, che in generale contiene almeno un vettore in meno della precedente. Così facendo $m-1$, si ottiene la base $\{u_1, u_2, ..., u_{m-1}, v_\alpha, ...\}$, dove si mantiene almeno un v_α dei generatori iniziali, perchè se così non fosse potremmo determinare u_m come combinazione lineare dei precedenti contro l'ipotesi di indipendenza lineare tra questi.
Negli $m-1$ passaggi abbiamo scartato almeno un v_i originario, per un totale di almeno $m-1$; dalle considerazioni precedenti possiamo dunque dire che $m-1 \leq n-1$, cioè $m \leq n$.

A questo lemma segue un importante e interessante corollario:

Corollario 1 *Due basi di uno spazio vettoriale V/K hanno lo stesso numero di elementi.*

Dimostrazione 15 *Siano $v_1, v_2, ..., v_n$ e $u_1, u_2, ..., u_m$ due basi di un comune spazio vettoriale V. Per il lemma precedente possiamo affermare che $m \leq n$. Tuttavia possiamo procedere con un discorso espressamente simmetrico, dicendo che è vero che anche $n \leq m$. Ma ciò implica necessariamente che $m = n$.*

6.10 Dimensione di uno spazio vettoriale

Definizione 34 *Si definisce* dimensione *di uno spazio vettoriale V su un campo K, e si indica con $dim_k V$, il numero degli elementi di una base di V.*

Esempio.
Sia K^n/K: $dim_k K^n = n$ poichè la base è costituita dagli n vettori canonici (le basi di R^n.

Esempio.
Sia C/C: $dim_C C = 1$, di generatore qualunque non nullo.

Esempio.
Sia C/R: $dim_R C = 2$, di base $(1, i)$ o in generale di due numeri complessi linearmente indipendenti.

Osservazione.
Sia V finitamente generato su K e siano $v_1, v_2, ..., v_n$ linearmente indipendenti. Allora $L(v_1, v_2, ..., v_n)$ o coincide con V o è contenuto in esso. Nel primo caso allora i vettori dati sono generatori; nel secondo caso ne segue che esistono in V vettori che non sono combinazione lineare dei vettori dati. Scelto $v_{r+1} \in V/L(v_1, v_2, ..., v_r)$, si ha che $v_1, v_2, ..., v_r, v_{r+1}$ è libero. In questo modo si può continuare il procedimento che tuttavia deve arrestarsi dato che V è finitamente generato.

Osservazione.
Sia V finitamente generato su K e sia U un sottospazio di V. Allora $dimU = dimV \Longrightarrow U = V$, poichè una base in U è anche base in V e $U \subseteq V$.

6.11 Cambiamenti di base

Teorema 22 *Sia V uno spazio vettoriale su K e $B = \{v_1, v_2, ..., v_n\}$ una sua base. Allora ogni $v \in V$ si esprime in maniera unica come combinazione lineare degli elementi di B.*

Dimostrazione 16 *La scelta di B ci fornisce un'applicazionedi $V \longrightarrow K^n$: $v \longmapsto (x_1, x_2, ..., x_n)$. Tale applicazione è una funzione lineare, poichè presi*

due vettori e due scalari si ha:

$$v' = \sum_i x'_i v_n \Longrightarrow av + bv' = a(\sum_i x_i v_n) + b(\sum_i x'_i v_n) = \\ = (a_1 x_1 + b_1 x'_1)v_1 + ... + (a_n x_n + b_n x'_n)v_n$$

Si nota subito dunque la combinazione lineare degli n scalari che esprimono tale somma come combinazione lineare dei vettori di B. Inoltre è anche un isomorfismo*, ovvero una funzione lineare biettiva.*

In questo modo abbiamo appena mostrato come tutti gli spazi vettoriali V così definiti siano isomorfi a K^n e sono strettamente dipendente dalla base B (ovvero questa proprietà non è intrinseca di ogni spazio vettoriale...).

Ci si pone adesso il problema di come operare un cambiamento da una base assegnata ad un'altra.
Sia V/K uno spazio vettoriale e $B = \{v_1, v_2, ..., v_n\}$, $B' = \{v'_1, v'_2, ..., v'_n\}$ due sue basi. Allora è ovvio che ogni vettore di V si può scrivere come combinazione lineare degli elementi di entrambe le basi:

$$v = \sum_{i=1}^{n} x_i v_i = \sum_{j=1}^{n} x'_j v'_j$$

Prendiamo un v'_j: poichè $v'_j \in V$ e $v'_j \in B'$, allora si può determinare come combinazione lineare dei vettori della base B a meno di precisi scalari:

$$v = \sum_{h=1}^{n} a'_{hj} v'_{hj}$$

poichè esso appartiene a $L(B)$. Visto che $j = 1, 2, ..., n$ ci sono in totale n^n scalari per esprimere i vettori di B' (n per ogni $1 \leq j \leq n$) come combinazione lineare di quelli di B.
Si è soliti scrivere questi coefficienti a doppio indice, in un quadrato diviso in righe e colonne:

$$\begin{matrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ & & & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{matrix}$$

dove gli scalari che esprimono v'_j sono nella $j-$esima colonna. A tale notazione è dato il nome di **matrice**, e sarà argomento che verrà approfondito

successivamente.
Adesso possiamo provare che è possibile scrivere le componenti di v in B rispetto a quelle di v in B':

$$v = \sum_{j=1}^{n} x'_j v'_j = \sum_{j=1}^{n} x'_j (\sum_{i=1}^{n} a_{ij} v_i) = \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} x'_j) v_i = \sum_{i=1}^{n} x_i v_i \Longrightarrow$$

$$\Longrightarrow x_i = \sum_{j=1}^{n} a_{ij} x'_j$$

dove l'implicazione è giustificata dal fatto che i vettori si esprimono in maniera unica come combinazione lineare dei vettori della base.
In questo modo abbiamo visto come sia possibile passare da una base all'altra di uno spazio vettoriale.
Procedendo analogamente si trova che

$$x'_j = \sum_{i=1}^{n} a'_{ij} x_i$$

con una matrice analoga alla precedente e strettamente legata ad essa.

6.12 Formula di Grasmann

Teorema 23 *Sia $V/K/$ uno spazio vettoriale finitamente generato e siano U, W due suoi sottospazi, tali che $dim U = n$ e $dim W = m$. Allora vale la relazione, detta* di Grasmann*:*

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W)$$

Dimostrazione 17 *Diamo per scontato che $U \cap W$ sia finitamente generato, poichè $(U \cap W) \subseteq U$ e $(U \cap W) \subseteq W$, entrambi sottospazi di V.*
Sia $u_1, u_2, ..., u_r$ una base di $U \cap W$.
Sia $u_1, u_2, ..., u_r, u_{r+1}, ..., u_n$ una base di U (che contiene la base di $U \cap W$).
Sia $u_1, u_2, ..., u_r, w_{r+1} ..., w_m$ una base di W (che contiene la base di $U \cap W$).
Si deve dimostrare che $A = u_1, u_2, ..., u_r, u_{r+1}, ..., u_n, w_{r+1} ..., w_m$ è una base di $U + W$, ovvero che il numero di tali vettori è $n + (m - r)$.
Vediamo che per ogni $(u + w) \in U + W$ vale:

$$u + w = \sum_{i=1}^{n} a_i u_i + \sum_{j=1}^{r} b_j u_j + \sum_{h=r+1}^{m} b_h w_h =$$

$$= (a_1 + b_1)u_1 + ... + (a_r + b_r)u_r +$$

$$+ a_{r+1} u_{r+1} + ... + a_n u_n +$$

$$+ b_{r+1} w_{r+1} + ... + b_m w_m$$

Sia ora

$$\sum_{i=1}^{n} a_i u_i + \sum_{h=r+1}^{m} b_h w_h = \underline{0}$$

da cui

$$\sum_{i=1}^{n} a_i u_i = -\sum_{h=r+1}^{m} b_h w_h$$

Il primo membro si trova in U e il secondo in W, per cui $a_1u_1 + ... + a_nu_n$ è un vettore di $U \cap W$. Di conseguenza, poichè $r < n$, $a_{r+1} = ... = a_n = 0$, per cui

$$a_1u_1 + ... + a_ru_r = -b_{r+1}w_{r+1} - ... - b_mw_m$$

cioè

$$a_1u_1 + ... + a_ru_r + b_{r+1}w_{r+1} + ... + b_mw_m = \underline{0}$$

Poichè $u_1, ..., u_r, w_{r+1}, ..., w_m$ è una base di W, per l'unicità dell'espressione di un vettore (in questo caso $\underline{0}$) rispetto alla base, si ha $a_1 = ... = a_r = b_{r+1} = ... = b_m = 0$, pertanto i vettori sono linearmente indipendenti. Questo conclude la dimostrazione.

6.13 Prodotto scalare

Definizione 35 *Sia V/K uno spazio vettoriale su R.Si definisce* prodotto scalare *(o interno) su V, un'applicazione $V \times V \longrightarrow R : (u, v) \longmapsto u \circ v$, tale che:*

- $u \circ v = v \circ u \ \forall u, v \in V$ *(**proprietà pseudo-commutativa**);*
- $(au) \circ v = u \circ (av) = a(u \circ v) \ \forall u, v \in V, \ \forall a \in R$ *(**proprietà di omogeneità**);*
- $u \circ (v + w) = u \circ v + u \circ w \ \forall u, v, w \in V$ *(**proprietà distributiva rispetto alla somma di vettori**);*
- $u \circ u \geq 0 \ \forall u \in V$ *e* $u \circ u = 0 \Longleftrightarrow u = \underline{0}$ *(**proprietà di positività**).*

Il lettore attento si starà chiedendo perchè scegliere R come campo scalare. La risposta è che in generale su nessun campo è definito un ordinamento parziale o totale e si può dimostrare che se un campo esistesse con tale definizione esso dovrebbe necessariamente coincidere con quello dei reali (esiste cioè un isomorfismo di tale campo in R[6]).

Nota.
Quanto vale $u \circ \underline{0}$? Poichè $\underline{0}$ è l'elemento neutro della somma, si ha, per la proprietà distributiva appena definita, che

$$u \circ \underline{0} = u \circ (\underline{0} + \underline{0}) = u \circ \underline{0} + u \circ \underline{0}$$

Ovvero il numero reale risultante è uguale a due volte sè stesso; come sappiamo l'unico numero con tale proprietà è proprio zero. Se ne deduce dunque che l'elemento neutro della somma è ancora annullatore.

Definizione 36 *Due vettori $u, v \in V^2/R$ si dicono* perpendicolari*, e si indicano con $u \perp v$, se $u \circ v = \underline{0}$*

Esempio.
Sia $V = R^n/R$ uno spazio vettoriale e siano

$$x = (x_1, ..., x_n)$$
$$y = (y_1, ..., y_n)$$

due sue $n-$uple ordinate.
Definiamo come prodotto scalare la relazione

$$x \circ y = \sum_{i=1}^{n} x_i y_i$$

ovvero il prodotto scalare standard: definita una base $\{ \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k} \}$ ortonormale, la definizione di prodotto appena dato corrisponde a quello tra vettori. In questo caso si parla anche di ***metrica euclidea***.

Ma potremmo definire un prodotto scalare anche con la relazione

$$x \circ y = \sum_{i=1}^{n} a_i x_i y_i \quad a_i \in R, a_i > 0$$

[6]Si noti per esempio che già il campo C dei numeri complessi non è più ordinato.

poichè tale relazione soddisfa le prime tre proprietà di un prodotto scalare e per la quarta si nota che $x \circ x \sum a_i x^2$ che è sempre non negativo. Tuttavia qui si richiede che sia almeno $a_i \geq 0$. La positività in senso stretto ci serve per dimostrare che $x \circ x = \underline{0} \Longleftrightarrow x = 0$. Infatti posto $a_i \neq 0$ e inoltre sempre positivo, si ha che il prodotto scalare su sè stesso è nullo se e solo se il vettore stesso è nullo, altrimenti per qualche $a_i = 0$ poteva accadere che $x \circ x = 0$ anche per $x \neq 0$.

Esempio.
Sia $V = R^2/R$ e sia $(x_1, x_2) \diamond (y_1, y_2) = x_1 y_1 + 2x_2 y_2$ una relazione. Si verifica facilmente che essa è ancora un prodotto scalare, ma con proprietà diverse da quello che siamo abituati ad utilizzare. Infatti presa la coppia $(2, 3)$, si verifica facilmente dalla definizione di perpendicolarità che $(2, 3) \perp (3, -1)$. Riportando tali coppia su un riferimento cartesiano si nota subito che l'angolo compreso tra i due vettori non è retto... Ciò significa che in tale prodotto scalare non è valida la metrica euclidea, pertanto se ne deduce che la relazione di perpendicolarità varia al variare della definizione di prodotto scalare.

Esempio.
Sia $V = \{f : [a, b] \longrightarrow R, continuein[a, b]\}$, ovvero sia V lo spazio vettoriale delle funzioni continue in un intervallo limitato assegnato.
Un prodotto scalare per tale spazio può essere definito come:

$$f * g = \int_a^b f(t)g(t)dt$$

il quale verifica le proprietà di prodotto scalare (il lettore dimostri ciò a scopo educativo...). Pertanto in V sono anche definite ***funzioni perpendicolari***, e questo avvalora la tesi secondo cui i vettori non sono semplici enti geometrici, ma a livello astratto sanno essere molto più generici e ricchi di proprietà affascinanti.

6.14 La disuguaglianza di Schwarz

Teorema 24 *Sia V/R uno spazio vettoriale con prodotto scalare $\circ$. Allora*

$$(x \circ y)^2 \leq (x \circ x)(y \circ y) \quad \forall x, y \in V$$

Dimostrazione 18 *Se $x = \underline{0}$ si verifica banalmente che la disuguaglianza è subito verificata.*

Supponiamo dunque $x \neq \underline{0}$.
Sia ha che, per ogni $\lambda \in R$ *e dagli assiomi dalle proprietà del prodotto scalare che*

$$(\lambda x + y) \circ (\lambda x + y) = \lambda^2 (x \circ x) + 2\lambda (x \circ y) + (y \circ y) \geq 0$$

Poichè $x \circ x > 0$, *segue* $\frac{\Delta}{4} \leq 0$, *ovvero*

$$(x \circ y)^2 - (x \circ x)(y \circ y) \leq 0$$

da cui la tesi spostando a secondo membro il termine negativo.
La tesi si può altresì esprimere come

$$|x \circ y| \leq \sqrt{x \circ x}\sqrt{y \circ y}$$

6.15 Norma, distanza, angolo, versore

Definizione 37 *Si dice* lunghezza *o* norma *di un vettore* $x \in V/R$, *il valore* $||x|| = \sqrt{x \circ x}$.

Definizione 38 *Si dice* distanza *di un vettore* $x \in V/R$ *da un vettore* $y \in V/R$, *il valore* $||x - y|| = \sqrt{(x-y) \circ (x-y)}$.

Definizione 39 *Si dice* angolo ϕ *di due vettori* $x, y \in V/R$ $(y \neq \underline{0})$, *quel valore tale che* $\cos(\phi) = \frac{x \circ y}{||x||||y||}$ $(0 \leq \phi \leq \pi)$.

Definizione 40 *Si dice* versore *di un vettore* $x \in V/R$ $(x \neq \underline{0})$, *il valore* $\frac{x}{||x||}$.

L'ultima definizione non è casuale, infatti:

$$||\frac{x}{||x||}|| = \sqrt{\frac{x}{||x||} \circ \frac{x}{||x||}} = \frac{1}{||x||} \cdot ||x|| = 1$$

6.16 La disuguaglianza triangolare

Teorema 25 *Siano* $x, y \in V$, *due vettori di uno spazio vettoriale su un campo* K. *Allora vale la relazione*

$$||x + y|| \leq ||x|| + ||y||$$

Dimostrazione 19 *Si nota che:*

$$||x+y||^2 = [\sqrt{(x+y)\circ(x+y)}]^2 = x\circ x + 2x\circ y + y\circ y \leq \\ \leq x\circ x + 2|x\circ y| + y\circ y$$

e per la disuguaglianza di Schwarz:

$$x\circ x + 2|x\circ y| + y\circ y \leq ||x||^2 + 2||x||||y|| + ||y||^2 = ||x+y||^2 \Longrightarrow \\ \Longrightarrow ||x+y|| \leq ||x|| + ||y||$$

che è la tesi.

6.17 Spazi metrici e normali

La ***distanza*** è un'applicazione tale che ad una coppia di vettori associa un numero reale:

$$d : V \times V \longrightarrow R : (x,y) \longmapsto d(x,y)$$

Definizione 41 *Si dice che $d(x,y)$ è una distanza se e solo se sono sempre verificate le tre proprietà per ogni $x, y \in V^2/R$:*

- ***Simmetria**: $d(x,y) = d(y,x)$;*
- *$d(x,x) = 0$;*
- ***Transitività**: $d(x,y) \leq d(x,z) + d(y,z)$.*

Da una distanza si può dedurre un prodotto scalare e viceversa.

Definizione 42 *Si definisce* spazio metrico *un qualunque spazio vettoriale nel quale sia assegnata una distanza d.*

Definizione 43 *Si definisce* spazio normale *un qualunque spazio vettoriale nel quale sia assegnata una norma $||x||$.*

6.18 Ortonormalità: relazioni tra vettori

Sia V/R uno spazio vettoriale con prodotto scalare $\circ$ e siano $v_1, v_2, ..., v_n$ vettori ortonormali[7], ovvero $v_i \circ v_j = \delta_{ij}$[8].

Teorema 26 *Uno spazio di vettori ortonormali è libero.*

Dimostrazione 20 *Sia* $\sum_i a_i v_i = \underline{0}$. *Moltiplicando per* v_1*:*

$$a_1 v_1 \circ v_1 + a_2 v_2 \circ v_1 + ... + a_n v_n \circ v_1 = \underline{0}$$

ma i vettori sono ortonormali, pertanto si ottiene

$$a_1 \cdot 1 + a_2 \cdot 0 + ... + a_n \cdot 0 = \underline{0}$$

da cui $a_1 = 0$.
Procedendo analogamente per ogni v_i *si dimostra che* $\sum_i a_i v_i = \underline{0} \Longrightarrow a_i = 0$.

Teorema 27 *Sia* $B = \{v_1, ..., v_n\}$ *una base di vettori ortonormali di* V *e sia ogni* $v \in V : v = \sum a_i v_i$. *Allora il coefficiente dell'*$i-$*esimo termine è dato da* $v \circ v_i$.

Dimostrazione 21 *Poichè* v_i *è un elemento della base, allora, per l'ortonormalità:*

$$v \circ v_i = a_1 v_1 \circ v_i + a_2 v_2 \circ v_i + ... + a_i v_i \circ v_i + ... + a_n v_n \circ v_i \Longrightarrow$$
$$v \circ v_i = a_1 \cdot 0 + a_2 \cdot 0 + ... + a_i \cdot 1 + ... + a_n \cdot 0 = a_i$$

che è la nostra tesi.

Osservazione.
Sempre utilizzando la stessa base B, si dimostra che

$$u \circ v = (\sum_i a_i u_i) \circ (\sum_i b_i v_i) = \sum_i a_i b_i$$

per il teorema appena dimostrato (ovvero molti prodotti si annullano e restano solo quelli della forma $a_i b_i (u_i \circ v_i) = a_i b_i$), che però non è il prodotto scalare standard in R^n, poichè i suoi vettori erano già $n-$uple e la sua base standard è ortonormale.

Esercizio.
Sia $V = R^4/R$ e siano V, W due suoi sottospazi non vuoti tali che $V \cap W = \{(x, y, z, t) : x - y = y + z = 0\}$.

[7] Versori (normali) e ortogonali a due a due.
[8] Delta di Kronecker o di Dirac, che dir si voglia. Essa vale zero se $i \neq j$, uno se $i = j$.

1. Determinare una base per $V \cap W$;
2. Determinare V, W tali che $V + W = R^4$;
3. Determinare $v \in V$ e $w \in W$ tali che $v + w = (1, 0, 0, 0)$.

Svolgimento.
Poniamo a sistema le nostre condizioni di intersezione:

$$\begin{cases} x = y \\ y = -z \end{cases}$$

da cui si nota subito che i vettori di $V \cap W$ sono della forma $(a, a, -a, b)$, ovvero dati dalla combinazione lineare

$$(a, a, -a, b) = (a, a, -a, 0) + (0, 0, 0, b) = a(1, 1, -1, 0) + b(0, 0, 0, 1)$$

ovvero i vettori

$$\begin{cases} u_1 = (1, 1, -1, 0) \\ u_2 = (0, 0, 0, 1) \end{cases}$$

sono una base di $V \cap W$. Possiamo anche scrivere $L(u_1, u_2) = V \cap W$.
Per generare V è necessario trovare un vettore linearmente indipendente da quelli della base di $V \cap W$, per cui verificato che, per esempio, $v_3 = (1, 0, 0, 0)$ è un possibile tale vettore, si ha $V = L(u_1, u_2, v_3)$, poichè aggiungendo un vettore linearmente indipendente dagli altri ad una base, si ottiene ancora un insieme libero. E' naturale che si verifica che v_3 non appartiene a $V \cap W$. Non aggiungiamo un ulteriore vettore libero poichè otterremmo tutto R^4 e non più V.
Analogamente per generare W si aggiunge $w = (0, 0, 1, 0)$, dimostrando che esso non appartiene a $V \cap W$ e non appartiene a V.
Infatti si ha:

$$V = L(u_1, u_2, v_3) = \{(x, y, z, t) : (x, y, z, t) = au_1 + bv_2 + cv_3; \forall a, b, c \in R\}$$

Adesso

$$(x, y, z, t) = (a * 1 + b * 0 + c * 1, a * 1 + b * 0 + c * 0, a * (-1) + b * 0 + c * 0, a * 0 + b * 1 + c * 0) = (a + c, a, -a, b) \iff \begin{cases} x = a + c \\ y = a \\ z = -a \\ t = b \end{cases}$$

Eliminando i parametri a, b, c:

$$\begin{cases} a = y \\ b = t \\ c = x - a = x - y \\ z = -y \end{cases}$$

ovvero $V = \{(x, y, z, t) : y + z = 0\}$, che è l'equazione cartesiana di V. Tutto questo per poter verificare che $w_3 = (0, 0, 1, 0)$ non appartiene a V,e si nota subito che esso non può essere generato dai vettori di V.
Per la relazione di Grasmann: $dim(V + W) = 3 + 3 - 2 = 4$, e poichè $dimR^4 = 4$, siamo autorizzati a dire che abbiamo trovato due sottospazi il cui congiungente è lo spazio totale.
Adesso dobbiamo trovare $v \in V, w \in W : v + w = (1, 0, 0, 0)$. Siano

$$\begin{aligned} v &= au_1 + bu_2 + cv_3 \\ w &= a'u_1 + b'u_2 + c'w_3 \end{aligned}$$

da cui

$$v + w = (a + a')u_1 + (b + b')u_2 + cv_3 + c'w_3 = (1, 0, 0, 0)$$

da cui, moltiplicando per le componenti dei vettori, discende

$$(a + a' + c, a + a', -a - a' + c', b + b') = (1, 0, 0, 0)$$

che dal sistema equivalente porta alle soluzioni

$$\begin{cases} c = 1 \\ a' = -a \\ c' = 0 \\ b' = -b \end{cases}$$

per cui

$$\begin{aligned} v &= au_1 + bu_2 + v_3 \\ w &= -au_1 - bu_2 \end{aligned}$$

ovvero infinite coppie (u, v) la cui somma restituisce la base canonica di R^4 richiesta.

6.19 Il prodotto scalare: relazioni particolari

In uno spazio metrico il prodotto scalare, come notato già in precedenza, 'misura' distanze, ed è stato anche definito rigorosamente in maniera generale e nel particolare caso dello spazio euclideo. Comunque questo sarà argomento del prossimo capitolo.

6.19.1 Perpendicolarità

Definizione 44 *Sia dato uno spazio vettoriale V/K con prodotto scalare $\circ$. Si dice che due vettori $u, v \in V$ sono perpendicolari, e si indica con $u \perp v$, se e solo se il loro prodotto scalare è nullo.*

A partire da tale definizione, nessuno ci vieta di costruire due sottoinsiemi di uno spazio vettoriale costituiti rispettivamente da tutti i vettori perpendicolari tra loro, cioè tali che i vettori dell'uno sono perpendicolari a quello dell'altro e di definire di conseguenza la*perpendicolarità tra sottospazi.*

Definizione 45 *Sia V uno spazio vettoriale e siano $S \subseteq V$, $S' \subseteq V$. Si dice che $S \perp S'$ se $u \circ v = 0 \ \forall u \in S, \forall v \in S'$.*

Definizione 46 *Si definisce $S^\perp$ l'insieme dei vettori*

$$v \in V : v \circ s \quad \forall s \in S$$

Si dimostra facilmente che $S^\perp$ è un sottospazio di V. Infatti presi $v, v' \in S^\perp$, dobbiamo dimostrare semplicemente che $(av + bv') \in S^\perp$ e che vale $(av + bv') \circ s$:

$$(av + bv') \circ s = a(v \circ s) + b(v' \circ s) = a0 + b0 = 0$$

cioè la combinazione lineare di due vettori di $S^\perp$ sta ancora in $S^\perp$.

Valgono i seguenti teoremi di facile dimostrazione:

Teorema 28 *Sia U un sottospazio di V e siano $dimV = n$ e $dimU = r$. Allora $dimU^\perp = n - r$ (ovvero U e $U^\perp$ sono complementari).*

Teorema 29 *Sia U un sottospazio di V. Allora*

- $(U^\perp)^\perp = U$;
- $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$;
- $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.

6.19.2 Procedimento di Gram-Schmidt

Sia V/R uno spazio vettoriale con prodotto scalare $\circ$ e sia $u_1, ..., u_n$ una base per tale spazio.
Il nostro obiettivo è di procurare una base ortonormale.
Sia $v_1 = u_1$ e $v_2 = u_2 + \lambda v_1$ ($\lambda \in R : v_2 \circ v_1 = 0$).
Dimostriamo dapprima che un tale λ esiste. Sia $(u_2 + \lambda u_1) \circ u_1 = 0$. Ne segue $u_2 \circ u_1 + \lambda u_1 \circ u_1 = 0$. Adesso essendo elemento della base, u_1 è non nullo, pertanto $u_1 \circ u_1 \neq 0$. Allora $\lambda = -\frac{u_2 \circ u_1}{u_1 \circ u_1}$.
Dimostrato ciò procediamo come segue:

$$\begin{aligned} v_3 &= u_3 + (av_1 + bv_2) \\ &\dots \\ v_n &= u_n + \sum_{i=1}^{n-1} a_i v_i \end{aligned} \tag{6.1}$$

dove i vettori v_j sono tutti ortogonali. Per la normalizzazione si assume $w_j = ||v_j||$.

Capitolo 7

Vettori geometrici

7.1 Definizioni di base

Sia S^3 lo spazio euclideo, ovvero l'insieme dei punti dove valgono gli assiomi euclidei.
Sia S^{3*} (S^3 segnato) l'insieme di tutti i segmenti orientati dello spazio euclideo, ovvero l'insieme di tutte le coppie ordinate di punti (A, B) congiungenti un segmento di estremi A e B nello spazio stesso.

Si deduce che sulla retta si può definire un orientamento, cioè si può stabilire la comune proprietà di *precedenza* tra A e B. Se si assume l'esistenza di una parallela al segmento AB, la proiezione di AB su tale parallela deve coincidere con AB stesso, a parità di orientamento (comunemente definito ***verso***). Tuttavia in S^{3*} tutto ciò è privo di senso, a meno che prima non si definiscano alcune proprietà fondamentali.

Lo studente ricorderà sicuramente cos'è una relazione di equivalenza tra due oggetti. Esse ci permettono di operare su un insieme una ***partizione***, cioè di costruire l'insieme dei sottoinsiemi di quello di partenza, con la particolarità che tali sottoinsiemi sono a due a due **disgiunti**[1], e che $\bigcup P(E)$ (cioè l'insieme unione di tutti i sottoinsiemi, detti ***parti*** di E) è uguale all'insieme E di partenza. In questo modo si è creato un insieme di ***classi di equivalenza***. Come si ricorderà, l'insieme di tali classi mediante una relazione $\Re$ è chiamato **insieme quoziente**, indicato con $E/\Re$.

Definizione 47 *Due segmenti orientati* (A, B) *e* (C, D) *si dicono* **Equipollenti** *(equivalenti) se:*

[1] Ovvero a due a due l'intersezione è sempre uguale all'insieme vuoto.

- *AB è parallela a CD;*
- *la lunghezza di AB è uguale a quella di CD;*
- *se la retta orientata da A verso B coincide con quella da C a D ($\overrightarrow{AB} \equiv \overrightarrow{CD}$), cioè se le rette sono **orientate concordemente**.*

La relazione di equipollenza è una relazione di equivalenza.
In S^{3*} abbiamo perciò definito una relazione di equivalenza, per cui l'insieme quoziente $S^{3*}/\Re$ è l'insieme delle ***classi dei vettori geometrici***.

Definizione 48 *Se $\Re$ è la relazione di equipollenza tra segmenti orientati, diremo $V = S^{3*}/\Re$ l'insieme dei vettori geometrici.*
Se $x, y \in V$, diciamo 'somma' $x + y$, il vettore congiungente l'estremo A e l'estremo D, assunto $B \equiv C$, operando una traslazione, se necessario, in questo spazio.

Lo studente noterà che è possibile 'scegliere' infiniti rappresentanti di una classe di equipollenza, pertanto si dovrebbe dimostrare che il risultato finale non varia al variare dei rappresentanti scelti (dimostrazione molto semplice).

Definizione 49 *Sia α un numero reale e x un vettore appartenente allo spazio V. Si dice αx la classe di equipollenza del segmento lungo α volte x.*

Quella data è una *buona definizione*.

A partire dagli assiomi e dalle proprietà di campo e di spazio vettoriale quale è V, si dimostrano facilmente alcune particolarità, come per esempio l'unicità del vettore nullo e dell'esistenza e unicità del vettore opposto.

7.2 Rappresentazione di vettori

Definizione 50 *I vettori $\{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$ si dicono* base *di V se essi non sono nulli, sono linearmente indipendenti e complanari a due a due.*

Ne segue subito che $dim V = 3$. Per le considerazioni precedenti possiamofare in modo che i tre vettori della base siano orientati a partire da un medesimo punto O, che per comodità possiamo definire *origine*, e tale che la definizione appena data non venga violata.
Allora, una volta definita la *lunghezza* di un vettore, potremmo mettere in corrispondenza biunivoca il nostro *sistema* con quello dei numeri reali. E' inoltre necessario fissare una terna in V per poter parlare di *componenti* di

un vettore nel nostro sistema di riferimento vettoriale.
Per procedere possiamo intraprendere due vie: da un lato si possono fissare 3 numeri, un origine e un vettore e prendere le coordinate di un generico punto P, che andranno a rappresentare le *componenti* del vettore v generico. Altrimenti si può utilizzare la nozione di piano passante per P e parallelo ad ogni coppia di vettori della base, trovare le intersezioni prese con il giusto segno e definirle ***componenti*** di v nel sistema di riferimento $\{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$.
Avremo che tale terna è determinata univocamente in modo che $\forall v \in V$: $v = x\widehat{i} + y\widehat{j} + z\widehat{k}$ $(x, y, z \in R)$, e le componenti siano proprio x, y, z rispetto al sistema $\{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$[2].
In generale, per le componenti valgono le consuete operazioni di somma e prodotto, in rispetto dell'ordine. Al contrario non è sempre definito un prodotto scalare. Procedendo, e ricordando che lo spazio dei vettori geometrici altro non è che un caso particolare di spazio vettoriale definito nel capitolo precedente, avremo:

Definizione 51 *Se le classi v e w formano un 'angolo' ϑ, definiamo* prodotto scalare *la relazione $v \circ w = ||v||||w|| \cos\vartheta$.*

Definizione 52 *Due classi di vettori si definiscono* ortogonali *se il loro prodotto scalare è nullo.*

Come il lettore può verificare, le definizioni o coincidono, o sono casi particolari di quelle date in precedenza.

7.3 Il prodotto scalare

Il prodotto scalare appena definito, gode delle seguenti proprietà:

- **Bilinearità**: ovvero dipende da entrambi i vettori interessati ed è associativa per l'operazione di somma rispetto a quella di prodotto esterno e viceversa;
- **Simmetria**: ovvero non importa l'ordine con il quale il prodotto viene eseguito.

[2]Si ricorda che esse sono classi di equipollenza di segmenti orientati: assunto un altro sistema $\{\overrightarrow{i'}, \overrightarrow{j'}, \overrightarrow{k'}\}$ **parallelo e concorde**, a $\{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$, otterremo le medesime componenti per il medesimo $v \in V$. Questa è una *buona definizione*, ovvero invariante al variare dei rappresentanti delle classi.

Da quello che è stato appena detto, e dalle definizioni precedenti, si può affermare senza ambiguità che il prodotto scalare tra due vettori può essere calcolato a partire dalle sole componenti degli stessi.
Assumiamo adesso che la base di V^3 introdotto in questo capitolo, sia *ortonormale*[3]: possiamo allora indicare tale base con la notazione $\{\widehat{i}, \widehat{j}, \widehat{k}\}$, per distinguere dal precedente caso.
Allora presi

$$||v|| = x_1\widehat{i} + y_1\widehat{j} + z_1\widehat{k}$$
$$||w|| = x_2\widehat{i} + y_2\widehat{j} + z_2\widehat{k}$$

si ha

$$v \circ w = x_1x_2 + y_1y_2 + z_1z_2$$

da cui risulta banale la dimostrazione della proprietà di bilinearità e di simmetria.
Infatti eseguendo il prodotto tra le componenti, si nota che i prodotti che contengono il prodotto tra i *versori*[4] valgono zero se essi non sono complanari: da cui la notazione sopra.

Osservazione.
Data per scontata la conoscenza della definizione di distanza tra due punti, si può dimostrare il ***teorema di Pitagora generalizzato***, più comunemente noto come ***teorema di Carnot***.

7.4 Il prodotto vettoriale

Diciamo fin da subito che il prodotto vettoriale così come verrà definito, è strettamente legato alla dimensione dello spazio vettoriale, cioè 3.

Definizione 53 *Dato uno spazio vettoriale V si distinguano due sole classi di segmenti orientati. Tutte le basi che si trovano nella medesima classe vengono definite* orientate ordinatamente e conocordemente.

La definizione appena data si giustifica banalmente, notando che due sistemi di riferimento differiscono a meno di rotazioni e operazioni di *ribaltamento*. Tuttavia quest'ultima operazione nasconde un'insidia: non appena

[3]Il termine assume qui il medesimo significato che ha assunto nel capitolo precedente.
[4]La terna dei vettori ortonormali.

decidiamo di ribaltare il sistema rispetto a un versore, dovremo necessariamente passare per il piano contenente gli altri due, ovvero per un istante astratto i tre versori non sono linearmente indipendenti. Tutto ciò per dimostrare, che al contrario del prodotto scalare, il prodotto vettoriale dipende strettamente dall'ordine di composizione.

Definizione 54 *Dati due vettori*[5] v *e* w *di uno spazio vettoriale* V*, si definisce* prodotto vettoriale $v \wedge w$ *un vettore di modulo* $||v||||w||sen\vartheta$ $(0 \leq \vartheta \leq \pi)$*, direzione ortogonale*[6] *al piano su cui giacciono* v *e* w *e di verso tale che la terna* $\{\overrightarrow{v}, \overrightarrow{w}, \overrightarrow{v} \wedge \overrightarrow{w}\}$ *sia orientata positivamente*[7]*.*

Se v e w non sono due classi linearmente indipendenti, significa che sono parallele, e dunque $v \wedge w = 0$, in modo che la terna $\{\overrightarrow{v}, \overrightarrow{w}, \underline{0}\}$ sia un particolare sistema di riferimento.
Geometricamente, il modulo di $v \wedge w$ rappresenta l'area del parallelogrammo di lati v e w.

7.4.1 Proprietà

Il prodotto vettoriale gode di particolari proprietà:

- **Distributività** rispetto alla somma: $(x + y) \wedge z = x \wedge z + y \wedge z$;
- **Associatività** rispetto al prodotto esterno: $(\alpha x) \wedge y = x \wedge (\alpha y) = \alpha(x \wedge y)$ $(\forall x, y \in V, \forall \alpha in R)$.

Il prodotto $\wedge$ come definito, è bilineare ma non simmetrico, poichè $x \wedge y = -(y \wedge x)$[8].

Esercizio.
Il lettore rifletta sul significato analitico e geometrico delle espressioni $(x \wedge y) \wedge z$ e $x \wedge (y \wedge z)$ e sulle loro relazioni.

Il prodotto vettoriale, può essere definito in maniera più rigorosa, ma che lascia meno spazio all'intuizione e alla rappresentazione geometrica, come segue, se ci troviamo nelle medesime ipotesi precedenti:

[5] Si intende *vettori geometrici* necessariamente, per le considerazioni fatte precedentemente.

[6] In generale ciò non è necessario a livello astratto.

[7] Generalmente il verso viene determinato con la **regola della mano destra**, per questo si parla anche di sistemi *destrorsi* o *sinistrorsi*.

[8] Proprietà di **antisimmetria**.

Definizione 55 *Si definisce prodotto vettoriale di due vettori $v, w \in V$, il vettore il cui modulo è*

$$||v \wedge w|| = (v_1\widehat{i} + v_2\widehat{j} + v_3\widehat{k}) \wedge (w_1\widehat{i} + w_2\widehat{j} + w_3\widehat{k}) = det\begin{pmatrix} \widehat{i} & \widehat{j} & \widehat{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$$

ovvero uguale al determinante della matrice associata[9].

che, come si verifica facilmente, verifica le proprietà prima enunciate.

7.5 Minima distanza

Definizione 56 *Data una retta r e un punto P_0 al di fuori di essa, si definisce* distanza $d(P_0, r)$ *la distanza $d(P_0, P_0')$, dove P_0' è la* proiezione ortogonale[10] *di P_0 su r.*

Da questa definizione si dimostra che $d(P_0, r) \leq d(P_0, P)$, ovvero essa è la **minima distanza** tra il punto e la retta. Ciò può essere provato assumendo una seconda distanza P_0P, dove P è un altro qualunque punto sulla retta r e considerando l'angolo ϑ che si viene a creare tra P_0P e r.
Infatti preso un altro vettore su P_0P, si ha che $d(P0, P)$ è minima quando è minimo il suo prodotto scalare con un vettore su r, cioè quando $\cos\vartheta$ è nullo, ovvero quando i vettori sono ortogonali.
Se $\widehat{r}$ è il versore di r, si ha $d(P_0, r) = ||\overrightarrow{PP_0} \wedge \widehat{r}||$ $(\forall P \in r)$.

7.6 Prodotti misti

Definizione 57 *Si definisce* prodotto misto *di una terna di vettori $x, y, z \in V$, la relazione*

$$x \circ (y \wedge z) = det\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

Ne segue subito un teorema di semplice dimostrazione, basta una immediata verifica a partire dalle definizioni di prodotto:

[9] Questa definizione, seppur più efficace, fa uso del concetto di matrice non ancora affrontato e oggetto del prossimo capitolo.

[10] Ciò significa che preso un qualunque vettore giacente su r e uno giacente sul segmento congiungente P_0 e P_0', il loro prodotto scalare è nullo.

Teorema 30 *Condizione sufficiente e necessaria perchè x, y, z siano complanari, è che il loro prodotto misto sia nullo.*

Dalla definizione di prodotto misto si verifica geometricamente che esso rappresenta il 'volume' di un parallelepipedo di lati x, y, z.
Il volume di tale parallelepipedo è costante, dunque ruotando il risultato non deve cambiare, e di fatto si verifica che

$$x \circ (y \wedge z) = y \circ (z \wedge x) = z \circ (x \wedge y)$$

Al contrario si osserva come il doppio prodotto vettoriale $x \wedge (y \wedge z)$, non è associativo.

7.7 Identità fondamentali

Quelle che seguono, sono identità fondamentali, che tuttavia utilizzano anche il concetto di matrice, che verrà approfondito nel prossimo capitolo.

- **Identità di Jacobi**: $x \wedge (y \wedge z) + z \wedge (x \wedge y) + y \wedge (z \wedge x) = 0$;
- $x \wedge (y \wedge z) = (x \circ z)y - (x \circ y)z$[11];
- **Identità di Laplace**:

$$(x \wedge y) \circ (z \wedge t) = det\begin{pmatrix} x \circ z & x \circ t \\ y \circ z & y \circ t \end{pmatrix}$$

7.8 Spazio vettoriale quoziente

Definizione 58 *Sia V uno spazio vettoriale su R, e sia W un suo sottospazio. Il sottospazio W induce la seguente relazione di equivalenza su V:*
I vettori $x, y \in V$ si dicono equivalenti modulo W se $(x - y) \in W$, ossia a meno di vettori di W.
Si verifica che lo l'insieme di tali classi di equivalenza così definite è uno spazio vettoriale $V_{/W}$ definito quoziente.

La relazione di equivalenza data è davvero tale:

- E' *riflessiva*: x è equivalente a sè stesso modulo W poichè $x - x = \underline{0} \in W$, $\forall x \in V$;

[11] Il secondo membro non dipende più dall'orientemento dei vettori.

- E' *simmetrica*: $(x-y) \in W \Longrightarrow -(x-y) \in W$ perchè W è un sottospazio vettoriale; ma allora $(y-x) \in W$;
- E' *transitiva*: $(x-y) \in W, (y-z) \in W \Longrightarrow [(x-y)+(y-z)] \in W \Longrightarrow (x-z) \in W$.

Se $W = \{\underline{0}\}$ allora $V_{/W} = V$; se $W = V$ allora $V_V = \underline{0}$.
E' necessario definire una *somma* tra gli elementi dello spazio quoziente e un *prodotto esterno.*

Definizione 59 *Siano $\overline{x}, \overline{y} \in V_{/W}$: si definisce* somma $\overline{x} + \overline{y} = \overline{x+y}$[12].

Definizione 60 *Sia $\overline{x} \in V_{/W}$: si definisce* prodotto *(esterno) $\alpha \cdot \overline{x} = \overline{\alpha x}$*[13] *$\forall \alpha \in R$.*

Teorema 31 *Siano V e W due spazi vettoriali finitamente generati, e sia il secondo sottospazio del primo. Allora $dim(V_{/W}) = dimV - dimW$.*

Per trovare una base di $V_{/W}$, partiamo dall'ipotesi che $\{w_1, ..., w_k\}$ è una base di W, e completiamola ad una base di V data da $\{w_1, ..., w_k, v_{k+1}, ..., v_n\}$. Allora $\{\overline{v_{k+1}}, ..., \overline{v_n}\}$ è una base di $V_{/W}$.

Esempio.
Due matrici si definiscono equivalenti, se la loro differenza è una matrice simmetrica[14]. Se $R^{2,2}$ è lo spazio vettoriale su R delle matrici quadrate 2×2, si dimostra che il sottoinsieme S di tutte le matrici simmetriche 2×2 è un sottospazio di $R^{2,2}$, a partire dagli assiomi di spazio.
Si ha che $dimR^{2,2} = 4$, $dimS = 3$, $dim(R^{2,2}_{/S}) = 1$. Generalizzando: $dimR^{n,n} = n^2$, $dimS = \frac{n(n+1)}{2}$, $dimA = \frac{n(n-1)}{2}$[15]. Le basi sono quelle canoniche.

7.9 Spazi di polinomi

Abbiamo incontrato nel precedente capitolo la definizione di spazio vettoriale dei polinomi su un campo K nella indeterminata x. Sia ora tale campo quello dei reali. Si nota che se $R[x]_n$ indica il sottospazio dei polinomi di ordine n, allora la sua dimensione è $n+1$, poichè una base è data da $\{1, x, x^2, ..., x^n\}$.

[12] Si verifica che essa è ben definita.
[13] Si verifica che essa è ben definita.
[14] Una matrice si dice **simmetrica** se e solo se gli elementi al di sotto della diagonale principale coincidono con quelli che vi stanno sopra. Se sono uguali e opposti, si parla di matrice **antisimmetrica**.
[15] Sottospazio delle matrici antisimmetriche.

Teorema 32 *Sia $R[x]$ lo spazio di tutti i polinomi a coefficienti in R. Allora $dim R[x] = \infty$.*

Dimostrazione 22 *Supponiamo per assurdo che esista una base finita di polinomi $p_1, ..., p_n$ tali che $R[x] = L(p_1, ..., p_n)$.*
Se d_i indica il grado dell'$i-$esimo polinomio, sia $d = max(d_1, ..., d_n)$. Sia adesso

Preso $D = d + 1 > d$, non esiste nessuna combinazione lineare dei vettori della base data da cui si ottiene un polinomio di grado D, dunque la base ipotizzata è incompleta. Completiamola con il metodo del completamento, aggiungendo p_{n+1}. Possiamo dunque ripetere l'operazione precedente prendendo $D' = (d + 1) + 1$ e dunque ritrovarci costretti a dover nuovamente aggiungere un elemento alla base. Tuttavia il procedimento esposto può ripetersi infinite volte, sempre per il medesimo motivo, di conseguenza la base non può contenere un numero finito di polinomi, da cui ne segue che la dimensione non può essere finita: deve dunque essere infinita.

Esempi.

- Si indichi con $deg(p)$ il grado di $p \in R[x]$. Si dimostra che $X_n = \{p \in R[x] : deg(p) = n\} \cup \{\underline{0}\}$ non è un sottospazio vettoriale;
- Si dimostra che l'insieme S di tutte le successioni reali è uno spazio vettoriale con le operazioni di somma e prodotto definite su R, inoltre $dim S = \infty$;
- L'insieme delle successioni limitate è un sottospazio vettoriale di S;
- L'insieme delle successioni infinitesime è un sottospazio vettoriale di S;
- Si dimostra che gli insiemi delle funzioni continue o delle funzioni derivabili o integrabili finitamente o indefinitamente, sono tutti spazi vettoriali.

Gli esempi riportati su R, valgono per un campo K qualsiasi.

Esercizio.
Dato $R[x]_2$, si consideri il sottoinsieme $X_2 = \{p \in R[x]_2 : p(1) = 0\}$. Si dimostri che X_2 è un sottospazio (cioè è chiuso rispetto alle operazioni di somma e prodotto esterno indotte da R).
Per la proprietà di linearità dei polinomi, $(p + q)(x) = p(x) + q(x)$, dunque:

$$(p + q)(1) = p(1) + q(1) = 0 + 0 = 0$$

poichè $p, q \in R[x]_2$.
Più rigorosamente si nota che per i polinomi dati:

$$p(x) = (x-1)r_{p_1}(x)$$
$$q(x) = (x-1)r_{q_1}(x)$$

da cui $x = 1$ implica la tesi.
La chiusura rispetto al prodotto si dimostra in maniera analoga.
Si nota che $dimX_2 = 2$ e non 3, poichè è stata imposta una condizione di linearità, che ha annullato la necessità dei 3 parametri. La base può essere assunta come $\{(x-1), (x^2-x)\}$. Generalizzando: $dimX_n = n$ se è stata posta una condizione di linearità, e più in generale $dimX_n = (n+1) - m$ se sono imposte m condizioni di linearità indipendenti.
Si dimostra anche che $Z_{x_0} = \{p \in R[x]_n : p(x_0) = 0\}$ è un sottospazio di dimensione n, o dimensione $(n+1) - m$ con m condizioni lineari imposte. Una base per tale sottospazio può essere assunta dai polinomi di molteplicità $i = 1, 2, ..., n$ che si annullano in x_0.
Al contrario si verifica che $Y_n = \{p \in R[x]_n : p(x_0) = k, k \in R - \{0\}\}$ non è un sottospazio[16].

Esercizio.
Verificare che il sottoinsieme $Y_2 = \{p \in R[x]_2 : p'(1) = 0\}$ è un sottospazio.
Generalizzare, dimostrando che $Y_n = \{p \in R[x]_n : p'(x_0) = 0\}$ è un sottospazio e che $dimY_n = n$ o che $dimY_n = (n+1) - m$ se sono imposte m condizioni lineari.

Esercizio.
Verificare che $Z_2 = \{p \in R[x]_2 : p(1) = p'(1) = 0\}$ è un sottospazio di $dim = 1$.
Suggerimento: considerare separatamente le due condizioni, e considerare l'intersezione l'insieme dato. Poi considerare come caso particolare dell'esercizio precedente generalizzato.
Osservazione.
Si può notare come una volta considerate le imposizioni lineari, il sistema associato è di $n+1$ equazioni in m incognite, le cui soluzioni sono in funzione

[16]Si può dimostrare banalmente sostenendo l'affermazione secondo cui occorrono le condizioni di **linearità** sui coefficienti del polinomio e di **omogeneità**, ossia l'annullamento del termine noto, affinchè l'imposizione di condizione lineare suddetta sussista. In questo caso, seppure sia rispettata la linearità, manca l'omogeneità.
Tuttavia si può notare che tale insieme altro non è che una 'traslazione' di Z_{x_0}, e pertanto con opportune trasformazioni può essere ricondotto a esso.

di $(n+1)-m$ parametri.
In particolare, per $n=2$ si ha che

$$p'(x) = [(x-1)r(x)]' = r(x) + (x-1)r'(x) \Longrightarrow$$
$$\Longrightarrow p'(1) = r(1) = 0$$

dove allora $r(1) = (x-1)r_2(x)$. Ma per le considerazioni precedenti:

$$p(x) = (x-1)(x-2)r_2(x) = (x-1)^2 r_2(x)$$

dove $r_2(x)$ è l'unico parametro variante[17], in aggiunta costante, poichè il grado massimo è 2. Dunque $Z_2 = L[(x-1)^2]$, che sta a significare che p ha molteplicità 2 in $x=1$.
Generalizzando, se $Z_{x_0} = \{p \in R[x]_n : p(x_0) = p'(x_0) = ... = p^{(n)}(x_0) = 0\}$, allora p ammette $x = x_0$ come radice con molteplicità $n+1$.

Esercizio.
Particolarmente istruttivo: dimostrare che se p è un polinomio non nullo tale che $deg(p) = n$, allora p ammette al più n radici ciascuna con la propria molteplicità.
Suggerimento:
Procedere per induzione o con la divisione tra polinomi (vedi esercizi precedenti).

[17] Infatti la dimensione è proprio (2+1)-2=1.

Capitolo 8

Teoria delle matrici

8.1 Definizioni fondamentali

Definizione 61 *Sia K un campo. Si definisce* matrice *un oggetto di elementi $a_{ij} \in K$ $(i = 1, 2, ..., m, j = 1, 2, ..., n)$ e si indica con*

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \ddots & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix}$$

o in forma contratta $A = (a_{ij})$. Si dice che gli elementi sono disposti secondo m righe *e n* colonne*.*

Definizione 62 *Si definisce* matrice nulla*, una matrice in cui $a_{ij} = 0$ $(i = 1, 2, ..., m, j = 1, 2, ..., n)$.*

Definizione 63 *Si definisce $K^{m,n}$ lo spazio vettoriale delle matrici di m righe e n colonne a cofficienti in K* [1]*.*

Osservazione.
Se $m = 1$, allora le matrici appartenenti a $K^{1,n}$ altro non sono che le $n-$uple ordinate di K^n.

Definizione 64 *Si definisce* somma *di due matrici $A, B \in K^{m,n}$, con $A = (a_{ij})$ e $B = (b_{ij})$, la matrice $C \in K^{m,n} : C = A + B = (a_{ij} + b_{ij})$.*

[1] Tale proposizione va dimostrata: una volta definite le operazioni di composizione tra matrici ciò sarà possibile e ne viene lasciata al lettore la prova.

Definizione 65 *Si definisce* prodotto *di uno scalare $a \in K$ per una matrice $A \in K^{m,n}$, con $A = (a_{ij})$, la matrice $C \in K^{m,n} : C = a \cdot A = (aa_{ij})$.*

Adesso è possibile dimostrare facilmente che per la struttura $(K^{m,n}, +, \cdot)$ valgono gli assiomi e le proprietà si spazio vettoriale.

Definizione 66 *Si definisce* matrice quadrata *una matrice $A \in K^{n,n}$ a coefficienti in K, $A = (a_{ij})$ ($i = 1, 2, ..., n$,$j = 1, 2, ..., n$).*

Definizione 67 *Sia $A \in K^{n,n}$: si definisce* diagonale principale*, la linea costituita da tutti gli elementi $a_{ij} \in A : i = j$.*
Si definisce diagonale secondaria*, la linea costituita da tutti gli elementi $a_{ij} \in A : i + j = n + 1$.*

Definizione 68 *Una matrice $A \in K^{n,n}$ si dice* diagonale *se $a_{ij} \in A : a_{ij} = 0, \forall i \neq j$.*

Definizione 69 *Si definisce* matrice scalare *una matrice diagonale i cui elementi coincidono.*

Definizione 70 *Si definisce* matrice identica (unità) *$I_n \in K^{n,n}$, la matrice di elemento scalare 1.*

Teorema 33 *Ogni matrice scalare è della forma $A = aI_n$ ($\forall a \in K, \forall A \in K^{n,n}$).*

Dimostrazione 23 *La dimostrazione discende dalla definizione di prodotto e di matrice identica.*

Definizione 71 *Sia $A \in K^{m,n}$: si definisce* matrice trasposta *di $A = (a_{ij})$ la matrice $A_t \in K^{n,m} : A_t = (a_{ji})$.*

Osservazione.
La trasposta di una matrice, si ottiene 'ruotando' i suoi elementi in senso orario sulla diagonale principale.

Definizione 72 *Sia $A \in K^{n,n}$: si definisce* traccia *di $A = (a_{ij})$ la somma degli elementi della diagonale principale.*

Definizione 73 *Sia $A \in K^{n,n}$: $A = (a_{ij})$ si definisce* triangolare *se gli elementi al di sotto della diagonale principale sono nulli.*

8.2 Prodotto tra matrici

Il prodotto tra matrici è in generale meno intuitivo della somma. E' importante sottolineare che la definizione di tale prodotto è stata fatta in modo che fossero rispettate diverse proprietà comuni in uno spazio vettoriale.

Definizione 74 *Siano* $A \in K^{m,n} : A = (a_{ij})$ *e* $B \in K^{n,r} : B = (b_{i,j})$. *Si definisce* matrice prodotto $C \in K^{m,r} : C = AB = (c_{ij})$, *con*

$$c_{ij} = \sum_{h=1}^{m} a_{ih}b_{hj} = (a_{i1}b_{1j}) + (a_{i2}b_{2j}) + ... + (a_{im}b_{mj})$$

Spesso si dice che tale prodotto è riga per colonna.

Osservazione.
Come appena detto, il prodotto è stato definito in modo tale da far valere determinate proprietà, come quella associativa.
Infatti se $A \in K^{m,n}$, $B \in K^{n,r}$, $D \in K^{r,s}$, allora $AB \in K^{m,r}$, $BD \in K^{n,s}$ e $(AB)D \in K^{m,s}$; d'altra parte $A(BD) \in K^{m,s}$.

Esempio.

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & -1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1*1+0*2 & 1*0+0*3 & 1*1+0*(-1) \\ 2*1+1*2 & 2*0+1*3 & 2*1+1*(-1) \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 1 \\ 4 & 3 & 1 \end{pmatrix}$$

Da notare che il prodotto non è assolutamente un'operazione interna: almeno in generale, poichè lo è invece nelle matrici quadrate.
Di conseguenza $K^{n,n}$ è un semigruppo e si verifica facilmente che se $A \in K^{n,n}$:

$$AI_n = I_nA = A$$

Si dimostra anche con qualche esempio particolare, che in generale il prodotto tra due matrici non è per nulla commutativo. Abbiamo parlato per questo motivo di semigruppo, in quanto in generale non è vero che ogni matrice ammette inversa B tale che

$$Ab = BA = I_n$$

Come vedremo successivamente, ci sono determinate condizioni perchè una matrice ammetta inversa.
Per costruire un gruppo su $K^{n,n}$ rispetto al prodotto esterno, si prendano le matrici che hanno inversa (di cui esiste almeno ${I_n}^2$).
Indicando con $GL_n(K)$ il gruppo lineare delle matrici invertibili di ordine n nel campo $K^{n,n}$[3], che è un gruppo poichè prese A, B in esso, la matrice AB è invertibile con inversa $(AB)^{-1} = B^{-1}A^{-1}$[4].
Da notare che l'inversa di una matrice inversa è invertibile: infatti nessuno vieta di considerare l'inversa dell'inversa come l'inversa come inversa della matrice inversa.
L'inversa del prodotto precedentemente definito, è dunque tale che, ricordando l'associatività:

$$(AB)(AB)^{-1} = (AB)(B^{-1}A^{-1}) = (AI_n)A^{-1} = AA^{-1} = I_n$$
$$(BA)(BA)^{-1} = (BA)(A^{-1}B^{-1}) = (BI_n)B^{-1} = BB^{-1} = I_n$$

8.3 Il determinante

Definizione 75 *Si definisce* prodotto dedotto *della matrice $A \in K^{n,n}$, il prodotto di n elementi di A in modo tale che ogni fattore compaia una e una sola volta come rappresentante dell'i−esima riga e della j−esima colonna.*

Il prodotto dedotto è spesso definito anche con l'*annullamento riga-colonna*.
La definizione appena data indica che il prodotto risultante contiene una e una sola volta un elemento appartenente ad una riga e colonna: ciò significa brevemente, per esempio, che $a_{15} * a_{a13} * ... * a_{23} * a_{54}$ non è un prodotto dedotto, poichè esistono due elementi della stessa riga, la 1, e due della stessa colonna, la 3.
Per l'estrazione di tutti i prodotti dedotti, che in numero sono $n!$[5], si procede permutando un insieme di indici: in generale $P = \{i_1, i_2, ..., i_n\}$ può essere messo in corrispondenza biunivoca con $\{1, 2, ..., n\}$, ma anche ogni permutazione degli elementi di P può esserlo, dunque si può assumere come insieme

[2]Infatti $I_n \cdot I_n = I_n$.
[3]Si noti che $GL_n(K) \subseteq K^{n,n}$.
[4]Si ricorda che il prodotto tra matrici non è commutativo e che pertanto l'ordine dei moltiplicandi è fondamentale.
[5]Questa scrittura indica il fattoriale di n, che è il prodotto $1 * 2 * 3 * ...(n-1) * n$. Ciò è vero poichè scelto un elemento in una riga, se ne possono scegliere tra $n^2 - (2n-1)$, vista l'eliminazione riga-colonna; tuttavia scelto un secondo termine, si annullano altre due linee di n elementi e così via: in generale il primo termine può essere scelto tra n, il secondo tra $n-1$, il terzo tra $n-2$, etc. In totale allora il loro numero è proprio $n!$.

dei prodotti dedotti, quello costituito dai prodotti del tipo

$$\prod_{h=1}^{n} a_{hi_h} = a_{1i_1} * a_{2i_2} * ... * a_{ni_n}$$

per tutte le permutazioni degli indici.
Ad ogni prodotto dedotto viene associata una parità (o segno), in generale restituito da $(-1)^r$, dove r è il numero di scambi tra indici, necessari per risalire alla permutazione fondamentale (quella ordinata). Per esempio $a_3a_2a_1$ ha classe dispari, poichè sono necessari gli scambi $a_3a_1a_2$, $a_1a_3a_2$, $a_1a_2a_3$, per arrivare alla permutazione ordinata; in totale gli scambi sono 3. In generale se $r = 2q$ allora si parlerà di **classe pari**, se $r = 2q + 1$ si parlerà di **classe dispari**.
Premesso ciò è possibile dare la seguente

Definizione 76 *Si definisce* determinante *di una matrice $A \in K^{n,n}$, l'elemento di K indicato come $det(A)$ tale che*

$$det(A) = \sum (-1)^r a_{1i_1} * a_{2i_2} * ... * a_{ni_n}$$

cioè definito come la somma di tutti i possibili prodotti dedotti.

Teorema 34 *Sia $A \in K^{n,n}$. Allora $det(A_t) = det(A)$.*

Dimostrazione 24 *Il teorema afferma che il determinante di una matrice e della sua trasposta coincidono. Ciò è vero dalla definizione di trasposta, in quanto essa muta l'ordine (ma non la parità) degli elementi della matrice, lasciando invariati i prodotti dedotti.*

Indicando con il termine *linea* una qualsiasi colonna o riga:

Teorema 35 *Se una matrice quadrata A contiene una linea nulla, $det(A) = 0$.*

Dimostrazione 25 *Ciò discende dalla definizione di determinante: poichè ogni prodotto dedotto deve contenere ogni elemento di una linea, anche se uno e un solo, esso avrà come fattore sempre almeno uno zero, da cui la sommatoria sarà nulla.*

Teorema 36 *Scambiando due linee parallele di una matrice quadrata A, il nuovo determinante cambia di segno rispetto a quello originale.*

Dimostrazione 26 *Sempre dalla definizione di determinante, abbiamo che i prodotti dedotti restano identici, a variare è solo la parità, in quanto tutte le classi dispari diventano pari e viceversa. Di conseguenza il determinante cambia di parità.*

Teorema 37 *Se una matrice quadrata A contiene due linee parallele uguali, allora $det(A) = 0$.*

Dimostrazione 27 *Infatti scambiando queste due linee, otteniamo l'equazione $det(A) = -det(A)$, che è vera se e solo se $det(A) = 0$.*

Teorema 38 *Moltiplicando gli elementi di una linea per uno scalare a, si ottiene una nuova matrice il cui determinante è a volte quello della matrice originale.*

Dimostrazione 28 *Dalla definizione di prodotto dedotto, avremo che ogni termine della somma per il calcolo del determinante, conterrà una e una sola volta lo scalare a. Poichè vale la proprietà distributiva tra gli scalari, possiamo mettere in evidenza a ottenendo le somme dei prodotti dedotti della matrice originale. Dunque $det(B) = a * det(A)$. In particolare si ha che $det(\lambda A) = \lambda^n det(A)$.*

Teorema 39 *Se due linee di una matrice A sono proporzionali, allora $det(A) = 0$.*

Dimostrazione 29 *Sia $L_1 = \mu L_2$. Per il teorema precedente si ha che $det(B) = \mu det(A)$, poichè la matrice B privata della linea con μ, diventa una matrice A. Tuttavia A ha due linee parallele uguali, dunque $det(A) = 0$, da cui la tesi.*

Teorema 40 *Se una linea L di una matrice A è somma di due $n-uple$ $L' + L''$, allora $det(A) = det(A') + det(A'')$, dove A' è la matrice che si ottiene sostituendo L' a L e A''quella ottenuta sostituendo L'' a L.*

Dimostrazione 30 *Dalla definizione di prodotto dedotto, abbiamo che ogni prodotto conterrà uno e un solo termine della forma $a'_{ji_j} + a''_{ji_j}$ nella $j-esima$ riga per esempio. Applicando la proprietà distributiva, si ottengono due sommatorie S' e S'' di termini, ciascuna contenente rispettivamente i termini della forma a'_{ji_j} e a''_{ji_j}. Ma queste due sommatorie possono essere interpretate come i determinanti di due nuove matrici, A' e A'', per cui $det(A) = det(A') + det(A'')$.*

Teorema 41 *Se ad una linea di A si somma una linea parallela moltiplicata per uno scalare, il determinante non cambia.*

Dimostrazione 31 *La dimostrazione discende dall'applicazione dei due teoremi precedenti: si scompone il determinante in quello di due matrici, una contenente l'addendo non moltiplicato per lo scalare e l'altra che invece lo contiene. Ma il det della seconda matrice è nullo per i teoremi precedenti, dunque il determinante non cambia.*

Teorema 42 *Se una linea di A è combinazione lineare di altre linee parallele allora* $det(A) = 0$.

Dimostrazione 32 *Applicando il teorema precedente si nota invece che i determinanti delle due matrici, ottenute scindendo la linea di quella originale, sono entrambi nulli.*

Teorema 43 *Il determinante di una matrice triangolare è uguale al prodotto dei termini che si trovano sulla diagonale principale.*

Dimostrazione 33 *La dimostrazione discende dalle definizioni di matrice triangolare e di prodotto dedotto.*

Si può generalizzare il teorema precedente come segue.

Definizione 77 *Si definisce* elemento speciale*, quell'elemento di posto tale che al di sotto abbia tutti elementi nulli.*

Teorema 44 *Il determinante di una matrice è uguale al prodotto degli elementi speciali.*

Si nota subito, che quello della matrice triangolare è un caso particolare, dove gli elementi speciali si trovano tutti sulla diaginale principale.

Se si applica un **procedimento di diagonalizzazione**, cioè un algoritmo tale, che a partire dalle proprietà del determinante appena dimostrate, fornisca una matrice di elementi speciali, si esegue una ***tecnica standard*** *di calcolo del determinante.*
Il procedimento di diagonalizzazione vale anche se applicato per colonna o da destra verso sinistra.

8.4 Il complemento algebrico

Definizione 78 *Si definisce* sottomatrice *di una matrice $A \in K^{m,n}$, una matrice $A_s \in K^{p,q}$ $(m \leq p, n \leq q)$, i cui elementi sono dati dall'intersezione delle p righe con le q colonne.*

Una volta scelta una sottomatrice, la 'restante', costituita dagli elementi che non fanno parte di A_s, viene chiamata **complementare**, ed è anch'essa una sottomatrice $\overline{A_s}$.
Se $A \in K^{n,n}$ e A_s è quadrata, allora anche $\overline{A_s}$ è quadrata.

Particolarmente interessanti sono quelle sottomatrici di un solo elemento, tali da avere complementare di ordine $n-1$. Il numero totale di tali matrici è ovviamente n^2.

Definizione 79 *Si definisce complemento algebrico di a_{ij}, il determinante della sua matrice complementare moltiplicato per $(-1)^{i+j}$.*

Vedremo che il complemento algebrico è strettamente legato al determinante.

Definizione 80 *Si definisce matrice* aggiunta A_a *di una matrice A, quella matrice che ha per elementi i complementi algebici di degli elementi di A.*

8.5 Teorema di Laplace

Il teorema di Laplace, per la sua importanza, merità di essere trattato distintamente.

Teorema 45 *La somma dei prodotti degli elementi di una linea di una matrice quadrata $A \in K^{n,n}$ per il loro rispettivo complemento algebrico è uguale al determinante di A:*

$$det(A) = \sum_{k=1}^{n} a_{ik}A_{ik} = \sum_{k=1}^{n} a_{kj}A_{kj} \quad (\forall i, j = 1, 2, ..., n)$$

Più che una vera e propria dimostrazione, tenteremo di dare un'idea di essa.

Dimostrazione 34 *Il determinante si può scrivere come*

$$det(A) = \prod_{k=1}^{n} a_{ji_j}$$

Se $i_1 = 1$ allora

$$det(A) = a_{11}[\sum(-1)^r \prod_{j=2}^{n} a_{ji_j}] + a_{12}[\sum(-1)^r \prod a_{ji_j}] + ...$$

Ma i termini in parentesi quadre, si verifica che coincidono con i complemento algebrici degli elementi anteposti ad ogni prodotto. Si verifica che la parità è comunque rispettata e pertanto il teorema è stato 'dimostrato'.

Il lettore capirà che sviluppare tale dimostrazione in maniera meno rigorosa è un grande risparmio di calcoli.

Corollario 2 *La somma dei prodotti degli elementi di una linea per i complementi algebrici di un'altra linea è nulla:*

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \sum_{k=1}^{n} a_{ki}A_{kh} = 0 \quad (\forall i, j = 1, 2, ..., n : i \neq j \neq h)$$

8.6 Teorema di Binet

Il teorema di Binet è di importanza fondamentale almeno quanto quello di Laplace, ma di dimostrazione molto più complessa che peranto viene omessa.

Teorema 46 *Siano $A, B \in K^{n,n}$: allora $det(AB) = det(A)det(B)$.*

Teorema 47 *Data una matrice $A \in K^{n,n}$, essa è invertibile se e solo se $det(A) \neq 0$.*

Dimostrazione 35 *Dimostriamo l'implicazione $\Longrightarrow$.*
Se A è invertibile, allora

$$\exists A^{-1} \in K^{n,n} : A^{-1}A = AA^{-1} = I_n$$

Allora per il teorema di Binet

$$det(A^{-1}A) = det(I_n) = 1$$
$$det(AA^{-1}) = det(I_n) = 1$$

da cui

$$det(A^{-1})det(A) = 1$$
$$det(A)det(A^{-1}) = 1$$

che in entrambi casi è soddisfatta se e solo se $det(A) \neq 0$*, per la legge di annullamento del prodotto. Ovviamente anche* $det(A^{-1}) \neq 0$ *per lo stesso motivo, ma ciò significa, come già detto in precedenza, che una matrice inversa è essa stessa invertibile, e questa ne è una dimostrazione da porre come corollario a quella appena data.*
Dimostriamo adesso l'implicazione $\Longleftarrow$*.*
Se $det(A) \neq 0$*, costruiamo* A^{-1}*. Sia* $(A_a)_t$[6]*, che dalle precedenti definizioni, una volta moltiplicata per A, vale di elementi* $(A_a)_i \cdot A_j$*, che rappresentano uno scambio di righe per colonne per via della trasposizione della matrice che aveva per elementi i complementi algebrici degli elementi di* a_{ij} *di A. Ma gli elementi sono i prodotti dei complementi algebrici per elementi di linee parallele (gli* A_j*) per ogni* $i \neq j$*, per cui sono tutti nulli. Per* $i = j$*, ovvero per gli elementi della diagonale principale, si ha il risultato costante di* $det(A)$*, che portato fuori dalla matrice per la ben nota proprietà, restituisce* $(A_a)_t \cdot A = (det(A)) \cdot I_n$*. Se* $det(A) \neq 0$*, allora*

$$[\frac{1}{det(A)}(A_a)_t]A = I_n$$

dove la matrice tra parentesi quadre è proprio tale che moltiplicata per A dia l'identica, cioè è l'inversa di A.
Analogamente si verifica che vale ugualmente il teorema nel caso in cui i moltiplicandi sono scambiati di posto, cioè vale la proprietà commutativa.

8.7 Il rango

Il *rango* è un altro elemento fondamentale nello studio delle matrici di ordine $m \times n$, come vedremo in questo paragrafo.

Definizione 81 *Si definisce* rango *(o* caratteristica*)* $rg(A)$ *di una matrice* $A \in K^{m,n}$*, l'ordine massimo dei minori dal determinante non nullo che si possono estrarre da A.*

Il numero di tutti i possibili minori estraibili, sono dati dal coefficiente binomiale $\binom{m}{n}$. Una tecnica standard per il calcolo del rango è quella di ridurre una matrice rettangolare ad una triangolare o in generale di una matrice di elementi speciali.

Osservazione.
Si ha che $rg(A) = r$ se e solo se:

[6] Ovvero la trasposta della matrice aggiunta.

- Esiste un minore di ordine r non nullo[7];
- Tutti i minori di ordine $> r$ sono nulli.

equivalentemente alla definizione appena data.

Osservazione.
Anche in questo caso il teorema di Laplace risulta utile. Possiamo restringere il campo di ricerca del rango ai minori di ordine $r+1$ nulli: infatti visto che i minori di ordine $r+2$ contengono quelli di ordine $r+1$, i cui determinanti sono nulli, si ha che anche i determinanti dei minori di ordine $r+2$ sono nulli per via dei complementi algebrici nulli (dal teorema di Laplace).
Possiamo definire rigorosamente questo procedimento, enunciando il seguente teorema di cui è stata appena data una dimostrazione rapida:

Teorema 48 ***(di Lapace per il rango)*** *Sia $A \in K^{m,n}$ una matrice. Se i minori di ordine $r+1$ sono nulli, allora $rg(A) = r$.*

Oltre a definire un rango per una matrice intera, possiamo definirlo solo per righe o solo per colonne, assumendo lo spazio vettoriale

$$R^* = L(A^1, A^2, ..., A^m) \subseteq K^n$$

generato dalle righe di A, e quello delle colonne

$$C^* = L(A_1, A_2, ..., A_n) \subseteq K^m$$

dove vale la relazione

$$A = \begin{pmatrix} A^1 \\ A^2 \\ \vdots \\ A^m \end{pmatrix} = (A_1, A_2, ..., A_n) \quad (A^i \in K^n; A_j \in K^m)$$

Applicando il metodo degli scarti agli elementi di A che potrebbero essere nulli o linearmente dipendenti, si trova una base di A, che genera una matrice A' il cui rango è $r = rg(A^i) \leq n$, e analogamente per le colonne: $c = rg(A_j) \leq m$, da cui $dim(R^*) = r$ e $dim(C^*) = c$. Dimostreremo che $dim(R^*) = dim(C^*) = rg(A)$, ovvero che il rango coincide con il numero di linee parallele linearmenti indipendenti.

Lemma 1 *Sia $A \in K^{m,n}$, con $m \leq n$. I minori di ordine m sono tutti nulli se e solo se le righe di A sono linearmente dipendenti.*

[7] E' sottointeso, che ci riferiamo a *i minori che hanno determinante non nullo.*

Dimostrazione 36 *Dimostriamo l'implicazione* $\Longleftarrow$.
I minori di ordine m, contengono linee linearmente indipendenti, e dunque il loro determinante è sempre nullo.
Dimostriamo l'implicazione $\Longrightarrow$.
Procediamo per assurdo, ammettiamo che le m righe siano linearmente indipendenti. Dobbiamo dimostrare che almeno un minore è non nullo.
Estendiamo l'insieme libero $A^1, A^2, ..., A^m$ *ad una base di* K^n, *trovando* $A^1, A^2, ..., A^m, A^{m+1}, ..., A^n$. *Siano*

$$\begin{array}{c} e_1 = (1, 0, ..., 0) \\ e_2 = (0, 1, ..., 0) \\ ... \\ e_n = (0, 0, ..., 1) \end{array}$$

i vettori della base canonica. Possiamo esprimerli come combinazione lineare degli A_i*:*

$$e_i = \sum_{h=1}^{n} m_{ih} A^h \quad (i = 1, 2, ..., n)$$

che in totale sono $n \times n = n^2$, *che danno luogo ad una matrice quadrata di ordine* n $M = (m_{ij})$. *Posti per righe gli* e_i, *otteniamo ovviamente la matrice identica, che è uguale al prodotto riga per colonna di due matrici,* M, *e un'altra vhe chiamiamo* A^*, *dove*

$$A^* = \begin{pmatrix} A^1 \\ A^2 \\ ... \\ A^n \end{pmatrix}$$

Dal teorema di Binet si ha che $1 = det(M)det(A^*)$, *da cui* $det(A) \neq 0$.
Ora se i minori di ordine r di $(A^1, A^2, ..., A^r)^t$ *fossero tutti nulli, lo devono essere anche quelli di ordine superiore, per il teorema di Laplace, ottenuti aggiungendo le righe* $A^{r+1}, A^{r+2}, ..., A^n$, *per cui* $det(A^*) = 0$, *mentre abbiamo dimostrato il contrario.*
Senza ambiguità, si può estendere allo spazio delle colonne questo lemma[8].

Teorema 49 *Sia* $A \in K^{m,r}$. *Allora il rango per righe r è uguale al rango per colonne c.*

[8]Suggerimento: basta prendere la trasposta delle righe.

Dimostrazione 37 *Sia $r = dim_K L(A^1, ..., A^m)$: cioè siano r le righe linearmente indipendenti in modo che si possa estrarre da A un minore di ordine r non nullo (per il lemma precedente ciò è possibile).*
Le colonne di A che contengono tale minore sono linearmente indipendenti, per il lemma precedente, pertanto si può constatare che, se
$c = dim_K L(A_1, ..., A_m)$, allora è $c \geq r$.
Tuttavia il punto di partenza può essere assunto dalle colonne per giungere alla relazione $c \leq r$.
Ma allora $c = r$.

Teorema 50 *Sia $A \in K^{m,n}$: allora posto $r = dim_K(R^*)$ e $c = dim_K(C^*)$, vale la relazione $rg(A) = c = r$.*

Dimostrazione 38 *Sia ρ l'ordine massimo dei minori non nulli di A. Allora esiste un minore di ordine r, non nullo, e tutti i minori di ordine maggiore di ρ sono nulli.*
Per il lemma che abbiamo dimostrato, abbiamo che il numero massimo di righe linearmente indipendenti è ρ, poichè se fosse un $\rho + \epsilon$ avrei un minore nullo. Ne segue necessariamente che $\rho = r$, e dunque, per il teorema precedente, è anche $\rho = c$, da cui $rg(A) = r = c$.

Teorema 51 ***(di Kronecker)*** *Sia $A \in K^{m,r}$. Allora $rg(A) = r$ se e solo se esiste un minore M di ordine r non nullo e sono nulli tutti i minori di ordine $r + 1$ che lo contengono.*

Il teorema in sè sembra essere una ripetizione del teorema di Laplace per il rango: tuttavia Kronecker qui restringe il campo di ricerca *ai soli minori di ordine $r+1$ che contengono M*, e non a tutti i minori di ordine $r+1$ estraibili da A.

Dimostrazione 39 *Dimostriamo l'implicazione $\Longrightarrow$.*
Questa implicazione è ovvia e scaturisce dalle definizioni e dal teorema di Laplace per il rango.
Dimostriamo l'implicazione $\Longleftarrow$.
Le righe che contengono M sono linearmente indipendenti, dunque ogni altra riga è combinazione lineare di queste perchè se così non fosse troverei minori di ordine maggiore di r che conterrebbero M, contro le ipotesi.

8.8 Applicazioni lineari

Definizione 82 *Si definisce* omomorfismo *un'applicazione tra due strutture dello stesso tipo che conserva le proprietà.*

Definizione 83 *Siano V, W spazi vettoriali su K. Si definisce* applicazione lineare *$f : V \longrightarrow W$ una f tale che $f(av + bv') = af(v) + bf(v')$, $\forall v, v' \in V$, $\forall a, b \in K$.*

Definizione 84 *Si definisce* Immagine *di f $Im(f)$, la totalità dei vettori $\{w : w = f(v), v \in V\}$.*

L'immagine non è sicuramente vuota, poichè contiene almeno l'elemento nullo $w_0 = f(\overline{0})$. Si verifica facilmente che essa è un sottospazio di W.

Definizione 85 *Si definisce* Nucleo *(o* kernel*) di f $Ker(f)$, l'insieme $\{v \in V : f(v) = \overline{0}\}$.*

Si verifica che anche $Ker(f)$ è un sottospazio di V.

Sia $Hom_K(V, W) = \{f : V \longrightarrow W\}$ l'insieme di tutte le f lineari definite in V, W.

Definizione 86 *Definiamo come* somma *di $f, g \in Hom_K(V, W)$, l'applicazione binaria $(f + g)(v) = f(v) + g(v)$*[9].
Definiamo prodotto di $a \in K$ per $f \in Hom_K(V, W)$, l'applicazione binaria $(a \cdot g)(v) = a \cdot g(v)$.

Sia $dimV = n$ e $dimW = m$. Siano

$$B = \{v_1, ..., v_n\}$$
$$C = \{w_1, ..., w_n\}$$

rispettivamente una base di V e W.
Presa una matrice $A \in K^{m,n}$, costruiamo una funzione lineare di $V \longrightarrow W$. Sia $v \in V : v = \sum x_i v_i = (\underline{x})_B$[10], dove $\underline{x} = (x_1, ..., x_n)^t$, cioè è una $n-$upla per colonna.
Se

$$f_A : V \longrightarrow W : f_A(v)(A\underline{x})_C$$

dove quella tra parentesi è una matrice $m \times n * n \times 1 = m \times 1$. In pratica si scelgano gli f_A che sono in numero m, come combinazione lineare per descrivere tutti i $w \in W$.

[9] Si verifica facilmente che l'insieme di tali funzioni in cui sia definita tale applicazione, è un gruppo abeliano, una volta definita anche la *funzione nulla*.

[10] Quest'ultima è un'altra notazione generalmente utilizzata per indicare un vettore in funzione delle componenti rispetto ad una base specifica.

Allora f_A è lineare. E andiamo a dimostrarlo.
Presi $v, v' \in V$, con $v = (\underline{x})_B$ e $v = (\underline{x'})_C$, sia dato $av + bv'$ rispetto alla base B. Avremo:

$$av + bv' = a(\sum_{i=1}^{n} x_i v_i) + b(\sum_{i=1}^{n} x'_i v_i) =$$
$$= (ax_1 + bx'_1)v_1 + ... + (ax_n + bx'_n)v_n$$

cioè $av + bv' = (a\underline{x} + b\underline{x'})_B$, da cui deriva subito che $f_A(av + bv') = (A(a\underline{x} + b\underline{x'}))_C$. Poichè il prodotto riga per colonna è distributivo, tale relazione può essere scritta come $f_A(av + bv') = (A(a\underline{x}) + A(b\underline{x'}))_C$. Ma le costanti, per proprietà delle matrici, possono essere messe in evidenza, dunque:

$$f_A(av + bv') = (A(a\underline{x}) + A(b\underline{x'}))_C = (a(A\underline{x}) + b(A\underline{x'}))_C = a(A\underline{x})_C + b(A\underline{x'})_C =$$
$$= af_A(v) + bf_A(v')$$

che completa la dimostrazione.

In questo modo siamo partiti da $V^n \longrightarrow W^m$ per giungere a $K^{m,n}$ con indici invertiti, appunto...
Abbiamo costruito $A \longrightarrow f_A$ dove $A \in K^{m,n}$ e $f_A \in Hom_k(V, W)$, che a sua volta è una funzione tra due spazi vettoriali: $K^{m,n}$ e $Hom_k(V, W)$; bisogna verificare se anche tale applicazione è lineare, ovvero se realizza un isomorfismo $\Phi : \Phi(A) = f_A$.
Φ ha bisogno sia della base B che della base C di A, V, W. In termini concreti, al variare anche di un solo parametro, Φ varia, e questo suo comportamento si indica con la dicitura Φ *non è canonica*.
Siano $a, b \in K$ e $A, B \in K^{m,n}$. Chi è $\Phi(aA + bB)$? Essa è un'applicazione $f_{aA+bB} : V \longrightarrow W$ tale che:

$$f_{aA+bB}(v) = f_{aA+bB}(\underline{x})_B = ((aA + bB)\underline{x})_C =$$
$$= (aA\underline{x} + bB\underline{x})_C = a(A\underline{x})_C + b(B\underline{x})_C = af_A(v) + bf_B(v) =$$
$$= (af_A + bf_B)(v) = (a\Phi(A) + b\Phi(B))(v)$$

cioè Φ è lineare. Per dimostrare l'isomorfismo occorre infine dimostrare che valgono iniettività e suriettività.

Capitolo 9

Equazioni lineari

9.1 Nozioni fondamentali

In questo capitolo tratteremo la teoria delle equazioni lineari e soprattutto dei *sistemi* di equazioni lineari, adoperando le conoscenze fin qui acquisite.

Definizione 87 *Si definisce* equazione lineare *a n incognite o parametri, un'equazione del tipo*

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \qquad a_i, b \in K$$

dove gli scalari a_i del campo K prendono il nome di coefficienti *e i vettori x_i dello spazio vettoriale V_K prendono il nome di* variabili. *Lo scalare b prende il nome di* termine noto.

Definizione 88 *Un'equazione lineare in n incognite si dice* omogenea *se il suo termine noto è nullo.*

Poichè si parla di *equazioni*, è lecito chiedersi se esistono dei metodi algebrici per determinare le loro soluzioni.
Tali metodi, come vedremo tra poco, esistono e sono strettamente legate al numero di equazioni e al numero di incognite.
Per esempio l'equazione $3x_1 + 2 = 0$ ammette una e una sola soluzione ottenibile con le proprietà che regolano le somme e i prodotti all'interno di uno spazio vettoriale.
Al contrario, l'equazione lineare $3x_1 + 2x_2 = 0$, ammette infinite coppie di soluzioni[1], a meno che non sia ben determinata un'altra equazione, indipendente da quella precedente, in cui sia data un'altra relazione tra le due

[1]Da notare che le infinite soluzioni esistono solo se il campo K è infinito. E' questo il caso del campo dei numeri razionali, dei numeri reali e dei numeri complessi.

incognite.
In realtà il lettore non dovrebbe essere sorpreso di questo, poichè quando abbiamo trattato gli spazi vettoriali, abbiamo dimostrato come per una soluzione univoca occorrono tante condizioni di linearità quante sono le variabili.
Diamo ora la seguente

Definizione 89 *Si definisce* sistema lineare *di m equazioni lineari in n incognite, l'intersezione delle m $n-$uple risolventi ciascuna equazione.*

Un sistema lineare è generalmente indicato con la notazione

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases}$$

In questa scrittura è facilmente determinabile un qualunque coefficiente dell'$i-$esima equazione nella $j-$esima incognita, in quanto è denotato con gli indici doppi a_{ij}.
Come abbiamo detto in precedenza, ci interessano le soluzioni comuni di queste equazioni, pertanto operiamo come segue:

- Costruiamo la **matrice dei coefficienti del sistema** $A = (a_{ij})$, con $A \in K^{m,n}$ (anche detta **matrice incompleta**);
- Costruiamo la **matrice completa** $B = (A, \underline{b})$, identica a quella incompleta ma con l'aggiunta della colonna dei termini noti a destra;
- Costruiamo la **matrice colonna delle incognite** $\underline{x} = (x_1, x_2, ..., x_n)_t$, ovvero la trasposta della matrice riga, con $\underline{x} \in K^{n,1}$, e definita in questo modo in modo da poter procedere con il prodotto con A;
- Costruiamo la **matrice colonna dei termini noti** $\underline{b} = (b_1, b_2, ..., b_m)_t$, con $\underline{b} \in K^{m,1}$.

A questo punto la risoluzione del sistema coincide con la soluzione dell'equazione matriciale

$$A\underline{x} = \underline{b} \qquad A\underline{x} \in K^{m,1}$$

Il sistema $A\underline{x} = \underline{0}$ è detto **omogeneo associato** a $A\underline{x} = \underline{b}$, e ci servirà nella risoluzione.
Infatti A può essere interpretata come una matrice della funzione lineare

$$f_A : K^m \longrightarrow K^n : \underline{x} \longrightarrow A\underline{x}$$

rispetto alle basi canoniche per entrambi.
Le soluzioni di $A\underline{x} = \underline{0}$ sono elementi di $Ker(f_A)$, e le soluzioni di $A\underline{x} = \underline{b}$ sono gli elementi di K^n che stanno in $f_A^{-1}(\underline{b})$, perchè gli elementi del dominio di f_A sono le controimmagini del dominio.

9.2 Risoluzione di sistemi lineari: teorema di Cramer

Sia $\underline{x_0}$ una soluzione del sistema $A\underline{x_0} = \underline{b}$. Se $\underline{y}$ è un'altra soluzione tale che $A\underline{y} = \underline{b}$, allora $A\underline{x_0} = A\underline{y}$, da cui ne segue che $A\underline{x_0} - A\underline{y} = \underline{0}$ per l'operazione permessa in uno spazio vettoriale.
Dimostriamo il seguente teorema:

Teorema 52 *Dato un sistema lineare di m equazioni lineari in n incognite, la differenza di due sue soluzioni è ancora una soluzione del sistema.*

Dimostrazione 40 Infatti, $A\underline{x_0} - A\underline{y} = A(\underline{x_0} - \underline{y}) = \underline{0}$ per le proprietà delle matrici.

Inoltre vale il seguente

Teorema 53 *Dato un sistema lineare di m equazioni lineari in n incognite, se z è soluzione dell'omogeneo associato, allora $x_0 + z$ è una soluzione.*

Dimostrazione 41 Infatti $A(\underline{x_0} + z) = A\underline{x_0} + A\underline{y}$, che per le ipotesi è uguale al vettore nullo.

Poniamoci nel caso in cui $m = n$, cioè il sistema ha n equazioni in n incognite. Un sistema di questo tipo è detto **sistema quadrato**.

Teorema 54 (Teorema di Cramer) *Sia $A \in K^{m,n}$ e sia $A\underline{x} = \underline{b}$ un sistema lineare. Se $det(A) \neq 0$ allora il sistema ammette una e una sola soluzione.*

Dimostrazione 42 Se $\underline{\alpha} \in K^n$ è una soluzione, allora $A\underline{\alpha} = \underline{b}$. Poichè $det(A) \neq 0$, esiste $A^{-1} = \frac{det(A)}{(A_{ij})_t}$[2], da cui

$$A^{-1}(A\underline{\alpha}) = A^{-1}\underline{b} \Longrightarrow I_n\underline{\alpha} = A^{-1}\underline{b} \Longrightarrow \underline{\alpha} = A^{-1}\underline{b}$$

Se α esiste, allora è $A^{-1}\underline{b}$, che sostituito a $\underline{x}$ in $A\underline{x} = \underline{b}$ porta a un'identità.

[2]Ovvero la trasposta della matrice aggiunta per l'inverso del determinante di A.

Poichè $\underline{\alpha}$ è un'$n-$upla, per specificare un α_r $(1 \leq r \leq n)$, scriviamo

$$\alpha_r = \frac{1}{det(A)} \sum_{i=1}^{n} A_{ir} b_i$$

poichè l'$r-$esimo elemento è l'$r-$esima riga della matrice A, che è l'$r-$esima colonna della sua trasposta. Se

$$B_r = \begin{pmatrix} a_{11} & ... & b_1 & ... & a_{1n} \\ a_{12} & ... & b_2 & ... & a_{2n} \\ \vdots & & & & \\ a_{n1} & ... & b_n & ... & a_{nn} \end{pmatrix}$$

dove la colonna dei termini noti è la $r-$esima, si interpreta $\sum A_{ir} b_i$ come il determinante della matrice B_r appena costruita.
Operativamente dunque, per trovare una generica α_r, si sostituisce alla $r-$esima colonna della matrice A la colonna dei termini noti, si calcola il determinante di questa nuova matrice e lo si divide per quello di A.

Osservazione.
Per applicare il teorema di Cramer, è necessario che tutti i termini noti si trovino al secondo membro di ogni equazione lineare del sistema.

9.3 Il teorema di Rouchè-Capelli

Il teorema di Cramer ci dice come individuare l'$n-$upla risolvente del sistema. Tuttavia abbiamo bisogno di un criterio che ci dica quando ha senso procedere con il teorema di Cramer: per esempio è inutile andare a cercare soluzioni che non esistono.
Per questo motivo dimostriamo il seguente

Teorema 55 (Teorema di Rouchè-Capelli) *Sia $A \in K^{m,n}$, $\underline{b} \in K^n$ e $A\underline{x} = \underline{b}$ un sistema lineare qualunque. Tale sistema ha soluzioni se e solo se $rg(A) = rg(A, \underline{b})$*[3].

Dimostrazione 43 Dimostriamo l'implicazione $\Longrightarrow$.
Il sistema può essere scritto come $\sum_i A_i x_i = \underline{b}$, dove A_i è la matrice colonna

[3]Ovvero quando il rango della matrice completa e di quella incompleta coincidono.

dei coefficienti a_i.
Sia $\underline{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$ una soluzione del sistema: allora $\sum_i A_i \alpha_i = \underline{b}$, cioè

$$\underline{b} \in L(A_1, A_2, ..., A_n) \subseteq K^m$$

Il rango di A è

$$rg(A) = dim[L(A_1, A_2, ..., A_n)]$$

Il rango di $(A, \underline{b})$ è

$$rg(A, \underline{b}) = dim[L(A_1, A_2, ..., A_n, \underline{b})] = dim[L(A_1, A_2, ..., A_n)]$$

poichè è stato supposto che α è soluzione.

Dimostriamo l'implicazione $\Longleftarrow$.
Viceversa sappiamo per ipotesi che $rg(A) = rg(A, \underline{b})$. Inoltre

$$L(A_1, A_2, ..., A_n) \subseteq L(A_1, A_2, ..., A_n, \underline{b})$$

ma per le ipotesi allora

$$dim[L(A_1, A_2, ..., A_n)] = dim[L(A_1, A_2, ..., A_n, \underline{b})]$$

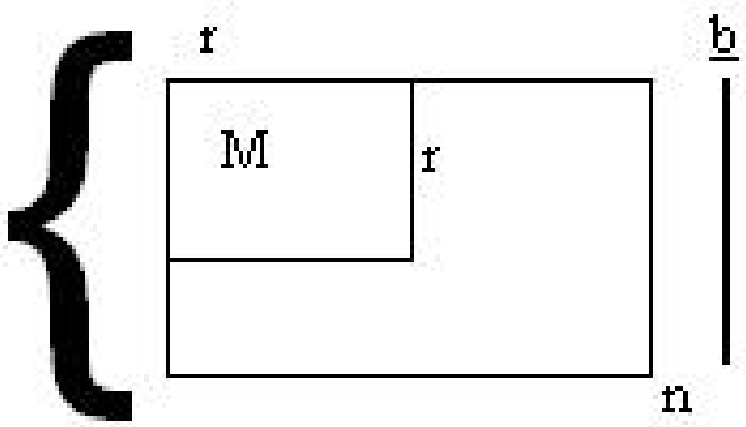
Ne segue che sono uguali e $\underline{b} \in L(A_1, A_2, ..., A_n)$, cioè

$$\exists \alpha_1, \alpha_2, ..., \alpha_n : \underline{b} = \sum_{i=1}^{n} \alpha_i A_i$$

e il teorema risulta dunque dimostrato.

9.4 Procedura di risoluzione

Sia $rg(A) = rg(A, \underline{b})$. Se vogliamo trovare le soluzioni del sistema $A\underline{x} = \underline{b}$, esisterà un minore M di ordine r non nullo e tutti i minori di ordine maggiore di r saranno nulli.
Pur di scambiare le equazioni, si può immaginare che M è contenuto nelle prime r righe. Inoltre per la proprietà commutativa della somma, $ax + by = by + ax = c$ sono equivalenti, cioè le soluzioni non cambiano se scambiamo le variabili. Cambiando l'ordine delle incognite dunque M è contenuto nelle prime r colonne.
Poichè i minori da ordine $r + 1$ in su sono nulli, significa che le righe restanti sono combinazione lineare delle prime r, e di conseguenza ogni soluzione delle prime r equazioni è soluzione delle altre $n - r$.
Per avere un'idea grafica:



Sia $\overline{A}\underline{x} = \underline{b}$ il sistema delle prime r equazioni, con $\overline{A} \in K^{r,n}$ e $\underline{b} \in K^r$. Riscrivendo il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1r}x_r = b_1 - (a_{1(r+1)}x_{r+1} + a_{1(r+2)}x_{r+2} + ... + a_{1n}x_n) \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2r}x_r = b_2 - (a_{2(r+1)}x_{r+1} + a_{2(r+2)}x_{r+2} + ... + a_{2n}x_n) \\ ... \\ a_{r1}x_1 + a_{r2}x_2 + ... + a_{rr}x_r = b_r - (a_{r(r+1)}x_{r+1} + a_{r(r+2)}x_{r+2} + ... + a_{rn}x_n) \end{cases}$$

Assegnamo adesso alle incognite del 2° membro valori ad arbitrio in K. Poichè $det(M) \neq 0$, i valori delle variabili $x_1, x_2, ..., x_r$ sono determinati dal teorema di Cramer e sono uniche se vengono ogni volta fissate le restanti da x_{r+1} in poi.
In totale ci sono tante soluzioni quanti sono gli elementi di K elevato a $n-r$, dove se K è infinito si dice che ci sono ∞^{n-r} soluzioni.