

Appunti
di

Geometria I

Philippe Ellia



Pitagora Editrice Bologna

V

Prefazione.

Partendo dal principio che è sempre preferibile una testa ben fatta piuttosto che una testa ben piena, questi appunti presentano una scelta sintetica di argomenti del programma standard di Geometria I (per il corso di laurea in Matematica).

Il primo capitolo è una specie di precorso che richiama nozioni generali di logica e di algebra.

Il secondo capitolo, quello più importante, introduce le nozioni di base dell'algebra lineare.

Il lettore osserverà che i capitoli successivi (geometria affine e geometria euclidea) possono essere visti come applicazioni, più o meno dirette, del secondo capitolo.

Il testo contiene più di 250 esercizi, che fanno parte integrante del corso; ritengo infatti che il solo apprendimento della teoria, senza il confronto con la pratica, sia un esercizio sterile.

Vari testi, elencati nella bibliografia alla fine del volume, sono stati consultati nella preparazione di queste note e potranno fornire al lettore interessato alcuni approfondimenti degli argomenti trattati.

Ringrazio vivamente Maria Chiara Gavioli per l'aiuto prezioso fornito durante la preparazione del manoscritto finale.

Indice

I) PRELIMINARI.

§1)	Elementi di logica, metodi di dimostrazione.....	1
§2)	Insiemi.....	12
§3)	Applicazioni.....	17
§4)	Relazioni d'equivalenza	26
§5)	Gruppi.....	32
§6)	Anelli, corpi, campi.....	40

II) ALGEBRA LINEARE.

§1)	Spazi vettoriali.....	47
§2)	Applicazioni lineari: definizione e prime proprietà.....	65
§3)	Spazi finitamente generati.....	69
§4)	Indipendenza lineare, basi.....	71
§5)	Teorema delle dimensioni, teorema di Grassmann; supplementari, quozienti	86
§6)	Anello degli endomorfismi, gruppo lineare.....	95
§7)	Forme lineari e dualità	97
§8)	Sistemi lineari omogenei e dualità.....	110
§9)	Scrittura matriciale delle applicazioni lineari.....	113
§10)	$M_{m,p}(k)$, $M_n(k)$, $GL_n(k)$ ed alcune matrici notevoli.....	122
§11)	Cambiamenti di basi.....	128
§12)	Rango di una matrice.....	134
§13)	Determinanti.....	137
§14)	Calcolo di un determinante.....	151
§15)	Rango e determinanti.....	161
§16)	Diagonalizzazione.....	166
§17)	Sistemi lineari.....	181

III) SPAZI AFFINI.

§1)	Sottospazi affini di uno spazio vettoriale.....	196
§2)	Equazioni degli sottospazi affini.....	205
§3)	Parallelismo ed incidentez.....	210
§4)	Incidenze nel piano e nello spazio.....	219
§5)	Riferimenti affini, affinità.....	224
§6)	Teoria generale.....	230

IV) SPAZI EUCLIDEI.

§1)	Forme bilineari.....	236
§2)	Forme bilineari simmetriche.....	240
§3)	Ortogonalità rispetto ad una forma bilineare simmetrica.....	244
§4)	Basi ortogonali (diagonalizzazione delle forme quadratiche)....	248
§5)	Basi ortonormali; teorema di Sylvester.....	251
§6)	Spazi metrici (cenni).....	258
§7)	Spazi vettoriali normati (cenni).....	261
§8)	Spazi vettoriali euclidei.....	263
§9)	Isometrie vettoriali.....	271
§10)	Angoli.....	278
§11)	Spazio affine euclideo.....	284
§12)	Teorema spettrale.....	291
§13)	Applicazioni alla classificazione delle coniche.....	295

1) LOGICA, METODI DI DEMOSTRAZIONE.

Tra le scienze esatte, la matematica è l'unica scienza ipotetico-deduttiva (la fisica, la chimica, ecc... sono scienze sperimentali). Quindi, partendo da enunciati "veri per ipotesi" (assiomi) e procedendo per "deduzioni logiche" (dimostrazioni), si ottengono altri enunciati veri (teoremi); poi si ricomincia: usando gli assiomi e i teoremi si dimostrano altri teoremi. Le deduzioni logiche sono regolate dalla logica formale. A priori gli assiomi non devono sottostare a nessuna condizione di verità "fisica" ma devono solo essere compatibili logicamente (non contraddittori). Pertanto esistono diversi sistemi ipotetico-deduttivi (o sistemi assiomatici), ossia diverse matematiche (con o senza l'assioma della scelta ecc...). Non è il caso qui di entrare nei dettagli, diciamo soltanto che il sistema più usato è quello di Zermelo-Freukel con l'assioma della scelta non numerabile. Questo sistema (chiamato nel seguito "la matematica") risponde bene alla nostra intuizione e permette di dimostrare il maggior numero di teoremi. Si osserverà tuttavia che un celebre risultato di Gödel afferma che non si può dimostrare (con gli strumenti della matematica) che la matematica non è contraddittoria! Però niente paura, sono più di duemila anni che si tira avanti così, nessuno ha trovato la contraddizione, i teoremi si accumulano e le applicazioni della matematica sono sotto gli occhi di tutti. Pertanto abbandoniamo le delicate questioni relative ai fondamenti della matematica e adottiamo il punto di vista più terra a terra "dell'addetto ai lavori".

Da questo punto di vista una dimostrazione è esattamente quello che avevano già stabilito gli Antichi Greci: uno sforzo cosciente per ordinare le argomentazioni in una successione tale che il passaggio da una tappa all'altra non lasci alcun dubbio; in modo che un virtuale interlocutore non potrebbe che acconsentire.

La presentazione formale non deve ingannare. L'attività matematica consiste nel risolvere problemi, teorici e pratici (dimostrare teoremi, risolvere esercizi). Le varie teorie non sono costruzioni formali sterili ma servono ad inquadrare, in un colpo solo, tutta una serie di problemi e quindi a facilitarne la comprensione e la risoluzione. Inoltre la dimostrazione di un teorema (o la risoluzione di un esercizio) non procede in modo razionale, formale, partendo dalle ipotesi e proseguendo in modo automatico fino a raggiungere la tesi. Sembra invece che si distinguano tre fasi: due razionali e una (la seconda, forse la più importante) non razionale. Nella prima fase (di analisi) si cerca di capire bene la domanda, il problema (tesi). Una volta che si è capito bene cosa si deve dimostrare, si guarda che cosa si ha a disposizione (ipotesi). A questo punto si conoscono sia il punto d'arrivo (tesi) che quello di partenza (ipotesi). Nella seconda fase, si cercherà di collegare questi due punti. In generale (ma non sempre!) il nostro cervello procederà per analogia. Partendo simultaneamente sia dal punto

d'arrivo che da quello di partenza, comincerà a prendere strade già note (esercizi già fatti, dimostrazioni viste a lezione, ecc...), finché (si spera) la connessione si stabilirà tra la tesi e l'ipotesi. A questo punto c'è un'idea di dimostrazione. La terza e ultima fase è di verifica, di aggiustamento. Si ripercorre il cammino in modo razionale e, partendo dalle ipotesi e procedendo per deduzioni (aggiustando magari i passaggi poco chiari), si raggiunge la tesi. Se salta fuori un errore non bisogna avere paura di ricominciare! L'ultima fase si concretizza tramite la redazione. La redazione finale deve essere ordinata, concisa, e limpida, senza passaggi oscuri. E' raccomandabile rileggersi con occhio critico e ricordarsi sempre che si scrive non per se stessi ma per comunicare con gli altri, "in modo che un virtuale interlocutore non potrebbe che acconsentire". Per raggiungere questa pratica (perché anche questo si impara) si raccomanda allo studente di esercitarsi a risolvere gli esercizi proposti e di redigerne completamente alcuni. Inutile dirlo: è più importante sapere risolvere gli esercizi che conoscere la teoria (anche perché è spesso impossibile risolvere gli esercizi senza conoscere la teoria).

In questo primo paragrafo introduciamo, senza tanti dettagli e giustificazioni, i connettori logici, i quantificatori universali ed esistenziali e le principali regole per utilizzarli. Il paragrafo si conclude con il richiamo(?) di alcuni metodi di dimostrazione, illustrati da esempi.

GENERALITÀ.

1: Dimostrazione: Una proposizione A è dimostrata (si dice anche che A è vera) se può essere inserita in un testo formalizzato che inizia con un assioma della teoria e che si sviluppa secondo le regole della teoria.

2: Negazione: Se A è una proposizione, la negazione di A è la proposizione $\neg A$ (si scrive anche $\text{non } A$).

Se $\neg A$ è vera (i.e. dimostrata), si dice che A è falsa.

3: Consistenza: Se in una teoria formalizzata, T, esiste una proposizione A tale che A sia vera e $\neg A$ sia vera si dice che T è contraddittoria (o inconsistente). In questo caso ogni proposizione di T è vera e falsa (cfr 12) e la teoria perde ogni interesse. Benché non sia dimostrato (e manca parecchio allo scopo!) si pensa generalmente che la teoria degli insiemi (i.e. "la matematica") sia consistente (i.e. non contraddittoria).

4: Decidibilità: Sia A una proposizione di una teoria assiomatizzata T. Se A non è vera (i.e. A non è dimostrata), ciò non vuol dire che A sia falsa (i.e. che $\neg A$ sia vera). Infatti può succedere che non si possa dimostrare in T

che una almeno delle due relazioni A, nonA sia vera. In questo caso la proposizione A si dice indecidibile.

Dire che una proposizione in T è indecidibile significa che gli strumenti a disposizione in T sono insufficienti per dirci quale valore logico (vero, falso) assegnare ad A (A è estranea, indipendente da T). Se A è indecidibile in T, possiamo considerare la teoria $T' = T + A$ i cui assiomi sono quelli di T più A. La teoria T' è consistente quanto T. In T', A è vera (è un assioma). Però possiamo anche considerare la teoria $T'' = T + \text{non } A$, T'' è consistente quanto T ma in T'' A è falsa (nonA è vera perché è un assioma).

Nella teoria degli insiemi esistono relazioni indecidibili ("l'ipotesi del continuo", per esempio).

Nella pratica non incontreremo relazioni indecidibili e, se mai questo succedesse, ne faremmo subito un assioma.

5: Valore logico: Ad ogni proposizione assegniamo un valore logico: V se A è vera, F se A è falsa.

SIMBOLI E RAGIONAMENTI LOGICI ELEMENTARI.

6: Il simbolo \vee ("o").

Se A e B sono due relazioni, $A \vee B$ (che si legge, e si scrive anche, A o B) è una proposizione.

La proposizione $A \vee B$ è vera se almeno una delle due relazioni A, B è vera. Questo si può esprimere con la seguente "tavola di verità":

A	B	$A \vee B$
V	V	V
V	F	V
F	V	V
F	F	F

7: Il simbolo \wedge ("e").

La proposizione: $\text{non}[(\text{non } A) \vee (\text{non } B)]$ si abbrevia in: $A \wedge B$; si legge (e si scrive) A e B.

La proposizione $A \wedge B$ è vera se entrambe le relazioni A, B sono vere.

A	B	$A \wedge B$
V	V	V
V	F	F
F	V	F
F	F	F

8: L'implicazione logica: \Rightarrow .

La proposizione: nonA o B si abbrevia in $A \Rightarrow B$ (che si legge A implica B).

A	B	non A	$A \Rightarrow B$
V	V	F	V
V	F	F	F
F	V	V	V
F	F	V	V

Notare che $A \Rightarrow B$ è vera tranne se A è vera e B è falsa. Questo rispecchia il senso comune: "il vero non implica il falso".

9: La dimostrazione diretta.

Da quanto precede se A è vera e se $A \Rightarrow B$ è vera allora B è vera. Il procedimento di dimostrazione diretta usa questa particolarità dell'implicazione: si parte da una proposizione vera A ("ipotesi") e si vuol dimostrare B ("la tesi"). Si cercano (eventualmente) relazioni intermedie B_1, \dots, B_n e si dimostrano le implicazioni:

$$\begin{array}{ll}
 A \Rightarrow B_1 & (\text{quindi } B_1 \text{ è vera}) \\
 B_1 \Rightarrow B_2 & (\text{quindi } B_2 \text{ è vera}) \\
 \dots & \dots \\
 B_{n-1} \Rightarrow B_n & (\text{quindi } B_n \text{ è vera}) \\
 B_n \Rightarrow B & \text{e B è dimostrata.}
 \end{array}$$

10: L'equivalenza logica: \Leftrightarrow .

La proposizione $[(A \Rightarrow B) \wedge (B \Rightarrow A)]$ si abbrevia in $A \Leftrightarrow B$ (che si legge A è equivalente a B).

Osserviamo che $A \Leftrightarrow B$ è vera se: A e B sono entrambe vere, o A e B sono entrambe false. Quindi $A \Leftrightarrow B$ vera significa che A e B hanno lo stesso valore logico.

Si ha: $A \Leftrightarrow \text{non}(\text{non}A)$.

11: La dimostrazione per contrapposizione:

Questo procedimento usa il seguente teorema: se A e B sono due relazioni, la proposizione: $[(A \Rightarrow B) \Leftrightarrow (\text{non}B \Rightarrow \text{non}A)]$ è vera.

Si può "dimostrare" questo teorema con le "tavole di verità":

A	B	$A \Rightarrow B$	$\neg A$	$\neg B$	$\neg B \Rightarrow \neg A$	$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$
V	V	V	F	F	V	V
V	F	F	F	V	F	V
F	V	V	V	F	V	V
F	F	V	V	V	V	V

Quindi dimostrare l'implicazione $A \Rightarrow B$ è equivalente a dimostrare $\neg B \Rightarrow \neg A$.

In pratica questo è molto utile perché può succedere che la dimostrazione dell'implicazione "contrapposta" $\neg B \Rightarrow \neg A$ risulti più facile di quella dell'implicazione diretta $A \Rightarrow B$.

11.1: Esempio : Sia n un numero naturale e A la proposizione " n^2 è pari", B la proposizione " n è pari".

Vogliamo dimostrare $A \Rightarrow B$. Da quanto detto questo è equivalente a dimostrare $\neg B \Rightarrow \neg A$. In questo caso l'ipotesi è $\neg B$: " n è dispari" e la tesi è $\neg A$: " n^2 è dispari". Infatti se $n = 2k+1$ allora $n^2 = 4k^2+4k+1$ che è dispari. L'implicazione $A \Rightarrow B$ è dimostrata.

12: Il ragionamento per assurdo.

12.1: Teorie inconsistenti : Se una teoria è inconsistente, ogni proposizione della teoria è vera (e anche falsa). Infatti supponiamo A vera e $\neg A$ vera. Sia B una proposizione. La proposizione $\neg A \Rightarrow (B \lor \neg A)$ è vera (fare la tavola di verità o considerare che $\neg A$ è falsa). Siccome $\neg A$ è vera ("per ipotesi"), anche $(B \lor \neg A)$ è vera, ma $(B \lor \neg A)$ è l'abbreviazione di $A \Rightarrow B$. Quindi $A \Rightarrow B$ è vera. Siccome A è vera, anche B è vera. Nello stesso modo si dimostra che $\neg B$ è vera...

12.2: Il ragionamento per assurdo : L'osservazione precedente è alla base del ragionamento per assurdo. Si vuole dimostrare la proposizione R . Si aggiunge momentaneamente ($\neg R$) agli assiomi della matematica ottenendo così una "nuova matematica". Si stabilisce poi che questa "nuova matematica" è contraddittoria. Da 12.1 segue che ogni proposizione di questa "nuova matematica" è vera. In particolare R è vera nella "nuova matematica". Quindi possiamo dire, senza entrare troppo nei dettagli, che R

è conseguenza logica della matematica e di $(\text{non}R)$, ossia l'implicazione: $(\text{non}R) \Rightarrow R$ è vera nella matematica.

Se $\text{non}R$ è falsa allora R è vera (2), e siamo a posto. Se $\text{non}R$ è vera, siccome l'implicazione $(\text{non}R) \Rightarrow R$ è vera, anche R è vera (8). Quindi R è dimostrata.

Nella pratica il ragionamento per assurdo si svolge così: "supponiamo R falsa" (stiamo aggiungendo $\text{non}R$ agli assiomi della matematica); si ragiona con questa ipotesi fino ad individuare una contraddizione. Si conclude dicendo: "ma questo è assurdo quindi R è vera".

12.3: Esempio : Vogliamo dimostrare R : " $\sqrt{2}$ è irrazionale". Supponiamo $\sqrt{2}$ razionale. Quindi $\sqrt{2} = m/n$. Possiamo supporre n e m primi tra di loro (i.e. il più grande comune divisore di m e n è 1). Infatti se $m = km'$, $n = kn'$ allora $m/n = m'/n'$. Elevando al quadrato: $2n^2 = m^2$. Quindi m^2 è pari. Sappiamo (cfr 11.1) che questo implica m pari. Quindi $m = 2t$ e $m^2 = 4t^2$. Ne segue che anche n^2 è pari, il ché, sempre per 11.1, implica che n è pari. In conclusione abbiamo dimostrato che 2 divide m e n ma questo è assurdo perché avevamo supposto $\sqrt{2} = m/n$ con m e n primi tra di loro. Quindi $\sqrt{2}$ è irrazionale.

Per un altro esempio classico di dimostrazione per assurdo, vedere Es. 5.

13: Induzione:

13.1: Gli assiomi di Peano : Nel tentativo⁽¹⁾ di assiomatizzare la costruzione dei numeri naturali il matematico G.Peano ha dato la seguente lista di assiomi. Innanzitutto ci sono tre termini primitivi: numero, zero, successore. Gli assiomi sono:

- (1) ogni numero ha un successore
- (2) zero non è il successore di nessun numero
- (3) due numeri distinti non possono avere lo stesso successore
- (4) sia F un insieme di numeri tale che: zero appartiene a F e se un numero appartiene a F anche il suo successore appartiene a F . Allora ogni numero appartiene a F .

Se indichiamo con $x+1$ il successore di x , l'assioma 4 (detto di induzione) si può riformulare così:

$$[(0 \in F) \wedge (x \in F \Rightarrow x+1 \in F)] \Rightarrow (F = \mathbb{N}).$$

13.2: Dimostrazione per induzione : La dimostrazione per induzione, uno dei metodi più usati in matematica, si basa sull'assioma 4. Supponiamo di voler dimostrare una proposizione $P(n)$ che dipenda dal numero naturale n . Allora:

⁽¹⁾ Questi assiomi non sono sufficienti per costruire \mathbb{N} in quanto usano, senza definirlo, il termine insieme.

$[P(0) \wedge P(n) \Rightarrow P(n+1)] \Rightarrow P(n)$ per ogni n . In altre parole per dimostrare $P(n)$ basta verificare $P(0)$ e dimostrare l'implicazione $P(n) \Rightarrow P(n+1)$ (se $P(n)$ è vera allora anche $P(n+1)$ è vera). Infatti basta considerare l'insieme $F = \{n \in \mathbb{N} / P(n) \text{ è vera}\}$ e applicare l'assioma 4.

13.3: Esempio : Supponiamo di volere dimostrare che: $1+2+3+\dots+n = n(n+1)/2$ per ogni naturale n . Chiamiamo $P(n)$ questa proposizione. Chiaramente $P(0)$ è vera. Supponiamo $P(n)$ vera (ipotesi d'induzione) e facciamo vedere che questo implica che anche $P(n+1)$ è vera. Abbiamo: $1+2+\dots+(n+1) = (1+2+\dots+n) + (n+1)$. Per ipotesi d'induzione il primo termine vale $n(n+1)/2$ quindi $1+2+\dots+(n+1) = n(n+1)/2 + (n+1)$ che è uguale a $(n+1)(n+2)/2$, quindi $P(n+1)$ è vera e la proposizione è dimostrata per ogni numero naturale.

Grossolanamente possiamo rappresentare l'insieme dei naturali come una scala a pioli infinita. Per essere sicuri di percorrerla tutta, passando per ogni piolo, basta sapere fare due cose: 1) fare il primo piolo, 2) quando si è ad un dato piolo (il piolo n^0), passare a quello successivo (il piolo $n^0 + 1$).

Il principio di induzione permette di dimostrare (cfr Es. 6) che \mathbb{N} è un insieme ben ordinato, ossia ogni sottinsieme non vuoto di \mathbb{N} ha un più piccolo elemento: sia $Z \subseteq \mathbb{N}$, $Z \neq \emptyset$, allora esiste $z_0 \in Z$ tale che: $x \in Z \wedge x \leq z_0 \Rightarrow x = z_0$.

Questa proprietà ("evidente") è fondamentale (per esempio rende possibile la divisione euclidea, il teorema di Bezout, il lemma di Gauss e la decomposizione in fattori primi, Es.8, 9, 10). Si osserverà che \mathbb{Z} , \mathbb{Q} , \mathbb{R} non godono di questa proprietà (vedere anche Es.7).

Usando questa proprietà possiamo formulare il secondo principio di induzione:

13.4: Teorema: Sia $a \in \mathbb{N}$ e $A = \{n \in \mathbb{N} / n \geq a\}$. Sia inoltre $P(n)$ una proprietà di n per n che varia in A . Supponiamo $P(a)$ vera, e: $P(m)$ vera per $a \leq m < n$, implica $P(n)$ vera. In queste condizioni $P(n)$ è vera per ogni n in A .

Dim: Siano $X = \{x \in A / P(x)$ è vera $\}$ e $Z = A \setminus X$. Supponiamo per assurdo $Z \neq \emptyset$. Allora Z ha un più piccolo elemento, z_0 . Esistono degli x tali che: $x \in A$ e $x < z_0$ (per esempio $x = a$). Se x è un tale elemento, per minimalità di z_0 , $x \notin Z$. Quindi $x \in X$ e $P(x)$ è vera. Pertanto, per ogni $a \leq x < z_0$, $P(x)$ è vera. Segue dall'ipotesi che $P(z_0)$ è vera. Dunque $z_0 \in X$, assurdo ♦

Ogni volta che vi trovate a dover dimostrare una proposizione del tipo $P(n)$, $n \in \mathbb{N}$, provate per induzione (funziona molto spesso!).

QUANTIFICATORI:

14: Il quantificatore esistenziale: \exists .

Sia A una proposizione che fa intervenire il termine x ("A dipende dalla variabile x") allora scriveremo A(x) per mettere in evidenza il termine x.

La proposizione " $\exists x, A(x)$ " si legge "esiste almeno un x tale che A(x)".

Il simbolo \exists si chiama quantificatore esistenziale. Si scrive anche $\exists x / A(x)$.

14.1: Esempio : Sia A(n) la proposizione "n è un numero naturale divisibile per 3". Per far vedere che la proposizione $\exists n, A(n)$ è vera basta trovare un numero naturale divisibile per 3; per es. 3 (anche se sappiamo che l'insieme dei multipli di tre è infinito).

14.2: Osservazione : Certi autori usano il simbolo $\exists!$ per significare "esiste uno e uno solo". Per esempio se B(n) è la proposizione "n è un numero primo e $4 \leq n \leq 6$ " allora $\exists! / B(n)$ è vera. Infatti 5 è primo ed è l'unico numero primo compreso tra 4 e 6.

15: Il quantificatore universale, \forall :

La proposizione "non[$\exists x, \text{non } A(x)$] " si abbrevia in: " $\forall x, A(x)$ " che si legge: "per ogni x, A(x)". Il simbolo \forall si chiama quantificatore universale.

16: Negazione di " $\forall x, A(x)$ ":

Per definizione $\forall x, A(x)$ è l'abbreviazione di non[$\exists x, \text{non } A(x)$]; stiamo quindi cercando di esprimere non(non[$\exists x, \text{non } A(x)$]). D'altra parte, se B è una proposizione: non(nonB) \Leftrightarrow B (cfr 10). Sostituendo B con ($\exists x, \text{non } A(x)$), otteniamo: $\exists x, \text{non } A(x) \Leftrightarrow \text{non}(\forall x, A(x))$

Per esempio la negazione logica della frase: "Tutte le macchine sono rosse" è "esiste una macchina non rossa".

17: Negazione di " $\exists x, A(x)$ ":

Usando 10 abbiamo: $[\exists x, A(x)] \Leftrightarrow [\exists x, \text{non}(\text{non}(A(x)))]$ e la negazione di $[\exists x, A(x)]$ è: non[$\exists x, \text{non}(\text{non}(A(x)))$]; per definizione del quantificatore universale quest'ultima proposizione è: $\forall x, \text{non}(A(x))$.

18: Abusi di linguaggio:

Nella pratica le dimostrazioni vengono scritte usando il linguaggio usuale e molte abbreviazioni (simboli matematici). L'impiego dei quantificatori, in particolare, può diventare assai delicato. Per esempio si usa dare la seguente definizione della continuità nel punto x_0 della funzione, f, di una variabile reale:

(i) $\forall \varepsilon > 0, \exists \eta > 0, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon$.

La discontinuità di f in x_0 si esprime negando (i). La negazione di (i) è:

(ii) $\exists \varepsilon > 0, \forall \eta > 0, \exists x, |x-x_0| \leq \eta \wedge |f(x)-f(x_0)| > \varepsilon$.

Infatti, in (i) si è sottinteso un quantificatore universale; sarebbe stato più rigoroso scrivere:

(i') $\forall \varepsilon > 0, \exists \eta > 0, \forall x, |x-x_0| \leq \eta \Rightarrow |f(x)-f(x_0)| \leq \varepsilon$.

L'omissione del quantificatore universale in (i) è giustificata dal fatto che x è inteso come variabile che può prendere ogni valore. Quindi questo abuso di linguaggio è accettato.

Esercizi:

- 1.1)** Dimostrare: (i) $\text{non}(A \wedge B) \Leftrightarrow (\text{non}A \vee \text{non}B)$
(ii) $\text{non}(A \vee B) \Leftrightarrow (\text{non}A \wedge \text{non}B)$.

1.2) Abbiamo dimostrato: $1+2+\dots+n = f(n)$ con $f(n) = n(n+1)/2$. In modo analogo dimostrare: $0^2+1^2+2^2+\dots+n^2 = q(n)$, per ogni naturale $n \geq 0$, dove $q(n)$ è un polinomio di grado tre (prima si espliciterà $q(n)$ e poi si dimostrerà la proposizione).

1.3) Sia n un numero naturale. Dimostrare che se il resto della divisione di n per 4 è uguale a 2 o 3 allora n non è un quadrato.

1.4) (Dimostrazione per assurdo). Da una favola orientale:

Per scegliere un ministro tra tre candidati, un re procede come segue: sulla schiena di ogni candidato appende un pezzo di stoffa. Nessun candidato può vedere il proprio pezzo di stoffa, ma vede quello degli altri. I candidati sanno che i pezzi di stoffa sono stati scelti tra tre di colore nero e due di colore bianco. Il primo a dire di che colore è il suo pezzo di stoffa sarà ministro ma in caso di risposta sbagliata, verrà ucciso. Uno dei candidati vedendo che i suoi due concorrenti hanno un pezzo di stoffa nera e vedendo che non dicono niente, dopo un po' afferma con sicurezza: "ho un pezzo di stoffa nera" e così diventò ministro (bei tempi!).

- (i) Esplicitate il suo ragionamento
(ii) Per non favorire nessuno come deve il re disporre i pezzi di stoffa?

1.5) Un numero naturale p si dice primo se: $p \geq 2$ e se gli unici numeri naturali che lo dividono sono 1 e p . Per esempio 2, 3, 5, 7, sono tutti i numeri primi minori di 10.

- (i) Fare la lista dei numeri primi ≤ 20 .
(ii) Dimostrare che se $n, n > 1$, non è primo allora esiste p primo che divide n .

(ii) Dimostrare il seguente:

Teorema: L'insieme dei numeri primi è infinito.

(Suggerimento: ragionare per assurdo, se $P = \{p_1, p_2, \dots, p_k\}$, $p_1 = 2, \dots$, è l'insieme di tutti i numeri primi, considerare $N = p_1 \cdot p_2 \cdots \cdot p_k + 1$, e guardare (usando (ii)) se N è primo o meno).

Il teorema precedente e la sua dimostrazione si trovano negli "Elementi" di Euclide. È uno dei teoremi più forti della matematica perché non esiste nessun "metodo" semplice per fabbricare un'infinità di numeri primi.

1.6) Si tratta di dimostrare che ogni sottinsieme non vuoto, X , di \mathbb{N} ha un elemento minimo. Si ricorda che se x, y sono due naturali allora $x \leq y \Leftrightarrow \exists k \in \mathbb{N}$ tale che $x+k=y$. Inoltre $x < y \Leftrightarrow x \leq y$ e $x \neq y$.

Supponiamo per assurdo che X non abbia un elemento minimo. Sia $V = \{n / \forall x \in X, n \leq x\}$.

(i) Dimostrare che $0 \in V$.

(ii) Dimostrare che se $n \in V$ allora $n+1 \in V$ (osservare che se $x \in V$ allora $x \notin X$ perché, per ipotesi, X non ha un elemento minimo).

(iii) Per il principio di induzione, da (ii), segue che $V = \mathbb{N}$. Sia $x_0 \in X$ (X è non vuoto). Osservare che $x_0+1 \in \mathbb{N} = V$. Dedurne la contraddizione $x_0+1 \leq x_0$.

1.7) Sia $I = \{1/n : n \in \mathbb{N}, n > 0\}$. Mostrare che I non ammette un elemento minimo.

1.8) (Divisione euclidea) Si tratta di dimostrare che dati due naturali a, n , $n \neq 0$, esiste un'unica coppia di naturali (q, r) tale che: $a = nq + r$, $e: 0 \leq r < n$; q è il quoziente e r è il resto nella divisione euclidea di a per n .

(i) Sia $A = \{x \in \mathbb{N} / \exists t \in \mathbb{N}$ tale che $a = nt + x\}$. Mostrare che A è non vuoto. Da Es.6, A ha un elemento minimo r .

(ii) Mostrare $r < n$ (ragionare per assurdo)

(iii) Siccome $r \in A$ abbiamo $a = nq + r$ per qualche q . Sia $a = ns + t$ con $0 \leq t < n$. Osservare che $r \leq t$. Dedurne $q \geq s$. Si ha quindi $q = s + m$, $m \in \mathbb{N}$. Concludere che $m = 0$.

(iv) Dedurre da quanto precede che se $n \in \mathbb{N}$ e $a \in \mathbb{Z}$ allora esistono q e r in \mathbb{Z} tali che $a = nq + r$, $0 \leq r < n$.

1.9) (Il teorema di Bezout) Il maggior comun divisore di due numeri naturali a, b si nota (a, b) . Se $d = (a, b)$ allora d divide a (da) e $d \mid b$, inoltre se $v \mid a$ e $v \mid b$ allora $v \mid d$ (in particolare $v \leq d$). Si dice che a e b sono primi tra di loro se $(a, b) = 1$ (i.e. a e b non hanno divisorì comuni non banali). Scopo dell'esercizio è di dimostrare:

Teorema: Se $d = (a, b)$ allora esistono due interi (positivi o negativi) u, v tali che $ua + vb = d$.

Per esempio $(3, 5) = 1$ e $2 \cdot 5 - 3 \cdot 3 = 1$.

(i) Sia $A = \{xa+yb / x \in \mathbb{Z}, y \in \mathbb{Z}\}$. Si consideri $A_+ = \{z \in A / z > 0\}$. Mostrare che A_+ è non vuoto. Sia d il più piccolo elemento di A_+ (**Es.6**). Osservare che $d = ua + vb$ per opportuni u, v .

(ii) Ogni elemento di A è multiplo di d (usando **Es.8** scrivere $t = qd + r$, $0 \leq r < d$, adesso se $t = ae + bf$, considerare $t - qd$)

(iii) Dedurre da (ii) che $d = (a, b)$.

1.10) (Il lemma di Gauss). Si tratta del seguente enunciato: "Siano a, b due naturali e p un numero primo. Se $p \nmid ab$ allora $p \nmid a$ o $p \nmid b$ ".

(i) Osservare che $(p, a) = 1$ o p .

(ii) Se $(p, a) = p$, il lemma è dimostrato. Nell'altro caso usare il teorema di Bezout (**Es.9**).

(iii) Usando **Es.5** (ii) e il lemma di Gauss dimostrare che ogni naturale $n \geq 2$ si scrive in modo unico come un prodotto di numeri primi (i.e. $n = p_1^{a_1} \dots p_r^{a_r}$ con p_i numeri primi distinti, $a_i \geq 1$ naturali; inoltre se $n = q_1^{b_1} \dots q_s^{b_s}$ con q_j numeri primi distinti, $b_j \geq 1$ naturali, allora $r = s$ e, dopo eventuale riordino degli indici, $q_i = p_i$, $a_i = b_i$).

1.11) Sia $p \geq 2$ un numero primo. Dimostrare che \sqrt{p} è irrazionale (usare il lemma di Gauss).

2) INSIEMI.

In questo paragrafo introduciamo rapidamente (e informalmente) alcune nozioni di base della teoria degli insiemi che costituiscono i rudimenti del linguaggio matematico. Non è il caso di dare qui una presentazione assiomatica della teoria degli insiemi.

Pertanto ammetteremo le nozioni di insieme, elemento, appartenenza ad un insieme come nozioni primitive.

1: Appartenenza: La proposizione " $x \in E$ " si legge "x è elemento di E" o "x appartiene ad E". Intuitivamente E è la collezione degli oggetti x tali che $x \in E$ (questa non è una definizione!).

2: Inclusione: Siano E e F degli insiemi. Si scrive $E \subseteq F$ e si dice che E è contenuto (o incluso) in F se ogni elemento di E è elemento di F.

Questo si può anche esprimere nel modo seguente: $E \subseteq F \Leftrightarrow \forall x, x \in E \Rightarrow x \in F$.

Se $E \subseteq F$ si dice che E è un sottoinsieme di F o una parte di F.

3: Uguaglianza: Due insiemi A, B sono uguali (si scrive $A = B$) se e soltanto se: $A \subseteq B$ e $B \subseteq A$; ossia $A = B \Leftrightarrow (\forall x, x \in A \Leftrightarrow x \in B)$. Quindi due insiemi sono uguali se e soltanto se hanno gli stessi elementi...

4: L'insieme delle parti di un insieme: Sia E un insieme allora esiste un insieme i cui elementi sono precisamente i sottoinsiemi di E. Questo insieme si chiama l'insieme delle parti di E e si nota $\wp(E)$.

5: Unione: Siano E, F due insiemi allora esiste un insieme, chiamato unione di E e F e notato $E \cup F$ tale che: $\forall x, x \in E \cup F \Leftrightarrow (x \in E \vee x \in F)$.

L'esistenza di $E \cup F$ è un assioma.

Non si ha nessun modo effettivo per decidere se una data collezione di oggetti è un insieme. Quindi la teoria procede per costruzioni, partendo da termini di cui si sa (o si è ammesso) che sono insiemi. Acceniamo rapidamente ad alcuni di questi procedimenti.

6: Parte di un insieme definita da una proposizione: Sia E un insieme e $A(x)$ una proposizione. La proposizione $(x \in E \wedge A(x))$ si dice proposizione su E. Data una proposizione su E esiste un insieme F tale che si abbia l'equivalenza:

$(x \in F) \Leftrightarrow (x \in E \wedge A(x))$. L'esistenza di F è un assioma.

Notare che: " $x, x \in F \Rightarrow x \in E$, quindi $F \subseteq E$. Intuitivamente F è l'insieme degli elementi di E che verificano la proposizione A .

Nel seguito scriveremo $F = \{x / x \in E \wedge A(x)\}$ o anche $F = \{x \in E / A(x)\}$.

6.1: Esempio : L'insieme $F = \{0, 1, 2, 3\}$ è definito da $F = \{x \in \mathbb{N} / x \leq 3\}$.

7: Complementare di un sottoinsieme: Siano E, F due insiemi tali che $E \subseteq F$. L'insieme: $\{x / x \in F \text{ e } x \notin E\}$ si chiama il complementare di E in F e si scrive $F \setminus E$ (attenzione all'ordine!). Chiaramente: $(F \setminus E) \subseteq F$; gli elementi di $F \setminus E$ sono gli elementi di F che non appartengono a E .

8: Osservazione : Sia $A(x)$ una proposizione. In generale non esiste un insieme E tale che si abbia l'equivalenza: $(x \in E) \Leftrightarrow A(x)$. Questo spiega perché ci siamo ristretti a considerare relazioni su degli insiemi.

L'insieme vuoto:

9: Proposizione: (i) *esiste un insieme, indicato con \emptyset , tale che: $\forall x, x \notin \emptyset$.*

(ii) *per ogni insieme X : $\emptyset \subseteq X$.*

(iii) *l'insieme \emptyset , chiamato ovviamente insieme vuoto, è l'unico insieme soddisfacente (ii) (e anche (i)).*

Dim: (i) Sia E un insieme qualsiasi. Tra i sottoinsiemi di E c'è E stesso. Poniamo $E \setminus E := \emptyset$. Se $x \in \emptyset$ allora $x \in E \setminus E = \{y / y \in E \text{ e } y \notin E\}$. Sia $A(y)$ la proposizione " $y \in E$ "; la proposizione $\neg A(y)$ è " $y \notin E$ ". La proposizione $A(y) \wedge \neg A(y)$ è sempre falsa (cf §1.7). Quindi $x \notin \emptyset$.

(ii) Bisogna dimostrare l'implicazione: $\forall x, x \in \emptyset \Rightarrow x \in X$. Questa implicazione è vera perché per ogni x , $x \in \emptyset$ è falsa (cf §1.8).

(iii) Sia Y un insieme soddisfacente (ii); quindi per ogni insieme X : $Y \subseteq X$. In particolare $Y \subseteq \emptyset$. Ma da (ii): $\emptyset \subseteq Y$, quindi (cf 3) $Y = \emptyset$ ♦

10: Intersezione: Siano E, F due insiemi. L'insieme: $\{x / x \in E \text{ e } x \in F\}$ si nota $E \cap F$ e si chiama l'intersezione di E e F .

Il fatto che $E \cap F$ sia un insieme segue da 6: sull'insieme E consideriamo la proposizione $A(x)$ data da: " $x \in F$ ", allora: $E \cap F = \{x \in E / A(x)\}$. Dalla definizione un elemento appartiene a $E \cap F$ se e solo se appartiene a E e appartiene a F : $x \in E \cap F \Leftrightarrow (x \in E \wedge x \in F)$.

10.1: Proprietà dell'intersezione: Se E, F, G sono degli insiemi allora: (i) $E \cap F = F \cap E$.

- (ii) $(E \cap F) \cap G = E \cap (F \cap G)$; questo è l'associatività dell'intersezione.
 Risulta che si può scrivere $(E \cap F) \cap G = E \cap (F \cap G) =: E \cap F \cap G$.
- (iii) $E \subseteq F \Leftrightarrow E \cap F = E$ (cf Es. 1).
- (iv) $E \cap \emptyset = \emptyset$ (cf Es. 2).

11: Coppie: Finora abbiamo introdotto i simboli \in e $=$ che servono a costruire delle relazioni. Adesso introduciamo una operazione che serve a costruire oggetti matematici.

Questa operazione consiste nel costruire, partendo da due oggetti matematici x e y , presi in quest'ordine, un terzo oggetto matematico notato (x, y) e che si chiama coppia (x, y) .

L'operazione che consiste nel formare coppie è definita dall'unica legge:

$$(x, y) = (a, b) \Leftrightarrow [x = a \text{ e } y = b].$$

In particolare $(x, y) = (y, x)$ se e solo se $x = y$; quindi l'ordine nel quale si scrivono i due termini di una coppia è essenziale. L'elemento x si chiama la prima proiezione della coppia (x, y) , y è la seconda proiezione.

11.1: Osservazione : Si avrà cura di non confondere la coppia (x, y) e l'insieme $\{x, y\}$. Infatti, se $x \neq y$: $\{x, y\} = \{y, x\}$ mentre $(x, y) \neq (y, x)$; "una coppia ha un ordine mentre un insieme no".

11.2: n-uple: Si può estendere la nozione di coppia. Siano x, y, z tre oggetti. Si pone:

$(x, y, z) := ((x, y), z)$ e si dice che (x, y, z) è la 3-upla (o terna) x, y, z .
 Si ha $(x, y, z) = (x', y', z')$ se e solo se $x = x'$, $y = y'$, $z = z'$ (cf Es. 3).

Ovviamente questa nozione si può estendere ulteriormente. Per n in N^* si definisce, per induzione, la nozione di n -upla. Questa nozione è già definita per $n \leq 3$; supponiamola definita per $n-1$ e cerchiamo di definirla per n . Siano x_1, \dots, x_n n oggetti. Si pone:

$(x_1, \dots, x_n) := ((x_1, \dots, x_{n-1}), x_n)$ e si dice che (x_1, \dots, x_n) è la n -upla x_1, \dots, x_n (nell'ordine!).

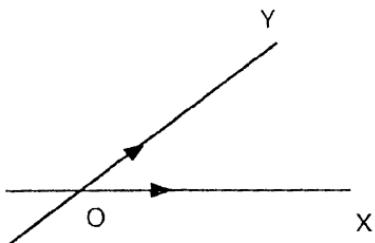
Si ha $(x_1, \dots, x_n) = (x'_1, \dots, x'_n)$ se e solo se $x_1 = x'_1, \dots, x_n = x'_n$ (cf Es. 3).

12: Prodotto cartesiano di due insiemi: Siano E, F due insiemi. Esiste un insieme, Z , tale che: $(z \in Z) \Leftrightarrow (\exists x \in E \wedge \exists y \in F / (x, y) = z)$. Si dice che Z è il prodotto cartesiano di E e F e si scrive $Z = E \times F$.

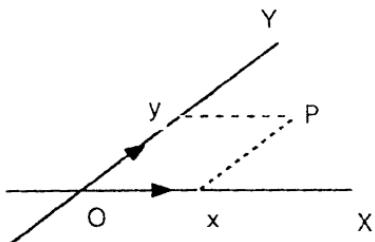
Se $E = F$ si scrive: $E \times E = E^2$.

12.1: Diagonale: Il sottoinsieme, D , di $E \times E$ definito da $D := \{(x, x) / x \in E\}$ si chiama diagonale di $E \times E$.

12.2: Osservazione : Il riferimento a Descartes nella definizione precedente si può giustificare così: nel "piano della geometria elementare" siano Ox e Oy due assi. Scegliamo delle unità di lunghezza su questi assi:



Fatto ciò ogni punto, P , del piano ha un ascisse e un'ordinata:



Si identifica così il punto P con la coppia (x, y) (le "coordinate cartesiane" di P). Se P' è un altro punto del piano di coordinate (x', y') allora: $P = P'$ se e solo se $x = x'$ e $y = y'$. Se assimiliamo l'asse Ox (risp. Oy) alla retta reale R , vediamo che la scelta di un sistema di coordinate permette di identificare il piano della geometria elementare a $R \times R = R^2$.

12.3: Osservazione : Si ha: $A \times B = \emptyset \Leftrightarrow (A = \emptyset \vee B = \emptyset)$.

12.4: Prodotto cartesiano di n insiemi: Seguendo il procedimento usato per le n -uple si definisce il prodotto cartesiano di n insiemi E_1, \dots, E_n . Si definisce $E_1 \times \dots \times E_n$ per induzione tramite: $E_1 \times \dots \times E_n = (E_1 \times \dots \times E_{n-1}) \times E_n$. Ne segue che z è un elemento di $E_1 \times \dots \times E_n$ se e solo se $z = ((x_1, \dots, x_{n-1}), x_n)$ con $((x_1, \dots, x_{n-1}), x_n)$ appartenente a $E_1 \times \dots \times E_{n-1}$ e x_n appartenente a E_n . In particolare x_i appartiene a E_i , $1 \leq i \leq n$.

Sia z' un altro elemento di $E_1 \times \dots \times E_n$ allora $z' = ((x'_1, \dots, x'_{n-1}), x'_n)$. Si ha $z = z'$ se e solo se $x_1 = x'_1, \dots, x_n = x'_n$.

12.5: Convenzione: Notare che a priori $(E_1 \times E_2) \times E_3$ e $E_1 \times (E_2 \times E_3)$ sono due insiemi diversi. Infatti un elemento, z , di $(E_1 \times E_2) \times E_3$ è della forma $((x_1, x_2), x_3)$ mentre uno, Z , di $E_1 \times (E_2 \times E_3)$ è della forma: $Z = (x_1, (x_2, x_3))$; e niente ci permette di affermare che: $((x_1, x_2), x_3) = (x_1, (x_2, x_3))$ perché (x_1, x_2) e x_1 sono due oggetti diversi.

Però se $z' = ((x'_1, x'_2), x'_3)$ (risp. $Z' = (x'_1, (x'_2, x'_3))$) allora $z = z'$ se e solo se $x_1 = x'_1, \dots, x_3 = x'_3$. Nello stesso modo $Z = Z'$ se e solo se $x_1 = x'_1, \dots, x_3 = x'_3$. Pertanto si vede che quello che importa è l'ordine e non le parentesi. Questo spiega che nella pratica, per abuso di linguaggio, si identificano $(E_1 \times E_2) \times E_3$ e $E_1 \times (E_2 \times E_3)$. Per riassumere:

$$E_1 \times (E_2 \times E_3) = (E_1 \times E_2) \times E_3 =: E_1 \times E_2 \times E_3 = \{(x_1, x_2, x_3) / x_i \in E_i, 1 \leq i \leq 3\}.$$

Questa convenzione si estende al caso di un prodotto con n fattori: $E_1 \times \dots \times E_n = \{(x_1, \dots, x_n) / x_i \in E_i, 1 \leq i \leq n\}$.

12.6: Osservazioni : Se $n \geq 1$ è un numero naturale e se E è un insieme si pone: $E \times \dots \times E = E^n$ (n fattori nel primo membro).

Si ha: $(E_1 \times \dots \times E_n = \emptyset) \Leftrightarrow (\exists i, 1 \leq i \leq n / E_i = \emptyset)$. In particolare: $\emptyset^n = \emptyset$.

Esercizi:

2.1) Siano E, F degli insiemi dimostrare: $E \subseteq F \Leftrightarrow E \cap F = E$.

2.2) Sia E un insieme. Dimostrare: $E \cap \emptyset = \emptyset$.

2.3) Dimostrare che due n -upli $(x_1, \dots, x_n), (y_1, \dots, y_n)$ sono uguali se e solo se $x_1 = y_1, \dots, x_n = y_n$.

2.4) Fare la lista degli elementi dell'insieme $\wp(\wp(\wp(\emptyset)))$ ($\wp(E)$ è l'insieme delle parti di E).

2.5) Siano A, B, C degli insiemi. Dimostrare che: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

2.6) Si pone $A = \{(x, y) \in \mathbf{R}^2 / x - 1\}, B = \{(x, y) \in \mathbf{R}^2 / 2x - y + 2 = 0\}, C = \{(x, y) \in \mathbf{R}^2 / x \leq y\}$. Rappresentare graficamente $A, B, C, A \cap B, A \cap B \cap C, A \cup B \cup C$. E' $A \cup B \cup C$ uguale a $A \cup C$?

3) APPLICAZIONI.

Intuitivamente un'applicazione (o funzione) da un insieme E in un insieme F è una legge che ad ogni elemento di E fa corrispondere uno (e uno solo) elemento di F.

Questa non è una definizione matematica rigorosa (non abbiamo definito che cosa sia una "legge" e cosa voglia dire "far corrispondere"), però, prese le dovute cautele, è "l'idea" giusta!

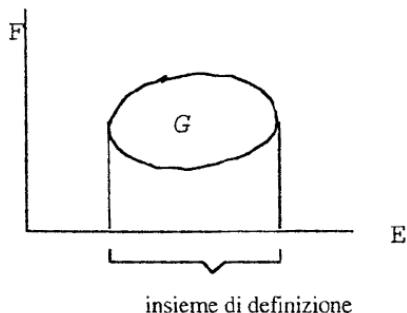
Tra le nozioni fondamentali di questo paragrafo segnaliamo quelle di applicazione iniettiva, suriettiva, biiettiva.

1: Grafici e corrispondenze:

Siano E, F due insiemi. Un grafico, G, di E in F è un sottinsieme dell'insieme prodotto $E \times F$.

Una corrispondenza di E in F è una terna (E, F, G) dove G è un grafico di E in F. Si dice che E è l'insieme di partenza e F l'insieme d'arrivo della corrispondenza.

2: Insieme di definizione di una corrispondenza: Sia $c = (E, F, G)$ una corrispondenza di E in F. L'insieme (proiezione di G su E): $\{x \in E / \exists y, (x, y) \in G\}$ si chiama insieme di definizione della corrispondenza c.



L'insieme di definizione è vuoto se e solo se G è l'insieme vuoto.

3: Applicazioni: Siano E e F due insiemi. Un'applicazione di E in F è una corrispondenza $f = (E, F, G)$ di E in F tale che:

(i) l'insieme di definizione di f è uguale a E.

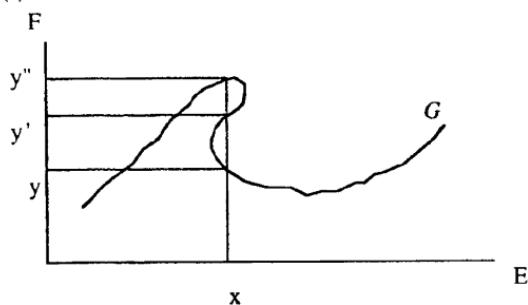
(ii) per ogni x appartenente a E l'insieme $\{y / y \in F \text{ e } (x, y) \in G\}$ ha un unico elemento.

NB: Si usa anche il termine funzione al posto di applicazione.

Sia $f = (E, F, G)$ un'applicazione di E in F . Sia x un elemento di E , l'unico elemento dell'insieme $\{y \mid y \in F \text{ e } (x, y) \in G\}$ si nota $f(x)$ e si chiama l'immagine di x tramite f (si dice anche che $y = f(x)$ è il valore della funzione f in x).

3.1: Esempi :

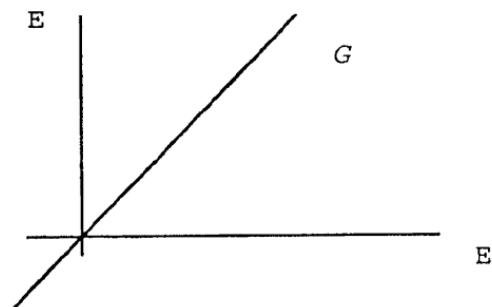
(i)



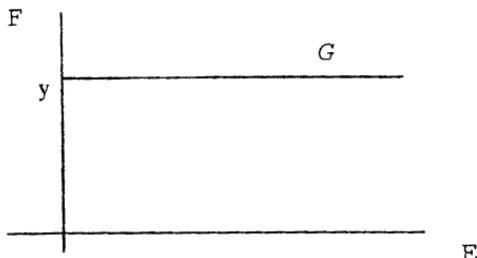
Questa corrispondenza non è un'applicazione perché $\{y \mid y \in F \text{ e } (x, y) \in G\}$ ha più di un elemento.

(ii) Sia E un insieme e D la diagonale di $E \times E$; si ricorda (§2, 11.1) che $D := \{(x, y) \in E^2 \mid x = y\}$. La terna (E, E, D) è un'applicazione da E in E chiamata identità di E e notata Id_E (o anche 1_E). Per ogni x in E : $Id_E(x) = x$.

Grafico di Id_E :



(iii) Siano E, F due insiemi e y' un elemento di F . Sia G il sottinsieme di $E \times F$ definito da: $G := \{(x, y') \mid x \in E\}$. La terna (E, F, G) è un'applicazione da E in F chiamata applicazione costante di valore y' e notata (per esempio) $C_{y'}$; per ogni x in E : $C_{y'}(x) = y'$.



3.2: Notazioni: Sia $f = (E, F, G)$ un'applicazione da E in F . La frase "f è un'applicazione da E in F " si abbrevia in "sia $f : E \rightarrow F$ un'applicazione" o anche, quando il contesto sia chiaro, "sia $f : E \rightarrow F$ ".

Inoltre si dice che f è definita su E e prende i suoi valori in F (o: f è a valori in F).

4: Insieme immagine di un'applicazione: Sia $f = (E, F, G)$ un'applicazione da E in F. L'insieme (proiezione su F del grafico G): $\{y / y \in F \text{ e } \exists x \in E, (x, y) \in G\} = \{y \in F / \exists x \in E \text{ e } y = f(x)\}$ si nota $f(E)$ e si chiama l'insieme immagine di f (o anche l'immagine di f). Si scrive $\text{Im}(f)$ o anche $\text{Im}(f)$ ("image" in inglese) al posto di $f(E)$. Notare che $\text{Im}(f) = \{f(x) / x \in E\}$.

5: Composizione di applicazioni: Siano $f : X \rightarrow Y$ e $g : Y \rightarrow Z$ due applicazioni. L'applicazione composta $g \circ f : X \rightarrow Z$ è definita da $(g \circ f)(x) = g(f(x))$.

Per comporre due applicazioni bisogna che l'insieme d'arrivo della prima sia uguale all'insieme di partenza della seconda.

5.1: Esempio : Siano $f : R \rightarrow R : x \rightarrow 2x$ e $g : R \rightarrow R : x \rightarrow x^2$. Abbiamo $g \circ f : R \rightarrow R : x \rightarrow 4x^2$ mentre $f \circ g : R \rightarrow R : x \rightarrow 2x^2$. Quindi $f \circ g \neq g \circ f$.

5.2: Lemma: Siano $f : X \rightarrow Y$, $g : Y \rightarrow Z$ e $h : Z \rightarrow T$ tre applicazioni allora: $h \circ (g \circ f) = (h \circ g) \circ f$.

Dim: Per ogni x appartenente a X:

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) \text{ e } ((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x))) \blacklozenge$$

6: Definizione: (Applicazioni suriettive) Un'applicazione $f : E \rightarrow F$ si dice suriettiva (o f è una suriezione) se $\text{Im}(f) = F$.

In altre parole f è suriettiva se per ogni y in F esiste almeno un x in E tale che $y = f(x)$.

In simboli: f suriettiva $\Leftrightarrow \forall y \in F, \exists x \in E / f(x) = y$.

6.1: Esempi : (i) Sia $f: \mathbb{R} \rightarrow \mathbb{R}: x \rightarrow 2x+1$. Allora f è suriettiva. Infatti per ogni y in \mathbb{R} si ha $f((y-1)/2) = y$.

(ii) Sia $R_+ = \{x \in \mathbb{R} / x \geq 0\}$ e $f: \mathbb{R} \rightarrow R_+: x \rightarrow x^2$, allora f è suriettiva.

Invece $g: \mathbb{R} \rightarrow \mathbb{R}: x \rightarrow x^2$ non è suriettiva (determinare $\text{Im}(g)$).

7: Definizione: (Applicazioni iniettive) Sia $f: E \rightarrow F$ un'applicazione. Si dice che f è iniettiva (o f è un'iniezione) se due elementi diversi di E hanno due immagini diverse. In simboli:

f è iniettiva $\Leftrightarrow \forall x \in E, \forall x' \in E: x \neq x' \Rightarrow f(x) \neq f(x')$.

Questo è equivalente a: $\forall x \in E, \forall x' \in E: f(x) = f(x') \Rightarrow x = x'$.

7.1: Esempi : Sia $f: \mathbb{R}_+ \rightarrow \mathbb{R}: x \rightarrow x^2$, allora f è iniettiva. Invece $g: \mathbb{R} \rightarrow \mathbb{R}: x \rightarrow x^2$, non è iniettiva.

8: Definizione: (Applicazioni biettive) Sia $f: E \rightarrow F$ un'applicazione. Si dice che f è biettiva se f è iniettiva e se f è suriettiva. Quindi f è biettiva se, per ogni y in F , esiste uno ed un unico elemento x di E tale che $f(x) = y$.

9: Insiemi finiti, infiniti: Intuitivamente un insieme si dice finito se ha un numero finito di elementi. Un insieme non finito è infinito. Più precisamente:

9.1: Definizione: Un insieme E è infinito se esiste un sottinsieme $F \subseteq E$, $F \neq E$, e una biiezione $f: F \rightarrow E$. Un insieme non infinito si dice finito.

Per esempio sia $P \subseteq \mathbb{N}$, $P = \{0, 2, 4, 6, \dots\}$ l'insieme dei numeri pari. E' chiaro che $P \neq \mathbb{N}$. Consideriamo l'applicazione $f: P \rightarrow \mathbb{N}: 2n \rightarrow n$. L'applicazione f è chiaramente iniettiva e suriettiva, quindi \mathbb{N} è un insieme infinito. Questo esempio è dovuto a Galileo Galilei.

Una delle regole della logica nell'antica Grecia era: "Il tutto è più grande della parte". Questa regola non vale per gli insiemi infiniti: ad ogni numero naturale possiamo associare uno, ed un'unico, numero pari. Ci sono "altrettanti" numeri pari che numeri naturali!

CARATTERIZZAZIONI DELLE APPLICAZIONI *-IETTIVE.

Diamo adesso delle condizioni necessarie e sufficienti affinché un'applicazione sia *-iettiva (* = in, sur, bi).

10: Proposizione: Sia $f: E \rightarrow F$ un'applicazione. Sono equivalenti:

(i) f è iniettiva

(ii) esiste un'applicazione $h : F \rightarrow E$ tale che $h \circ f = \text{Id}_E$.

Dim: (i) \Rightarrow (ii) Sia x_0 un elemento qualsiasi di E . Se $y \in F \setminus f(E)$ poniamo $h(y) = x_0$. Se $y \in f(E)$ allora $y = f(x)$ e poniamo $h(y) = x$. Questo definisce bene un'applicazione perché se $y = f(x')$ allora $x = x'$ visto che f è iniettiva. Verifichiamo $h \circ f = \text{Id}_E$: per ogni x in E abbiamo: $(h \circ f)(x) = h(f(x)) = h(x) = x$.

(ii) \Rightarrow (i) Supponiamo $f(x) = f(x')$, allora $h(f(x)) = h(f(x'))$ ossia $(h \circ f)(x) = (h \circ f)(x')$. Siccome $h \circ f = \text{Id}_E$ per ipotesi, viene $x = x'$ e quindi f è iniettiva ♦

10.1: Osservazione : In generale l'applicazione h della proposizione 10 non è unica.

11: Proposizione: Sia $f : E \rightarrow F$ un'applicazione. Sono equivalenti:

(i) f è suriettiva

(ii) esiste un'applicazione $h : F \rightarrow E$ tale che $f \circ h = \text{Id}_F$.

La dimostrazione di questa proposizione necessita l'uso dell'assioma della scelta:

Assioma della scelta: Siano X, Y due insiemi e $s : X \rightarrow \wp(Y)$ un'applicazione tale che: $\forall x \in X, s(x) \neq \emptyset$. Allora esiste un'applicazione $r : X \rightarrow Y$ tale che $r(x) \in s(x)$, per ogni x in X .

Il significato di questo assioma è il seguente: l'applicazione r "sceglie" un elemento in ognuno dei sottinsiemi $s(x)$, al variare di x in X . Questa "scelta" è evidente quando X è un insieme finito ma diventa più problematica quando X è un insieme infinito.

Dimostrazione di 11: (i) \Rightarrow (ii) Intuitivamente quello che si deve fare è chiaro (fare un disegno): sia y in F . Siccome f è suriettiva l'insieme $\{x \in E / f(x) = y\}$ è non vuoto. Scegliamo un x in questo insieme (che può anche essere infinito) e poniamo $h(y) = x$. Abbiamo allora $(f \circ h)(y) = f(h(y)) = f(x) = y$. Ecco una presentazione più rigorosa:

per ogni y in F sia $E_y := \{x \in E / f(x) = y\}$; siccome f è suriettiva, per ogni y in F si ha $E_y \neq \emptyset$. Consideriamo l'applicazione: $s : F \rightarrow \wp(E) : y \rightarrow E_y$. Come già detto $s(y) \neq \emptyset$ per ogni y in F . Dall'assioma della scelta segue che esiste un'applicazione $h : F \rightarrow E$ tale che $h(y) \in E_y$ per ogni y in F . Adesso $(f \circ h)(y) = f(h(y)) = y$ perché $h(y) \in E_y$.

(ii) \Rightarrow (i) Sia y in F , per ipotesi $y = (f \circ h)(y)$ quindi $y = f(h(y))$ e pertanto f è suriettiva ♦

11.1: Osservazione : In generale l'applicazione h della proposizione 11 non è unica.

12: Teorema: Sia $f : E \rightarrow F$ un'applicazione. Sono equivalenti:

(i) f è biiettiva

(ii) esistono delle applicazioni h, g tali che: $h : F \rightarrow E$ e $f \circ h = \text{Id}_F$; $g : F \rightarrow E$ e $g \circ f = \text{Id}_E$.

Inoltre se le condizioni (i) e (ii) sono verificate allora $g = h$ e g è l'unica applicazione da F in E tale che $g \circ f = \text{Id}_E$, e $f \circ g = \text{Id}_F$. L'applicazione g si dice applicazione reciproca di f e si nota f^{-1} .

Dim: (i) \Leftrightarrow (ii) Questo segue da 10 e 11 perché f è biiettiva se e solo se è iniettiva e suriettiva.

Mostriamo l'unicità: sia (r,s) una coppia soddisfacente $r : F \rightarrow E$ e $f \circ r = \text{Id}_F$; $s : F \rightarrow E$ e $s \circ f = \text{Id}_E$. Per (5.2) abbiamo $s \circ (f \circ r) = (s \circ f) \circ r$. Ma: $s \circ (f \circ r) = s \circ \text{Id}_F = s$ mentre: $(s \circ f) \circ r = \text{Id}_E \circ r = r$; quindi $s = r$. Siccome la coppia (s,r) era qualsiasi, questo dimostra l'unicità dell'applicazione reciproca♦

12.1: Osservazione : In altri termini $f : E \rightarrow F$ è biiettiva se e solo se esiste un'applicazione $f^{-1} : F \rightarrow E$ tale che: $f^{-1} \circ f = \text{Id}_E$; $f \circ f^{-1} = \text{Id}_F$. E' chiaro (perché?) che anche f^{-1} è biiettiva.

13: Contr'immagine: Sia $f : X \rightarrow Y$ un'applicazione e F un sottinsieme di Y . L'insieme: $\{x \in X / f(x) \in F\}$ si chiama la contr'immagine di F tramite f e si nota $f^{-1}(F)$. Questo non significa che f sia biiettiva.

In tutta generalità se G è il grafico di f consideriamo il sottinsieme T di G dato da: $T = \{(x,y) \in G / y \in F\}$. La controimmagine di F è la proiezione di T su X : $f^{-1}(F) = \{x \in X / \exists y \in F, (x,y) \in T\}$.

13.1: Esempio : Se a, b sono due numeri reali si nota $[a, b]$ l'insieme (detto anche intervallo chiuso): $\{x \in \mathbb{R} / x \geq a \wedge x \leq b\}$.

Sia $f : \mathbb{R} \rightarrow \mathbb{R} : x \rightarrow x^2$. Allora $f^{-1}([1,2]) = [-\sqrt{2}, -1] \cup [1, \sqrt{2}]$.

14: Famiglie: Si usa ogni tanto una terminologia diversa per parlare di applicazioni. Sia $f : I \rightarrow X$ un'applicazione e scriviamo $f(i) = x_i$ si dice allora che f è una famiglia di elementi di X con indici in I e si indica f nel modo seguente: $(x_i)_{i \in I}$. Se $I = \mathbb{N}$ (insieme dei numeri naturali) si dice che $(x_n)_{n \in \mathbb{N}}$ è una successione.

Sia $f : I \rightarrow \wp(E)$ un'applicazione e poniamo $f(i) = A_i$ (A_i è un sottinsieme di E); allora $(A_i)_{i \in I}$ è una famiglia di sottinsiemi di E con indici in I . Nella pratica non si precisa chi sia l'insieme

E si dice solo che $(A_i)_{i \in I}$ è una famiglia d'insiemi con indici in I . In realtà si può sempre prendere per E l'insieme riunione della famiglia notato: $\cup_{i \in I} A_i$ e definito da:

$$\forall x, x \in \cup_{i \in I} A_i \Leftrightarrow \exists i \in I / x \in A_i.$$

15: Proposizione: Sia $(A_i)_{i \in I}$ una famiglia d'insiemi e $A = \cup_{i \in I} A_i$. Sia E un insieme. Sono equivalenti:

- (i) $\forall i \in I, A_i \subseteq E$
- (ii) $A \subseteq E$

Dim: (i) \Rightarrow (ii) Sia x in A . Per definizione esiste i in I tale che $x \in A_i$. Siccome, per ipotesi $A_i \subseteq E$, $x \in E$.

(ii) \Rightarrow (i) Segue dal fatto che $A_i \subseteq A$ per ogni i ♦

16: Intersezione di una famiglia non vuota: Sia $(A_i)_{i \in I}$ una famiglia non vuota (i.e. I non vuoto) di insiemi. L'intersezione di $(A_i)_{i \in I}$ è l'insieme notato $\cap_{i \in I} A_i$ e definito da:

$$\forall x, x \in \cap_{i \in I} A_i \Leftrightarrow \forall i \in I, x \in A_i.$$

17: Proposizione: Sia $(A_i)_{i \in I}$ una famiglia non vuota d'insiemi e $A = \cap_{i \in I} A_i$. Sia E un insieme. Sono equivalenti:

- (i) $\forall i \in I, E \subseteq A_i$
- (ii) $E \subseteq A$

Dim: Es. 6♦

Esercizi:

3.1) Sia $f : X \rightarrow Y$ un'applicazione. Siano $X' \subseteq X$ un sottinsieme, e $f' : X' \rightarrow Y$ l'applicazione definita tramite: $\forall x \in X', f'(x) = f(x)$. L'applicazione f' è la restrizione di f a X' e si nota $f|_{X'}$.

- (i) Dimostrare che se f è iniettiva allora $f|_{X'}$ è iniettiva
- (ii) Se f è suriettiva è $f|_{X'}$ suriettiva?

Sia adesso l'applicazione $f' : X' \rightarrow f(X) : x \rightarrow f(x)$.

- (iii) Dimostrare che f' è suriettiva; inoltre se f è iniettiva, f' è biettiva.

3.2) Dare degli esempi esplicativi per mostrare che in generale l'applicazione h della Prop.10 (risp. Prop.11) non è unica.

3.3) Sia $f : X \rightarrow Y$ un'applicazione. Per $y \in Y$ si pone $X_y := \{x \in X / f(x) = y\}$ e si dice che X_y è la fibra di f in y . Dimostrare:

- (i) $y \neq y' \Rightarrow X_y \cap X_{y'} = \emptyset$.
- (ii) X è uguale alla riunione della famiglia $(X_y)_{y \in f(X)}$ (X è "fibrato" dalle fibre di f ; fare un disegno)

3.4) Sia $f : X \rightarrow Y$ un'applicazione e siano $A \subseteq X$, $B \subseteq \text{Im}(f)$ dei sottinsiemi. Dimostrare:

(i) $A \subseteq f^{-1}(f(A))$; (ii) $B = f(f^{-1}(B))$

(iii) Mostrare con un esempio che in generale non si ha uguaglianza in (i).

(iv) E' ancora vera (ii) senza l'ipotesi $B \subseteq \text{Im}(f)$? Mostrare che si ha sempre $f(f^{-1}(B)) \subseteq B$.

3.5) Se A è un insieme finito si nota $\text{card}(A)$ il numero dei suoi elementi. Siano X, Y due insiemi finiti con $\text{card}(X) = \text{card}(Y)$. Sia $f : X \rightarrow Y$ un'applicazione. Dimostrare che (i), (ii) e (iii) sono equivalenti:

(i) f è iniettiva

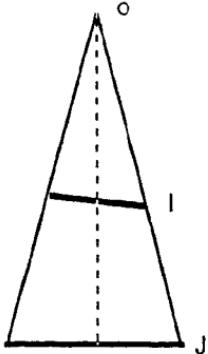
(ii) f è suriettiva

(iii) f è biiettiva

Per risolvere l'esercizio basta dimostrare (per esempio) (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i). Perché?

3.6) Dimostrare la proposizione 17.

3.7) Siano I, J i seguenti intervalli chiusi di \mathbb{R} : $I = [0, 1]$, $J = [2, 4]$. Sia $f : I \rightarrow J$: $x \rightarrow 2x + 2$. Dimostrare che f è biiettiva. Ci sono quindi "altrettanti punti" in un intervallo di lunghezza uno che in un intervallo di lunghezza due; questo si può anche vedere usando il seguente disegno:



Proiettando dal punto O ad ogni punto di I si fa corrispondere un punto di J; viceversa ogni punto di J è la proiezione di un punto di I.

3.8) Siano E, F due insiemi finiti. Si nota $A(E,F)$ l'insieme delle applicazioni di E in F . Dimostrare che se $\text{card}(E) = n$ e $\text{card}(F) = p$ allora $\text{card}(A(E,F)) = p^n$. (ragionare per induzione su n).

3.9) Siano E, F_1, \dots, F_n degli insiemi non vuoti. Per i , $1 \leq i \leq n$, si consideri l'applicazione $p_i : F_1 \times \dots \times F_n \rightarrow F_i$: $(x_1, \dots, x_n) \rightarrow x_i$.

(i) Mostrare che p è suriettiva. L'applicazione p_i si chiama la i -esima proiezione.

Se $f : E \rightarrow F_1 \times \dots \times F_n$ è un'applicazione si nota f_i l'applicazione $p_i \circ f$. Se X, Y sono due insiemi si nota $A(X, Y)$ l'insieme delle applicazioni da X in Y .

- (ii) Mostrare che l'applicazione $\phi : A(E, F_1 \times \dots \times F_n) \rightarrow A(E, F_1) \times \dots \times A(E, F_n)$: $f \mapsto (f_1, \dots, f_n)$ è una biiezione. In altri termini darsi un'applicazione da un insieme E in un prodotto cartesiano $F_1 \times \dots \times F_n$ è equivalente a darsi n applicazioni $f_i : E \rightarrow F_i$, $1 \leq i \leq n$.
- (iii) Sia $f : E \rightarrow F_1 \times \dots \times F_n$: $x \mapsto (f_1(x), \dots, f_n(x))$ un'applicazione. Mostrare che se esiste j , $1 \leq j \leq n$, tale che f_j sia iniettiva allora f è iniettiva. È vero che: f iniettiva \Rightarrow esiste j , $1 \leq j \leq n$, tale che f_j sia iniettiva?
- (iv) Sia $f : E \rightarrow F_1 \times \dots \times F_n$: $x \mapsto (f_1(x), \dots, f_n(x))$ un'applicazione. È vero che f_i è suriettiva, $\forall i$, $1 \leq i \leq n \Rightarrow f$ è suriettiva?
- (v) Dire se le seguenti applicazioni sono iniettive, suriettive:
- $f : \mathbf{R} \rightarrow \mathbf{R}^3$: $x \mapsto (1, 2x, x^2)$, $g : \mathbf{R} \rightarrow \mathbf{R}^3$: $x \mapsto (1, x^2 - x, x^2)$, $h : \mathbf{R} \rightarrow \mathbf{R}$: $x \mapsto x^2 - x$,
 $q : \mathbf{R} \rightarrow \mathbf{R}$: $x \mapsto x^2$.

4) RELAZIONI DI EQUIVALENZA.

Le relazioni di equivalenza e i relativi insiemi quozienti sono tra i concetti più importanti della matematica; la loro importanza in geometria è notevole (geometria proiettiva, algebra omologica, topologia algebrica, ...).

Intuitivamente una relazione binaria, R , su un insieme E è una proprietà riguardante le coppie (x,y) , $x \in E$, $y \in E$. Alcune coppie sono messe in relazione da R (o legate da R), altre no. Per esempio " $>$ " è una relazione su \mathbb{N} : $3 > 2$ e scriviamo $3 R 2$. Invece non si ha $2 R 3$.

Un momento di riflessione ci permette di vedere che una relazione binaria su E non è nient'altro che una corrispondenza di E in E . Infatti ad una relazione binaria, R , sull'insieme E possiamo associare la corrispondenza $c = (E, E, G)$ dove G è il sottinsieme di $E \times E$ dato da: $(x,y) \in G \Leftrightarrow x R y$; altrimenti detto, $G = \{(x,y) \in E \times E / x R y\}$.

Viceversa, data una corrispondenza $c = (E, E, G)$ possiamo definire una relazione, R , su E tramite: $x R y \Leftrightarrow (x,y) \in G$.

1: Definizione: Una relazione binaria, R , sull'insieme E è una corrispondenza, $c = (E, E, G)$, da E verso E . Si dice che G è il grafico della relazione R .

Adesso passiamo a studiare una classe particolare di relazioni binarie, le relazioni di equivalenza. Con le relazioni d'ordine queste sono tra le più importanti in matematica.

2: Definizione: Una relazione binaria, R , sull'insieme E si dice relazione di equivalenza se soddisfa:

- (1) $\forall x \in E, x R x$ (riflessività)
- (2) $\forall (x,y) \in E^2, x R y \Rightarrow y R x$ (simmetria)
- (3) $\forall (x,y,z) \in E^3, (x R y) \wedge (y R z) \Rightarrow x R z$ (transitività).

Se R è una relazione di equivalenza, invece di $x R y$ si usa anche scrivere: $x \equiv y$ (mod R) che si legge: "x è equivalente a y modulo R" o "x è congruo a y modulo R". Quando il contesto è chiaro si scrive anche $x \equiv y$.

2.1: Osservazione : La condizione (1) di riflessività significa che la diagonale di $E \times E$ è contenuta nel grafico di R . La condizione (2) di simmetria significa che il grafico è simmetrico rispetto alla diagonale.

2.2: Esempi : (i) La relazione di uguaglianza: $x R y \Leftrightarrow x = y$. Il suo grafico è la diagonale (i.e. il grafico dell'applicazione Id_E).

(ii) La relazione di uguaglianza è un caso particolare della situazione seguente:

sia $f : X \rightarrow Y$ un'applicazione. La relazione $x R y \Leftrightarrow f(x) = f(y)$ è una relazione di equivalenza.

(iii) Il procedimento più generale per costruire una relazione di equivalenza è il seguente: sia $(X_i)_{i \in I}$ una famiglia di sottinsiemi (non vuoti) di E tale che:

$$(a) \cup_{i \in I} X_i = E$$

$$(b) \forall (i,j) \in I^2, X_i \neq X_j \Rightarrow X_j \cap X_i = \emptyset$$

Se queste due condizioni sono verificate si dice che la famiglia $(X_i)_{i \in I}$ è una partizione dell'insieme E .

Definiamo una relazione binaria su E tramite: $x R y \Leftrightarrow \exists i \in I / y \in X_i \text{ e } x \in X_i$. Allora R è una relazione di equivalenza. Infatti:

(1) $x R x$: siccome $\cup_{i \in I} X_i = E$ (cfr a), abbiamo (cfr §3, 15) $x \in X_i$ per qualche i in I .

(2) $x R y \Rightarrow y R x$: è chiaro.

(3) $(x R y) \wedge (y R z) \Rightarrow x R z$: da $x R y$ segue che esiste i in I tale che $x \in X_i$ e $y \in X_i$. Nello stesso modo $y R z$ significa che esiste j in I tale che: $y \in X_j$ e $z \in X_j$. Quindi $y \in X_i \cap X_j$. Dalla condizione (b) segue $X_i = X_j$. Quindi z appartiene a X_i e pertanto $x R z$.

2.3: Esempio : Sia $n > 0$ un numero naturale. Su Z consideriamo la relazione:

$$x R y \Leftrightarrow x-y \text{ è divisibile per } n.$$

Tradizionalmente questa relazione si nota: $x \equiv y \pmod{n}$; x è congruo a y modulo n .

E' chiaro che la congruenza \pmod{n} è una relazione di equivalenza.

2.4: Esempio : (numeri razionali)

Notiamo $Z^* = \{n \in Z / n \neq 0\}$. Sull'insieme $Z \times Z^*$ definiamo la relazione:

(a,b) $R(x,y) \Leftrightarrow ay = bx$. Si verifica (cfr Es. 1) che R è una relazione di equivalenza. Questa relazione di equivalenza permette di costruire Q partendo da Z .

3: Classe di equivalenza: Sia R una relazione di equivalenza su E e x un elemento di E . La classe di equivalenza di x , nella relazione R , è l'insieme, notato $R(x)$, degli elementi di E equivalenti (modulo R) a x : $R(x) = \{y \in E / x R y\}$.

3.1: Esempi : (a) Per quanto riguarda la relazione di uguaglianza (cfr 2.2) si ha ovviamente $R(x) = \{x\}$ per ogni x in E .

(b) Sia $f: E \rightarrow F$ un'applicazione e R la relazione di equivalenza associata (cfr 2.2(ii)). La classe di equivalenza di $x \in E$ è la fibra di f in $f(x)$. Infatti: $x' R x \Leftrightarrow f(x') = f(x)$ e $f(x') = f(x)$ se e solo se x' appartiene a $f^{-1}\{f(x)\}$.

(c) Nella relazione di 2.2(iii) sia $x \in X_i$ allora $R(x) = X_i$.

(d) Nella relazione di 2.3: $R(x) = \{y \in Z / \exists k \in Z, y = x + kn\}$.

(e) Nella relazione di 2.4: $R(a,b) = \{(x,y) \in Z \times Z^* / ay = bx\}$. Notare che ogni elemento di $R(a,b)$ "rappresenta" il numero razionale a/b (perché se $a/b \in Q$ e $x/y \in Q$ allora $a/b = x/y$ se e solo se $ay = bx$).

4: Classi di equivalenza e partizioni: Mostriamo adesso che data una relazione di equivalenza R su E la famiglia delle classi di equivalenza $(R(x))_{x \in E}$ è una partizione di E .

5: Lemma: Con le notazioni precedenti: $x R y \Leftrightarrow R(x) = R(y)$.

Dim: Supponiamo $R(x) = R(y)$. Siccome $y R y$ abbiamo $y \in R(y)$ e quindi $y \in R(x)$ perciò $x R y$.

Supponiamo $x R y$ e facciamo vedere $R(y) \subseteq R(x)$. Sia z in $R(y)$ allora $y R z$. Dalla transitività di R segue $x R z$ ossia $z \in R(x)$. Nello stesso modo si mostra $R(x) \subseteq R(y)$ ♦

Quindi due elementi sono equivalenti (mod R) se e solo se appartengono alla stessa classe di equivalenza.

6: Proposizione: Sia R una relazione di equivalenza su E . Allora $(R(x))_{x \in E}$ è una partizione di E .

Dim: Bisogna mostrare (a) $E = \bigcup_{x \in E} R(x)$ e (b) $R(x) \cap R(y) \neq \emptyset \Rightarrow R(x) = R(y)$.

(a) $\bigcup_{x \in E} R(x) \subseteq E$: questo è chiaro: se $z \in \bigcup_{x \in E} R(x)$ allora per definizione esiste y in E tale che $z \in R(y)$ ma $R(y) \subseteq E$ quindi $z \in E$.

$E \subseteq \bigcup_{x \in E} R(x)$: se x' è un elemento di E allora $x' R x'$ e quindi $x' \in R(x')$ e, per definizione, x' appartiene a $\bigcup_{x \in E} R(x)$.

(b) Sia z in $R(x) \cap R(y)$. Da 3: $z \in R(x) \Leftrightarrow x R z$; e $z \in R(y) \Leftrightarrow z R y$. Per transitività: $x R y$, e per 5 questo implica $R(x) = R(y)$ ♦

INSIEME QUOZIENTE PER UNA RELAZIONE DI EQUIVALENZA:

7: Teorema: Sia R una relazione di equivalenza sull'insieme E . Allora esistono un insieme Q e un'applicazione $f : E \rightarrow Q$ tali che: $x R y \Leftrightarrow f(x) = f(y)$, per ogni x, y in E .

Dim: Prendiamo per Q l'insieme i cui elementi sono le classi di equivalenza di R (Q è un sottinsieme di $\wp(E)$). Per f prendiamo l'applicazione che a x associa la sua classe di equivalenza: $f(x) = R(x)$. Abbiamo: $x R y \Leftrightarrow f(x) = f(y)$ in virtù di 5♦

8: Definizione: L'insieme Q è l'insieme quoziante di E per la relazione di equivalenza R .

!!! Benché la dimostrazione del teorema 7 sia di una facilità disarmante, il teorema 7 introduce la nozione, fondamentale, di insieme quoziante che è una tra le maggiori difficoltà concettuali della matematica.

8.1: Esempio : Su $Z \times Z^*$ abbiamo considerato la relazione: $(a,b) R (x,y) \Leftrightarrow ay = bx$. L'insieme quoziante di $Z \times Z^*$ modulo questa relazione di equivalenza è Q , l'insieme dei numeri razionali (cfr Es. 1).

8.2: Esempio : In 2.3 abbiamo considerato su Z la relazione $x \equiv y \pmod{n}$ se e solo se $x-y$ è divisibile per n (n numero naturale positivo). Abbiamo visto che $R(x) = \{y / y = kn + x\}$. D'altra parte possiamo sempre fare la divisione euclidea (Es. 1.8) di x per n : $x = kn + r$, $0 \leq r < n$. In modo simile sia: $y = tn + r'$, $0 \leq r' < n$. Allora $x \equiv y \pmod{n}$ se e solo se: $r = r'$. Quindi la classe di equivalenza di x è l'insieme degli y che hanno lo stesso resto di x nella divisione euclidea per n . Vediamo così che l'insieme quoziante ha n elementi; sono gli n resti possibili nella divisione per n ($0, 1, \dots, n-1$). Questo insieme quoziante si nota Z/nZ e si chiama l'insieme delle congruenze modulo n .

Per esempio $Z/2Z$ ha due elementi: $Z/2Z = \{[0], [1]\}$, $[0]$ è la classe dei numeri pari, $[1]$ quella dei numeri dispari.

Esercizi:

4.1) Sull'insieme $Z \times Z^*$ si considera la seguente relazione: $(a, b) R (x, y) \Leftrightarrow ay = bx$.

(i) Dimostrare che R è una relazione di equivalenza.

La classe di equivalenza della coppia $(a, b) \in Z \times Z^*$ si nota a/b . L'insieme quoziante di $Z \times Z^*$ per R si nota Q e si chiama l'insieme dei numeri razionali. Un numero razionale a/b è quindi una classe di equivalenza: $a/b = \{(x, y) \in Z \times Z^* / ay = bx\}$.

(ii) Siano a/b , x/y , due numeri razionali. Si pone $a/b + x/y = (ay+bx)/by$. Dimostrare che se $(a', b') R (a, b)$ e $(x', y') R (x, y)$, allora $(a'y'+b'x'), b'y') R (ay+bx, by)$. Questo

dimostra che + definisce un'applicazione: $\mathbf{Q} \times \mathbf{Q} \rightarrow \mathbf{Q}$: $(a/b, x/y) \rightarrow (ay+bx)/by$; infatti la classe $(ay+bx)/by$ non dipende dalla scelta dei rappresentanti nelle classi $a/b, x/y$.

(iii) Nello stesso modo dimostrare che $a/b \cdot x/y = ax/by$ definisce un'applicazione: $\mathbf{Q} \times \mathbf{Q} \rightarrow \mathbf{Q}$: $(a/b, x/y) \rightarrow ax/by$. Abbiamo così definito l'addizione e la moltiplicazione dei numeri razionali.

(iv) dimostrare che l'applicazione $\phi: \mathbf{Z} \rightarrow \mathbf{Q}: n \rightarrow n/1$, è iniettiva. L'applicazione ϕ permette di identificare \mathbf{Z} a un sottinsieme ($\{n/1 / n \in \mathbf{Z}\}$) di \mathbf{Q} . Mostrare che $\phi(n+m) = \phi(n) + \phi(m)$, $\phi(nm) = \phi(n) \cdot \phi(m)$, per ogni (n,m) in \mathbf{Z}^2 .

(v) fare la lista di tutti gli elementi di \mathbf{Z}^* che hanno un simmetrico (in \mathbf{Z}) rispetto alla moltiplicazione. Mostrare che ogni elemento di \mathbf{Q}^* ($= \mathbf{Q} \setminus \{0/1\}$) ha un simmetrico rispetto alla moltiplicazione (determinare prima il neutro per . in \mathbf{Q}^*). In particolare ogni elemento di \mathbf{Z}^* (considerato come sottinsieme di \mathbf{Q}^* tramite ϕ) ha un simmetrico (in \mathbf{Q}^*) per la moltiplicazione (i.e. "è invertibile" (per .)). Si può dire che \mathbf{Q} si costruisce per rendere invertibili gli elementi di \mathbf{Z}^* . Questo procedimento è molto generale.

4.2) Sia $n \in \mathbf{N}$, $n > 0$. Notiamo $\mathbf{Z}/n\mathbf{Z}$ l'insieme quoziente di \mathbf{Z} per la relazione: $x \equiv y \pmod{n}$ se e solo se $x-y$ è divisibile per n . Sia $p: \mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}$ l'applicazione canonica che ad ogni x associa la sua classe di equivalenza. Se $x \in \mathbf{Z}$ notiamo $\bar{x} = p(x)$.

Si pone $\bar{x} + \bar{y} := \overline{x+y}$, $\bar{x} \cdot \bar{y} := \overline{xy}$.

(i) Dimostrare che $+$, \cdot sono delle leggi di composizione interna su $\mathbf{Z}/n\mathbf{Z}$; ossia che $+$: $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}$: $(\bar{x}, \bar{y}) \rightarrow \bar{x} + \bar{y}$ e \cdot : $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}$: $(\bar{x}, \bar{y}) \rightarrow \bar{x} \cdot \bar{y}$, sono delle applicazioni.

(ii) fare le tabelle di $+$, \cdot per $n = 2, 3, 4$.

4.3) Sia P il piano della geometria elementare. Si fissa un'origine O e si definisce su P la relazione: $x R y$ se e solo se esiste una retta contenente x, y e O .

(i) Dire se R è una relazione di equivalenza su P .

(ii) Sia $P^* := P \setminus \{O\}$. Mostrare che R ristretta a P^* è una relazione di equivalenza su P^* . Descrivere la classe di equivalenza di un punto x di P^* , quindi descrivere l'insieme quoziente P^*/R .

NB: L'insieme quoziente P^*/R si nota $\mathbf{P}^1(\mathbf{R})$ (o anche $\mathbf{P}(\mathbf{R}^2)$) e si chiama retta proiettiva reale.

4.4) In \mathbf{R}^2 sia C la circonferenza di centro il punto $M = (1, 0)$ e di equazione $x^2 - 2x + y^2 = 0$.

(i) Rappresentare graficamente C .

(ii) Sia R una retta passante per l'origine $O = (0, 0)$, R diversa da OY , l'asse delle y . Mostrare che R interseca C in due punti distinti: $R \cap C = \{O, P_R\}$, $P_R \neq O$.

(iii) Sia $P \neq O$ un punto di C . Si nota $[P:O]$ la retta per O e P . Se si fa tendere P verso O lungo la circonferenza C , la retta $[P:O]$ tende verso l'asse delle y : si dice che OY è tangente a C in O . In questa situazione possiamo dire che l'intersezione di OY con C è il punto O "contato due volte". Giustificare algebricamente quest'ultima affermazione.

(iv) Sia \mathbf{P} l'insieme delle rette passanti per O . Se $R \in \mathbf{P}$, $R \neq OY$, abbiamo $R \cap C = \{O, P_R\}$ e poniamo $f(R) = P_R$. Questo definisce un'applicazione $f: \mathbf{P} \setminus \{OY\} \rightarrow C$. Poi, "per continuità" (cfr (iii)) estendiamo questa applicazione a tutto \mathbf{P} ponendo $f(OY) = O$. Mostrare che $f: \mathbf{P} \rightarrow C$ è una biiezione e concludere che la retta proiettiva $\mathbf{P}^1(\mathbb{R})$ (cfr Es. 3) è in biiezione con C .

4.5) In \mathbb{R}^2 sia S la circonferenza di centro $O = (0, 0)$ e di equazione $x^2 + y^2 = 1$. Su S si definisce la relazione: $P \equiv Q \Leftrightarrow P = Q$ o $P, Q \neq O$ sono allineati. Mostrare che \equiv è una relazione di equivalenza. L'insieme quoziante si nota $S/\{\pm 1\}$ (giustificare questa scrittura). Mostrare che $S/\{\pm 1\}$ è in biiezione con: (a) la semicirconferenza $S_+ = \{(x, y) \in S / y \geq 0\}$ con le estremità $(-1, 0), (1, 0)$ identificate, (b) l'insieme delle rette passanti per O , (c) $\mathbf{P}^1(\mathbb{R})$, (d) la circonferenza C di equazione $x^2 - 2x + y^2 = 0$ (cfr Es. 4).

5) GRUPPI.

Notiamo O il punto $(0, 0)$ di \mathbb{R}^2 e sia $G = \{f / f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, f \text{ è biiettiva, } e f(O) = O\}$. Abbiamo chiaramente: (a) $\text{Id}_{\mathbb{R}^2}$ appartiene a G , (b) se f e g appartengono a G allora anche $f \circ g$ appartiene a G , (c) se f appartiene a G anche f^{-1} appartiene a G . Da (a), (b) e (c) segue che (G, \cdot) è un "gruppo di trasformazioni".

Analogamente possiamo considerare l'insieme, M , delle applicazioni biiettive dal piano in se stesso che conservano la distanza tra due punti ($d(P, Q) = d(f(P), f(Q))$). Le proprietà (a), (b), (c) sono verificate per (M, \cdot) . Il gruppo (M, \cdot) gioca un ruolo fondamentale in geometria euclidea. Se Δ è una figura geometrica (per esempio un triangolo), possiamo considerare $M_\Delta = \{f \in M / f(\Delta) = \Delta\}$. Le proprietà (a), (b), (c) sono soddisfatte per M_Δ (M_Δ è un sottogruppo di M). Il gruppo M_Δ riflette le simmetrie della figura Δ (cfr Es. 5.8). Quindi, in qualche misura, studiare le proprietà eucleede delle figure geometriche è equivalente a studiare i sottogruppi di M . Per questo, ed altri motivi, la teoria dei gruppi è presente in molte questioni di matematica, fisica, chimica, ... (simmetrie delle equazioni algebriche: teoria di Galois, simmetrie delle leggi fisiche: gruppo di Lorentz, simmetrie dei cristalli, ...). Questa onnipresenza mette in evidenza che quello che conta veramente non è la specifica natura degli oggetti matematici considerati, ma piuttosto le relazioni che intercorrono tra di loro. Questo punto di vista introduce la nozione di struttura algebrica che, nel caso più semplice, si può descrivere così: su un insieme non meglio precisato sono definite delle "operazioni" che soddisfano certi assiomi (commutatività ecc...). Lo studio di queste situazioni astratte, oltre che essere un gran risparmio di tempo e quindi una semplificazione, permette spesso una maggior comprensione di problemi concreti.

La struttura di gruppi è tra le strutture algebriche più semplici. In questo paragrafo, dopo le definizioni, introduciamo le nozioni di base (sottogruppi, morfismi, nuclei).

1: Definizione: (legge di composizione interna):

Una legge di composizione interna su un insieme E è un'applicazione $f: E \times E \rightarrow E$.

Si tratta dunque di fare corrispondere ad una coppia di elementi di E uno (ed uno solo) elemento di E .

In pratica invece di scrivere $f(x,y)$ si usa scrivere $x \cdot y$ (o: $x \perp y$, $x \cdot y$, $x+y$, ecc....)

2: Definizione: (*Struttura di gruppo*) Sia $E \times E \rightarrow E: (x,y) \rightarrow x * y$ una legge di composizione interna sull'insieme E . Se le condizioni (G1), ..., (G3), qui sotto, sono verificate si dice che $(E, *)$ è una struttura di gruppo sull'insieme E (si dice anche che $(E, *)$ è un gruppo).

(G1) $\forall (x,y,z) \in E^3: x * (y * z) = (x * y) * z$ (associatività)

(G2) $\exists e \in E / \forall x \in E: x * e = e * x = x$ (e viene chiamato elemento neutro)

(G3) $\forall x \in E, \exists x' \in E / x * x' = x' * x = e$ (l'elemento x' si dice simmetrico, o inverso, di x).

3: Osservazioni : (i) Nella pratica quando la legge è precisata dal contesto si usa dire: "... E è un gruppo" però bisogna sempre ricordare che un gruppo non è un insieme ma una coppia $(E, *)$.

(ii) Il simmetrico di un elemento x si nota anche x^{-1} (specialmente quando $*$ = .) o anche $-x$ (quando $*$ = +).

Osservare che il simmetrico del simmetrico di x è x : $(x^{-1})^{-1} = x$.

4: Esempi : (a) Sull'insieme dei numeri interi abbiamo l'addizione che definisce una legge di composizione interna: $Z \times Z \rightarrow Z : (a,b) \rightarrow a+b$. Si verifica facilmente che $(Z, +)$ è un gruppo. Il neutro è 0 e il simmetrico di n è $-n$.

Notare che $(N, +)$ non è un gruppo (mancano i simmetrici), infatti Z si costruisce a partire da N "aggiungendo" i simmetrici per $+$.

Si verifica facilmente che $(Q, +)$, $(R, +)$, $(C, +)$ sono dei gruppi.

(b) Sia $Z^* := Z \setminus \{0\}$. La moltiplicazione definisce una legge di composizione interna su Z^* ma Z^* non è un gruppo: gli unici elementi che hanno un simmetrico sono 1 e -1. Per rendere gli elementi di Z^* invertibili si costruisce Q ; infatti (Q^*, \cdot) è un gruppo. Nello stesso modo (R^*, \cdot) , (C^*, \cdot) sono dei gruppi.

(c) Sia X un insieme e $\mathfrak{S}(X)$ l'insieme delle permutazioni di X . Su $\mathfrak{S}(X)$ abbiamo la legge di composizione interna: $\mathfrak{S}(X) \times \mathfrak{S}(X) \rightarrow \mathfrak{S}(X) : (f,g) \rightarrow fog$. Si verifica (cfr Es.3) che $(\mathfrak{S}(X), \circ)$ è un gruppo.

5: Neutro e simmetrici: unicità: Sia $(E, *)$ un gruppo. Si vede facilmente (cfr Es.1) che l'elemento neutro è unico cioè se: $\forall x \in E, a * x = x * a = x$ allora $a = e$. Nello stesso modo si vede che ogni elemento x di E ha un unico inverso.

6: Definizione: (*Gruppi abeliani, o commutativi*) Sia $(E, *)$ un gruppo. Si dice che $(E, *)$ è un gruppo commutativo (o abeliano, in memoria del matematico norvegese Abel) se:

$$\forall (x,y) \in E^2: x * y = y * x.$$

6.1: Esempi : Negli esempi di 4 (a), (b) i gruppi sono commutativi, nell'esempio di 4(c) questo non è sempre vero (cfr Es. 3).

7: Equazioni in un gruppo: In un gruppo $(E, *)$ l'equazione $x*a = b$ ha sempre una soluzione (tra l'altro unica): $(x*a)*a^{-1} = b*a^{-1}$ ma $(x*a)*a^{-1} = x*(a*a^{-1})$ per associatività; inoltre $x*(a*a^{-1}) = x*e = x$, quindi l'unica soluzione dell'equazione è: $x = b*a^{-1}$.

In particolare se $x*a = x*b$ allora si può cancellare e concludere che $a = b$. Infatti componendo a sinistra col simmetrico di x abbiamo: $x^{-1}*(x*a) = x^{-1}*(x*b)$, per associatività: $(x^{-1}*x)*a = (x^{-1}*x)*b$ da cui: $e*a = e*b$ e finalmente: $a = b$.

La cancellazione è possibile perché $(E, *)$ è un gruppo. Se non si è in un gruppo non è detto che si possa semplificare. Esempio: sia $E = \mathbb{R}$ e $* = \cdot$, la moltiplicazione usuale. Da una relazione $xa = xb$ non si può (sempre) dedurre $a = b$, per esempio $0.1 = 0.2$ ma $1 \neq 2$. Infatti (\mathbb{R}, \cdot) non è un gruppo (0 non ha un simmetrico per \cdot).

Invece (\mathbb{R}^*, \cdot) è un gruppo e in \mathbb{R}^* (quindi $x \neq 0$) la relazione precedente fornisce $a = b$.

8: Definizione: (Sottogruppi) Sia $(G, *)$ un gruppo e $H \subseteq G$ un sottinsieme. Supponiamo soddisfatte le condizioni:

$$(1) x \in H \wedge y \in H \Rightarrow x*y \in H$$

$$(2) e \in H$$

$$(3) x \in H \Rightarrow x^{-1} \in H.$$

Allora si dice che $(H, *)$ è un sottogruppo di $(G, *)$.

La condizione (1) permette di considerare l'applicazione: $H \times H \rightarrow H : (x, y) \mapsto x*y$ (detta anche restrizione di $*$: $G \times G \rightarrow G$ a $H \times H$); in particolare $*$ è interna su H . Le condizioni (2), (3) e il fatto che $(G, *)$ sia un gruppo permettono di verificare che $(H, *)$ è un gruppo. Siccome la legge di $(H, *)$ è indotta da quella su G è naturale dire che $(H, *)$ è un sottogruppo di $(G, *)$.

8.1: Osservazione : Un sottogruppo di un gruppo abeliano è abeliano.

9: Proposizione: (Criterio affinchè un sottinsieme sia un sottogruppo):

Sia $(G, *)$ un gruppo e $H \subseteq G$ un sottinsieme. Sono equivalenti:

(a) $(H, *)$ è un sottogruppo di $(G, *)$

(b) H è non vuoto e: per ogni x in H e ogni y in H : $x*y^{-1}$ appartiene a H .

Dim: (a) \Rightarrow (b) Se H è un sottogruppo di $(G, *)$ allora $e \in H$ e quindi $H \neq \Delta$.

Siano x e y due elementi di H da 8.(3): $y^{-1} \in H$ e poi da 8.(1): $x * y^{-1} \in H$.

(b) \Rightarrow (a) Verifichiamo (1), ..., (3) di 8: siccome H è non vuoto esiste $a \in H$. Abbiamo $a \in H$, e $a \in H$ quindi: $e = a * a^{-1} \in H$ e (2) è verificata. Adesso se $x \in H$ abbiamo: $e \in H$, e: $x \in H$ da cui: $e * x^{-1} = x^{-1} \in H$ e (3) è soddisfatta. Finalmente se x e y sono elementi di H , da quanto precede $y^{-1} \in H$, pertanto: $x * (y^{-1})^{-1} = x * y \in H$ e anche (1) è verificata♦

10: Definizione: (*Morfismi di gruppi*) Siano $(G, *)$ e (E, o) due gruppi. Un'applicazione $f : G \rightarrow E$ è un morfismo dal gruppo $(G, *)$ nel gruppo (E, o) se: $\forall (g, g') \in G^2$: $f(g * g') = f(g) o f(g')$.

10.1: Terminologia: Certi autori chiamano omomorfismo quello che noi chiamiamo morfismo.

11: Prime proprietà dei morfismi di gruppi: Notiamo e l'elemento neutro di $(G, *)$ e ϵ quello di (E, o) .

(i) Ponendo $g' = e$ nella definizione 10 otteniamo: $f(g * e) = f(g) o f(e)$ ma $g * e = g$ quindi viene: $f(g) = f(g) o f(e)$. Componendo a sinistra col simmetrico di $f(g)$:

$$\begin{aligned} (f(g))^{-1} o f(g) &= (f(g))^{-1} o [f(g) o f(e)] \\ \epsilon &= (f(g))^{-1} o [f(g) o f(e)], \quad \text{per definizione di simmetrico} \\ \epsilon &= [(f(g))^{-1} o f(g)] o f(e), \quad \text{per associatività} \\ \epsilon &= \epsilon o f(e), \quad \text{per definizione di simmetrico} \\ \epsilon &= f(e), \quad \text{per definizione di elemento neutro.} \end{aligned}$$

In conclusione: se f è un morfismo di gruppi allora l'immagine dell'elemento neutro è l'elemento neutro: $f(e) = \epsilon$.

(ii) Poniamo adesso $g' = g^{-1}$ nella definizione 10:

$$f(g * g^{-1}) = f(g) o f(g^{-1})$$

$$f(e) = f(g) o f(g^{-1}); \text{ ma dal passo precedente } f(e) = \epsilon \text{ quindi:}$$

$$\epsilon = f(g) o f(g^{-1}).$$

Nello stesso modo prendendo $g^{-1} = g$ e $g' = g$ otteniamo: $\epsilon = f(g^{-1}) o f(g)$. Risulta, dalla definizione del simmetrico che: $f(g^{-1}) = f(g)^{-1}$.

In un morfismo di gruppi l'immagine del simmetrico è il simmetrico dell'immagine.

11: Immagine di un morfismo di gruppi: Siano $(G, *)$ e (E, o) due gruppi e $f : G \rightarrow E$ un morfismo di gruppi. Si verifica (cfr Es. 2) che l'insieme immagine di f , $\text{Im}(f)$, è un sottogruppo di (E, o) . Si ricorda che $\text{Im}(f) = \{x \in E / \exists g \in G \text{ tale che } f(g) = x\}$.

12: Nucleo di un morfismo di gruppi: Siano $(G, *)$ e (E, o) due gruppi e $f : G \rightarrow E$ un morfismo di gruppi. Il nucleo del morfismo f è l'insieme: $\{g \in G / f(g) = \varepsilon\}$, dove ε indica il neutro di (E, o) .

Invece di nucleo si dice spesso Ker (che è l'abbreviazione di nucleo in tedesco) e si scrive $\text{Ker}(f)$.

Sia e l'elemento neutro di $(G, *)$, abbiamo (cfr 10.2): $f(e) = \varepsilon$ quindi: $e \in \text{Ker}(f)$; il Ker di un morfismo di gruppi non è mai vuoto, in realtà abbiamo molto di più:

12.1: Proposizione: Siano $(G, *)$ e (E, o) due gruppi e $f : G \rightarrow E$ un morfismo di gruppi. Allora $\text{Ker}(f)$ è un sottogruppo di $(G, *)$.

Dim.: Sappiamo già che $\text{Ker}(f)$ è non vuoto quindi (cfr 9) basta verificare: $g \in \text{Ker}(f)$ e $h \in \text{Ker}(f)$ implica: $g * h^{-1} \in \text{Ker}(f)$. Abbiamo:

$$\begin{aligned} f(g * h^{-1}) &= f(g) o f(h^{-1}) \\ &= f(g) o f(h)^{-1} \quad (\text{cfr 10.2 (ii)}) \\ &= \varepsilon o \varepsilon^{-1} \quad (\text{perché } g \text{ e } h \text{ sono in } \text{Ker}(f)) \\ &= \varepsilon \quad (\text{perché } \varepsilon^{-1} = \varepsilon) \end{aligned}$$

Quindi: $g * h^{-1} \in \text{Ker}(f)$ ♦

13: Proposizione: Siano $(G, *)$ e (E, o) due gruppi e $f : G \rightarrow E$ un morfismo di gruppi. Il morfismo f è iniettivo se e solo se $\text{Ker}(f) = \{e\}$ dove e indica il neutro di $(G, *)$.

Dim.: (i) Supponiamo $\text{Ker}(f) = \{e\}$ e mostriamo che f è iniettivo. Siano g, h elementi di G tali che $f(g) = f(h)$. Abbiamo:

$$\begin{aligned} f(g * h^{-1}) &= f(g) o f(h^{-1}) \\ &= f(g) o f(h)^{-1} \quad (\text{cfr 10.2(ii)}) \\ &= f(g) o f(g)^{-1} \quad (\text{perché } f(g) = f(h) \text{ per ipotesi}) \\ &= e. \end{aligned}$$

Quindi $g * h^{-1}$ appartiene a $\text{Ker}(f) = \{e\}$ pertanto: $g * h^{-1} = e$. Componendo a destra con h si ricava $g = h$; quindi f è iniettiva.

(ii) Supponiamo f iniettiva e mostriamo $\text{Ker}(f) = \{e\}$. Se $g \in \text{Ker}(f)$ allora $f(g) = f(e) = e$ per definizione di Ker . Essendo l'applicazione f iniettiva questo implica $g = e$ ♦

14: L'associatività e le parentesi: Sia $(G, *)$ un gruppo e x, y, z tre elementi di G .

Poniamo: $x * y * z := (x * y) * z$. Notare che a priori la scrittura $x * y * z$ non ha senso: non sappiamo comporre tre elementi di G ma solo due.

Per associatività di $*$ abbiamo: $x * y * z := (x * y) * z = x * (y * z)$. Quindi $x * y * z$ è un elemento ben definito del gruppo che si ottiene così: si mettono due termini consecutivi qualsiasi tra parentesi, si calcola il loro prodotto e poi si compone questo prodotto con l'elemento rimanente.

Se adesso abbiamo quattro elementi x_1, \dots, x_4 definiamo: $x_1 * x_2 * x_3 * x_4 := (x_1 * x_2 * x_3) * x_4$; e per induzione definiamo $x_1 * \dots * x_n := (x_1 * \dots * x_{n-1}) * x_n$, per ogni intero naturale n .

Sia adesso $(x_1 * \dots * (x_i * x_{i+1} * \dots * x_k) * \dots * x_n)$ un'espressione con n termini e contenente delle parentesi in un ordine qualsiasi. Dimostriamo, per induzione su n , che questa espressione è uguale a $x_1 * \dots * x_n$. Il caso $n = 3$ è già stato fatto; supponiamo quindi $n > 3$ e il risultato vero per ogni $m < n$ (cfr §1, 13.4). C'è almeno una parentesi che raccoglie più di un termine per esempio $(x_i * x_{i+1} * \dots * x_k)$. Poniamo $x_i * x_{i+1} * \dots * x_k = y$. L'espressione precedente è della forma: $x_1 * \dots * y * \dots * x_n$ e contiene meno di n termini e forse altre parentesi. Per ipotesi di induzione completa questa espressione è uguale a: $(x_1 * \dots * y * \dots * x_{n-1}) * x_n$ (senza parentesi all'interno). Inserendo il valore precedente di y , viene: $(x_1 * \dots * y * \dots) * x_n = (x_1 * \dots * (x_i * x_{i+1} * \dots * x_k) * \dots * x_{n-1}) * x_n$; la prima parentesi raccoglie $n-1$ termini quindi per ipotesi di induzione è uguale a $(x_1 * \dots * x_i * x_{i+1} * \dots * x_k * \dots * x_{n-1})$ (senza parentesi all'interno). Finalmente l'espressione iniziale è uguale a $(x_1 * \dots * x_{n-1}) * x_n = x_1 * \dots * x_n$.

In altre parole, le parentesi si possono togliere o mettere dove si vuole. La scrittura $x_1 * \dots * x_n$ ($n \in \mathbb{N}^*$) denota un elemento ben definito di G che si calcola raccogliendo tra parentesi gli elementi che si desiderano, facendone il prodotto e ripetendo questa operazione fino a rimanere con un elemento solo.

In particolare, la scrittura $x * \dots * x$ è ben definita. Osserviamo che $(x * \dots * x)^{-1} = x^{-1} * \dots * x^{-1}$ (con lo stesso numero di termini da entrambe le parti). Infatti:

$$(x * \dots * x) * (x^{-1} * \dots * x^{-1}) = x * \dots * (x * x^{-1}) * \dots * x^{-1} = \dots = e.$$

Sia $(G, +)$ un gruppo in notazione additiva. Invece di $x + \dots + x$ (n termini) si scrive nx .

Se il gruppo è scritto in notazione moltiplicativa, invece di $x \dots x$ (n termini) si scrive x^n .

Fissiamo un elemento x del gruppo $(G, *)$ e definiamo un'applicazione $f: \mathbb{Z} \rightarrow G$ nel modo seguente: $f(n) = (x * \dots * x)$, n termini, se $n \geq 1$; $f(0) = e$; $f(n) = (x * \dots * x)^{-1}$, $-n$ termini, se $n \leq -1$.

Si verifica subito che f è un morfismo dal gruppo $(\mathbb{Z}, +)$ nel gruppo $(G, *)$.

In notazione additiva, $(G, +)$, si scrive: $f(n) = nx$ (quindi $f(0) = 0x = e = 0_G$)

In notazione moltiplicativa, (G, \cdot) , si scrive: $f(n) = x^n$ (quindi $f(0) = x^0 = e = 1_G$).

L'immagine di f è il sottogruppo (cfr 8.9) generato da x , questo sottogruppo si nota: (x) . Abbiamo quindi: $(x) = \{ \dots, x^{-1} * x^{-1}, x^{-1}, e, x, x * x, \dots \}$.

Esercizi:

5.1) Sia $(G, *)$ un gruppo di elemento neutro e . Se $x \in G$ si nota x^{-1} il simmetrico di x per $*$.

(i) Dimostrare l'unicità dell'elemento neutro (ossia: se esiste $e' \in G$ tale che: per ogni x in G , $e' * x = x * e' = x$, allora $e' = e$).

(ii) Dimostrare l'unicità del simmetrico (ossia: sia $x \in G$, se esiste $y \in G$ tale che $x * y = y * x = e$, allora $y = x^{-1}$).

5.2) Siano $(G, *)$, $(E, +)$ due gruppi e $f: G \rightarrow E$ un morfismo di gruppi. Dimostrare che $\text{Im}(f)$ è un sottogruppo di $(E, +)$.

5.3) Siano $(G, +)$, $(E, .)$, $(F, *)$ tre gruppi e $f: G \rightarrow E$, $h: E \rightarrow F$ due morfismi di gruppi. Dimostrare che $h \circ f: G \rightarrow F$ è un morfismo di gruppi.

5.4) Su \mathbf{R}^n definiamo la legge di composizione interna, $+$, tramite:

$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$. Dimostrare che $(\mathbf{R}^n, +)$ è un gruppo.

5.5) Sia $n \in \mathbf{N}$, $n > 0$, e $\mathbf{Z}/n\mathbf{Z}$ l'insieme quoziente di \mathbf{Z} per la relazione: $x \equiv y \pmod{n}$ se e solo se $x-y$ è divisibile per n (§4, 8.2). Su $\mathbf{Z}/n\mathbf{Z}$ è stata definita l'addizione $+$ (Es.4.2). Mostrare che $(\mathbf{Z}/n\mathbf{Z}, +)$ è un gruppo abeliano.

5.6) Determinare tutti i gruppi $(E, *)$ con $\text{card}(E) \leq 3$ (si darà la tabella di $*$). Che cosa osservate?

5.7) Sia X un insieme e sia $\mathfrak{S}(X)$ l'insieme delle applicazioni biettive da X in X ("permutazioni dell'insieme X "). Consideriamo la legge di composizione interna: $\mathfrak{S}(X) \times \mathfrak{S}(X) \rightarrow \mathfrak{S}(X)$: $(f, g) \rightarrow f \circ g$.

(i) Mostrare che $(\mathfrak{S}(X), \circ)$ è una struttura di gruppo su $\mathfrak{S}(X)$.

(ii) Sia X un insieme finito, $\text{card}(X) = n$. Mostrare che $\mathfrak{S}(X)$ è un insieme finito e determinare $\text{card}(\mathfrak{S}(X))$.

(iii) Dire se $(\mathfrak{S}(X), \circ)$ è un gruppo abeliano (suggerimento: considerare X insieme finito con $\text{card}(X) \leq 3$)

5.8) Sia Δ un triangolo nel piano Π . Sia $M = \{f: \Pi \rightarrow \Pi / f \text{ è biettiva e } f \text{ conserva le distanze}\}$. Notiamo $M_\Delta = \{f \in M / f(\Delta) = \Delta\}$. Dimostrare che M_Δ è un sottogruppo di M .

Se Δ è un triangolo (a) non isoscele (risp. (b) isoscele, (c) equilatero), osservare che $\text{card}(M_\Delta) = 1$ (risp. 2, 6) e descrivere geometricamente gli elementi di M_Δ .

5.9) Siano $(G, +)$, $(E, .)$ due gruppi e $f: G \rightarrow E$ e f un morfismo di gruppi. Dimostrare che se l'applicazione f è biettiva allora l'applicazione reciproca, f^{-1} , è un morfismo.

NB: Un morfismo biettivo si dice isomorfismo.

5.10) Sia $\mathbf{R}_+^* = \{x \in \mathbf{R} / x > 0\}$.

- (i) Dimostrare che (\mathbb{R}_+^*, \cdot) e $(\mathbb{R}, +)$ sono dei gruppi abeliani.
(ii) E' Log: $\mathbb{R}_+^* \rightarrow \mathbb{R}$ un isomorfismo tra i gruppi (\mathbb{R}_+^*, \cdot) e $(\mathbb{R}, +)$?

5.11) Segue dal teorema fondamentale dell'algebra che l'equazione $X^n = 1$ (*), ha n radici (contate con molteplicità) in \mathbb{C} .

(i) Mostrare che queste radici sono distinte (osservare che 0 è radice con molteplicità n-1 della derivata di $P(X) = X^n - 1$, inoltre $0^n \neq 1$). Le radici dell'equazione (*) si chiamano radici n-esime dell'unità. L'insieme delle radici n-esime dell'unità si nota μ_n . Si ha $\mu_n \cap \mathbb{R} = \{1\}$ se n è dispari, $\{\pm 1\}$ se n è pari.

(ii) Sia $z = r(\cos\theta + \sin\theta) = r \cdot e^{i\theta}$ un numero complesso e cerchiamo sotto quali condizioni verifica (*). Abbiamo $z^n = r^n(\cos n\theta + \sin n\theta) = r^n e^{in\theta}$. Siccome $1 = 1 \cdot e^{i0}$, viene $r^n = 1$ e $n\theta = 0$ modulo 2π . Pertanto $r = 1$ (r è reale positivo) e $\theta = 2k\pi/n$. Ogni radice di (*) è della forma $e^{i2k\pi/n}$. In particolare ogni radice dell'unità è situata sulla circonferenza unità $U = \{z \in \mathbb{C} / |z| = 1\}$. Rappresentare graficamente (sulla circonferenza unità) gli elementi di μ_n per $n = 3, 4$.

(iii) Dimostrare che $\mu_n = \{\omega_k = e^{i2k\pi/n}, 0 \leq k \leq n-1\}$. Osservare che gli elementi di μ_n sono i vertici di un poligono con n lati iscritto nella circonferenza unità.

(iv) Dimostrare che (U, \cdot) è un gruppo (\cdot è la moltiplicazione tra numeri complessi), e che (μ_n, \cdot) è un sottogruppo di (U, \cdot) .

(v) Sia $f: \mu_n \rightarrow \mathbb{Z}/n\mathbb{Z}: e^{i2k\pi/n} \mapsto [k]$ ([k] è la classe di k modulo $n\mathbb{Z}$). Mostrare che f stabilisce un isomorfismo tra il gruppo moltiplicativo (μ_n, \cdot) e il gruppo additivo $(\mathbb{Z}/n\mathbb{Z}, +)$.

(vi) Sia Z un numero complesso non nullo. Se z è una radice n-esima di Z, mostrare che le n radici n-esime di Z sono $z\omega_k$, $\omega_k \in \mu_n$.

5.12) Sia (G, \cdot) un gruppo. Mostrare che l'applicazione $G \rightarrow G: x \mapsto x^2 = x \cdot x$ è un morfismo di gruppi se e solo se (G, \cdot) è abeliano.

5.13) Sia (G, \cdot) un gruppo e $f: G \rightarrow G: x \mapsto x^{-1}$.

(i) f è un morfismo di gruppi se e solo se (G, \cdot) è abeliano.
(ii) un automorfismo del gruppo (G, \cdot) è un'applicazione biiettiva $h: G \rightarrow G$ tale che h è l'applicazione reciproca h^{-1} siano dei morfismi di gruppi.
Con le notazioni precedenti mostrare che se f è un morfismo allora è un automorfismo.

5.14) Sia G un sottogruppo di $(\mathbb{Z}, +)$. Si pone $G_+ := \{x \in G / x > 0\}$.

Mostrare che se $G \neq \{0\}$, allora G_+ è non vuoto. Sia d il più piccolo elemento di G_+ (cfr §1, Es.6). Mostrare che ogni elemento di G è multiplo di d (cfr §1, Es.9). Concludere che se G è un sottogruppo di $(\mathbb{Z}, +)$, allora esiste un unico d $\in \mathbb{N}$, tale che $G = d\mathbb{Z} := \{m \in \mathbb{Z} / \exists k \in \mathbb{Z}, m = kd\}$.

6) Anelli, corpi, campi.

La struttura di campo è forse la struttura algebrica più "naturale" in quanto i numeri reali, con le operazioni di addizione e moltiplicazione, ne forniscono un esempio che conosciamo bene. Questa struttura si può riassumere così: $(\mathbb{R}, +)$ è un gruppo abeliano, (\mathbb{R}^, \cdot) è un gruppo, la moltiplicazione è distributiva rispetto all'addizione, la moltiplicazione è commutativa. Sono le regole del "calcolo usuale". Se non si richiede la commutatività della moltiplicazione, si ottiene la struttura di corpo. Inoltre se non si richiede che ogni elemento non nullo abbia un simmetrico rispetto alla moltiplicazione, si ottiene la struttura di anello. Per esempio $(\mathbb{Z}, +, \cdot)$ è un anello. Ci sono anelli in cui la relazione $x \cdot y = 0$ non implica $x = 0$ o $y = 0$. Noi, purtroppo, avremo a che fare con anelli di questo tipo (funzioni, matrici).*

1: Definizione: Sia A un insieme con due leggi di composizione interne notate $+$ e \cdot , la terna $(A, +, \cdot)$ è una struttura di anello su A se:

(A1) $(A, +)$ è un gruppo abeliano

(A2) esiste un neutro (notato 1) per \cdot :

$$\exists 1 \in A / \forall x \in A, 1 \cdot x = x \cdot 1 = x$$

(A3) la legge \cdot è associativa: $\forall (x, y, z) \in A^3: x \cdot (y \cdot z) = (x \cdot y) \cdot z$

(A4) la legge \cdot è distributiva rispetto a $+$: $\forall (x, y, z) \in A^3, x \cdot (y + z) = x \cdot y + x \cdot z$, e: $(x + y) \cdot z = x \cdot z + y \cdot z$.

1.1: Osservazione : Se $(A, +, \cdot)$ è una struttura di anello su A si dice che $(A, +, \cdot)$ è un anello. Se le leggi $+$, \cdot sono chiaramente specificate dal contesto, con abuso di linguaggio, si dice che A è un anello.

2: Esempi : $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$, dove $+$ e \cdot sono le operazioni naturali, sono degli anelli; $(\mathbb{N}, +, \cdot)$ non è un anello (perché?).

3: Elementi neutri : Sia $(A, +, \cdot)$ un anello. Abbiamo due elementi neutri: quello del gruppo abeliano $(A, +)$, notato 0, e quello della "moltiplicazione" \cdot , notato 1. A priori potrebbe essere $0 = 1$. Questa circostanza è poco interessante: se $0 = 1$ allora ogni elemento di A è uguale a 0. Per vederlo dimostriamo prima il:

4: Lemma: Sia $(A, +, \cdot)$ un anello. Se 0 è il neutro del gruppo abeliano $(A, +)$ allora per ogni x in A : $x \cdot 0 = 0 \cdot x = 0$.

Dim: Abbiamo: $x = x \cdot 1$ (A2)

$$= x \cdot (1+0) \text{ (per A1: } 1 = 1+0\text{)}$$

$$= x \cdot 1 + x \cdot 0 \text{ (per A4)}$$

$= x + x \cdot 0$ (per A2); sommando a sinistra con $-x$:
 $-x + x = -x + (x + x \cdot 0)$ da cui $0 = -x + (x + x \cdot 0)$, per associatività: $-x + (x + x \cdot 0) = (-x + x) + x \cdot 0 = 0 + x \cdot 0 = x \cdot 0$. In conclusione $x \cdot 0 = 0$. Nello stesso modo si dimostra $0 \cdot x = 0$ ♦

Tornando agli elementi neutri, se $0 = 1$, per ogni x in A abbiamo: $x = x \cdot 1 = x \cdot 0$, per il lemma precedente $x \cdot 0 = 0$, e pertanto $x = 0$. Quindi se $0 = 1$, $A = \{0\}$ e c'è ben poco da dire. D'ora in poi supporremo sempre $0 \neq 1$.

5: Definizione: (*Anelli commutativi*) Un anello $(A, +, \cdot)$ è commutativo se . è commutativa, cioè se:

$$\forall (x,y) \in A^2: x \cdot y = y \cdot x.$$

5.1: Esempi : Gli anelli dell'esempio 2 sono tutti commutativi, ma vedremo più avanti esempi importanti di anelli non commutativi.

6: Definizione: Siano $(A, +, \cdot)$ un anello e $A' \subseteq A$ un sottinsieme. Si dice che A' è un sottanello di A se:

(i) $(A', +)$ è un sottogruppo di $(A, +)$

(ii) la moltiplicazione induce una legge interna su A' : $\forall (x,y) \in A'^2, x \cdot y \in A'$.

(iii) l'elemento neutro rispetto alla moltiplicazione appartiene a A' : $1 \in A'$.

6.1: Osservazione : Come nel caso dei gruppi, se A' è un sottanello di $(A, +, \cdot)$, si verifica che $(A', +, \cdot)$ è una struttura di anello su A' .

Corpi: Siano $(A, +, \cdot)$ un anello e x un elemento di A . In generale x non ha un simmetrico rispetto alla moltiplicazione; ossia non esiste nessun y in A tale che: $x \cdot y = y \cdot x = 1$. Questo è senz'altro vero se $x = 0$ (cfr lemma 4) ma può succedere anche per elementi non nulli; per esempio nell'anello $(\mathbb{Z}, +, \cdot)$ gli unici elementi che abbiano un simmetrico rispetto a . sono 1 e -1.

7: Definizione: (*Corpi, campi*) Sia $(A, +, \cdot)$ un anello. Notiamo A^* l'insieme $A \setminus \{0\}$. Se (A^*, \cdot) è un gruppo si dice che $(A, +, \cdot)$ è una struttura di corpo su A . Un corpo commutativo (i.e. la moltiplicazione è commutativa) si chiama un campo (in altre parole A è un campo se e solo se (A^*, \cdot) è un gruppo abeliano).

7.1: Osservazione : Per riassumere $(k, +, \cdot)$ è un campo se e solo se:

(i) $(k, +)$ è un gruppo abeliano

(ii) (k^*, \cdot) è un gruppo abeliano ($k^* = k \setminus \{0\}$ dove 0 è il neutro di +)

(iii) la legge . è distributiva rispetto a +: $\forall (x,y,z) \in k^3: x \cdot (y+z) = x \cdot y + x \cdot z$, e $(x+y) \cdot z = x \cdot z + y \cdot z$.

Osservare che (ii) implica $0 \neq 1$.

7.2: Esempi : $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sono dei campi; $(\mathbb{Z}, +, \cdot)$ non è un campo. Un esempio più esotico di campo è $(\mathbb{Z}/2\mathbb{Z}, +, \cdot)$ (cfr Es. 4).

Morfismi: Siano $(A, +, \cdot)$, $(B, +, \cdot)$ due anelli (attenzione: prendiamo le stesse notazioni per le leggi in A e B ma ovviamente queste leggi sono diverse (sono definite su insiemi diversi). Si tratta solo di semplificare la scrittura, il contesto indicherà sempre in quale insieme stiamo operando). Tra le applicazioni da A in B ci sono quelle che "trasportano" la struttura di anello: sono i morfismi di anelli.

8: Definizione: Siano A, B due anelli e $f: A \rightarrow B$ un'applicazione, f è un morfismo di anelli se:

- (i) f è un morfismo di gruppi tra $(A, +)$ e $(B, +)$: $\forall (x,y) \in A^2: f(x+y) = f(x) + f(y)$
- (ii) $\forall (x,y) \in A^2: f(x \cdot y) = f(x) \cdot f(y)$
- (iii) $f(1_A) = 1_B$.

Le proprietà dei morfismi di anelli sono analoghe a quelle dei morfismi di gruppi. Per esempio se $f: A \rightarrow B$, e $g: B \rightarrow C$ sono due morfismi di anelli, allora $g \circ f: A \rightarrow C$ è un morfismo di anelli.

Se $f: A \rightarrow B$ è un morfismo di anelli, il nucleo di f, $\text{Ker}(f)$ è il nucleo del morfismo di gruppi tra i gruppi $(A, +)$, $(B, +)$: $\text{Ker}(f) = \{x \in A / f(x) = 0_B\}$. Un morfismo di anelli è iniettivo se e solo se $\text{Ker}(f) = \{0_A\}$.

Se A, B sono due corpi (risp. campi), un morfismo di anelli $f: A \rightarrow B$ si dice anche morfismo di corpi (risp. campi). Un morfismo non nullo di corpi è sempre iniettivo (cfr Es. 4).

ALCUNE REGOLE DI CALCOLO IN UN CAMPO:

Abbiamo già visto che in un anello $(A, +, \cdot)$: $x \cdot 0 = 0 \cdot x = 0$ per ogni x in A. Però non è sempre vero che: $x \cdot y = 0 \Rightarrow x = 0$ o $y = 0$, in particolare in un anello non si può sempre "semplificare" (cfr Es. 1, 2, 3). Invece in un campo si può semplificare e le regole di calcolo sono simili a quelle "usuali" (i.e. sui numeri reali). Nel seguito (cfr Spazi vettoriali) useremo liberamente le regole di calcolo in un campo.

9: Lemma: Siano k un campo, e x, y due elementi di k. Se $x \cdot y = 0$ allora $x = 0$ o $y = 0$.

Dim: Sia $x \cdot y = 0$. Se $x = 0$, non c'è niente da dimostrare. Supponiamo $x \neq 0$. Quindi $x \in k^*$, ma (k^*, \cdot) è un gruppo quindi x ha un simmetrico, x^{-1} , rispetto alla moltiplicazione. Componendo con x^{-1} viene: $x^{-1} \cdot (x \cdot y) = x^{-1} \cdot 0$. Abbiamo $x^{-1} \cdot 0 = 0$ (cfr lemma 4). D'altra parte $x^{-1} \cdot (x \cdot y) = (x^{-1} \cdot x) \cdot y = 1 \cdot y = y$ (per associatività e proprietà del simmetrico e del neutro). In conclusione $y = 0$ ♦

Se k è un campo, $(k,+)$ e (k^*, \cdot) sono dei gruppi abeliani e possiamo considerare i morfismi:

$f: \mathbf{Z} \rightarrow k: n \rightarrow nx, g: \mathbf{Z} \rightarrow k^*: n \rightarrow x^n$ (cfr I.6), con le solite conclusioni: $0x = 0, (-n)x = -(nx), x^0 = 1, x^{-n} = (x^n)^{-1}$, ecc...

COEFFICIENTI BINOMIALI, LA FORMULA DEL BINOMIO:

10: Definizione: Sia $n \in \mathbf{N}^*$, un numero naturale positivo. Il numero $1.2.3....n$ si chiama n fattoriale e si scrive $n!$

Per convenzione si pone $0! = 1$

10.1: Esempio : Abbiamo $2! = 1.2 = 2, 3! = 1.2.3 = 6, 4! = 1.2.3.4 = 24$, ecc...

11: Definizione: Siano p, n due numeri naturali con $0 \leq p \leq n$. Il numero

$$\frac{n!}{p!(n-p)!} \quad \text{si nota } [n; p] \quad (\text{o anche } \binom{n}{p}) \quad \text{e si legge "n su p".}$$

Che cosa rappresenta questo numero? osservare che a priori non è chiaro che sia ancora un numero naturale.

12: Teorema: Siano p, n come sopra e X un insieme con n elementi. Il numero di sottinsiemi di X aventi p elementi è $[n; p]$. In particolare $[n; p]$ è un numero naturale.

Dim: cfr Es. 6♦

I numeri $[n; p]$ vengono chiamati coefficienti binomiali e godono di alcune proprietà particolari:

13: Lemma: Siano $n \geq p$ due naturali:

- (i) $[n; p] = [n-1; p-1] + [n-1; p]$
- (ii) $[n; p] = [n; n-p]$.

La dimostrazione non presenta nessuna difficoltà ed è lasciata al lettore.

La relazione (i) del lemma 13 fornisce un modo pratico per calcolare di volta in volta $[n; p]$. Questo metodo viene chiamato a seconda del paese nel quale ci si trovi: triangolo di Tartaglia, Pascal, Newton, Leibniz, ... comunque sia, vediamo come funziona:

Per $n = 1$, $[n; p]$ si riduce a $[1; 0]$ e $[1; 1]$. Usando la convenzione $0! = 1$, otteniamo $[1; 0] = [1; 1] = 1$.

Per $n = 2$ dobbiamo calcolare $[2; p]$ per $0 \leq p \leq 2$. Abbiamo $[2; 0] = 1$ (più generalmente $[n; 0] = 1$ per ogni $n \geq 1$); poi, se $p \geq 1$, dal lemma 13: $[2; p] = [1; p-1] + [1; p]$ e il lavoro è già fatto:

n	0	1	2
p			
1	1	1	0
2	1	2	1

Per ottenere l'elemento all'incrocio della riga n e della colonna p basta sommare l'elemento immediatamente sopra nella tabella a quello che lo precede sulla stessa riga ($2 = 1+1$, $1 = 0+1$ fornisce la riga per $n = 2$). Quindi la tabella dei coefficienti binomiali si completa automaticamente. Per i primi valori di n viene:

n	0	1	2	3	4	5
p						
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	

Adesso possiamo enunciare la formula del binomio:

14: Proposizione: (*Formula del binomio*) Siano $(k, +, \cdot)$ un campo e x, y due elementi di k . Per ogni naturale $n \geq 1$: $(x+y)^n = \sum_{k=0}^n [n:k] x^{n-k} y^k$.

14.1: Osservazione : Si ricorda il significato del simbolo di sommatoria: $\sum_{k \in I} a_k$ (o $\sum_{k \geq x} a_k$, $\sum_{k=0}^n a_k$, ecc ...) significa che bisogna fare la somma di tutti gli a_i con l'indice variante nell'insieme I . Quando non si specifichi l'insieme I , si sottintende che l'indice sia un numero naturale. Per esempio: $\sum_{k=0}^n a_k = a_0 + a_1 + a_2 + a_3 + \dots + a_n$.

Dimostrazione della Prop.14: Si procede per induzione su n . Per $n = 1$, la formula è chiara. Supponiamo la formula vera per $n-1$ (ipotesi di induzione) e mostriamo che questo implica che la formula è vera per n .

Abbiamo $(x+y)^n = (x+y).(x+y)^{n-1} = x.(x+y)^{n-1} + y(x+y)^{n-1}$. Per ipotesi di induzione $(x+y)^{n-1} =$

$$\sum_{k=0}^{n-1} [n-1;k].x^{n-1-k}.y^k. \text{ Quindi } x.(x+y)^{n-1} = \sum_{k=0}^{n-1} [n-1;k].x^{n-k}.y^k (*), \text{ e } y.(x+y)^{n-1} =$$

$$\sum_{k=0}^{n-1} [n-1;k].x^{n-1-k}.y^{k+1} (**). \text{ Facendo la somma: } (x+y)^n = \sum_{k=0}^{n-1} [n-1;k].x^{n-k}.y^k +$$

$$\sum_{k=0}^{n-1} [n-1;k].x^{n-1-k}.y^{k+1}. \text{ Sia } r \text{ tale che } 0 < r < n. \text{ Il termine } x^{n-r}.y^r \text{ compare in}$$

(*) col coefficiente $[n-1;r]$, e in (**) col coefficiente $[n-1;r-1]$. Pertanto $x^{n-r}.y^r$ compare col coefficiente $[n-1;r] + [n-1;r-1] = [n;r]$ (cfr lemma 13) nell'espressione di $(x+y)^n$. Se $r=0$, x^n compare in (*) col coefficiente $[n-1;0] = 1 = [n;0]$. Se $r=n$, y^n compare in (**) col coefficiente $[n-1;n-1] = 1 = [n;n]$.

Concludiamo che $(x+y)^n = \sum_{k=0}^n [n:k] x^{n-k} y^k$, e la formula è dimostrata ♦

15: Esempi : Usando la tabella dei coefficienti binomiali si ottiene subito: $(x+y)^2 = x^2 + 2xy + y^2$; $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$, ecc ...

16: Osservazione : Con la stessa dimostrazione si ottiene l'enunciato seguente, più generale: *Siano $(A, +, \cdot)$ un anello e x, y due elementi di A che commutino tra di loro (i.e. $xy = yx$). Per ogni naturale $n \geq 1$: $(x + y)^n = \sum_{k=0}^n [n:k] x^{n-k} y^k$.*

Esercizi:

6.1 Un anello $(A, +, \cdot)$ si dice integro se: $\forall (x, y) \in A^2, x.y = 0 \Rightarrow x = 0, \text{ o } y = 0$. Dimostrare che un corpo (e a fortiori un campo) è un anello integro.

6.2 Sia A l'insieme di tutte le applicazioni di \mathbf{R} in \mathbf{R} . Su A definiamo due leggi $+$ e \cdot nel modo seguente:

Se f, g sono due elementi di A allora $f+g$ è l'applicazione: $f+g: \mathbf{R} \rightarrow \mathbf{R}: x \rightarrow f(x)+g(x)$, e $f.g$ è l'applicazione: $f.g: \mathbf{R} \rightarrow \mathbf{R}: x \rightarrow f(x).g(x)$.

(i) Dimostrare che $(A, +, \cdot)$ è un anello commutativo

(ii) Mostrare che $(A, +, \cdot)$ non è integro (e quindi non è un corpo).

(iii) Sono ancora veri (i) e (ii) con A uguale all'insieme delle applicazioni continue da \mathbf{R} in \mathbf{R} ?

6.3) Una matrice 2×2 reale (o a coefficienti reali), M , è una tabella: $\begin{pmatrix} ab \\ cd \end{pmatrix}$, con a, b, c, d , numeri reali. Si nota $M_2(\mathbb{R})$ l'insieme delle matrici reali 2×2 . Se $M = \begin{pmatrix} ab \\ cd \end{pmatrix}$, $N = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$, sono due elementi di $M_2(\mathbb{R})$, si definisce $M+N := \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}$ e $M.N := \begin{pmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{pmatrix}$.

(i) Dimostrare che $(M_2(\mathbb{R}), +, .)$ è un anello. E' commutativo?

(ii) Mostrare che $(M_2(\mathbb{R}), +, .)$ non è un anello integro.

6.4) Sia $n \in \mathbb{N}$, $n > 0$. Notiamo $\mathbb{Z}/n\mathbb{Z}$ l'insieme quoziente di \mathbb{Z} per la relazione: $x \equiv y \pmod{n}$ se e solo se $x-y$ è divisibile per n . Sia $p: \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ l'applicazione canonica che ad ogni x associa la sua classe d'equivalenza. Poniamo $\bar{x} = p(x)$. Su $\mathbb{Z}/n\mathbb{Z}$ sono state definite un'addizione e una moltiplicazione che noteremo $+, \cdot$ (cfr Es. 4.2). Inoltre $(\mathbb{Z}/n\mathbb{Z}, +)$ è un gruppo abeliano (Es. 5.5).

(i) Dimostrare che $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ è un anello commutativo.

(ii) Si tratta di determinare gli n tali che $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ sia un campo. Sia $\alpha \neq 0$ un elemento di $\mathbb{Z}/n\mathbb{Z}$. Definiamo $f_\alpha: \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}: \beta \mapsto \alpha\beta$. Mostrare che se f_α è iniettiva allora α è invertibile (usare Es. 3.5).

(iii) Sia $\alpha = \bar{x}$, $\beta = \bar{y}$. Osservare che si può assumere $1 \leq x \leq n-1$, $0 \leq y \leq n-1$. La relazione $\alpha\beta = 0$ è equivalente a: $n \mid xy$. Mostrare che, se n non è primo, esiste $\alpha \neq 0$ tale che f_α non sia iniettiva (usare Es. 1.5).

(iv) Concludere che $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ è un campo se e solo se n è primo (usare il lemma di Gauss, Es. 1.10).

6.5) Siano k, K due campi e $f: k \rightarrow K$ un morfismo di anelli.

(i) Dimostrare che f è iniettivo (si ricorda che in un anello A si ha sempre, per convenzione, $0_A \neq 1_A$).

(ii) Osservare che $f(k)$ con le restrizioni delle operazioni di K è un campo isomorfo a k . Quindi se esiste $f: k \rightarrow K$, morfismo di anelli tra i due campi k, K , possiamo identificare k a un sottocampo di K tramite l'applicazione iniettiva f . Per esempio \mathbb{Q} è un sottocampo di \mathbb{R} .

(iii) Dimostrare che non esiste nessun morfismo di anelli $f: \mathbb{Z}/2\mathbb{Z} \rightarrow \mathbb{R}$ tra i campi $\mathbb{Z}/2\mathbb{Z}$ e \mathbb{R} ($\mathbb{Z}/2\mathbb{Z}$ non è un sottocampo di \mathbb{R}).

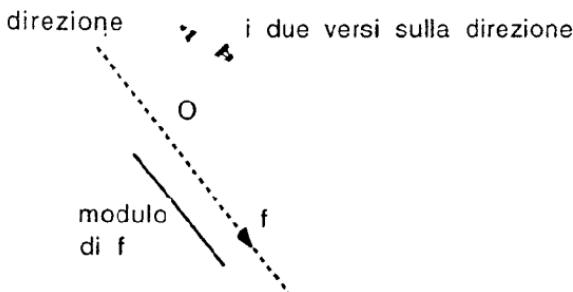
6.6) Dimostrare il teorema 12.

6.7) Sia k un campo. Si considera l'applicazione $f: \mathbb{Z} \rightarrow k: n \mapsto n \cdot 1_k$ ($n \cdot 1_k = 1_k + \dots + 1_k$). Mostrare che f è un morfismo d'anelli. Osservare che $\text{Ker}(f)$ è un sottogruppo di $(\mathbb{Z}, +)$ quindi della forma $d\mathbb{Z}$ (d il più piccolo elemento di $\text{Ker}(f)_+$, cfr §5, Es. 1.4); d si chiama la caratteristica del campo k e si indica $\text{ch}(k)$. Mostrare che se $\text{ch}(k) \neq 0$, allora $\text{ch}(k)$ è un numero primo (suggerimento: se $d = rs$ allora $0 = d \cdot 1 = (r \cdot 1)(s \cdot 1)$, ma k è un anello integro). Se p è un numero primo, qual è la caratteristica del campo $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$?

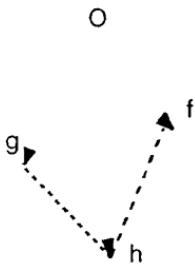
1) SPAZI VETTORIALI.

Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ una funzione derivabile. Localmente la funzione f è approssimata dalla sua derivata, ossia f è localmente "equivalente" alla sua retta tangente. Siccome le rette sono le curve più semplici, la derivata presenta una grossa semplificazione nello studio locale della funzione f . Questo è un procedimento generale in matematica: dato un problema che dipende da dati "complicati" si cerca di approssimarli con dati lineari per i quali la risoluzione sia più facile. Questo spiega l'importanza dell'algebra lineare che consiste, appunto, nello studio dei fenomeni lineari. L'algebra lineare (studio degli spazi vettoriali e delle applicazioni lineari) è onnipresente in matematica pura ed applicata. Come al solito sono più importanti le strutture degli oggetti e, dopo una motivazione "fisica" (forze applicate in un punto) dei vettori, seguiremo una presentazione più formale.

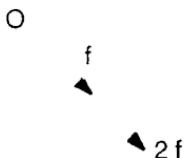
Sia O un punto del piano. Una forza applicata nel punto O si può rappresentare con un vettore, f . Per darsi il vettore f bisogna darsi: (i) una direzione (una retta passante per O) (ii) un verso (verso "destra" o verso "sinistra" rispetto ad O) (iii) una lunghezza (o modulo): la direzione indica la direzione in cui viene applicata la forza; il verso il verso in cui viene applicata (alto, basso; destra, sinistra) e il modulo indica l'intensità della forza.



Se applichiamo due forze f, g nel punto O , l'effetto ottenuto è quello della forza risultante h . La forza risultante, $h = f+g$, è rappresentata dal vettore h ottenuto tramite la "regola del parallelogramma" da f e g :



Se applichiamo in O la forza f ma con intensità doppia ($2f$), questa forza sarà rappresentata da un vettore, $2f$, con stessa direzione e verso di f ma con modulo doppio:



Nello stesso modo possiamo rappresentare λf per λ numero reale positivo qualsiasi.

Se applichiamo nel punto O una forza f e una forza, $-f$, avente la stessa direzione e lo stesso modulo, ma di verso opposto, allora, la forza risultante è nulla (regola del parallelogramma!):



In questo modo possiamo definire $-\lambda f$ dove $\lambda > 0$ è un numero reale: $-\lambda f$ è il vettore che ha stessa direzione e stesso modulo di λf ma verso opposto.

Abbiamo quindi definito l'addizione di due vettori applicati in O e la moltiplicazione di un vettore applicato in O con un numero reale λ . In entrambi i casi il risultato ottenuto è un vettore applicato in O ($f+g$, λf).

Identifichiamo il piano con \mathbb{R}^2 e O con il punto $(0, 0)$. Un vettore applicato in O, OP , è determinato dalla sua estremità $P = (x, y)$. Se $f = OP$, $g = OQ$, con $Q = (x', y')$, un po' di geometria elementare mostra che il vettore risultante $h =$

$f+g$ ha per estremità il punto $M = (x+x', y+y')$. Nello stesso modo il vettore λf ha per estremità il punto $(\lambda x, \lambda y)$. Questo suggerisce di introdurre due operazioni su \mathbb{R}^2 : la somma di due elementi di \mathbb{R}^2 è definita tramite: $(x, y) \oplus (x', y') := (x+x', y+y')$; la moltiplicazione di un elemento di \mathbb{R}^2 per un elemento di \mathbb{R} è definita tramite: $\lambda.(x, y) := (\lambda x, \lambda y)$. Abbiamo appena visto che queste operazioni sono "naturali". Osserviamo che l'addizione \oplus consiste nel fare la somma (in \mathbb{R}) componente per componente e nello stesso modo la moltiplicazione per λ consiste nel moltiplicare ogni coordinata per λ . E' chiaro inoltre che queste operazioni hanno tutte le buone proprietà che uno si aspetta (cfr (SV1), ..., (SV4) qui sotto). In particolare (\mathbb{R}^2, \oplus) è un gruppo abeliano (l'elemento neutro è il vettore nullo OO , cfr Es.I.5.4).

Queste considerazioni suggeriscono un'ulteriore generalizzazione: perché non fare la stessa cosa su \mathbb{R}^n ? L'addizione di due elementi $(x_1, \dots, x_n), (y_1, \dots, y_n)$ di \mathbb{R}^n è definita tramite $(x_1, \dots, x_n) \oplus (y_1, \dots, y_n) = (x_1+y_1, \dots, x_n+y_n)$ e la moltiplicazione tramite: $\lambda.(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$ (sempre coordinata per coordinata). Ancora una volta queste operazioni hanno tutte le buone proprietà. Osserviamo infine che la generalizzazione a \mathbb{R}^n non è una semplice questione formale: per esempio, in un esperimento pratico possiamo avere più di due o tre parametri (coordinate nello spazio, altitudine, temperatura, pressione, ecc...) e quindi voler rappresentare un dato sperimentale come un punto in uno spazio a n dimensioni.

Quanto fatto finora definisce la struttura "naturale" di spazio vettoriale reale su \mathbb{R}^n . In tutta generalità abbiamo:

1: Definizione: Siano (E, \oplus) un gruppo abeliano e $(k, +, \cdot)$ un campo. Se λ, μ , sono elementi di k , noteremo $\lambda\mu$ invece di $\lambda\cdot\mu$.

Sia $f: kxE \rightarrow E: (\lambda, v) \rightarrow f(\lambda, v) = \lambda \cdot v$, un'applicazione. Se le condizioni (SV1), ..., (SV4) qui sotto sono verificate si dice che $((E, \oplus), (k, +, \cdot), f)$ è una struttura di k -spazio vettoriale su E ; con abuso di linguaggio (e per alleggerire le notazioni!) si dice più semplicemente che E è un k -spazio vettoriale.

$$(SV1) \forall \lambda \in k, \forall (v, w) \in E^2, \lambda.(v \oplus w) = (\lambda \cdot v) \oplus (\lambda \cdot w)$$

$$(SV2) \forall (\lambda, \mu) \in k^2, \forall v \in E, (\lambda + \mu) \cdot v = (\lambda \cdot v) \oplus (\mu \cdot v)$$

$$(SV3) \forall (\lambda, \mu) \in k^2, \forall v \in E, (\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$$

$$(SV4) \forall v \in E, 1 \cdot v = v \text{ (qui } 1 = 1_k \text{ è il neutro per la moltiplicazione in } k\text{).}$$

2: Notazioni, convenzioni: Sia E un k -spazio vettoriale (più precisamente $((E, +), (k, +, \cdot), f)$ una struttura di k -spazio vettoriale su E). Di solito indicheremo gli elementi del campo k con delle lettere dell'alfabeto greco: α, β, \dots e gli elementi del gruppo abeliano $(E, +)$ con delle lettere del nostro alfabeto: x, y, z, v, w, \dots

Gli elementi di k vengono chiamati **scalari**, mentre quelli di E vengono chiamati **vettori**.

Come già detto scriveremo $\lambda\mu$ invece di $\lambda x\mu$ (moltiplicazione in k).

L'applicazione $f: kxE \rightarrow E: (\lambda, v) \rightarrow \lambda \cdot v$ è la **moltiplicazione esterna**.

D'ora in poi, scriveremo λv invece di $\lambda \cdot v$ per indicare la moltiplicazione esterna del vettore v per lo scalare λ .

Si avrà cura di non confondere $\lambda\mu$ con λv : il primo è il risultato della moltiplicazione interna su k , il secondo è il risultato della moltiplicazione esterna tra un elemento di k e uno di E : $\lambda\mu \in k$ mentre $\lambda v \in E$.

La legge interna, \oplus , su E verrà notata $+$. Si avrà cura di non confonderla con l'addizione in k : $\lambda+\mu \in k$ mentre $u+v \in E$.

Osservare che la scrittura $\lambda+v$ **non ha senso**: non sappiamo addizionare un vettore con uno scalare.

Sottrazione: Sia $(G, +)$ un gruppo in notazione addittiva. Abbiamo già visto che il simmetrico di un elemento, x , di G si indica $-x$. Invece di scrivere $x+(-y)$ (x composto col simmetrico di y), si scrive $x-y$, definendo così l'operazione di sottrazione in G .

Siccome $(k,+)$ e $(E,+)$ sono dei gruppi abeliani useremo le notazioni: $\lambda-\mu, v-w$.

Elementi neutri: Come nel caso generale, il neutro di $(k,+)$ viene indicato con 0_k , quello di (k^*, \cdot) , con 1_k , e quello di $(E,+)$ con 0_E . Vedremo più avanti che non c'è nessun inconveniente a tralasciare gli indici e scrivere $0, 1, 0$ al posto di $0_k, 1_k, 0_E$.

Notare che, se può esserci un'ambiguità sullo zero ($0 \in k$ o $0 \in E$?), è invece chiaro che $1 \in k$.

Finalmente, come già detto, invece di dire che $((E, +), (k, +, \cdot), f)$ è una struttura di k -spazio vettoriale su E , diremo che E è un k -spazio vettoriale. Bisogna però sempre tenere presente che su uno stesso insieme, E , possono esistere diverse strutture di gruppo abeliano e, a fortiori, diverse strutture di k -spazio vettoriale. Vedremo inoltre esempi di insiemi E con strutture di k -spazio vettoriale per diversi campi k (cfr 3 (ii)).

Le convenzioni qui sopra verranno introdotte gradualmente nel seguito.

3: Esempi :

(i). Un campo è uno spazio vettoriale su se stesso (k è un k -spazio vettoriale); poniamo $(E, +) = (k, +)$: in questo caso un vettore non è altro che uno scalare. Sia $f: kxE \rightarrow E: (\lambda, u) \rightarrow \lambda u$ (moltiplicazione in k : $u \in k$): quindi la moltiplicazione esterna coincide con la moltiplicazione di k . È facile verificare gli assiomi (SV1), ..., (SV4).

(ii). Un campo è uno spazio vettoriale su ogni suo sottocampo:

Sia K un campo e $k \subseteq K$ un sottocampo di K (cfr Es. I.6.4). Sia $(E, +) = (K, +)$ e consideriamo $f: kxE \rightarrow E: (\lambda, v) \rightarrow \lambda v$ (moltiplicazione in K : λ e v sono elementi di K). Si verifica facilmente che gli assiomi (SV1), ..., (SV4) sono soddisfatti.

Per esempio Q è un sottocampo di R e quindi R è un Q -spazio vettoriale. Abbiamo anche che R è un sottocampo di C e quindi C è un R -spazio vettoriale. Ma C è anche, per gli stessi motivi, un Q -spazio vettoriale. Abbiamo così almeno tre strutture di spazio vettoriale sull'insieme C (più precisamente sul gruppo abeliano $(C, +)$): C è un k -spazio vettoriale, $k = C, R, Q$. Vedremo (cfr 3, 3(c)) che queste strutture sono "diverse". Infine ci sono molti altri sottocampi di C oltre a Q e R .

(iii). Se k è un campo, k^n è un k -spazio vettoriale (detto n -spazio numerico su k):

Sia $E = k^n$ il prodotto cartesiano $kx...xk$ (n fattori). Su $k^n = \{(x_1, \dots, x_n) / x_i \in k, 1 \leq i \leq n\}$ definiamo una legge $+$ tramite: $(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1+y_1, \dots, x_n+y_n)$. Osservare che siccome k è un campo (in particolare $(k, +)$ è un gruppo abeliano), ogni coordinata $x_i + y_i$ è un elemento ben definito di k .

Poi definiamo la moltiplicazione esterna $f: kxE \rightarrow E: (\lambda, (x_1, \dots, x_n)) \rightarrow (\lambda x_1, \dots, \lambda x_n)$. Come prima, siccome k è un campo (in particolare un anello) ogni coordinata λx_i è un elemento ben definito di k .

Si verifica che queste operazioni conferiscono a k^n una struttura di k -spazio vettoriale (cfr Es.1).

(iv). Più generalmente se E_1, \dots, E_n sono dei k -spazi vettoriali, si può munire il prodotto cartesiano $E_1x...xE_n$ di una struttura ("prodotto") di k -spazio vettoriale. L'addizione è definita coordinata per coordinata: $(e_1, \dots, e_n) + (f_1, \dots, f_n) := (e_1+f_1, \dots, e_n+f_n)$.

$\dots, f_n := (e_1 + f_1, \dots, e_n + f_n)$. Questo ha senso perché, per ogni i , e_i e f_i sono elementi dello stesso spazio vettoriale E_i , e la somma $e_i + f_i$ è fatta in E_i (abbiamo quindi n operazioni di somma diverse, fatte in spazi diversi se $E_i \neq E_j$). La moltiplicazione esterna è definita, anche lei, coordinata per coordinata tramite: $\lambda(e_1, \dots, e_n) := (\lambda e_1, \dots, \lambda e_n)$; anche questa volta abbiamo operazioni fatte in spazi diversi: λe_i è la moltiplicazione esterna $k \times E_i \rightarrow E_i$. In particolare se $E_1 = E_2 = \dots = E_n = k$, ritroviamo l'esempio (iii).

(v). Ogni spazio vettoriale su K è uno spazio vettoriale su ogni sottocampo di K .

Siano E un K -spazio vettoriale e $k \subseteq K$ un sottocampo di K . Consideriamo $f: Exk \rightarrow E: (\lambda, v) \rightarrow \lambda v$ (si considera λ come un elemento di K). Allora $((E, +), k, f)$ è una struttura di k -spazio vettoriale su E ("ottenuta per restrizione degli scalari a k "). In particolare ogni C -spazio vettoriale può essere considerato come un R (risp. Q)-spazio vettoriale.

(vi). L'insieme delle applicazioni da un insieme E in un campo k è un k -spazio vettoriale.

Siano E un insieme e k un campo. Notiamo $A(E, k)$ l'insieme delle applicazioni da E in k : $A(E, k) := \{f / f: E \rightarrow k \text{ è un'applicazione}\}$. Per semplificare la scrittura noteremo A invece di $A(E, k)$.

Su A definiamo una legge $+$ nel modo seguente: se f, g sono elementi di A allora $(f+g)$ è l'applicazione: $(f+g): E \rightarrow k: x \rightarrow f(x)+g(x)$. Questo ha senso perché sappiamo addizionare i due elementi $f(x), g(x)$ di k (notare che $f+g$ è l'addizione in A mentre $f(x)+g(x)$ è l'addizione in k).

La moltiplicazione esterna su A è definita tramite: $\phi: k \times A \rightarrow A: (\lambda, f) \rightarrow \lambda f$ dove $\lambda f: A \rightarrow k$ è l'applicazione definita da: $(\lambda f)(x) = \lambda f(x)$ (sappiamo moltiplicare λ e $f(x)$ in k). Si verifica (cfr Es.1) che $((A, +), k, \phi)$ è una struttura di k -spazio vettoriale su A . Per esempio l'insieme delle applicazioni da R^n in R è un R -spazio vettoriale.

(vii). Polinomi a coefficienti in un campo:

Sia k un campo, un polinomio, $P(X)$, nella variabile (o indeterminata) X , a coefficienti in k è un'espressione della forma: $P(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n + \dots$, dove gli a_i sono elementi di k che sono tutti nulli tranne al più un numero finito. In modo più rigoroso un polinomio, P , in una variabile è una successione $(a_0, a_1, \dots, a_n, \dots)$ di elementi di k , tutti nulli, tranne al più un

numero finito. Usando il simbolo di sommatoria: $P(X) = \sum_{i \geq 0} a_i X^i$. Si ricorda che $\sum_{i \in I} b_i$ significa che bisogna fare la somma degli elementi b_i per ogni i appartenente a I . Quando I non è specificato si sottintende $I = \mathbb{N}$ (o \mathbb{Z}). Per esempio $\sum_{i=1}^n b_i = b_1 + b_2 + \dots + b_n$.

Sia $P(X) = \sum_{i \geq 0} a_i X^i$, gli a_i sono i coefficienti di P . Due polinomi sono uguali se e solo se hanno gli stessi coefficienti. Il polinomio nullo è l'unico polinomio i cui coefficienti siano tutti nulli, si nota 0. Se $P(X) = \sum_{i \geq 0} a_i X^i$ è un polinomio non nullo il suo grado è: $\text{gr}(P) = \max\{n / a_n \neq 0\}$; essendoci solo un numero finito di coefficienti non nulli, il grado è ben definito. Un polinomio di grado zero è una costante (i.e. un elemento di k). Il polinomio nullo non ha grado (o, per convenzione, ha grado $-\infty$).

L'insieme dei polinomi nella variabile X , a coefficienti in k , si nota $k[X]$.

Su $k[X]$ definiamo una legge $+$ tramite:

$$(\sum_{i \geq 0} a_i X^i) + (\sum_{i \geq 0} b_i X^i) = \sum_{i \geq 0} (a_i + b_i) X^i \quad (\text{addizione coordinata per coordinata per le successioni}).$$

Per esempio se $k = \mathbb{R}$ e se $P(X) = 2 + 3X + X^2 + 4X^4$, $Q(X) = 5X + 4X^2 + X^3$, allora $(P+Q)(X) = 2 + 8X + 5X^2 + X^3 + 4X^4$.

Si verifica che $(k[X], +)$ è un gruppo abeliano.

Sia poi $f: k \times k[X] \rightarrow k[X]: (\lambda, P(X)) \mapsto \lambda P(X)$, definita nel modo seguente: se $P(X) = \sum_{i \geq 0} a_i X^i$, allora $\lambda P(X) = \sum_{i \geq 0} (\lambda a_i) X^i$. Con queste due operazioni $K[X]$ è un k -spazio vettoriale (cfr Es.1).

(viii). Matrici (m,n) a coefficienti in un campo:

Una matrice, M , (m,n) a coefficienti in un campo k è una tabella con m righe e n colonne di elementi di k :

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{con } a_{ij} \in k, 1 \leq i \leq m, 1 \leq j \leq n.$$

Gli a_{ij} sono i coefficienti della matrice M . Notare che la scrittura a_{ij} indica l'elemento che sta all'incrocio della i -esima riga con la j -esima colonna; per convenzione scriveremo sempre per primo l'indice della riga.

In forma più condensata la matrice M si può scrivere $M = (a_{ij})$, $1 \leq i \leq m$, $1 \leq j \leq n$.

La i -esima riga si può identificare alla matrice $(1, n)$: (a_{i1}, \dots, a_{in}) , la j -esima

colonna alla matrice $(m, 1)$: $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

L'insieme delle matrici (m, n) a coefficienti in k si nota $M_{m,n}(k)$ (prima l'indice delle righe!).

Definiamo una legge $+$ su $M_{m,n}(k)$ nel modo seguente:

se $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix}$, sono due matrici (m, n) a

coefficienti in k , la matrice somma $A+B$ è la matrice (m, n) a coefficienti in

k : $A+B = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$. Si verifica (cfr anche Oss. 3.1) che

$(M_{m,n}(k), +)$ è un gruppo abeliano (chi è l'elemento neutro?). Poi si definisce la moltiplicazione esterna tra una matrice e un elemento di k . Se $A =$

$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, e se $\lambda \in k$, la matrice λA è, per definizione, la matrice

$\begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}$.

Con queste operazioni si verifica (cfr anche Oss. 3.1) che $M_{m,n}(k)$ è un k -spazio vettoriale.

3.1: *Osservazione* : Una matrice (m, n) a coefficienti in k , $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$,

dopo la scelta di un ordine nel quale scrivere gli $m \cdot n$ coefficienti di A , non è nient'altro che un elemento di k^{mn} , per esempio: $A = (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn})$. Dopo questa scelta dell'ordine nel quale scrivere i coefficienti, possiamo definire un'applicazione: $\varphi: M_{m,n}(k) \rightarrow k^{mn}: A \rightarrow (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn})$. E' chiaro che φ è biiettiva. Inoltre si ha anche $\varphi(A+B) = \varphi(A)+\varphi(B)$ dove $\varphi(A)+\varphi(B)$ è la somma in k^{mn} definita in 3 (iii). Si ha pure

$\varphi(\lambda A) = \lambda\varphi(A)$, dove $\lambda\varphi(A)$ è la moltiplicazione esterna definita in 3 (iii). Anche l'applicazione reciproca:

$$\varphi^{-1}: k^{mn} \rightarrow M_{m,n}(k): \underline{x} = (x_{11}, \dots, x_{1n}, \dots, x_{m1}, \dots, x_{mn}) \rightarrow \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \dots & x_{mn} \end{pmatrix}, \text{ verifica}$$

$\varphi^{-1}(\underline{x} + \underline{y}) = \varphi^{-1}(\underline{x}) + \varphi^{-1}(\underline{y})$ e $\varphi^{-1}(\lambda \underline{x}) = \lambda \varphi^{-1}(\underline{x})$. Quindi vediamo che le strutture di k -spazio vettoriale su $M_{m,n}(k)$ e su k^{mn} sono in perfetta corrispondenza tramite φ ; φ ci permette di identificare queste due strutture di k -spazio vettoriale. Vedremo più avanti che φ è un isomorfismo lineare tra i k -spazi vettoriali $M_{m,n}(k)$ e k^{mn} e pertanto questi due k -spazi vettoriali sono isomorfi (i.e. possono essere identificati).

4) Alcune regole di calcolo in uno spazio vettoriale:

Sia $((E, \oplus), (k, +, .), f)$ un k -spazio vettoriale. Ricordiamo le convenzioni introdotte in 2: i vettori (i.e. gli elementi di E) sono rappresentati con lettere del nostro alfabeto: $x, y, z, t, u, v, w, \dots$ mentre gli scalari (i.e. gli elementi di k) sono rappresentati con lettere greche. L'operazione \oplus viene notata $+$; la sottrazione viene notata $-$ (anche in k), e la moltiplicazione esterna λu ($\lambda \in k, u \in E$).

$$4.1: \forall \lambda \in k, \forall (u, v) \in E^2: \lambda(u - v) = \lambda u - \lambda v.$$

$$\begin{aligned} \text{Infatti: } \lambda(u-v) + \lambda v &= \lambda[(u-v) + v] \quad (\text{cfr SV1}) \\ &= \lambda[u+(-v+v)] \quad (\text{associatività in } (E, +)) \\ &= \lambda[u+0_E] \quad (\text{proprietà del simmetrico}) \\ &= \lambda[u] \quad (\text{proprietà dell'elemento neutro}) \end{aligned}$$

In conclusione: $\lambda(u-v) + \lambda v = \lambda u$.

Componendo a destra col simmetrico di λv in E :

$$[\lambda(u-v) + \lambda v] - \lambda v = \lambda u - \lambda v.$$

Per associatività: $[\lambda(u-v)+\lambda v] - \lambda v = \lambda(u-v) + [\lambda v - \lambda v] = \lambda(u-v)$.

Finalmente, mettendo tutto insieme: $\lambda(u-v) = \lambda u - \lambda v$.

$$4.2: \forall \lambda \in k: \lambda 0_E = 0_E.$$

Poniamo $u = v$ in 4.1: $\lambda(u-u) = \lambda u - \lambda u = 0_E$. D'altra parte $u-u = 0_E$.

4.3: $\forall \lambda \in \mathbf{k}, \forall \mathbf{v} \in E: \lambda(-\mathbf{v}) = -(\lambda\mathbf{v}).$

Poniamo $\mathbf{u} = 0_E$ in 4.1: $\lambda(0_E - \mathbf{v}) = \lambda(-\mathbf{v})$ (proprietà dell'elemento neutro). Quindi da 4.1: $\lambda(-\mathbf{v}) = \lambda(0_E - \mathbf{v}) = \lambda 0_E - \lambda \mathbf{v}$, ma da 4.2 otteniamo $\lambda 0_E = 0_E$, quindi $\lambda(-\mathbf{v}) = -\lambda \mathbf{v}$.

4.4: $\forall (\lambda, \mu) \in \mathbf{k}^2, \forall \mathbf{v} \in E: (\lambda\mathbf{v}) + (\mu\mathbf{v}) = (\lambda + \mu)\mathbf{v}.$

$$\begin{aligned} \text{Infatti: } (\lambda\mathbf{v}) + (\mu\mathbf{v}) &= [(\lambda + \mu)\mathbf{v}], \text{ (cfr SV.2)} \\ &= [\lambda + (\mu + \lambda)\mathbf{v}], \text{ (associatività in } (\mathbf{k}, +)) \\ &= [\lambda + 0_K]\mathbf{v} \\ &= \lambda\mathbf{v} \end{aligned}$$

Quindi $(\lambda\mathbf{v}) + (\mu\mathbf{v}) = \lambda\mathbf{v}$. Si prova l'asserto componendo con il simmetrico di $\mu\mathbf{v}$.

4.5: $\forall \mathbf{v} \in E: 0_{\mathbf{k}} \cdot \mathbf{v} = 0_E.$

Facciamo $\lambda = \mu$ in 4.4: $(\lambda - \lambda)\mathbf{v} = \lambda\mathbf{v} - \lambda\mathbf{v}$; ossia $0_K \cdot \mathbf{v} = 0_E$.

4.6: $\forall \lambda \in \mathbf{k}, \forall \mathbf{v} \in E: \lambda(-\mathbf{v}) = -(\lambda\mathbf{v}) = (\lambda\mathbf{v}).$

La prima uguaglianza è stata dimostrata in 4.3. Abbiamo $(-\lambda + \lambda)\mathbf{v} = (-\lambda)\mathbf{v} + \lambda\mathbf{v}$ (SV2). Ma $(-\lambda + \lambda)\mathbf{v} = 0_K \mathbf{v} = 0_E$. Quindi $0_E = (-\lambda)\mathbf{v} + \lambda\mathbf{v} = \lambda\mathbf{v} + (-\lambda)\mathbf{v}$ ($(E, +)$ abeliano). Pertanto $(-\lambda)\mathbf{v}$ è il simmetrico di $\lambda\mathbf{v}$, ossia $(-\lambda)\mathbf{v} = -(\lambda\mathbf{v})$, e l'ultima uguaglianza è dimostrata.

4.7: Osservazione : Da 4.2, 4.5 si vede che non ci sono inconvenienti a notare con 0 sia 0_K che 0_E . Inoltre, tenuto conto degli assiomi e di 4.1, ..., 4.6, il calcolo in espressioni contenenti scalari e vettori si effettua secondo le regole usuali (per esempio del calcolo con i numeri reali), purché non si addizionino un vettore e uno scalare e non si moltiplichino due vettori tra di loro (queste operazioni non sono definite); per il resto si possono assimilare scalari e vettori a numeri reali ed usare le solite regole ($(\mathbf{R}, +, \cdot)$ è un R-spazio vettoriale! cfr 3(i)).

4.8: $\forall \lambda \in \mathbf{k}, \forall \mathbf{v} \in E: \lambda\mathbf{v} = 0 \text{ se e solo se } \lambda = 0 \text{ o } \mathbf{v} = 0.$

Sia $\lambda\mathbf{v} = 0$ con $\lambda \neq 0$. Siccome $\lambda \neq 0$, esiste λ^{-1} in \mathbf{k} tale che $\lambda^{-1}\lambda = 1_K$. Abbiamo:

$$\begin{aligned}\lambda^{-1}(\lambda v) &= (\lambda^{-1}\lambda)v \text{ (SV3)} \\ &= 1_k \cdot v = v \text{ (SV4)}\end{aligned}$$

D'altra parte: $\lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$. In conclusione $v = 0$. Questo dimostra 4.8♦

5: Sottospazi vettoriali.

Sia E un k -spazio vettoriale e $F \subseteq E$ un sottinsieme. Diremo che F è un sotto k-spazio vettoriale di E se le leggi di E ristrette a F definiscono una struttura di k -spazio vettoriale su F .

Questo significa che: (i) F è un sottogruppo di $(E, +)$, (ii) la restrizione della moltiplicazione esterna $f_1: F \times k \rightarrow F: (u, \lambda) \mapsto \lambda u$ ha la sua immagine contenuta in F ; ossia se $u \in F$ e $\lambda \in k$ allora $\lambda u \in F$.

Se le condizioni (i), (ii) sono soddisfatte allora è facile verificare che $((F, +), k, f_1)$ è una struttura di k -spazio vettoriale su F .

Questa struttura è indotta da quella di E e quindi si dice che F è un sotto k -spazio vettoriale di E (o sottospazio vettoriale se il campo k è fissato dal contesto).

Abbiamo già visto (I, §5, Prop. 9) che la condizione (i) è equivalente a:

(i) F è non vuoto e: per ogni u in F e per ogni v in F , $u - v$ appartiene a F .

Invece la condizione (ii) è equivalente a:

(ii)' $\forall u \in F, \forall \lambda \in k: \lambda u \in F$.

Possiamo riassumere quanto precede nella seguente:

5.1: Proposizione: *Sia E un k -spazio vettoriale e $F \subseteq E$ un sottinsieme. Sono equivalenti:*

(a) F è un sottospazio vettoriale di E .

(b) $0 \in F$ e: $\forall (\lambda, \mu) \in k^2, \forall (u, v) \in F^2, \lambda u + \mu v \in F$.

Dim: (a) \Rightarrow (b): chiaro.

(b) \Rightarrow (a): se $0 \in F$ allora $F \neq \emptyset$. Ponendo $\lambda = 1$ e $\mu = -1$ in (b) si ottiene (i)'.

Ponendo $\mu = 0$ in (b) si ottiene (ii)'♦

5.2: Esempio : Sia $E = \mathbb{R}^n$ con la sua struttura naturale di \mathbb{R} -spazio vettoriale (3 (iii)). Sia $F = \{(x_1, \dots, x_n) \in \mathbb{R}^n / x_n = 0\}$. Mostriamo che F è un \mathbb{R} -sottospazio vettoriale di E . Abbiamo $0 = (0, \dots, 0) \in F$. La prima condizione di 5.1 (b) è verificata. Siano $u = (u_1, \dots, u_{n-1}, 0)$ e $v = (v_1, \dots, v_{n-1}, 0)$ due elementi di F . Se λ, μ sono dei numeri reali, $\lambda u + \mu v = (\lambda u_1 + \mu v_1, \dots, \lambda u_{n-1} + \mu v_{n-1}, 0)$ appartiene

ancora a F. La seconda condizione di 5.1 (b) è verificata e pertanto F è un sottospazio vettoriale di E.

6: Combinazioni lineari:

6.1: Definizione: Sia E un k-spazio vettoriale e A un sottinsieme di E. Una combinazione lineare di elementi di A è una somma finita:

$\lambda_1x_1 + \dots + \lambda_nx_n$ dove $\lambda_1, \dots, \lambda_n$ sono elementi di k e dove x_1, \dots, x_n appartengono ad A.

6.1.1: Osservazione : Sia $v \in A$ allora $v = 1.v$, quindi v appartiene all'insieme delle combinazioni lineari di elementi di A.

6.2: Proposizione: Sia E un k-spazio vettoriale e $A \subseteq E$ un sottinsieme non vuoto. Notiamo F l'insieme delle combinazioni lineari di elementi di A: $F = \{w \in E / w = \alpha_1a_1 + \dots + \alpha_ka_k, \alpha_1, \dots, \alpha_k \text{ in } k \text{ e } a_1, \dots, a_k \text{ in } A\}$. Allora:

(i) F è un sottospazio vettoriale di E.

(ii) se F' è un sottospazio vettoriale di E contenente A allora F' contiene F. In altre parole F è il più piccolo sottospazio vettoriale di E contenente l'insieme A.

Dim: (i) Secondo 5.1 basta dimostrare: (α) $0 \in F$, (β) $\forall (v, w) \in F^2, \forall (\lambda, \mu) \in k^2, \lambda v + \mu w \in F$.

(α) basta osservare che preso $x \in A$ (x esiste perché A è non vuoto per ipotesi), $0.x = 0$ appartiene a F.

(β) siano $u = \alpha_1x_1 + \dots + \alpha_nx_n, v = \beta_1y_1 + \dots + \beta_ky_k$ due combinazioni lineari di elementi in A (quindi $x_i \in A, 1 \leq i \leq n$, e $y_j \in A, 1 \leq j \leq k$). Abbiamo $\lambda u + \mu v = \lambda(\alpha_1x_1 + \dots + \alpha_nx_n) + \mu(\beta_1y_1 + \dots + \beta_ky_k) = (\lambda\alpha_1)x_1 + \dots + (\lambda\alpha_n)x_n + (\mu\beta_1)y_1 + \dots + (\mu\beta_k)y_k$. Siccome $\lambda\alpha_1, \dots, \lambda\alpha_n, \mu\beta_1, \dots, \mu\beta_k$, sono scalari e $x_1, \dots, x_n, y_1, \dots, y_k$, sono elementi di A, $\lambda u + \mu v$ è una combinazione lineare di elementi di A. Quindi F è un sottospazio vettoriale di E.

(ii) Sia $v = \lambda_1x_1 + \dots + \lambda_nx_n$ un elemento di F (quindi $x_i \in A, 1 \leq i \leq n$). Siccome per ipotesi F' contiene A, $x_1 \in F', \dots, x_n \in F'$. Pertanto $\lambda_i x_i \in F'$ perché F' è un sottospazio vettoriale (cfr (ii)', 5.1). Per finire, siccome F' è un sottospazio vettoriale, $(F', +)$ è un sottogruppo di $(E, +)$, e $\lambda_1x_1 + \dots + \lambda_nx_n \in F'$. Quindi ogni elemento, v, di F appartiene a F'♦

6.3: Definizione: Con le notazioni di 6.2 si dice che F è il sottospazio vettoriale di E generato da A .

Se $A = \{e_1, \dots, e_k\}$ allora si nota $F = \langle e_1, \dots, e_k \rangle$ e si dice che e_1, \dots, e_k , sono generatori di F .

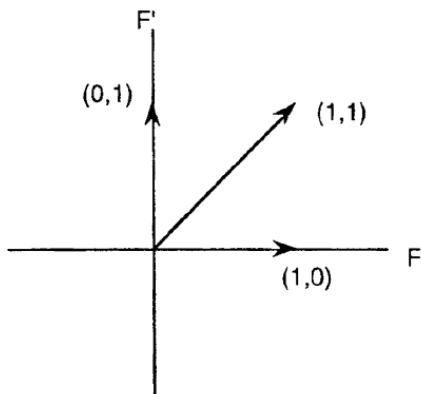
6.3.1: Osservazione : Se $F = \langle e_1, \dots, e_k \rangle$ allora ogni elemento di F si può scrivere nella forma $\lambda_1 e_1 + \dots + \lambda_k e_k$ (con eventualmente alcuni λ_i uguali a zero): ogni elemento di F è combinazione lineare dei generatori e_1, \dots, e_k .

7: Unione, somma, intersezione.

Sia E un k -spazio vettoriale, F, F' due sottospazi vettoriali di E . In generale $F \cup F'$ **non** è un sottospazio vettoriale di E (cfr Es.5).

7.1: Esempio : Sia $E = \mathbb{R}^2$, $F' = \{(x,y) \in \mathbb{R}^2 / x = 0\}$, $F = \{(x,y) \in \mathbb{R}^2 / y = 0\}$. Si verifica facilmente, con 5.1 (cfr 5.2), che F e F' sono sottospazi vettoriali di E . Abbiamo, per esempio, $u = (0, 1) \in F'$, e $v = (1, 0) \in F$. Quindi, a fortiori, u e v appartengono a $F \cup F'$. Se $F \cup F'$ fosse un sottospazio vettoriale dovremmo avere $u+v = (1, 1) \in F \cup F'$. Ossia $(1, 1) \in F$, o $(1, 1) \in F'$; ma chiaramente nessuna di queste due possibilità è verificata. Quindi $F \cup F'$ non è un sottospazio vettoriale di E .

Facendo un disegno si capisce facilmente quello che succede:



Possiamo, però, considerare il sottospazio vettoriale generato dal sottinsieme $F \cup F'$, ossia il più piccolo sottospazio vettoriale contenente $F \cup F'$:

7.2: Definizione: Il sottospazio vettoriale di E generato da $A = F \cup F'$ si nota $F+F'$ e si chiama la somma di F e F' . Si ha $F+F' = \{v+v' / v \in F \text{ e } v' \in F'\}$.

7.2.1: Osservazione : A priori (cfr 6.3) $F+F'$ è l'insieme delle combinazioni lineari: $\lambda_1x_1+\dots+\lambda_nx_n, x_i \in F \cup F', 1 \leq i \leq n$. Siccome $x_i \in F \cup F' \Leftrightarrow x_i \in F \text{ o } x_i \in F'$, riordinando semmai gli indici, possiamo supporre x_1, \dots, x_k appartenenti a F , e x_{k+1}, \dots, x_n appartenenti a F' . Quindi $\lambda_1x_1+\dots+\lambda_kx_k \in F$, e $\lambda_{k+1}x_{k+1}+\dots+\lambda_nx_n \in F'$ (perché F e F' sono dei sottospazi vettoriali). In definitiva la combinazione lineare iniziale è del tipo $v+v'$, con $v \in F$ e $v' \in F'$. Viceversa è chiaro che $v+v'$, con $v \in F$, $v' \in F'$, è una combinazione lineare di elementi di $F \cup F'$.

Contrariamente all'unione, l'intersezione di due sottospazi vettoriali è ancora un sottospazio vettoriale, più generalmente l'intersezione di un numero qualsiasi (anche infinito) di sottospazi vettoriali è un sottospazio vettoriale:

7.3: Proposizione: Sia E un k -spazio vettoriale e $(F_i)_{i \in I}$ una famiglia di sottospazi vettoriali di E . Allora $\bigcap_{i \in I} F_i$ è un sottospazio vettoriale di E .

Dim: Verifichiamo (b) di 5.1. Siccome per ogni i in I , F_i è un sottospazio vettoriale di E , abbiamo $0 \in F_i$ per ogni i in I . Quindi per definizione (cfr I, §3, 17) $0 \in \bigcap_{i \in I} F_i$.

Siano λ, μ , degli scalari e u, v , degli elementi di $\bigcap_{i \in I} F_i$. Allora per ogni i , u e v appartengono a F_i . Siccome, per ogni i , F_i è un sottospazio vettoriale: $\lambda u + \mu v$ appartiene a F_i per ogni i . Quindi (cfr I, §3, 17) $\lambda u + \mu v \in \bigcap_{i \in I} F_i$ ♦

8: Somma diretta:

Abbiamo visto (cfr 7.1) che in generale l'unione di due sottospazi vettoriali non è un sottospazio vettoriale. Possiamo però considerare il sottospazio vettoriale, $F+F'$, generato da $F \cup F'$ (cfr 7.2). Il sottospazio $F+F'$ si chiama il sottospazio somma di F e F' . Vogliamo adesso precisare questa nozione di somma di due (o più) sottospazi.

8.1: Definizione: Siano E un k -spazio vettoriale, F e H due sottospazi vettoriali di E . Se $F \cap H = \{0\}$ si dice che la somma $F+H$ è diretta e si scrive $F \oplus H$ invece di $F+H$.

La proposizione seguente spiega perché sia utile distinguere tra somma e somma diretta (se F e H sono in somma diretta, ogni elemento di $F+H$ si scrive in modo unico come somma di un elemento di F e di un elemento di H):

8.2: Proposizione: Siano E un k -spazio vettoriale, e F, H due sottospazi vettoriali di E . Sono equivalenti:

- (i) $F \cap H = \{0\}$
- (ii) ogni vettore di $F+H$ si scrive in modo unico come somma di un vettore di F e di un vettore di H
- (iii) se $x \in F$ e $y \in H$ allora: $x + y = 0 \Leftrightarrow x = 0$ e $y = 0$.

Dim: (i) \Rightarrow (ii) supponiamo $v = x + y = x' + y'$, con x, x' in F e y, y' in H . Allora $x - x' = y' - y$. Il vettore $w = x - x'$ appartiene a F (perché x e x' appartengono a F , e F è un sottospazio vettoriale); nello stesso modo $w \in H$ (perché $w = y' - y$). Da (i) segue $w = 0$, ossia $x = x'$ e $y = y'$.

(ii) \Rightarrow (iii) se $x + y = 0$ allora $x + y = 0 + 0$ con $0 \in H$ e $0 \in F$. Da (ii) segue $x = 0$ e $y = 0$.

(iii) \Rightarrow (i) sia $v \in F \cap H$. Abbiamo $v \in F$ e $v \in H$. In particolare $-v \in H$. Ma $(-v) + v = 0$ con $v \in F$ e $-v \in H$; da (iii) segue $v = 0$ ♦

8.3: Definizione: Siano E un k -spazio vettoriale, e F, H , due sottospazi vettoriali di E . Se $E = F \oplus H$, i due sottospazi vettoriali F, H , si dicono supplementari (o ancora F (risp. H) è un supplementare di H (risp. F)).

8.3.1: Osservazione : Da 8.2 segue che $E = F \oplus H$ se e solo se ogni vettore di E si scrive in modo unico come somma di un vettore di F e di un vettore di H .

8.3.2: Esempio : Sia $E = \mathbb{R}^2$, E è un \mathbb{R} -spazio vettoriale. Il sottospazio vettoriale, F , generato dal vettore $u = (1, 0)$ è $F = \{\lambda(1, 0) / \lambda \in \mathbb{R}\} = \{(\lambda, 0) / \lambda \in \mathbb{R}\}$. Graficamente F si può rappresentare come l'asse degli x nel piano reale cartesiano xOy .

Mostriamo che se $v \in \mathbb{R}^2$ è un vettore non appartenente a F (quindi $v = (a, b)$ con $b \neq 0$), allora $\mathbb{R}^2 = F \oplus H$ dove $H = \langle v \rangle$ è il sottospazio vettoriale generato da v .

Vediamo prima che la somma $F+H$ è diretta. Da 8.2 (i) basta far vedere che $F \cap H = \{0\}$. Se $w \in F \cap H$ allora $w = (\lambda, 0)$ e $w = \mu(a, b) = (\mu a, \mu b)$. Segue che $\mu b = 0$. Siccome $b \neq 0$ per ipotesi, questo implica $\mu = 0$. Quindi $w = 0(a, b) = (0, 0) = 0$ (il vettore nullo).

Abbiamo quindi $F+H = F \oplus H$. Per concludere mostriamo che $F+H = \mathbb{R}^2$, ossia che ogni vettore di \mathbb{R}^2 si può scrivere come somma di un vettore di F con un vettore di H . Sia (x, y) un vettore qualsiasi di \mathbb{R}^2 . Allora $(x, y) = (x - ay/b, 0) + (ay/b, y)$ (*). Il vettore $(x - ay/b, 0)$ appartiene a F perché la sua seconda coordinata è nulla; inoltre $(ay/b, y) = y/b \cdot (a, b)$ quindi il vettore $(ay/b, y)$ appartiene a H . Infine osserviamo che la scrittura precedente è lecita perché $b \neq 0$ per ipotesi.

Questo dimostra che $\mathbb{R}^2 = F + H = F \oplus H$.

Trovare la decomposizione (*) è un semplice problema di geometria analitica nel piano. Supponiamo $a \neq 0$. Dato il punto $P = (x, y)$ del piano proiettandolo su F parallelamente ad H si ottiene il punto $f = (x - ay/b, 0)$ di F ; proiettando P su H parallelamente ad F si ottiene il punto $h = (ay/b, y)$ (fare un disegno). Infatti la retta, H , passante per l'origine e (a, b) ha equazione $Y = bX/a$. La retta, H' , parallela ad H e passante per (x, y) ha equazione $Y = bX/a + y - bx/a$. Il punto f è dato dall'intersezione delle rette H' e F quindi dal sistema: $Y = 0$ e $Y = bX/a + y - bx/a$. Si ricava $X = x - ay/b$, $Y = 0$, che sono proprio le coordinate di f . Il punto h è dato dall'intersezione della retta H con la retta F parallela ad F e passante per (x, y) . L'equazione di F è $Y = y$ da cui il sistema $Y = 0$ e $Y = bX/a$. Si ricava $X = ay/b$, $Y = y$. Finalmente se $a = 0$ si ha chiaramente $(x, y) = (x, 0) + (0, y)$.

Da quanto precede (cfr 8.2 (ii)) la relazione (*) fornisce l'unico modo in cui si può scrivere (x, y) come somma di un elemento di F e di un elemento di H .

Il sottospazio vettoriale H è dunque un supplementare di F . Risulta in particolare che il sottospazio vettoriale F di \mathbb{R}^2 ammette un'infinità di supplementari. Questo non è più vero se si lavora su un campo finito (cfr Es.6).

La nozione di somma diretta si può estendere al caso di $n > 2$ sottospazi vettoriali.

8.4: Proposizione: Siano E un k -spazio vettoriale, e F_1, \dots, F_n dei sottospazi vettoriali di E . Sono equivalenti:

$$(i) F_i \cap (\sum_{j \neq i} F_j) = \{0\} \text{ per ogni } i, 1 \leq i \leq n.$$

(ii) ogni vettore x di $F_1 + \dots + F_n$ si scrive in modo unico nella forma $x = x_1 + \dots + x_n$, con $x_i \in F_i, 1 \leq i \leq n$.

(iii) se $x_i \in F_i, 1 \leq i \leq n$, allora: $x_1 + \dots + x_n = 0 \Leftrightarrow x_i = 0$, per ogni $i, 1 \leq i \leq n$.

Dim: (i) \Rightarrow (ii) Sia $x \in F_1 + \dots + F_n$ e supponiamo $x = x_1 + \dots + x_n = y_1 + \dots + y_n$, con $x_i \in F_i, y_i \in F_i, 1 \leq i \leq n$. Abbiamo $x_i - y_i = (y_1 - x_1) + \dots + (y_{i-1} - x_{i-1}) + (y_{i+1} - x_{i+1}) + \dots + (y_n - x_n)$. Il membro di sinistra è un vettore di F_i mentre quello di destra è un vettore di $F_1 + \dots + F_{i-1} + F_{i+1} + \dots + F_n = \sum_{j \neq i} F_j$. Dall'ipotesi (i) segue

che $(y_1 - x_1) + \dots + (y_{i-1} - x_{i-1}) + (y_{i+1} - x_{i+1}) + \dots + (y_n - x_n) = 0$ e $x_i - y_i = 0$. In particolare $x_i = y_i$. Siccome questo vale per ogni $i, 1 \leq i \leq n$, (ii) è dimostrato.

(ii) \Rightarrow (iii) Sia $x_1 + \dots + x_n = 0$, con $x_i \in F_i, 1 \leq i \leq n$. Abbiamo $x_1 + \dots + x_n = 0 = 0 + \dots + 0$. Da (ii) segue che $x_i = 0$ per ogni i .

(iii) \Rightarrow (i) Sia $z \in F_i \cap (\sum_{j \neq i} F_j)$. Allora $z = x_i \in F_i$ e $z = x_1 + \dots + x_{i-1} + x_{i+1} + \dots + x_n$

$\in \sum_{j \neq i} F_j$. Abbiamo $(-z) + z = 0$. Ossia $(-x_i) + x_1 + \dots + x_{i-1} + x_{i+1} + \dots + x_n = 0$, o ancora

$x_1 + \dots + x_{i-1} + (-x_i) + x_{i+1} + \dots + x_n = 0$. Da (iii): $x_1 = 0, \dots, x_{i-1} = 0, -x_i = 0, x_{i+1} = 0, \dots, x_n = 0$. Quindi $x_i = z = 0 \blacklozenge$

8.4.1: Osservazione : La condizione (i) di 8.4 è più forte della richiesta $F_i \cap F_j = \{0\}$ se $i \neq j$ (cfr Es.7).

8.5: Definizione: Se F_1, \dots, F_n soddisfano le condizioni equivalenti della Proposizione 8.4 si dice che F_1, \dots, F_n sono in somma diretta e si nota $F_1 \oplus \dots \oplus F_n$ la somma $F_1 + \dots + F_n$.

Esercizi:

1.1) Fare tutte le verifiche negli esempi 3 (i), ..., (viii).

1.2) Dire se E è un sottospazio vettoriale di \mathbb{R}^n :

(a) $E = \{(x, y, z) \in \mathbb{R}^3 / x + 2y - 3z = 0\}$

- (b) $E = \{(x, y, z) \in \mathbf{R}^3 / x + 2y - 3z = 0, \text{ e } 2x - y + 3z = 0\}$
 (c) $E = \{(x, y, z) \in \mathbf{R}^3 / x + 2y - 3z = 1\}$
 (d) $E = \{(x, y) \in \mathbf{R}^2 / x^2 - y = 1\}$
 (e) $E = \{(x, y) \in \mathbf{R}^2 / x \geq 0\}$.

Rappresentare graficamente gli insiemi E.

1.3) Dire se A è un sottospazio vettoriale di $M_2(\mathbf{R})$ dove:

- (i) $A = \left\{ \begin{pmatrix} ab \\ ba \end{pmatrix} / a \in \mathbf{R}, b \in \mathbf{R} \right\}$; (ii) $A = \left\{ \begin{pmatrix} a & a+b \\ b & 0 \end{pmatrix} / a \in \mathbf{R}, b \in \mathbf{R} \right\}$;
 (iii) $A = \left\{ \begin{pmatrix} a & ab \\ b & 0 \end{pmatrix} / a \in \mathbf{R}, b \in \mathbf{R} \right\}$; (iv) $A = \left\{ \begin{pmatrix} a & 1 \\ a & b \end{pmatrix} / a \in \mathbf{R}, b \in \mathbf{R} \right\}$

1.4) Sia E un C-spazio vettoriale. Siano x, y due elementi di E. Si pone: $u = x+y$, $v = x-y$.

- (i) Dimostrare che il sottospazio generato da {x, y} è uguale a quello generato da {u, v}.
 (ii) Vale ancora (i) se E è un $\mathbf{Z}/2\mathbf{Z}$ -spazio vettoriale?

1.5) Sia E un k-spazio vettoriale. Dare una condizione necessaria e sufficiente affinché la riunione $F \cup H$ dei due sottospazi vettoriali F, H di E sia un sottospazio vettoriale di E.

1.6) (i) Siano E un k-spazio vettoriale, F un sottospazio vettoriale di E e $v \in E$ un vettore non nullo. Mostrare che la somma $\langle v \rangle + F$ è diretta se e solo se $v \notin F$.

(ii) Siano $k = \mathbf{Z}/2\mathbf{Z}$ e $E = k^2$ con la sua struttura naturale di k-spazio vettoriale (3 (iii)). Siano $e = (1, 0)$ e F il sottospazio vettoriale di E generato da e. Quanti sono i vettori v di E tali che la somma $\langle v \rangle + F$ sia diretta? Quanti sono i supplementari di F?

1.7) Sia \mathbf{R}^3 con la sua struttura naturale di \mathbf{R} -spazio vettoriale. Si ricorda che $\langle x_1, \dots, x_n \rangle$ indica il sottospazio generato dai vettori x_1, \dots, x_n .

(a) Siano $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (2, 1, 0)$. Si pone $E_i = \langle e_i \rangle$, $1 \leq i \leq 3$. Mostrare che $E_i \cap E_k = \{0\}$ se $i \neq k$. Sia $v = (3, 2, 0)$. Scrivere v come combinazione lineare di e_1 , e_2 (risp. di e_1 , e_3 , e di e_2 , e_3). Determinare la somma $E_1 + E_2 + E_3$, è questa somma diretta?

(b) Sia $x = (1, 1, 0)$, $y = (1, 2, 0)$, $z = (0, 0, 3)$. Dire se i sottospazi $\langle x, y \rangle$ e $\langle z \rangle$ sono in somma diretta.

(c) Dire se i sottospazi $\langle x \rangle$, $\langle y \rangle$, $\langle z \rangle$ sono in somma diretta dove:

- (i) $x = (1, 1, 0)$, $y = (1, 2, 0)$, $z = (0, 0, 3)$
 (ii) $x = (2, 1, 1)$, $y = (0, 1, 3)$, $z = (1, 0, 2)$

2) APPLICAZIONI LINEARI: DEFINIZIONE E PRIME PROPRIETA'.

Come nel caso dei gruppi, anelli, campi, dopo gli oggetti (k -spazi vettoriali) introduciamo i morfismi (applicazioni k -lineari), ossia le applicazioni che rispettano la struttura considerata.

1: Definizione: Siano E, F due k -spazi vettoriali e $f: E \rightarrow F$ un'applicazione. Si dice che f è un'applicazione k -lineare se:

(i) f è un morfismo di gruppi, dal gruppo $(E, +)$ nel gruppo $(F, +)$, ossia:

$$\forall (x, y) \in E^2, f(x+y) = f(x) + f(y)$$

(ii) f rispetta la moltiplicazione esterna:

$$\forall x \in E, \forall \lambda \in k: f(\lambda x) = \lambda f(x).$$

1.1: Osservazione : (i) Si dice anche morfismo k -lineare, o operatore k -lineare invece di applicazione lineare. Inoltre se il campo k è precisato dal contesto si dirà più semplicemente morfismo (risp. applicazione, operatore) lineare.

(ii) Se $f: E \rightarrow F$ è un'applicazione lineare allora $f(0) = 0$ (il primo zero denota il vettore nullo di E , il secondo quello di F). Infatti f è un morfismo di gruppi quindi manda l'elemento neutro di $(E, +)$ sull'elemento neutro di $(F, +)$ (cfr I, §5, 11).

(iii) Per ripetute applicazioni delle proprietà (i), (ii) di linearità, se $f: E \rightarrow F$ è lineare allora $f(\lambda_1 v_1 + \dots + \lambda_r v_r) = \lambda_1 f(v_1) + \dots + \lambda_r f(v_r)$.

2: Lemma: Siano E, F due k -spazi vettoriali e $f: E \rightarrow F$ un'applicazione. L'applicazione f è un morfismo k -lineare se e solo se:

$$\forall (x, y) \in E^2, \forall (\lambda, \mu) \in k^2: f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

Dim: Se f soddisfa le condizioni del lemma ponendo $\lambda = \mu = 1$, si ritrova (i) della definizione; ponendo $\mu = 0$, si ritrova (ii).

Viceversa se f è un morfismo lineare abbiamo: $f(\lambda x + \mu y) = f(\lambda x) + f(\mu y)$ per (i) e $f(\lambda x) + f(\mu y) = \lambda f(x) + \mu f(y)$ per (ii) ♦

3: Esempi : (i) Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ un'applicazione \mathbb{R} -lineare (qui il campo \mathbb{R} è considerato come spazio vettoriale su se stesso, cfr §1, 3(i)). Per la proprietà (ii) della definizione, per ogni x e ogni y in \mathbb{R} : $f(xy) = xf(y)$. In particolare $f(x)$

$= f(x \cdot 1) = xf(1)$. Se $f(1) = a$ allora: $\forall x \in \mathbb{R}$: $f(x) = ax$. L'applicazione f è completamente determinata dal suo valore in 1.

Il grafico di f in $\mathbb{R} \times \mathbb{R}$ è la retta (passante per l'origine) di equazione $y = ax$.

(ii) Più generalmente, se E è un k -spazio vettoriale, ogni applicazione k -lineare $f : k \rightarrow E$ è completamente determinata dal suo valore per un elemento non nullo di k . Infatti sia $\alpha \neq 0$ allora: $\forall \lambda \in k$, $f(\lambda) = f(\lambda\alpha^{-1}\alpha) = \lambda\alpha^{-1}f(\alpha)$ (α è invertibile perché non nullo). In particolare si può prendere $\alpha = 1$.

(iii) L'applicazione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$: $(x, y) \rightarrow (ax + by, cx + dy)$ dove a, b, c, d , sono dei numeri reali, è un'applicazione \mathbb{R} -lineare (cfr Es. 1).

4: Proposizione: Siano E, F, V tre k -spazi vettoriali e $f : E \rightarrow F, g : F \rightarrow V$, due morfismi lineari. L'applicazione composta $g \circ f : E \rightarrow V$ è un'applicazione k -lineare.

Dim: Per x, y in E e per ogni λ, μ in k abbiamo:

$$\begin{aligned} (g \circ f)(\lambda x + \mu y) &= g(f(\lambda x + \mu y)) \\ &= g(\lambda f(x) + \mu f(y)), \text{ per linearità di } f \text{ (cfr 2),} \\ &= \lambda g(f(x)) + \mu g(f(y)), \text{ per linearità di } g, \\ &= \lambda(g \circ f)(x) + \mu(g \circ f)(y) \blacklozenge \end{aligned}$$

5: Definizione: (Ker) Siano E, F , due k -spazi vettoriali e $f : E \rightarrow F$ un'applicazione k -lineare. Il nucleo (o Ker) di f è il nucleo del morfismo di gruppi corrispondente:

$\text{Ker}(f) := \{x \in E / f(x) = 0\}$ (qui 0 indica il neutro di $(F, +)$, i.e. il vettore nullo di F).

6: Proposizione: Sia $f : E \rightarrow F$ un morfismo di k -spazi vettoriali, allora $\text{Ker}(f)$ è un sottospazio vettoriale di E .

Dim: In virtù di §1. 5.1, basta verificare: (i) $0 \in \text{Ker}(f)$, (ii) $\forall u \in \text{Ker}(f), \forall v \in \text{Ker}(f), \forall (\lambda, \mu) \in k^2, \lambda u + \mu v \in \text{Ker}(f)$.

(i) Dall'osservazione 1.1, abbiamo $f(0) = 0$, quindi $0 \in \text{Ker}(f)$.

(ii) Bisogna verificare che $f(\lambda u + \mu v) = 0$. Per linearità (cfr 2) abbiamo: $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$. Per ipotesi u e v sono elementi di $\text{Ker}(f)$ quindi $f(u) = f(v) = 0$. Pertanto $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v) = \lambda 0 + \mu 0 = 0 \blacklozenge$

La proposizione seguente fornisce un criterio molto comodo per vedere se un'applicazione lineare sia o meno iniettiva. Questo risultato verrà usato frequentemente nel seguito.

7: Proposizione: *Un morfismo di k-spazi vettoriali, $f : E \rightarrow F$, è iniettivo se e solo se $\text{Ker}(f) = \{0\}$.*

Dim: Segue dalla proposizione analoga per i gruppi (I. §5, 13)♦

8: Proposizione: *Sia $f : E \rightarrow F$ un morfismo di k-spazi vettoriali. L'immagine di f , $\text{Im}(f)$, è un sottospazio vettoriale di F .*

Dim: Verifichiamo le condizioni di §1, 5.1: (i) $0 \in \text{Im}(f)$, (ii) $\forall (x, y) \in \text{Im}(f)^2$, $\forall (\lambda, \mu) \in k^2$, $\lambda x + \mu y \in \text{Im}(f)$. Si ricorda che $\text{Im}(f) = \{x \in F / \exists u \in E \text{ tale che } f(u) = x\}$.

(i) Abbiamo $f(0) = 0$ (1.1), quindi $0 \in \text{Im}(f)$.

(ii) Siccome x e y sono elementi di $\text{Im}(f)$ esistono u e v in E tali che $f(u) = x$, $f(v) = y$. Abbiamo $\lambda x + \mu y = \lambda f(u) + \mu f(v) = f(\lambda u + \mu v)$ (per linearità di f). Quindi $\lambda x + \mu y = f(w)$ (con $w = \lambda u + \mu v$) appartiene a $\text{Im}(f)$ ♦

9: Proposizione: *Siano E, F due k-spazi vettoriali, $f : E \rightarrow F$, un morfismo k-lineare. Se l'applicazione f è biettiva allora l'applicazione reciproca $f^{-1} : F \rightarrow E$ è un morfismo k-lineare.*

Dim: Es. 4♦

10: Definizione: Nella situazione della Proposizione 9 si dice che f è un isomorfismo tra gli spazi vettoriali E e F .

Un automorfismo dello spazio vettoriale E è un'applicazione lineare, biettiva, da E in se stesso.

10.1: Osservazione : Se esiste un isomorfismo da E in F gli spazi vettoriali E e F vengono detti isomorfi. Due spazi vettoriali isomorfi hanno le stesse proprietà vettoriali (e quindi possono essere identificati): un teorema di algebra lineare dimostrato per E si traduce, via f , in un teorema per F .

Esercizi:

2.1) Si ricorda che un polinomio, a coefficienti reali, omogeneo di grado uno nelle variabili x_1, \dots, x_n , è un polinomio del tipo $P(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n$. Sia $f : \mathbf{R}^n \rightarrow$

$R^k: (x_1, \dots, x_n) \rightarrow (P_1(x_1, \dots, x_n), \dots, P_k(x_1, \dots, x_n))$, dove $P_1(x_1, \dots, x_n), \dots, P_k(x_1, \dots, x_n)$ sono dei polinomi omogenei di grado uno. Dimostrare che f è lineare.

2.2) Dire quali delle seguenti applicazioni sono lineari:

- (a) $f: R^2 \rightarrow R^2: (x, y) \rightarrow (2x-y+1, x+y)$
- (b) $g: R^2 \rightarrow R^2: (x, y) \rightarrow (2x-y, x+y)$
- (c) $h: R^2 \rightarrow R^2: (x, y) \rightarrow (2x-y, x^2+y)$

2.3) Siano $f: R^2 \rightarrow R^2: (x, y) \rightarrow (3x-y, x+y)$, $g: R^2 \rightarrow R^2: (x, y) \rightarrow (2x-y, 4x-2y)$. Dimostrare che sono applicazioni lineari, determinarne il nucleo e l'immagine. Dire se f e g sono iniettive, suriettive, biiettive.

2.4) Dimostrare la Proposizione 9.

2.5) Sia E un k -spazio vettoriale e F un insieme. Sia $f: E \rightarrow F$ una biiezione.

Per ogni (x, y) in F^2 si pone: $x+y := f(f^{-1}(x) + f^{-1}(y))$.

Poi si definisce un'applicazione: $m: kxF \rightarrow F: (\lambda, x) \rightarrow m(\lambda, x) =: \lambda x$ tramite $\lambda x = f(\lambda f^{-1}(x))$.

Dimostrare che queste operazioni definiscono su F una struttura di k -spazio vettoriale (struttura ottenuta per "trasporto" tramite f), e che per questa struttura f è un isomorfismo lineare tra E e F .

3) SPAZI FINITAMENTE GENERATI.

Questo paragrafo contiene un primo approccio alla nozione di dimensione di un k-spazio vettoriale. Ci limitiamo qui alla distinzione tra spazi di dimensione finita (o finitamente generati) e spazi di dimensione infinita. Nel seguito considereremo esclusivamente spazi di dimensione finita.

1: Definizione: Sia E un k -spazio vettoriale e $A \subseteq E$ un sottinsieme. Gli elementi di A costituiscono un sistema di generatori di E se ogni vettore di E è combinazione lineare di vettori di A . In altri termini A è un sistema di generatori di E se il sottospazio generato da A (cfr §1, 6.3) è E tutt'intero.

1.1: Esempio : Sia $A \subseteq \mathbb{R}^2$, $A = \{u, v\}$ con $u = (1, 0)$, $v = (a, b)$ con $b \neq 0$. Allora A è un sistema di generatori di \mathbb{R}^2 (cfr §1, 8.3.2).

2: Definizione: Sia E un k -spazio vettoriale. Se esiste un sottinsieme finito $A \subseteq E$ che genera E allora E è un k -spazio vettoriale di dimensione finita (o finitamente generato). Se invece non esiste nessun sistema di generatori finito, E è un k -spazio vettoriale di dimensione infinita (si nota $\dim_k(E) = +\infty$).

Nel seguito svilupperemo la teoria essenzialmente per spazi vettoriali di dimensione finita. Però, è bene sapere che esistono spazi vettoriali di dimensione infinita che sono oggetti matematici naturali. Ecco alcuni esempi:

3: Esempi : (a) Lo spazio vettoriale reale, $\mathbb{R}[X]$, dei polinomi a coefficienti reali, in una variabile. Si ha $\dim_{\mathbb{R}} \mathbb{R}[X] = +\infty$ (**Es.1**).

(b) Lo spazio vettoriale delle applicazioni (risp. delle applicazioni continue, delle applicazioni indefinitamente derivabili) da \mathbb{R} in \mathbb{R} è uno spazio vettoriale reale di dimensione infinita (**Es.1**).

(c) Un esempio forse più inaspettato: \mathbb{Q} è un sottocampo di \mathbb{R} , quindi possiamo considerare \mathbb{R} come un \mathbb{Q} -spazio vettoriale (cfr §1, 3 (ii)), allora \mathbb{R} è un \mathbb{Q} -spazio vettoriale di dimensione infinita: $\dim_{\mathbb{Q}} \mathbb{R} = +\infty$. Questo segue, tra l'altro, dal fatto che \mathbb{Q} è un insieme numerabile (in biiezione con \mathbb{N}) mentre \mathbb{R} non è numerabile.

Osserviamo però il fatto seguente: \mathbb{R} è un sottocampo di \mathbb{C} , quindi possiamo considerare \mathbb{C} come \mathbb{R} -spazio vettoriale. Questa volta siamo in una situazione del tutto diversa: infatti \mathbb{C} è un \mathbb{R} -spazio vettoriale di dimensione finita perché $A = \{1, i\}$ genera \mathbb{C} come \mathbb{R} -spazio vettoriale (ogni numero complesso, z , si scrive $z = a + ib$ con a e b reali). E' \mathbb{C} un \mathbb{Q} -spazio vettoriale di dimensione finita?

Esercizi:

3.1) Dimostrare che lo spazio vettoriale reale $\mathbb{R}[X]$ non è finitamente generato (ragionare per assurdo considerando i gradi). Dedurne che lo spazio vettoriale reale delle applicazioni (risp. delle applicazioni continue, risp. delle applicazioni indefinitamente derivabili) da \mathbb{R} in \mathbb{R} ha dimensione infinita.

3.2) Sia $k = \mathbb{Z}/2\mathbb{Z}$. Dire se il k -spazio vettoriale $k[X]$ è finitamente generato. Sia A l'insieme delle applicazioni da k in k . Dire se il k -spazio vettoriale A è finitamente generato.

3.3) Siano E, F due k -spazi vettoriali e $f: E \rightarrow F$ un'applicazione lineare suriettiva. Dimostrare che se E è finitamente generato allora anche F è finitamente generato.

4) INDIPENDENZA LINEARE, BASI.

Scopo di questo paragrafo è di mostrare l'esistenza di basi in uno spazio vettoriale di dimensione finita.

Per definizione, un k -spazio vettoriale di dimensione finita, E , è generato da un numero finito di vettori g_1, \dots, g_k . Quindi ogni vettore, v , di E si scrive come combinazione lineare dei g_i : $v = \lambda_1 g_1 + \dots + \lambda_k g_k$, $\lambda_i \in k$. Può essere però che ci sia più di un modo per scrivere v come combinazione lineare dei g_i (i.e. con altri scalari, cfr Es.4.1). Poniamo $G = \{g_1, \dots, g_k\}$. E' chiaro che se $G \subseteq H$ allora anche H genera E . Viceversa possiamo chiederci se esistono sottinsiemi stretti $G' \subseteq G$, $G' \neq G$, che generano E . Possiamo anche cercare tali sottinsiemi in modo che siano minimali (in questo contesto G' è minimale se $G'' \subseteq G'$ e G'' genera E implica $G'' = G'$).

Diremo che $G = \{g_1, \dots, g_k\}$ è un sistema minimale di generatori di E se:

(i) G genera E

(ii) non esiste nessun sottinsieme stretto di G che genera E .

La condizione (ii) è equivalente a: $\forall g_i \in G$, $G \setminus \{g_i\}$ non genera E .

Una condizione necessaria affinché G sia un sistema minimale di generatori è, chiaramente, che i g_i siano tutti non nulli e distinti.

Più generalmente una condizione necessaria è:

(t) se $\alpha_1 g_1 + \dots + \alpha_k g_k = 0$, $\alpha_i \in k$, allora dev'essere: $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$.
(si dice allora che g_1, \dots, g_k sono linearmente indipendenti, o liberi).

Mostriamo che questa condizione è necessaria. Se (t) non è verificata, esistono β_1, \dots, β_k , non tutti nulli tali che: $\beta_1 g_1 + \dots + \beta_k g_k = 0$. Eventualmente riordinando gli indici possiamo supporre $\beta_k \neq 0$. Abbiamo $\beta_k g_k = -\beta_1 g_1 - \dots - \beta_{k-1} g_{k-1}$. Componendo con β_k^{-1} (lecito perché $\beta_k \neq 0$): $g_k = \delta_1 g_1 + \dots + \delta_{k-1} g_{k-1}$ ($\delta_i = -\beta_i \beta_k^{-1}$). Adesso ogni vettore v è combinazione lineare dei g_i : $v = \lambda_1 g_1 + \dots + \lambda_{k-1} g_{k-1} + \lambda_k g_k$. Quindi $v = \lambda_1 g_1 + \dots + \lambda_{k-1} g_{k-1} + \lambda_k (\delta_1 g_1 + \dots + \delta_{k-1} g_{k-1}) = (\lambda_1 + \lambda_k \delta_1) g_1 + \dots + (\lambda_{k-1} + \delta_{k-1}) g_{k-1}$, e vediamo che ogni vettore è combinazione lineare di g_1, \dots, g_{k-1} , in contraddizione con l'ipotesi che G sia un sistema minimale di generatori.

Abbiamo quindi dimostrato che un sistema minimale di generatori soddisfa la condizione (t). Una base è un sistema minimale di generatori.

Dimostreremo che in uno spazio vettoriale di dimensione finita:

(1) esiste sempre una base

(2) tutte le basi hanno lo stesso numero di elementi. Questo numero è la dimensione del k-spazio vettoriale considerato.

(3) se e_1, \dots, e_n è una base di E, ogni vettore v di E si scrive in modo unico come combinazione lineare degli e_i : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Gli scalari $\lambda_1, \dots, \lambda_n$, sono le componenti o coordinate del vettore v nella base e_1, \dots, e_n .

(4) ogni applicazione lineare $f : E \rightarrow V$, V un k-spazio vettoriale qualsiasi, è completamente determinata dai suoi valori su una base di E: se e_1, \dots, e_n è una base di E, dare f è equivalente a dare gli n vettori di V $f(e_1), \dots, f(e_n)$ (cfr §2, 3(ii), 1 è una base del k-spazio vettoriale k).

(5) Ogni k-spazio vettoriale di dimensione n è isomorfo (§2, 10) allo spazio vettoriale k^n (detto "spazio numerico").

Il punto (5) risolve il problema della classificazione degli spazi vettoriali di dimensione finita, mentre (4) determina, almeno teoricamente, la struttura delle applicazioni lineari. Come abbiamo già visto uno strumento essenziale per realizzare questo programma è la nozione di indipendenza lineare che adesso introduciamo in modo più formale.

1: Definizione: Siano V un k-spazio vettoriale, v_1, \dots, v_n , n vettori di V. I vettori v_1, \dots, v_n , sono linearmente indipendenti (o liberi) se: per ogni $(\lambda_1, \dots, \lambda_n) \in k^n$, la relazione $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ implica che tutti i λ_i sono nulli: $\lambda_1 = 0, \dots, \lambda_n = 0$.

Se v_1, \dots, v_n , non sono linearmente indipendenti allora v_1, \dots, v_n , vengono detti linearmente dipendenti (o legati). Abbiamo quindi:

v_1, \dots, v_n , sono linearmente dipendenti (o legati) se e solo se esistono degli scalari $\alpha_1, \dots, \alpha_n$, non tutti nulli (i.e. $(\alpha_1, \dots, \alpha_n) \neq (0, \dots, 0)$), tali che $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$.

1.1: Esempi : (i) Sia v un vettore di V. Allora v è libero se e solo se v non è il vettore nullo. Infatti se $\lambda v = 0$ allora $\lambda = 0$ o $v = 0$ (§1, 4.8).

(ii) Sia $v_1 = 0, v_2, \dots, v_n$, dove v_2, \dots, v_n , sono n-1 vettori qualsiasi di V. Mostriamo che v_1, \dots, v_n , sono linearmente dipendenti. Infatti sia $\lambda_1 = 1, \lambda_2 = \dots$

$= \lambda_n = 0$. Abbiamo $(\lambda_1, \lambda_2, \dots, \lambda_n) = (1, 0, \dots, 0) \neq (0, \dots, 0)$ e $\lambda_1 v_1 + \dots + \lambda_n v_n = 1.0 + 0.v_2 + \dots + 0.v_n = 0$.

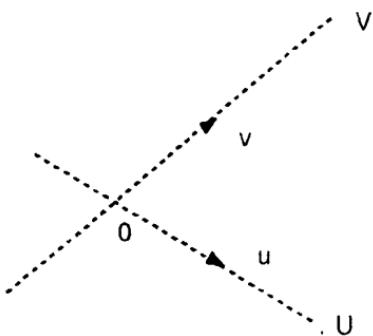
Quindi ogni insieme di vettori contenente il vettore nullo è legato.

(iii) Sia v un vettore non nullo di V . Poniamo $v_1 = v = v_2$, e siano $v_3, \dots, v_n, n-2$ vettori qualsiasi di V . I vettori $v_1, v_2, v_3, \dots, v_n$ sono linearmente dipendenti. Infatti sia $\lambda \in k^* = k \setminus \{0\}$ uno scalare non nullo e poniamo $\lambda_1 = \lambda, \lambda_2 = -\lambda, \lambda_3 = \dots = \lambda_n = 0$. Abbiamo $(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n) = (\lambda, -\lambda, 0, \dots, 0) \neq (0, \dots, 0)$ e $\lambda_1 v_1 + \dots + \lambda_n v_n = \lambda v - \lambda v + 0v_3 + \dots + 0v_n = 0$.

In conclusione se n vettori v_1, \dots, v_n sono linearmente indipendenti allora sono tutti non nulli ($v_i \neq 0, \forall i$) e a due a due distinti ($i \neq j \Rightarrow v_i \neq v_j$).

Queste condizioni, necessarie, non sono sufficienti per assicurare l'indipendenza lineare.

(iv) Siano u e v due vettori non nulli del piano vettoriale, con direzioni diverse:



Ricordiamo (§1, 8.3.2) che se $U = \{\lambda u / \lambda \in \mathbb{R}\}, V = \{\mu v / \mu \in \mathbb{R}\}$ sono le rette vettoriali generate da u e v , l'ipotesi: "u e v hanno direzioni diverse", è equivalente a $U \cap V = \{0\}$. Mostriamo che i vettori u e v sono linearmente indipendenti (o liberi).

Infatti supponiamo $\lambda u + \mu v = 0$ con λ e μ non tutti nulli, per esempio $\lambda \neq 0$. Abbiamo $\lambda u = -\mu v$, componendo con λ^{-1} : $u = (\lambda^{-1}\mu)v$. Quindi $u \in V$. Pertanto $u \in U \cap V$. Per ipotesi $U \cap V = \{0\}$, quindi $u = 0$, assurdo.

Osservare che u e v generano tutto il piano vettoriale (cfr §1, 8.3.2). Quindi, come vedremo più avanti, u e v formano una base del piano vettoriale (il quale, come ogni buon piano che si rispetti, ha quindi dimensione due).

(v) Due vettori u, v di un k -spazio vettoriale E sono proporzionali se esiste $\lambda \in k, \lambda \neq 0$, tale che $\lambda u = v$ (o $\lambda v = u$). Due vettori proporzionali sono linearmente dipendenti perché $\lambda u - v = 0$ ($(\lambda, -1) \neq (0, 0)$). Viceversa se u e v sono due vettori linearmente dipendenti allora $\alpha u + \beta v = 0$ con $(\alpha, \beta) \neq (0, 0)$, e u, v sono proporzionali.

Invece, se prendiamo tre vettori a due a due non proporzionali, questi vettori non sono necessariamente linearmente indipendenti (cfr Es. 1).

2: Definizione: Sia E un k -spazio vettoriale. Una famiglia $(x_i)_{i \in I}$ (non necessariamente finita) di elementi di E è libera se ogni sottofamiglia finita è libera. Altrimenti la famiglia è legata.

2.1: Esempio : Sia $E = R[X]$, lo spazio vettoriale reale dei polinomi in una variabile, a coefficienti reali. La famiglia $(X^n)_{n \in N}$ è libera (cfr Es.3.1).

3: Proposizione: Siano E un k -spazio vettoriale, v_1, \dots, v_n n vettori di E . Gli n vettori v_1, \dots, v_n sono linearmente dipendenti se e solo se: $n = 1, v_1 = 0$; $n > 1$ e esiste $j, 1 \leq j \leq n$, tale che v_j sia combinazione lineare di $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$.

Dim: Il caso $n = 1$ è chiaro (1.1 (i)). Supponiamo quindi $n > 1$.

Supponiamo v_1, \dots, v_n linearmente dipendenti. Esistono quindi degli scalari, $\lambda_1, \dots, \lambda_n$, non tutti nulli, tali che $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Siccome i λ_i non sono tutti nulli esiste $j, 1 \leq j \leq n$, tale che $\lambda_j \neq 0$. Abbiamo $\lambda_j v_j = -\lambda_1 v_1 - \dots - \lambda_{j-1} v_{j-1} - \lambda_{j+1} v_{j+1} - \dots - \lambda_n v_n$, e quindi, λ_j essendo non nullo: $v_j = \alpha_1 v_1 + \dots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \dots + \alpha_n v_n$, con $\alpha_i = -\lambda_i / \lambda_j$ ($\lambda_j \lambda_j^{-1}$).

Viceversa se v_j è combinazione lineare dei rimanenti v_i : $v_j = \alpha_1 v_1 + \dots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \dots + \alpha_n v_n$, allora $\alpha_1 v_1 + \dots + \alpha_{j-1} v_{j-1} - v_j + \alpha_{j+1} v_{j+1} + \dots + \alpha_n v_n = 0$. Visto che $(\alpha_1, \dots, \alpha_{j-1}, -1, \dots, \alpha_n) \neq (0, \dots, 0)$, i vettori v_1, \dots, v_n sono linearmente dipendenti ♦

4: Proposizione: Sia E un k -spazio vettoriale e v_1, \dots, v_n, n vettori linearmente indipendenti. Sia v un vettore di E tale che v, v_1, \dots, v_n , siano legati. Allora v è combinazione lineare di v_1, \dots, v_n .

Dim: Se v, v_1, \dots, v_n sono legati esistono degli scalari $\lambda, \lambda_1, \dots, \lambda_n$, non tutti nulli, tali che $\lambda v + \lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Se $\lambda = 0$ allora $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Essendo i v_i linearmente indipendenti per ipotesi, questo implica $\lambda_1 = \dots = \lambda_n = 0$, contro

l'ipotesi $\lambda, \lambda_1, \dots, \lambda_n$ non tutti nulli. Quindi $\lambda \neq 0$. Abbiamo $\lambda v = -\lambda_1 v_1 - \dots - \lambda_n v_n$, componendo con λ^{-1} : $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, con $\alpha_i = -\lambda_i / \lambda$ ♦

5: Definizione: Siano E un k -spazio vettoriale, e_1, \dots, e_n , n vettori di E tali che:

- (a) e_1, \dots, e_n siano linearmente indipendenti
- (b) e_1, \dots, e_n generino E .

Allora $B = (e_1, \dots, e_n)$ è una base di E .

5.1: Osservazione : (i) Nella definizione precedente risulta che E ha dimensione finita.

(ii) Da (b) segue che ogni vettore v di E è combinazione lineare di e_1, \dots, e_n : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Da (a) segue che ogni vettore è combinazione lineare degli e_i in modo unico: infatti se $v = \lambda_1 e_1 + \dots + \lambda_n e_n = \alpha_1 e_1 + \dots + \alpha_n e_n$, allora $(\alpha_1 - \lambda_1)e_1 + \dots + (\alpha_n - \lambda_n)e_n = 0$. Ma gli e_i sono linearmente indipendenti quindi $\alpha_1 = \lambda_1, \dots, \alpha_n = \lambda_n$.

(iii) Convenzione: notare che una base è un n -upla (c'è un ordine sull'insieme $\{e_1, \dots, e_n\}$). Se $B = (e_1, \dots, e_n)$ è una base anche $B' = (e_2, e_1, \dots, e_n)$ è una base (l'indipendenza lineare e la generazione non dipendono dall'ordine nel quale si prendono i vettori). Però B e B' sono due basi diverse (se $B = B'$ allora $e_1 = e_2$ e i vettori e_1, \dots, e_n non sono liberi). Più generalmente, per ogni permutazione, σ , di $\{1, \dots, n\}$, $(e_{\sigma(1)}, \dots, e_{\sigma(n)})$ è una base di E . L'ordine in una base permette di definire le coordinate di un vettore: se $v = \lambda_1 e_1 + \dots + \lambda_n e_n$, $(\lambda_1, \dots, \lambda_n)$ sono le coordinate di v nella base B (λ_1 è la prima coordinata, λ_2 la seconda, ecc ...). Per esempio $e_1 = (1, 0), e_2 = (0, 1)$ è una base di R^2 (la base "canonica", cfr 12.2); se $v = (x, y)$, cioè $v = x e_1 + y e_2$, allora le coordinate di v sono: x (prima coordinata), y (seconda coordinata).

Tuttavia, se non si vuole precisare l'ordine si dice, con abuso di linguaggio, che e_1, \dots, e_n formano una base.

(iv) Se E ha dimensione infinita, una base di E è una famiglia libera che genera E . Nell'esempio 2.1 $(X^n)_{n \in \mathbb{N}}$ è una base di $R[X]$ (cfr Es. 5).

Abbiamo la seguente caratterizzazione:

6: Proposizione: Siano $E \neq \{0\}$ un k -spazio vettoriale di dimensione finita, e_1, \dots, e_n n vettori di E . Sono equivalenti:

- (i) e_1, \dots, e_n formano una base
- (ii) $G = \{e_1, \dots, e_n\}$ è un sistema minimale di generatori di E .

(iii) G è un sistema massimale di vettori linearmente indipendenti di E (i.e. e_1, \dots, e_n sono liberi e, per ogni vettore v di E , v, e_1, \dots, e_n sono linearmente dipendenti).

Dim: (i) \Rightarrow (ii) Per ipotesi $\langle G \rangle = E$. Se G non è un sistema minimale di generatori, esiste j , $1 \leq j \leq n$, tale che $G \setminus \{e_j\}$ genera E . In particolare e_j è combinazione lineare degli elementi di $G \setminus \{e_j\}$. Dalla Proposizione 3 segue che i vettori di G sono legati e questo è assurdo perché i vettori di G formano una base.

(ii) \Rightarrow (iii) Siccome $E \neq \{0\}$, uno degli e_i è senz'altro non nullo. Per minimalità di G questo implica che nessuno degli e_i sia nullo (se $e_j = 0$, $G \setminus \{e_j\}$ genera quanto G). Se gli e_i sono linearmente dipendenti, dalla Proposizione 3, uno degli e_i è combinazione lineare degli altri. Riordinando eventualmente gli indici, possiamo supporre $e_n = \alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}$ (*). Per ipotesi ogni vettore è combinazione lineare degli e_i : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Usando (*): $v = \lambda_1 e_1 + \dots + \lambda_{n-1} e_{n-1} + \lambda_n (\alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}) = (\lambda_1 + \lambda_n \alpha_1) e_1 + \dots + (\lambda_{n-1} + \lambda_n \alpha_{n-1}) e_{n-1}$. Quindi $G \setminus \{e_n\}$ genera E . Questo è assurdo perché G è un sistema minimale di generatori. Quindi i vettori di G sono liberi.

Finalmente siccome $\langle G \rangle = E$, se $v \in E$, v è combinazione lineare degli e_i , dalla Proposizione 3 segue che v, e_1, \dots, e_n sono linearmente dipendenti. Quindi G è un sistema massimale di vettori liberi di E .

(iii) \Rightarrow (i) Per ipotesi e_1, \dots, e_n sono liberi. Rimane da vedere che $\langle G \rangle = E$. Sia $v \in E$, per ipotesi v, e_1, \dots, e_n sono legati. Essendo gli e_i linearmente indipendenti, la Proposizione 4 implica che v è combinazione lineare di e_1, \dots, e_n ♦

6.1: Osservazione : Se $E = \{0\}$ allora $\{0\}$ è un sistema minimale di generatori di E , ma 0 è legato ($1 \cdot 0 = 0$). Sistemeremo più avanti questa specie di gag.

Il risultato seguente è fondamentale:

7: Teorema: Siano E un k-spazio vettoriale, A e A' due sottinsiemi finiti, non vuoti, di E tali che $A \subseteq A'$. Supponiamo che i vettori di A siano linearmente indipendenti e che i vettori di A' generino E . Allora esiste una base di E , B , tale che: $A \subseteq B \subseteq A'$.

Dim: Consideriamo tutti i sottinsiemi X tali che: $A \subseteq X \subseteq A'$ e tali che i vettori di X siano linearmente indipendenti. Tali sottinsiemi esistono (per esempio A) e sono in numero finito (perché A' è un insieme finito). Scegliamo uno di

questi insiemi, B , col massimo numero di elementi. Per costruzione gli elementi di B sono linearmente indipendenti. Per mostrare che B è una base rimane da far vedere che B genera E . Sia $x \in A'$, $x \notin B$. I vettori di $B \cup \{x\}$ sono linearmente dipendenti (altrimenti B non avrebbe il massimo numero possibile di elementi). Dalla Proposizione 4 segue che x è combinazione lineare dei vettori di B . In conclusione ogni elemento di A' è combinazione lineare degli elementi di B . Siccome ogni vettore di E è combinazione lineare degli elementi di A' (A' genera E per ipotesi), ogni vettore di E è combinazione lineare degli elementi di B . Quindi B è una base♦

8: Corollario: *Sia E un k-spazio vettoriale di dimensione finita. Se $E \neq \{0\}$ allora E ammette una base.*

Dim: Per definizione E è generato da un insieme finito (e non vuoto), A' , di vettori. Siccome $E \neq \{0\}$ esiste $v \in A'$ tale che $v \neq 0$. Si conclude applicando il teorema 7 con $A = \{v\}$ e A' ♦

8.1: Osservazione : Continuazione di 6.1: sia E un k-spazio vettoriale. L'insieme $\{0\}$, costituito dal solo vettore nullo di E , è un sottospazio vettoriale di E , quindi un k-spazio vettoriale. Questo spazio vettoriale ha dimensione finita perché è generato dall'unico elemento 0, ma non ci sono vettori liberi in $\{0\}$ (cfr 6.1), quindi non possono esistere basi (non vuote) di questo spazio vettoriale! Si può fare la convenzione che $E = \{0\}$ ammetta come base l'insieme vuoto. Con questa convenzione ogni k-spazio vettoriale di dimensione finita ammette una base.

9: Corollario: ("teorema della base incompleta")

Siano E un k-spazio vettoriale di dimensione finita e v_1, \dots, v_p dei vettori linearmente indipendenti di E . Allora esiste una base di E contenente v_1, \dots, v_p .

Dim: Poniamo $A = \{v_1, \dots, v_p\}$. Poichè E ha dimensione finita, per definizione, esiste un sottinsieme finito, G , che genera E . Sia $A' = A \cup G$, allora A' genera E e $A \subseteq A'$. Si conclude col teorema 7♦

9.1: Osservazione : (i) Se $v \neq 0$ appartiene a E esiste sempre una base di E della forma $(e_1 = v, e_2, \dots, e_n)$.

(ii) Vedere negli esercizi (Es. 7, 8) un'altra versione del teorema della base incompleta.

Il lettore si convincerà facilmente che tutti i risultati di questo paragrafo discendono dal teorema 7 e dal lemma seguente:

10; Lemma: Siano E un k -spazio vettoriale, x_1, \dots, x_p dei vettori di E . Se i vettori y_1, \dots, y_{p+1} sono combinazioni lineari di x_1, \dots, x_p allora y_1, \dots, y_{p+1} sono linearmente dipendenti.

Dim: Per induzione su p .

Se $p = 1$ abbiamo $y_1 = \lambda x_1$, $y_2 = \mu x_1$. Se $\lambda = 0$ allora $y_1 = 0$ e y_1, y_2 sono senz'altro legati (1.1 (ii)). Se $\lambda \neq 0$ allora $x_1 = \lambda^{-1}y_1$, quindi $y_2 = \mu\lambda^{-1}y_1$ e y_1, y_2 sono proporzionali, quindi legati (1.1 (v)).

Sia $p > 1$, supponiamo (ipotesi di induzione) il lemma vero per $p-1$ e dimostriamolo per p . Per ipotesi gli y_j , $1 \leq j \leq p+1$, sono combinazioni lineari degli x_i , $1 \leq i \leq p$.

$$y_1 = \lambda_1^{-1}x_1 + \dots + \lambda_p^{-1}x_p$$

$$\dots$$

$$y_p = \lambda_1^p x_1 + \dots + \lambda_p^p x_p$$

$$y_{p+1} = \lambda_1^{p+1} x_1 + \dots + \lambda_p^{p+1} x_p.$$

Se $\lambda_1^{-1} = \dots = \lambda_p^{-1} = 0$, allora y_1, \dots, y_p sono combinazioni lineari dei $p-1$ vettori x_2, \dots, x_p . Per ipotesi di induzione, y_1, \dots, y_p sono legati. Questo implica (cfr Es.6) che anche y_1, \dots, y_p, y_{p+1} sono legati.

Possiamo quindi assumere che almeno uno degli scalari $\lambda_1^{-1}, \dots, \lambda_p^{-1}$ sia non nullo. Salvo riordinare gli indici possiamo supporre $\lambda_1^{-1} \neq 0$.

Se $\lambda_1^{-1} \neq 0$, x_1 è combinazione lineare di y_1, x_2, \dots, x_p . Per conseguenza i p vettori y_2, \dots, y_{p+1} sono combinazioni lineari di y_1, x_2, \dots, x_p :

$$y_2 = \lambda_2 y_1 + \mu_2^{-1} x_2 + \dots + \mu_p^{-1} x_p$$

$$\dots$$

$$y_{p+1} = \lambda_{p+1} y_1 + \mu_2^{-1} x_2 + \dots + \mu_p^{-1} x_p$$

Poniamo $w_2 = y_2 - \lambda_2 y_1, \dots, w_{p+1} = y_{p+1} - \lambda_{p+1} y_1$. I p vettori w_2, \dots, w_{p+1} sono combinazioni lineari dei $p-1$ vettori x_2, \dots, x_p . Per ipotesi di induzione w_2, \dots, w_{p+1} sono legati. Quindi esistono degli scalari, non tutti nulli, $\alpha_2, \dots, \alpha_{p+1}$, tali che: $\alpha_2 w_2 + \dots + \alpha_{p+1} w_{p+1} = 0$. Sostituendo w_i con $y_i - \lambda_i y_1$: $\alpha_2(y_2 - \lambda_2 y_1) + \dots + \alpha_{p+1}(y_{p+1} - \lambda_{p+1} y_1) = 0$, ossia: $y_1(-\alpha_2 \lambda_2 - \dots - \alpha_{p+1} \lambda_{p+1}) + \alpha_2 y_2 + \dots + \alpha_{p+1} y_{p+1} = 0$.

Siccome $\alpha_2, \dots, \alpha_{p+1}$ non sono tutti nulli, y_1, \dots, y_{p+1} sono linearmente dipendenti e il lemma è dimostrato♦

11: Teorema: Siano E un k-spazio vettoriale di dimensione finita e (x_1, \dots, x_n) , (y_1, \dots, y_p) due basi di E. Allora $n = p$.

Quindi tutte le basi di E hanno lo stesso numero di elementi.

Dim: Siccome (x_1, \dots, x_n) è una base di E, y_1, \dots, y_p sono combinazioni lineari di x_1, \dots, x_n . Se $p > n$, allora y_1, \dots, y_{n+1} sono combinazioni lineari di x_1, \dots, x_n . Il lemma 10 implica che y_1, \dots, y_{n+1} sono linearmente dipendenti. A fortiori (cfr Es. 6) anche y_1, \dots, y_p sono legati. Questo è assurdo perché (y_1, \dots, y_p) è una base. Pertanto $p \leq n$.

Scambiando gli x_i con gli y_j e ragionando allo stesso modo, si mostra $n \leq p$. Quindi $n = p$ ♦

12: Definizione: Sia E un k-spazio vettoriale di dimensione finita. Se $E = \{0\}$ si pone $\dim_k E = 0$. Altrimenti, se $E \neq \{0\}$, la dimensione di E su k (denotata con $\dim_k E$) è il numero di elementi di una base di E.

12.1: Osservazione : In virtù di 8, 11, la dimensione di un k-spazio vettoriale di dimensione finita è ben definita. Ogni k-spazio vettoriale ha una dimensione e $\dim_k E = 0$ se e solo se $E = \{0\}$. Quindi la convenzione $\dim_k \{0\} = 0$ permette di assegnare in modo coerente una dimensione a $\{0\}$, che ha dimensione finita, ma non possiede basi (non vuote).

12.2: Esempio : Lo spazio numerico, k^n , come k-spazio vettoriale, ha dimensione n. Una base di k^n è data dagli n vettori $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$. Gli e_i sono liberi perché $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$ è equivalente a $(\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$. Gli e_i generano tutto k^n perché se $v = (\alpha_1, \dots, \alpha_n)$ è un vettore di k^n allora $v = \alpha_1 e_1 + \dots + \alpha_n e_n$. La base (e_1, \dots, e_n) (in quest'ordine!) è detta base canonica di k^n .

13: Definizione: Siano E un k-spazio vettoriale, $B = (e_1, \dots, e_n)$ una base di E. Ogni vettore v di E si scrive in modo unico (5.1 (ii)) come combinazione lineare degli e_i : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Gli scalari $\lambda_1, \dots, \lambda_n$ sono le coordinate (o componenti) del vettore v rispetto alla base B.

13.1: Osservazione : Se $v = (\lambda_1, \dots, \lambda_n)$ è un vettore di k^n , le coordinate di v rispetto alla base canonica di k^n (cfr 12.2 (ii)) sono proprio $\lambda_1, \dots, \lambda_n$. Questo spiega la terminologia.

Il risultato seguente è molto utile nella pratica:

14: Proposizione: Sia E un k -spazio vettoriale di dimensione n , n vettori di E formano una base di E se e solo se sono linearmente indipendenti.

Dim: Se v_1, \dots, v_n formano una base sono linearmente indipendenti per definizione. Viceversa se v_1, \dots, v_n sono linearmente indipendenti allora costituiscono un sistema massimale di vettori linearmente indipendenti (perché $\dim_k E = n$) e quindi (6(iii)) formano una base♦

15: Teorema: Sia E un k -spazio vettoriale di dimensione n . Se F è un sottospazio vettoriale di E allora F ha dimensione finita e $\dim_k F \leq n$, con uguaglianza se e solo se $E = F$.

Dim: Se $F = \{0\}$, il teorema è vero. Supponiamo quindi $F \neq \{0\}$ (e quindi $n > 0$). Se x_1, \dots, x_p sono p vettori di F linearmente indipendenti allora x_1, \dots, x_p sono anche p vettori di E linearmente indipendenti. Per il teorema della base incompleta (9) esiste una base di E contenente x_1, \dots, x_p . Quindi $p \leq n$. Sia r il numero massimo di vettori di F linearmente indipendenti ($r \leq n$ da quanto precede). Se f_1, \dots, f_r sono r vettori linearmente indipendenti di F allora generano F . Infatti per ogni $f \in F$ i vettori f_1, \dots, f_r, f sono legati e quindi (4) f è combinazione lineare degli f_i . Questo mostra che F ha dimensione finita e $\dim_k F = r \leq n$. Se $r = n$ allora gli f_i formano una base di E (14) e quindi $F = \langle f_1, \dots, f_n \rangle = E$ ♦

15.1: Osservazione : Riassumiamo alcuni fatti semplici ma essenziali:

Sia E un k -spazio vettoriale di dimensione n .

(i) se v_1, \dots, v_k generano E allora $k \geq n$ ed esistono n elementi di $\{v_1, \dots, v_k\}$ che formano una base di E .

(ii) p vettori di E formano una base di E se e solo se $p = n$ e gli n vettori sono liberi.

(iii) siano w_1, \dots, w_r r vettori di E . Se $r > n$, w_1, \dots, w_r sono linearmente dipendenti.

Adesso passiamo allo studio delle applicazioni lineari.

16: Proposizione: Siano E, F , due k -spazi vettoriali, (e_1, \dots, e_n) una base di E , e f_1, \dots, f_n n vettori di F .

(i) Esiste una ed una sola applicazione lineare $u : E \rightarrow F$ tale che $u(e_1) = f_1, \dots, u(e_n) = f_n$.

(ii) u è iniettiva se e solo se i vettori f_1, \dots, f_n di F sono linearmente indipendenti.

(iii) u è suriettiva se e solo se i vettori f_1, \dots, f_n generano F .

(iv) u è biiettiva se e solo se f_1, \dots, f_n formano una base di F .

Dim: (i) Esistenza di u : si definisce u nel modo seguente: ogni vettore v di E si scrive in modo unico come combinazione lineare dei vettori della base $B = (e_1, \dots, e_n)$: $v = \lambda_1 e_1 + \dots + \lambda_n e_n$; poniamo $u(v) := \lambda_1 f_1 + \dots + \lambda_n f_n$.

E' chiaro che $u(e_i) = f_i$, $1 \leq i \leq n$. Mostriamo che u è lineare. Bisogna verificare $u(\lambda v + \mu w) = \lambda u(v) + \mu u(w)$. Siano $v = \alpha_1 e_1 + \dots + \alpha_n e_n$, $w = \beta_1 e_1 + \dots + \beta_n e_n$. Allora $\lambda v + \mu w = (\lambda\alpha_1 + \mu\beta_1)e_1 + \dots + (\lambda\alpha_n + \mu\beta_n)e_n$. Per definizione di u : $u(\lambda v + \mu w) = (\lambda\alpha_1 + \mu\beta_1)f_1 + \dots + (\lambda\alpha_n + \mu\beta_n)f_n$. Sempre per definizione: $u(v) = \alpha_1 f_1 + \dots + \alpha_n f_n$, $u(w) = \beta_1 f_1 + \dots + \beta_n f_n$. Quindi $\lambda u(v) + \mu u(w) = \lambda(\alpha_1 f_1 + \dots + \alpha_n f_n) + \mu(\beta_1 f_1 + \dots + \beta_n f_n) = (\lambda\alpha_1 + \mu\beta_1)f_1 + \dots + (\lambda\alpha_n + \mu\beta_n)f_n = u(\lambda v + \mu w)$ e u è lineare.

Unicità di u : Sia $u' : E \rightarrow F$ un'applicazione lineare tale che $u'(e_i) = f_i$. Sia $v \in E$ e scriviamo v sulla base B : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$.

Abbiamo $u'(v) = u'(\lambda_1 e_1 + \dots + \lambda_n e_n)$

$$= \lambda_1 u'(e_1) + \dots + \lambda_n u'(e_n) \quad (\text{per linearità, §2, 1.1(iii)})$$

$$= \lambda_1 f_1 + \dots + \lambda_n f_n \quad (\text{perché } u'(e_i) = f_i)$$

$$= u(v).$$

Quindi per ogni v in E , $u'(v) = u(v)$, pertanto $u' = u$.

(ii) (a) Supponiamo u iniettiva e mostriamo che f_1, \dots, f_n sono liberi. Sia $\lambda_1 f_1 + \dots + \lambda_n f_n = 0$ una relazione di dipendenza lineare. Abbiamo $\lambda_1 e_1 + \dots + \lambda_n e_n = u(\lambda_1 e_1 + \dots + \lambda_n e_n) = 0$ e $\lambda_1 e_1 + \dots + \lambda_n e_n \in \text{Ker}(u)$. Siccome u è iniettiva, $\text{Ker}(u) = \{0\}$, quindi $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$. Ma gli e_i sono liberi, e pertanto $\lambda_1 = \dots = \lambda_n = 0$. Abbiamo dimostrato che se $\lambda_1 f_1 + \dots + \lambda_n f_n = 0$ allora $\lambda_1 = \dots = \lambda_n = 0$, quindi f_1, \dots, f_n sono linearmente indipendenti.

(b) Supponiamo adesso i vettori f_1, \dots, f_n linearmente indipendenti e mostriamo che u è iniettiva. Sia $v \in E$, $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Se $u(v) = 0$ allora $\lambda_1 f_1 + \dots + \lambda_n f_n = 0$. Siccome gli f_i sono linearmente indipendenti questo

implica $\lambda_1 = \dots = \lambda_n = 0$, quindi $v = 0$. Questo dimostra $\text{Ker}(u) = \{0\}$ e quindi (§2, 7) u è iniettiva.

(iii) (a) Supponiamo u suriettiva e facciamo vedere che questo implica che f_1, \dots, f_n generano F . Sia $x \in F$. Siccome u è suriettiva esiste v in E tale $u(v) = x$. Se $v = \lambda_1 e_1 + \dots + \lambda_n e_n$ allora $u(v) = \lambda_1 f_1 + \dots + \lambda_n f_n = x$; quindi ogni vettore, x , di F è combinazione lineare di f_1, \dots, f_n .

(b) Supponiamo che i vettori f_1, \dots, f_n generino F e facciamo vedere che questo implica u suriettiva. Se $x \in F$, x è combinazione lineare degli f_i : $x = \lambda_1 f_1 + \dots + \lambda_n f_n$. Poniamo $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Allora, per definizione, $u(v) = x$. Quindi ogni vettore di F è nell'immagine di u e u è suriettiva.

(iv) u è biettiva se e solo se u è iniettiva e suriettiva; quindi se e solo se f_1, \dots, f_n sono linearmente indipendenti e generano F , ossia se e solo se f_1, \dots, f_n formano una base di F ♦

17: Corollario: Due k -spazi vettoriali di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione.

Dim: Siano E, F due k -spazi vettoriali di dimensione finita. Supponiamo $\dim_k E = \dim_k F = n$. Se $n = 0$, $E = F = \{0\}$ e il corollario è vero. Altrimenti sia (e_1, \dots, e_n) (risp. (f_1, \dots, f_n)) una base di E (risp. di F). Definiamo $f : E \rightarrow F$ tramite $f(e_i) = f_i$ (16 (i)). Da 16 (iv) segue che f è biettiva (ossia f è un isomorfismo, cfr §2, 10).

Viceversa se $f : E \rightarrow F$ è un isomorfismo tra due k -spazi vettoriali di dimensione finita, da 16(iv) segue che l'immagine di una base di E è una base di F , quindi $\dim_k E = \dim_k F$ ♦

17.1: Convenzione : Per convenzione si pone $k^0 := \{0\}$.

Si ha allora:

18: Corollario: Ogni k -spazio vettoriale di dimensione n è isomorfo a k^n .

Dim: Segue da 17 perché $\dim_k k^n = n$ (cfr 12.2) ♦

18.1: Osservazione : Modulo isomorfismo c'è un unico k -spazio vettoriale di dimensione n , è k^n .

Esercizi:

4.1) Siano in \mathbb{R}^3 i vettori $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (2, 1, 0)$. Mostrare che questi vettori sono due a due linearmente indipendenti. Mostrare che e_1, e_2, e_3 non sono linearmente indipendenti (cfr Es. 1.7). Sia $v = (4, 3, 0)$ scrivere v come combinazione lineare di e_1, e_2, e_3 in due modi diversi.

4.2) Sia E un k -spazio vettoriale e siano v_1, \dots, v_m m vettori di E . Mostrare che v_1, \dots, v_m sono linearmente indipendenti se e solo se i sottospazi $\langle v_1 \rangle, \dots, \langle v_m \rangle$ sono in somma diretta.

4.3) Nell' \mathbb{R} -spazio vettoriale \mathbb{R}^3 consideriamo i tre vettori x, y, z dati da:
 $x = (1, 2, 1)$; $y = (2, 3, 3)$; $z = (3, 7, 1)$. Dire se questi tre vettori sono linearmente indipendenti.

4.4) Dare un sistema di generatori di C come \mathbb{R} -spazio vettoriale. Dire se $\dim_{\mathbb{R}} C$ è finita o infinita.

4.5) Nell' \mathbb{R} -spazio vettoriale $\mathbb{R}[X]$ si consideri la famiglia $\{X^n\}_{n \in \mathbb{N}}$. Dimostrare che $\{X^n\}_{n \in \mathbb{N}}$ è una famiglia libera.

4.6) Sia E un k -spazio vettoriale e siano v_1, \dots, v_m m vettori di E .

- (i) Se v_1, \dots, v_m sono linearmente indipendenti e se $1 \leq t \leq m$, i vettori di $\{v_1, \dots, v_m\}$ sono linearmente indipendenti.
- (ii) Se v_1, \dots, v_m sono linearmente dipendenti e se $\{w_j\}_{j \in J}$ è una famiglia qualsiasi di vettori di E , la famiglia $\{v_1, \dots, v_m\} \cup \{w_j\}_{j \in J}$ è legata.

4.7) (Versione "forte" del teorema della base incompleta).

Sia E un k -spazio vettoriale e $B = (e_1, \dots, e_n)$ una base di E . Siano v_1, \dots, v_t , t ($1 \leq t \leq n$) vettori di E linearmente indipendenti. Mostrare che esistono $n-t$ vettori di B che insieme a v_1, \dots, v_t formano una base di E .

Dare un procedimento generale per trovare un supplementare di un sottospazio vettoriale F di E dove $E = \mathbb{R}^n, k^n, E$ un k -spazio vettoriale di dimensione n .

4.8) "Il metodo degli scarti successivi":

Scopo di questo esercizio è di dare un metodo effettivo per dimostrare:

Proposizione: Sia E un k -spazio vettoriale di dimensione n e v_1, \dots, v_m , m vettori che generano E . Allora esistono n vettori in $\{v_1, \dots, v_m\}$ che formano una base di E .

Si procede così. Scartiamo tutti i vettori nulli (perché?). D'ora in poi supponiamo tutti gli v_i non nulli. Al primo passo prendiamo v_1 . Poi, al secondo passo, si considera v_2 . Se v_1, v_2 sono legati, v_2 viene scartato e si tiene solo v_1 . Se v_1, v_2 sono linearmente indipendenti, vanno entrambi tenuti. Si continua così: dopo t passi sia A l'insieme dei vettori "tenuti"; se i vettori di $A \cup \{v_{t+1}\}$ sono legati allora v_{t+1} viene scartato, altrimenti viene tenuto. Alla fine (dopo m passi) i vettori tenuti formano una base. Perché?

Se $\dim(E) = n$, i primi n vettori "tenuti" formano una base (i. e. è inutile andare avanti). Dedurne:

Teorema: Sia $E \neq \{0\}$ un k -spazio vettoriale di dimensione finita allora E ammette una base.

Osservazione: combinato con l'esercizio precedente si ottiene un metodo generale effettivo per trovare un supplementare di un sottospazio vettoriale F di E .

4.9) In \mathbb{R}^3 sia $F = \{(x,y,z) / x+y+z = 0 \text{ e } 2x-z = 0\}$.

(i) Dimostrare che F è un sottospazio vettoriale di \mathbb{R}^3 e determinare $\dim(F)$.

(ii) Trovare un supplementare di F . (Suggerimento: usare quanto precede; ossia determinare una base F di F . Sia B la base canonica di \mathbb{R}^3 allora $F \cup B$ genera \mathbb{R}^3 . Usando il metodo degli scarti successivi trovare una base di \mathbb{R}^3 contenente F).

4.10) Siano, in \mathbb{R}^3 , $v_1 = (1,1,0)$, $v_2 = (1,0,2)$, $v_3 = (0,2,2)$. Inoltre sia (e_1, e_2, e_3) la base canonica di \mathbb{R}^3 .

(i) Quali sono le coordinate di v_1, v_2, v_3 rispetto alla base canonica?

(ii) Mostrare che $B = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 e dare le coordinate di e_2 rispetto alla base B .

4.11) Dimostrare che l'insieme degli (x,y,z) di \mathbb{C}^3 che soddisfano le equazioni:

$x+y+z = 0$, $2x+iy-z = 0$, è un sottospazio vettoriale di \mathbb{C}^3 . Determinare la dimensione di questo sottospazio. (nb: qui \mathbb{C}^3 è considerato come \mathbb{C} -spazio vettoriale).

4.12) Siano E, E' due k -spazi vettoriali di dimensione finita. Siano (e_1, \dots, e_n) una base di E e (v_1, \dots, v_t) una base di E' . Si considera $E \times E'$ con la sua struttura prodotto (§1, 3(iv)). Dimostrare che $(e_1, 0), \dots, (e_n, 0), (0, v_1), \dots, (0, v_t)$ è una base di $E \times E'$. Dedurne che i k -spazi vettoriali $k^n \times k^m$, k^{n+m} sono isomorfi (n, m numeri naturali).

4.13) Sia E un k -spazio vettoriale di dimensione finita. Si suppone $E = F \oplus H$ per due sottospazi F, H di E . Sia (f_1, \dots, f_n) una base di F e (h_1, \dots, h_k) una base di H .

Dimostrare che $(f_1, \dots, f_n, h_1, \dots, h_k)$ è una base di E .

4.14) Sia E un k -spazio vettoriale di dimensione finita. Si assume $E = E' \oplus E''$ dove E' e E'' sono due sottospazi vettoriali di E . Siano $x_1 = x'_1 + x''_1, \dots, x_n = x'_n + x''_n$, n vettori di E dove $x'_1 \in E', \dots, x'_n \in E'$, e $x''_1 \in E'', \dots, x''_n \in E''$. Dimostrare:

(i) Se x'_1, \dots, x'_n sono linearmente indipendenti allora x_1, \dots, x_n sono linearmente indipendenti.

(ii) Dare degli esempi in cui x_1, \dots, x_n siano linearmente indipendenti, x'_1, \dots, x'_n siano linearmente dipendenti e dove x''_1, \dots, x''_n siano: (a) linearmente dipendenti (b) linearmente indipendenti (suggerimento: fare un disegno).

(iii) Se x_1, \dots, x_n sono linearmente indipendenti allora: x'_1, \dots, x'_n sono linearmente indipendenti $\Leftrightarrow \langle x_1, \dots, x_n \rangle \cap E'' = \{0\}$. E' ancora vera questa equivalenza se x_1, \dots, x_n non sono linearmente indipendenti?

(iv) Dire se i seguenti vettori u, u', u'' di \mathbb{R}^{2000} sono linearmente indipendenti:

$u = (x_1, \dots, x_{1997}, 1, 0, 0)$, $x_1 = 1$, $x_{i+1} = x_i + 1$, $1 \leq i \leq 1996$; $u' = (y_1, \dots, y_{1997}, 0, 2, 0)$, $y_1 = 2$, $y_{i+1} = y_i - 1$, $1 \leq i \leq 1996$; $u'' = (z_1, \dots, z_{1997}, 0, 0, 3)$, $z_1 = 3$, $z_{i+1} = z_i + 2$, $1 \leq i \leq 1996$.

4.15) In questo esercizio \mathbb{R}^2 ha la sua struttura naturale di \mathbb{R} -spazio vettoriale.

Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x,y) \rightarrow (2x+4y, x+2y)$

(i) Dimostrare che f è un'applicazione \mathbb{R} -lineare

(ii) Dare un sistema di generatori di $\text{Ker}(f)$ e dire se f è iniettiva

(iii) Dare un sistema di generatori di $\text{Im}(f)$ e dire se f è suriettiva.

4.16) Sia $\mathbf{R}[X]$ l' \mathbf{R} -spazio vettoriale dei polinomi a coefficienti reali nella variabile X . Si ricorda che un elemento, $P(X)$, di $\mathbf{R}[X]$ si scrive $P(X) = \sum_{n \geq 0} a_n X^n = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n + \dots$ (si è posto $X^0 = 1$): gli a_i sono numeri reali e sono nulli tranne un numero finito. Se $a_n = 0$ per ogni $n \geq 0$, $P(X)$ è il polinomio nullo (notato 0). Se $P(X)$ non è il polinomio nullo, il grado di $P(X)$ è: $\deg(P(X)) := \max \{n / a_n \neq 0\}$. Osservare che il grado è ben definito perché gli a_n sono tutti nulli tranne un numero finito. Il polinomio nullo non ha grado.

(i) Sia $\mathbf{R}[X]_{\leq k} = \{0\} \cup \{P(X) \in \mathbf{R}[X] / P(X) \neq 0 \text{ e } \deg(P(X)) \leq k\}$, l'insieme dei polinomi di "grado al più k ". Dimostrare che $\mathbf{R}[X]_{\leq k}$ è un sottospazio vettoriale di $\mathbf{R}[X]$. Dare una base di questo sottospazio.

(ii) Sia $V = \{a_0 + \dots + a_k X^k \in \mathbf{R}[X]_{\leq k} / \sum_{0 \leq i \leq k} a_i = 0\}$. Dimostrare che V è un sottospazio vettoriale di $\mathbf{R}[X]$. Trovare una base di V .

4.17) Notiamo $\mathbf{F}_2 = \{0,1\}$, il corpo a due elementi, $\mathbf{Z}/2\mathbf{Z}$. Inoltre sia $E := \mathbf{F}_2 \times \mathbf{F}_2 = (\mathbf{F}_2)^2$. In questo esercizio E viene considerato come un \mathbf{F}_2 -spazio vettoriale.

(i) Qual è la dimensione di E ?

(ii) Calcolare $\text{card}(E)$ (il numero di elementi di E).

(iii) Siano v, w due vettori non nulli, qualsiasi di E . Dimostrare che v e w sono linearmente indipendenti.

(iv) Fare la lista di tutti i sottospazi vettoriali di E .

(v) Sia $u \in E$ il vettore $(1,0)$ e F la retta vettoriale generata da u . Fare la lista di tutti i supplementari di F in E (cfr Es.1.6).

5) IL TEOREMA DELLE DIMENSIONI, RELAZIONE DI GRASSMANN, SUPPLEMENTARI, SPAZI QUOZIENTI.

La prima parte di questo paragrafo presenta due risultati fondamentali di questo corso: il teorema delle dimensioni ($\dim_k E = \dim_k \text{Ker } f + \dim_k \text{Im } f$) e la relazione di Grassmann.

La seconda parte (spazi quoziensi) può essere omessa in prima lettura (tranne la proposizione 10 sull'esistenza di un supplementare), in quanto non verrà usata nel seguito del corso.

La lettura di questa seconda parte è però un ottimo esercizio: il lettore si familiarizzerà ulteriormente con una delle nozioni più difficili ed importanti della matematica (il passaggio al quoziente) e troverà (teorema 12) un risultato valido in dimensione infinita. "Last but not least" gli spazi quoziensi sono alla base dell'algebra omologica, disciplina ormai fondamentale per tutta la geometria "superiore". Si raccomanda quindi vivamente la lettura di questa seconda parte a chi intende proseguire gli studi in matematica "pura".

1: Lemma: Sia $f: E \rightarrow F$ un'applicazione lineare tra due k -spazi vettoriali. Sia $A \subseteq E$ un sottinsieme che genera E (i. e. $\langle A \rangle = E$). Allora $f(A)$ genera $\text{Im}(f)$.

Dim: Sia $y = f(x)$ un elemento di $\text{Im}(f)$. Siccome $\langle A \rangle = E$, x è combinazione lineare di elementi di A : $x = \lambda_1 a_1 + \dots + \lambda_r a_r$. Per linearità $f(x) = \lambda_1 f(a_1) + \dots + \lambda_r f(a_r)$, quindi y è combinazione lineare di elementi di $f(A)$ •

Il teorema delle dimensioni è uno dei risultati fondamentali di questo corso:

2: Teorema: ("delle dimensioni") Sia $f: E \rightarrow F$ un'applicazione lineare tra due k -spazi vettoriali. Si suppone E di dimensione finita. Allora $\text{Im}(f)$ ha dimensione finita e:

$$\dim_k E = \dim_k \text{Im}(f) + \dim_k \text{Ker}(f).$$

Dim: Se $B = (e_1, \dots, e_n)$ è una base qualsiasi di E allora (1) $f(e_1), \dots, f(e_n)$ generano $\text{Im}(f)$. Pertanto $\text{Im}(f)$ ha dimensione finita (§3, 2).

Adesso, siccome E ha dimensione finita, il sottospazio $\text{Ker}(f) \subseteq E$ ha dimensione finita (§4, 15). Se $\text{Ker}(f) = \{0\}$, f è iniettiva (§2, 7) e i vettori $f(e_1), \dots, f(e_n)$ sono linearmente indipendenti (§4, 16(ii)). Quindi formano una base di $\text{Im}(f)$ e $\dim_k \text{Im}(f) = \dim_k E$, $\dim_k \text{Ker}(f) = 0$.

Supponiamo $\text{Ker}(f) \neq \{0\}$. Sia (v_1, \dots, v_r) una base di $\text{Ker}(f)$ (cfr §4, 8). Per il teorema della base incompleta (§4, 9), esistono v_{r+1}, \dots, v_n tali che $B' = (v_1, \dots, v_n)$ sia una base di E . Mostriamo che $f(v_{r+1}), \dots, f(v_n)$ formano una base di $\text{Im}(f)$. Per (1), i vettori $f(v_i)$ generano $\text{Im}(f)$; ma siccome v_1, \dots, v_r sono in $\text{Ker}(f)$, $f(v_1) = \dots = f(v_r) = 0$. Quindi $\text{Im}(f) = \langle f(v_1), \dots, f(v_r), f(v_{r+1}), \dots, f(v_n) \rangle = \langle f(v_{r+1}), \dots, f(v_n) \rangle$. Rimane da vedere che i vettori $f(v_{r+1}), \dots, f(v_n)$ sono linearmente indipendenti. Se $\lambda_{r+1}f(v_{r+1}) + \dots + \lambda_nf(v_n) = 0$, per linearità $f(\lambda_{r+1}v_{r+1} + \dots + \lambda_nv_n) = 0$. Quindi $x := \lambda_{r+1}v_{r+1} + \dots + \lambda_nv_n$ appartiene a $\text{Ker}(f)$. Il vettore x di $\text{Ker}(f)$ è combinazione lineare dei vettori v_1, \dots, v_r che formano una base di $\text{Ker}(f)$: $x = \alpha_1v_1 + \dots + \alpha_rv_r$. Abbiamo quindi: $\lambda_{r+1}v_{r+1} + \dots + \lambda_nv_n = \alpha_1v_1 + \dots + \alpha_rv_r$, ossia $\alpha_1v_1 + \dots + \alpha_rv_r + \lambda_{r+1}v_{r+1} + \dots + \lambda_nv_n = 0$. Siccome (v_1, \dots, v_n) è una base di E questo implica $\alpha_i = 0$, $\lambda_j = 0$, per ogni i, j . Quindi $f(v_{r+1}), \dots, f(v_n)$ formano una base di $\text{Im}(f)$. Pertanto $\dim_k \text{Im}(f) = n - r = \dim_k E - \dim_k \text{Ker}(f)$ ♦

2.1: Osservazione : Nel teorema delle dimensioni non si suppone F di dimensione finita.

3: Corollario: *Sia $f: E \rightarrow F$ un'applicazione lineare tra due k -spazi vettoriali di dimensione finita con $\dim_k E = \dim_k F$. Sono equivalenti:*

- (i) f è biettiva (cioè f è un isomorfismo)
- (ii) f è iniettiva
- (iii) f è suriettiva.

Dim: (i) \Rightarrow (ii): Chiaro.

(ii) \Rightarrow (iii) Se f è iniettiva, $\text{Ker}(f) = \{0\}$. Dal teorema delle dimensioni (2) $\dim_k \text{Im}(f) = \dim_k E$. Per ipotesi $\dim_k E = \dim_k F$, quindi $\dim_k \text{Im}(f) = \dim_k F$. Da (§4, 15) segue che $\text{Im}(f) = F$ cioè f è suriettiva.

(iii) \Rightarrow (i) Se f è suriettiva allora $\text{Im}(f) = F$ e quindi $\dim_k \text{Im}(f) = \dim_k F = \dim_k E$. Dal teorema delle dimensioni segue $\dim_k \text{Ker}(f) = 0$, quindi (§4, 12.1) $\text{Ker}(f) = \{0\}$ e f è iniettiva ♦

3.1: Osservazione : Il corollario 3 non è più vero se E e F hanno dimensione infinita (cfr Es. 7. 2), o se f non è lineare.

Diamo adesso un'altra conseguenza importante di quanto precede:

4: Teorema: ("Relazione di Grassmann") *Siano E un k-spazio vettoriale, F, F' due sottospazi vettoriali di dimensione finita di E. Allora $\dim_k(F+F') = \dim_k F + \dim_k F' - \dim_k(F \cap F')$.*

Dim: Sia $V := F \times F'$ con la struttura di k-spazio vettoriale prodotto (**§1, 3(iv)**). Consideriamo l'applicazione $f: V \rightarrow E: (x, x') \mapsto x+x'$. L'applicazione f è k-lineare (**Es. 1**).

Determiniamo $\text{Ker}(f)$: abbiamo $f(x, x') = 0 \Leftrightarrow x+x' = 0 \Leftrightarrow x' = -x$, e perciò $\text{Ker}(f) = \{(x, -x) \in V / x \in F \cap F'\}$ (è chiaro viceversa che se $x \in F \cap F'$ allora $(x, -x) \in \text{Ker}(f)$). Quindi $\text{Ker}(f)$ è isomorfo a $F \cap F'$ (**Es. 1**). D'altra parte è chiaro che $\text{Im}(f) = F+F'$ (**§1, 7.2**).

Mostriamo adesso che V ha dimensione finita e, più precisamente, $\dim_k V = \dim_k F + \dim_k F'$. Infatti se (e_1, \dots, e_k) è una base di F e se (v_1, \dots, v_t) è una base di F' , si verifica (cfr **Es. 4. 12**) che $(e_1, 0), \dots, (e_k, 0), (0, v_1), \dots, (0, v_t)$ è una base di $F \times F'$.

Finalmente dal teorema delle dimensioni (2): $\dim_k V = \dim_k \text{Im}(f) + \dim_k \text{Ker}(f)$, risulta quindi che: $\dim_k F + \dim_k F' = \dim_k(F+F') + \dim_k(F \cap F')$ ♦

4.1: Osservazione : Nel teorema 4 non si è supposto E di dimensione finita. Vedere negli esercizi altre dimostrazioni della relazione di Grassmann.

5: Corollario: *Siano E un k-spazio vettoriale, F, F' due sottospazi vettoriali di dimensione finita di E. Se $F \cap F' = \{0\}$ e se $\dim_k F + \dim_k F' = \dim_k E$ allora $E = F \oplus F'$.*

Dim: Per ipotesi la somma $F+F'$ è diretta, basta far vedere che $F+F' = E$. Dalla relazione di Grassmann $\dim_k(F+F') = \dim_k F + \dim_k F' = \dim_k E$, quindi (**§4, 15**) $F+F' = E$ ♦

SPAZI QUOZIENTI.

6: Proposizione: *Siano E un k-spazio vettoriale, F un sottospazio vettoriale di E. La relazione binaria su E: $x R y \Leftrightarrow x - y \in F$, è una relazione d'equivalenza.*

Dim: Verifichiamo gli assiomi di relazione d'equivalenza (cfr **I, §4**).

Riflessività: $x R x \Leftrightarrow x - x = 0 \in F$, questo è vero perché F è un sottospazio vettoriale.

Simmetria: $x R y \Rightarrow y R x$: $x R y$ è equivalente a $u \in F$ dove $u = x - y$. Siccome F è un sottospazio vettoriale, anche $-u \in F$, ossia $y - x \in F$, o ancora $y R x$.

Transitività: $x R y$, e $y R z \Rightarrow x R z$: abbiamo $u = x - y \in F$, e $v = y - z \in F$. Siccome F è un sottospazio vettoriale $u + v = x - z \in F$, quindi $x R z$ ♦

7: Definizione: L'insieme quoziante di E per la relazione d'equivalenza: $x R y \Leftrightarrow x - y \in F$ si nota E/F .

Sia $p : E \rightarrow E/F$ l'applicazione che ad ogni x associa la sua classe d'equivalenza. Per facilitare la scrittura poniamo $p(x) = [x]$. Altre notazioni sono: $p(x) = \bar{x}$, $p(x) = x+F (= x \pmod{F}, \dots)$.

8: Proposizione: Siano E un k -spazio vettoriale, $F \subseteq E$ un sottospazio vettoriale. L'insieme quoziante E/F è munito, in modo naturale, di una struttura di k -spazio vettoriale tramite: $[x] + [y] = [x+y]$, $\lambda[x] = [\lambda x]$, $\forall (x, y) \in E^2$, $\forall \lambda \in k$. Con questa struttura l'applicazione $p : E \rightarrow E/F$ è k -lineare.

Dim: Prima di tutto bisogna verificare che la legge interna di addizione e la legge esterna di moltiplicazione siano ben definite, ossia non dipendano dalla scelta di un rappresentante delle classi. Guardiamo prima l'addizione. Sia $x R x'$ e sia $y R y'$ allora $[x] = [x']$ e $[y] = [y']$; dobbiamo quindi vedere che: $[x] + [y] = [x'] + [y']$. Per definizione $[x] + [y] = [x+y]$, quindi dobbiamo mostrare $[x+y] = [x'+y']$, ossia $(x+y) R (x'+y')$. L'ultima relazione è equivalente a $(x+y) - (x'+y') \in F$. Possiamo riscrivere questa relazione come $(x-x') + (y-y') \in F$. Per ipotesi $x R x'$, quindi $x - x' \in F$; nello stesso modo $y - y' \in F$. Finalmente F essendo un sottospazio vettoriale $(x-x') + (y-y') \in F$. Quindi l'addizione delle classi è ben definita. Passiamo adesso alla legge esterna. Dobbiamo vedere che se $x R x'$ allora $\lambda[x] = \lambda[x']$, o per definizione $[\lambda x] = [\lambda x']$. Questo è equivalente a $\lambda x - \lambda x' \in F$. L'ultima relazione si può riscrivere $\lambda(x - x') \in F$. Per ipotesi $x R x'$ ossia $x - x' \in F$, e F essendo un sottospazio vettoriale, si ha $\lambda(x - x') \in F$. Quindi anche la legge esterna è ben definita.

Adesso bisogna verificare che l'addizione e la moltiplicazione soddisfano gli assiomi di spazio vettoriale. Queste verifiche non presentano difficoltà e sono lasciate al lettore, indichiamo come procedere:

Con quanto precede abbiamo mostrato: $\forall (x, y) \in E^2$, $\forall (\lambda, \mu) \in k^2$, $p(\lambda x + \mu y) = \lambda p(x) + \mu p(y)$. Infatti $p(\lambda x + \mu y) = [\lambda x + \mu y]$ (per definizione!), per come è definita l'addizione: $[\lambda x + \mu y] = [\lambda x] + [\mu y]$, e per come è definita la moltiplicazione: $[\lambda x] + [\mu y] = \lambda[x] + \mu[y] = \lambda p(x) + \mu p(y)$. Quindi, in qualche modo (non sappiamo ancora

che E/F è uno spazio vettoriale), p è lineare. Adesso, visto che p è suriettiva, per verificare un assioma, basta "risalire" con p in E , dove l'assioma è vero, e per "linearità" di p , ridiscendere in E/F . Se per esempio vogliamo verificare $[x] + [y] = [y] + [x]$. Abbiamo $[x] + [y] = p(x) + p(y) = p(x+y)$ ("linearità"), $= p(y+x)$ ($x+y = y+x$ in E), $= p(y) + p(x)$ ("linearità"), $= [y] + [x]$. Osserviamo per concludere che lo zero di $(E/F, +)$ è $[0]$ (e $[f] = [0]$, $\forall f \in F$)♦

9: Teorema: Siano E un k -spazio vettoriale, F e G due sottospazi vettoriali di E . Se $E = F \oplus G$ allora esiste un isomorfismo: $E/F \xrightarrow{\cong} G$.

Dim: Sia $p: E \rightarrow E/F$ l'applicazione che ad ogni x in E associa la sua classe $[x]$. Notiamo p' la restrizione di p a G : $p': G \rightarrow E/F$: $y \rightarrow [y]$.

L'applicazione p' è lineare perché p lo è (8).

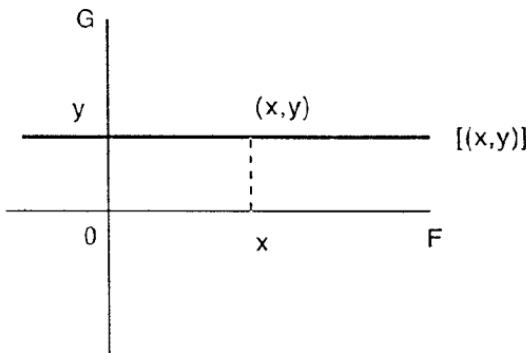
L'applicazione p' è iniettiva: $\text{Ker}(p') = \{y \in G / [y] = [0]\}$, $[y] = [0]$ è equivalente a $y - 0 = y \in F$, ma siccome F e G sono in somma diretta $F \cap G = \{0\}$, quindi $y = 0$.

L'applicazione p' è suriettiva: sia $[x] \in E/F$. Siccome $E = F \oplus G$, $x = f + g$, con $f \in F$, $g \in G$. Abbiamo $x - g = f \in F$, quindi $[x] = [g]$ e pertanto $p'(g) = [x]$.

In conclusione p' è lineare e biiettiva, quindi è un isomorfismo lineare tra G e E/F ♦

9.1: Osservazione : (i) Con le notazioni precedenti l'applicazione reciproca $q = p^{-1}: E/F \rightarrow G$ è data nel modo seguente: sia $[x] \in E/F$, siccome $E = F \oplus G$, $x = f + g$, $f \in F$, $g \in G$, in modo unico. Si definisce $q([x]) = g$.

9.2: Esempio : Sia $E = \mathbb{R}^2$, $F = \{(x, 0) / x \in \mathbb{R}\}$, $G = \{(0, y) / y \in \mathbb{R}\}$. Chiaramente F e G sono sottospazi vettoriali di E e $E = F \oplus G$. Si ha $(x, y) R (x', y') \Leftrightarrow (x-x', y-y') \in F$, cioè $(x, y) R (x', y') \Leftrightarrow y = y'$. L'insieme E/F è l'insieme delle classi d'equivalenza $[(x, y)]$, e $[(x, y)] = \{(x', y') / y' = y\}$.

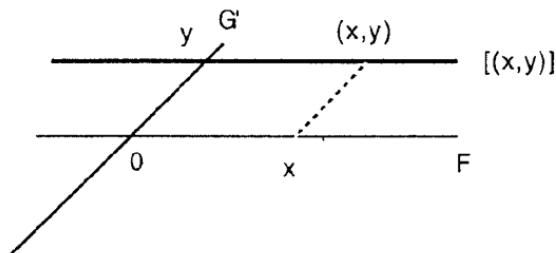


Vediamo così che E/F , l'insieme delle classi d'equivalenza, è l'insieme delle rette parallele a F . Con questa rappresentazione possiamo descrivere geometricamente le applicazioni p , q :

-la biiezione $q: E/F \rightarrow G$: dato un elemento di E/F , cioè una retta R parallela a F , gli si associa l'elemento di G dato da $G \cap R$.

-la biiezione $p: G \rightarrow E/F$: dato un elemento di G , g , gli si associa l'unica retta, R , passante per g e parallela a F ($R = F$ se $g = (0, 0)$).

9.3: Osservazione : Benché E/F sia isomorfo a G , E/F non è in modo "naturale" un sottospazio vettoriale di E ("naturale" = "in un modo più bello degli altri"!). Per convincersene prendere G' al posto di G :



e ripetere tutto quello che precede (non c'è nessuna differenza, E/F è "altrettanto" isomorfo a G che a G').

Per quanto riguarda l'esistenza di supplementari abbiamo:

10: Proposizione: Siano $E \neq \{0\}$ un k -spazio vettoriale di dimensione finita e $F \subseteq E$ un sottospazio vettoriale. Allora F ammette almeno un supplementare (cioè esiste un sottospazio vettoriale, G , di E tale che $F \oplus G = E$).

Dim: Osserviamo che F ha dimensione finita (§4, 15). Se $F = \{0\}$ (risp. $F = E$), la proposizione è vera con $G = E$ (risp. $G = \{0\}$). Supponiamo quindi $\dim_k F > 0$, $F \neq E$. Da (§4, 8), F ammette una base (f_1, \dots, f_p) . I vettori f_1, \dots, f_p sono linearmente indipendenti in E , e per il teorema della base incompleta (§4, 9) esistono dei vettori f_{p+1}, \dots, f_n tali che $(f_1, \dots, f_p, f_{p+1}, \dots, f_n)$ sia una base di E . Sia $G = \langle f_{p+1}, \dots, f_n \rangle$ il sottospazio di E generato da f_{p+1}, \dots, f_n . Mostriamo che $E = F \oplus G$. Siccome (f_1, \dots, f_n) è una base di E , se $x \in E$ allora $x = (\lambda_1 f_1 + \dots + \lambda_p f_p) + (\lambda_{p+1} f_{p+1} + \dots + \lambda_n f_n) = f + g$, quindi $F+G = E$. Vediamo adesso che $F \cap G = \{0\}$. Sia $x \in F \cap G$, allora $x = \alpha_1 f_1 + \dots + \alpha_p f_p$ (perché $x \in F$ e f_1, \dots, f_p formano una base di F). Nello stesso modo $x = \beta_1 f_{p+1} + \dots + \beta_n f_n$. Risulta: $\alpha_1 f_1 + \dots + \alpha_p f_p = \beta_1 f_{p+1} + \dots + \beta_n f_n$, ossia $\alpha_1 f_1 + \dots + \alpha_p f_p - \beta_1 f_{p+1} - \dots - \beta_n f_n = 0$. Siccome gli f_i sono linearmente indipendenti, questo implica $\alpha_k = 0, \beta_k = 0$, per ogni k , p . Pertanto $x = 0$ ♦

10.1: Osservazione : Se il campo k è infinito e se $F \neq \{0\}$, $F \neq E$, F ammette un'infinità di supplementari. Questo non è più vero se il campo k è finito (cfr Es. 2).

11: Corollario: Siano E un k -spazio vettoriale di dimensione finita e $F \subseteq E$ un sottospazio vettoriale. Allora E/F ha dimensione finita e $\dim_k E/F = \dim_k E - \dim_k F$.

Dim: Segue da 9, 10 ♦

12: Teorema: ("di fattorizzazione") Sia $f: E \rightarrow V$ un'applicazione lineare tra due k -spazi vettoriali.

(i) f fattorizza attraverso l'applicazione canonica $p: E \rightarrow E/\text{Ker}(f)$, cioè esiste un'applicazione lineare $\varphi: E/\text{Ker}(f) \rightarrow V$ tale che $f = \varphi \circ p$.

(ii) Inoltre $\text{Im}(\varphi) = \text{Im}(f)$ e $\varphi: E/\text{Ker}(f) \rightarrow \text{Im}(f)$ è un isomorfismo ($\varphi'(x) := \varphi(x)$, è cambiato solo il codominio).

Dim: Se $[x] \in E/\text{Ker}(f)$ poniamo $\varphi([x]) = f(x)$. Dobbiamo mostrare che l'applicazione φ è ben definita ossia che se x e x' sono equivalenti modulo $\text{Ker}(f)$ allora $\varphi([x]) = \varphi([x'])$. Se x e x' sono equivalenti mod $\text{Ker}(f)$ allora $x - x' \in$

$\text{Ker}(f)$ e quindi $f(x) = f(x')$, perciò φ è ben definita. La verifica che φ è lineare è lasciata al lettore.

E' chiaro dalle definizioni che $p \cdot \varphi = f$.

Mostriamo che $\text{Im}(\varphi) = \text{Im}(f)$. Se $v \in \text{Im}(f)$ allora $v = f(x) = \varphi([x])$ e $v \in \text{Im}(\varphi)$. E' altrettanto chiaro che $\text{Im}(\varphi) \subseteq \text{Im}(f)$.

Siccome $\text{Im}(\varphi) = \text{Im}(f)$ l'applicazione $\varphi': E/\text{Ker}(f) \rightarrow \text{Im}(f) : [x] \mapsto \varphi([x])$ è suriettiva. Osserviamo che φ' è lineare perchè φ lo è. Finalmente mostriamo che φ' è iniettiva. Supponiamo $\varphi([x]) = 0$. Per definizione $\varphi([x]) = f(x)$, quindi $x \in \text{Ker}(f)$ e pertanto $[x] = 0$, quindi φ e φ' sono iniettive♦

12.1: Osservazione : Il teorema 12 non richiede ipotesi sulle dimensioni di E , V .

13: Il teorema delle dimensioni (bis).

Per concludere vediamo come il teorema delle dimensioni (2) è conseguenza del teorema di fattorizzazione (12):

Siano E, F , due k -spazi vettoriali e $f: E \rightarrow F$ un morfismo. Da (12): $E/\text{Ker}(f) \cong \text{Im}(f)$ (" \cong " significa "sono isomorfi"). Se E ha dimensione finita, da (11): $\dim_k(E/\text{Ker}(f)) = \dim_k E - \dim_k \text{Ker}(f)$. Quindi $\dim_k \text{Im}(f) = \dim_k E - \dim_k \text{Ker}(f)$ ♦

Siccome (11) e (12) sono stati dimostrati senza usare il teorema delle dimensioni, questo fornisce una dimostrazione alternativa del teorema delle dimensioni. Inoltre l'isomorfismo $E/\text{Ker}(f) \cong \text{Im}(f)$ è valido anche se E ha dimensione infinita.

Esercizi:

5.1) Completare i dettagli della dimostrazione del teorema 4.

5.2) (i) Sia E un k -spazio vettoriale di dimensione $n > 0$. Dimostrare che E è un insieme infinito se e solo se k è un campo infinito. Qual è la cardinalità di E se $\text{card}(k) = p$?
(ii) Si suppone k infinito. Sia F un sottospazio non banale di E (i.e. $F \neq E$ e $F \neq \{0\}$). Dimostrare che F ammette un'infinità di supplementari (cfr Es. 1.6, 4.17).

(suggerimento: sia $E = F \oplus G$, (f_1, \dots, f_p) una base di F , (g_{p+1}, \dots, g_n) una base di G . Se $v \neq 0$ è un vettore di F si ponga: $g'_i = g_i + v$. Mostrare che $G' = \langle g'_{p+1}, \dots, g'_n \rangle$ è un supplementare di F e che $G' \neq G$).

5.3) Sia $M_2(\mathbb{R})$ lo spazio vettoriale reale delle matrici 2×2 . La matrice $M = (a_{ij})$, $1 \leq i \leq 2, 1 \leq j \leq 2$, si dice simmetrica se $a_{ij} = a_{ji}$ per ogni i, j ; e antisimmetrica se $a_{ij} = -a_{ji}$ per ogni i, j . Si nota $\text{Sim}_2(\mathbb{R})$ (risp. $A_2(\mathbb{R})$) l'insieme delle matrici simmetriche (risp. antisimmetriche).

- (i) Dimostrare che $\text{Sim}_2(\mathbf{R})$ e $A_2(\mathbf{R})$ sono due sottospazi vettoriali di $M_2(\mathbf{R})$ e calcolare la loro dimensione.
(ii) Usando la relazione di Grassmann, dimostrare che $\text{Sim}_2(\mathbf{R}) \oplus A_2(\mathbf{R}) = M_2(\mathbf{R})$.
(iii) Trovare un supplementare di $A_2(\mathbf{R})$ in $M_2(\mathbf{R})$ diverso da $\text{Sim}_2(\mathbf{R})$.

5.4) Siano E_i , $1 \leq i \leq n$, dei k -spazi vettoriali, $f_i: E_i \rightarrow E_{i+1}$, delle applicazioni lineari. Questi dati formano un complesso di spazi vettoriali (indicato $\dots \rightarrow E_k \cdot f_k \rightarrow E_{k+1} \cdot f_{k+1} \rightarrow E_{k+2} \rightarrow \dots$) se $\text{Im}(f_i) \subseteq \text{Ker}(f_{i+1})$ (i.e. $f_{i+1} \circ f_i = 0$). Gli spazi quoienti $\text{Ker}(f_{i+1})/\text{Im}(f_i)$ sono i gruppi d'omologia del complesso. Il complesso è esatto in E_{k+1} se $\text{Im}(f_k) = \text{Ker}(f_{k+1})$. Il complesso è esatto se è esatto in ogni E_i . Una successione esatta (corta) $0 \rightarrow E \rightarrow F \rightarrow G \rightarrow 0$ è un complesso esatto in E, F, G (0 sta ad indicare lo spazio vettoriale banale $\{0\}$).

- (i) Sia $0 \rightarrow E \rightarrow F \rightarrow G \rightarrow 0$ una successione esatta di spazi vettoriali di dimensione finita. Dimostrare che $E \rightarrow F$ è iniettiva e $F \rightarrow G$ è suriettiva. Dimostrare che $\dim(F) = \dim(E) + \dim(G)$.
(ii) Siano F, F' due sottospazi vettoriali di un k -spazio vettoriale E . Dimostrare che si ha una successione esatta: $0 \rightarrow F \cap F' \rightarrow F \times F' \rightarrow F + F' \rightarrow 0$. Dedurne la relazione di Grassmann.
(iii) Se F è un sottospazio vettoriale di E si ha una successione esatta:
 $0 \rightarrow F \rightarrow E \rightarrow E/F \rightarrow 0$.
(iv) Se $f: E \rightarrow F$ è un'applicazione lineare si ha un complesso esatto: $0 \rightarrow \text{Ker}(f) \rightarrow E \cdot f \rightarrow F \rightarrow F/\text{Im}(f) \rightarrow 0$.

6) ANELLO DEGLI ENDOMORFISMI, GRUPPO LINEARE.

In questo paragrafo vediamo che le operazioni naturali di somma, prodotto per uno scalare, composizione conferiscono, all'insieme delle applicazioni lineari da un k-spazio vettoriale in un altro, certe strutture algebriche.

Siano E, F due k-spazi vettoriali. Noteremo con $L_k(E,F)$ o anche con $\text{Hom}_k(E,F)$ l'insieme delle applicazioni k-lineari da E in F.

Se f, g sono due elementi di $L_k(E,F)$ definiamo $f+g$ tramite: per ogni x in E, $(f+g)(x) := f(x)+g(x)$.

Se $\lambda \in k$ definiamo λf tramite: $(\lambda f)(x) := \lambda \cdot f(x)$, per ogni x in E. Si verifica subito:

1: Proposizione: L'addizione e la moltiplicazione esterna definite qui sopra conferiscono a $L_k(E,F)$ una struttura di k-spazio vettoriale.

1.1: Osservazione : Se E ed F hanno dimensione finita allora anche $L_k(E,F)$ ha dimensione finita, più precisamente: $\dim(L_k(E,F)) = \dim(E) \cdot \dim(F)$ (cfr Es. 2)

Composizione delle applicazioni lineari:

Siano E, F, G tre k-spazi vettoriali. Se $f \in L_k(E,F)$ e $g \in L_k(F,G)$ allora $g \circ f \in L_k(E,G)$ (Es. 3). Inoltre se $f \in L_k(E,F)$ e $g' \in L_k(F,G)$ allora si verifica (Es. 3):

$$g' \circ (f+f') = g' \circ f + g' \circ f' ; \quad (g+g') \circ f = g \circ f + g' \circ f.$$

L'anello degli endomorfismi di uno spazio vettoriale:

Se $E = F$ invece di $L_k(E,E)$ si scrive $L_k(E)$ o ancora $\text{End}_k(E)$. Se poi il campo k è fissato dal contesto si scrive più semplicemente $\text{End}(E)$. Un elemento, f , di $\text{End}(E)$ è detto endomorfismo (o operatore lineare) di E. Un endomorfismo di E è quindi un'applicazione lineare da E in E: $f: E \rightarrow E$, f k-lineare.

Se f, g sono degli elementi di $\text{End}(E)$ allora anche $f \circ g$ appartiene ad $\text{End}(E)$; come pure $f+g$. Abbiamo così due leggi di composizione interne su $\text{End}(E)$: $+$ e \circ . Si verifica subito che $(\text{End}(E), +, \circ)$ è un anello. Quest'anello in generale non è commutativo (i.e. esistono f, g tali che $f \circ g \neq g \circ f$) e non è integro (i.e. esistono $f \neq 0, g \neq 0$ tali che $f \circ g = 0$: vedere Es. 4).

Abbiamo quindi due strutture su $\text{End}(E)$: $\text{End}(E)$ è un k-spazio vettoriale (per $+$ e la moltiplicazione esterna), inoltre $\text{End}(E)$ è un anello per $+$ e \circ . Le tre

operazioni $+$, \circ , e la moltiplicazione esterna sono compatibili tra di loro: $\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$, per ogni λ in k e per ogni f, g in $\text{End}(E)$. Questo conferisce a $\text{End}(E)$ una struttura di **k-algebra**.

Il gruppo degli automorfismi di uno spazio vettoriale (gruppo lineare $\text{GL}_k(E)$).

2: Definizione: Sia E un k -spazio vettoriale. Un automorfismo lineare di E è un'applicazione lineare $f: E \rightarrow E$, biettiva.

2.1: Osservazione: Abbiamo visto (cfr Es. 2.4) che se f è un automorfismo allora l'applicazione reciproca f^{-1} è anch'essa lineare.

Sia $\text{Aut}_k(E)$ (o $\text{GL}_k(E)$) l'insieme degli automorfismi lineari di E . Per la composizione delle applicazioni ($\text{GL}_k(E), \circ$) è un gruppo chiamato **il gruppo lineare di E** (del k -spazio vettoriale E).

In generale questo gruppo non è abeliano.

Esercizi:

6.1) Dimostrare la proposizione 1.

6.2) Siano E, F dei k -spazi vettoriali di dimensione finita. Scegliendo delle basi in E, F e usando il fatto che un'applicazione lineare è completamente determinata dalle immagini dei vettori di una base (cfr §4, 16 (i)), dimostrare che $\dim_k(L_k(E, F)) = \dim_k(E) \cdot \dim_k(F)$.

6.3) Verificare tutte le proprietà, asserite in questo paragrafo, della composizione delle applicazioni lineari.

6.4) Sia E un k -spazio vettoriale di dimensione finita ≥ 2 . Dimostrare che l'anello $\text{End}(E)$ è non commutativo (i.e. esistono f, g in $\text{End}(E)$ tali che $f \circ g \neq g \circ f$). Dimostrare che $\text{End}(E)$ contiene elementi nilpotenti non nulli (i.e. esiste $h \neq 0$ tale che $h^n = 0$ dove $h^n = h \circ h \dots \circ h$, n termini). Questo dimostra che $\text{End}(E)$ non è integro (i.e. esistono $s \neq 0, t \neq 0$ in $\text{End}(E)$ tali che $s \circ t = 0$). Trovare s, t , non nulli, $s \neq t$, tali che $s \circ t = 0$. (Suggerimento: iniziare con $\dim(E) = 2$ e usare §4, 16).

6.5) Sia E un k -spazio vettoriale di dimensione finita. E' ($\text{Gl}(E), +$) un gruppo?

7) FORME LINEARI E DUALITÀ.

Scopo di questo paragrafo è di introdurre il duale e il biduale di un k-spazio vettoriale, E. Il duale di E è l'insieme delle applicazioni lineari: $f: E \rightarrow k$ e si nota E^* (o E° o $\text{Hom}_k(E, k)$ ecc...). Da quanto precede (§6), E^* è un k-spazio vettoriale. Il biduale, notato E^{**} , è il duale di E^* . Il risultato principale è la corrispondenza ("dualità"), tramite l'ortogonalità e l'isomorfismo canonico tra E e E^{**} , tra sottospazi di E e sottospazi di E^* . Questa dualità è fondamentale in geometria proiettiva.

Sia E un k-spazio vettoriale. Si ricorda che k può essere considerato come spazio vettoriale (di dimensione uno) su se stesso.

1: Definizione: Un'applicazione lineare $f: E \rightarrow k$ (k con la sua struttura naturale di k-spazio vettoriale) si dice forma lineare su E. L'insieme di tutte le forme lineari su E si chiama il duale di E e si nota E^* .

Il duale di E^* si nota E^{**} e si chiama il biduale di E.

2: Osservazione : Risulta da §6, 1 che E^* è un k-spazio vettoriale. Inoltre se $\dim_k(E) = n$ allora (Es. 6.2) $\dim_k(E^*) = n$. Nello stesso modo E^{**} è un k-spazio vettoriale e $\dim_k(E) = \dim_k(E^*) = \dim_k(E^{**})$. Questo risulta anche da:

3: Lemma: Sia E un k-spazio vettoriale di dimensione finita e (e_1, \dots, e_n) una base di E.

- (i) Esiste una base (e^*_1, \dots, e^*_n) di E^* tale che $e^*_j(e_i) = \delta_{ij}$, per ogni j, i ($\delta_{ij} = 0$ se $i \neq j$, $\delta_{ii} = 1$)
- (ii) $\dim(E) = \dim(E^*) = \dim(E^{**})$.

Dim: (i) Useremo ripetutamente il fatto che un'applicazione lineare è completamente determinata dai suoi valori sui vettori di una base (cfr §4, 16).

Sia k , $1 \leq k \leq n$; da §4, 16 esiste un'applicazione lineare $f_k: E \rightarrow k$ tale che $f_k(e_j) = 0$ se $j \neq k$ e $f_k(e_k) = 1$. Abbiamo così definito f_1, \dots, f_n . Mostriamo adesso che le forme lineari f_1, \dots, f_n generano E^* . Sia $g: E \rightarrow k$ un vettore di E^* ; g è completamente determinato dagli scalari $g(e_i) = \lambda_i$, $1 \leq i \leq n$ (sempre per §4, 16). Consideriamo $G = \lambda_1 f_1 + \dots + \lambda_n f_n$; G è una forma lineare. Abbiamo $G(e_i) =$

$(\lambda_1 f_1 + \dots + \lambda_n f_n)(e_i) = \lambda_1 f_1(e_i) + \dots + \lambda_n f_n(e_i) = \lambda_i$. Quindi G e g coincidono sulla base (e_1, \dots, e_n) ; da §4, 16 segue che le applicazioni lineari G e g sono uguali. In conclusione $g = \lambda_1 f_1 + \dots + \lambda_n f_n$ e ogni forma lineare, g , è combinazione lineare di f_1, \dots, f_n .

Per concludere mostriamo che f_1, \dots, f_n sono liberi. Sia $\alpha_1 f_1 + \dots + \alpha_n f_n = 0$, una relazione di dipendenza lineare. Per ogni j , $1 \leq j \leq n$, abbiamo: $(\alpha_1 f_1 + \dots + \alpha_n f_n)(e_j) = 0$ ma $(\alpha_1 f_1 + \dots + \alpha_n f_n)(e_j) = \alpha_1(f_1)(e_j) + \dots + \alpha_n(f_n)(e_j) = \alpha_j$. Quindi $\alpha_j = 0$ per ogni j e f_1, \dots, f_n sono liberi. Ponendo $e_i^* = f_i$ si ottiene (i).

(ii) Da (i) segue che $\dim(E) = n = \dim(E^*)$. Ponendo $F = E^*$, $F^* = E^{**}$ possiamo ripetere la dimostrazione precedente con F al posto di E e concludere che $\dim(F^*) = \dim(E^{**})$. Come già osservato quest'ultima parte segue anche da Es. 6.2♦

4: Definizione: Con le notazioni del lemma 3, la base (e_1^*, \dots, e_n^*) di E^* si chiama la base duale di (e_1, \dots, e_n) . Si ha quindi $e_j^*: E \rightarrow k$, forma lineare con $e_j^*(e_p) = \delta_{jp}$.

4.1: Osservazione : Il simbolo δ_{jp} si chiama simbolo di Kronecker (L.Kronecker, 1823-1891, matematico tedesco noto per i suoi lavori in teoria dei numeri, tra le altre cose).

5: Esempio : Sia $E = \mathbb{R}^n$ e $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$ la base canonica di \mathbb{R}^n . Sia e_i^* , $1 \leq i \leq n$, la base duale della base canonica. Per definizione $e_k^*(e_i) = \delta_{ki}$. Se $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, allora $v = \alpha_1 e_1 + \dots + \alpha_n e_n$, e $e_k^*(v) = e_k^*(\alpha_1 e_1 + \dots + \alpha_n e_n) = \alpha_1 e_k^*(e_1) + \dots + \alpha_k e_k^*(e_k) + \dots + \alpha_n e_k^*(e_n) = \alpha_k$. In altri termini e_k^* è la funzione k -esima coordinata: $\mathbb{R}^n \rightarrow \mathbb{R}: (\alpha_1, \dots, \alpha_n) \rightarrow \alpha_k$. Consideriamo il polinomio $P(X_1, \dots, X_n) = X_k$; la funzione polinomiale associata non è altro che e_k^* . Quindi possiamo identificare e_k^* con il polinomio X_k . Siccome ogni elemento di \mathbb{R}^{n*} è combinazione lineare degli e_i^* , vediamo che ogni forma lineare su \mathbb{R}^n si identifica ad un polinomio omogeneo di grado uno nelle variabili X_1, \dots, X_n , ossia ad un polinomio del tipo $\lambda_1 X_1 + \dots + \lambda_n X_n$, $\lambda_i \in \mathbb{R}$ (omogeneo significa che il termine costante del polinomio è nullo, vedere 12). In conclusione:

$\mathbb{R}^{n*} \cong \{\text{polinomi omogenei di grado uno, a coefficienti reali, nelle variabili } X_1, \dots, X_n\} \cong \{\lambda_1 X_1 + \dots + \lambda_n X_n / \lambda_i \in \mathbb{R}\}$.

Quanto detto sopra è valido per un campo qualsiasi k : la funzione i -esima coordinata $k^n \rightarrow k: (\alpha_1, \dots, \alpha_n) \rightarrow \alpha_i$ è la funzione polinomiale associata al

polinomio $P_i(X_1, \dots, X_n) = X_i$, e pertanto: $k^{n*} \cong \{\text{polinomi omogenei di grado uno, a coefficienti in } k, \text{ nelle variabili } X_1, \dots, X_n\} \cong \{\lambda_1 X_1 + \dots + \lambda_n X_n / \lambda_i \in k\}$.

Finalmente, sia E un k -spazio vettoriale di dimensione n . Scegliendo una base (v_i) di E stabiliamo un isomorfismo, f , tra E e k^n (§4, 18) tramite $f(v_i) = e_i$. Pertanto, tramite scelta di una base di E (e della base duale di E^*) possiamo identificare E^* con lo spazio vettoriale dei polinomi omogenei di grado uno, a coefficienti in k , nelle variabili X_1, \dots, X_n .

6: Equazioni cartesiane: In R^2 l'equazione cartesiana di una retta, L , passante per l'origine è della forma $aX + bY = 0$, o ancora $aX_1 + bX_2 = 0$. La retta L è quindi l'insieme soluzione dell'equazione $P(X_1, X_2) = 0$ dove $P(X_1, X_2)$ è il polinomio omogeneo $aX_1 + bX_2$. Ogni retta ha un'equazione cartesiana. Inoltre se D è l'insieme delle soluzioni dell'equazione $c(aX_1 + bX_2) = 0$ ($c \in R$, $c \neq 0$), allora $D = L$. Abbiamo quindi una corrispondenza tra {rette passanti per l'origine} e {equazioni $aX+bY=0$ definite a meno di una costante}. Questa è la dualità tra rette ed equazioni che adesso generalizzeremo al caso di uno spazio vettoriale (di dimensione finita) qualsiasi.

7: Lemma: Sia E un k -spazio vettoriale di dimensione n e $f: E \rightarrow k$ una forma lineare non nulla. Allora f è suriettiva e $\dim(\ker f) = n-1$.

Dim: Se f non è nulla esiste $v \in E$ tale che $f(v) \neq 0$. Quindi $\text{Im}(f)$ contiene un vettore non nullo e pertanto $\dim(\text{Im } f) \geq 1$. D'altra parte $\dim(k) = 1$ quindi (§4, 15) $\text{Im } f = k$ ossia f è suriettiva. Dal teorema delle dimensioni (§5, 2) segue che $\dim(\text{Ker } f) = n-1$ ♦

8: Definizione: Sia E un k -spazio vettoriale di dimensione n e F un sotto spazio vettoriale di E .

Se $\dim(F) = 1$ si dice che F è una retta (vettoriale).

Se $\dim(F) = n-1$ si dice che F è un iperpiano (vettoriale); se $n = 3$ si dice anche piano invece di iperpiano.

9: Proposizione: Sia E un k -spazio vettoriale di dimensione n .

(i) Sia H un iperpiano di E , allora esiste una forma lineare $f: E \rightarrow k$ tale che $\text{Ker}(f) = H$.

(ii) Due forme lineari non nulle su E , f, g sono proporzionali (i.e. legate, i.e. $g = \alpha f$ per qualche α in k) se e solo se $\text{Ker}(f) = \text{Ker}(g)$.

Dim: (i) Sia P un supplementare di H (§5, 10), osservare che $\dim(L) = 1$. Siano (h_1, \dots, h_{n-1}, p) delle basi di H e P . Allora (h_1, \dots, h_{n-1}, p) è una base di E . Esiste un'applicazione lineare $f: E \rightarrow k$ tale che $f(h_1) = \dots = f(h_{n-1}) = 0$ e $f(p) = 1$ (§4, 16). Ogni h in H si scrive come combinazione lineare dei vettori della base: $h = \lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1}$, e per linearità di f : $f(h) = \lambda_1 f(h_1) + \dots + \lambda_{n-1} f(h_{n-1}) = 0$. Quindi H è contenuto in $\text{Ker}(f)$; siccome $f \neq 0$ ($f(p) = 1$), da 3 (e §4, 15) segue che $H = \text{Ker}(f)$.

(ii) Supponiamo $g = \alpha f$. Siccome f, g non sono identicamente nulle, $\alpha \neq 0$. Pertanto $g(x) = 0$ se e solo se $f(x) = 0$, quindi $\text{Ker}(f) = \text{Ker}(g)$.

Viceversa supponiamo $\text{Ker}(f) = \text{Ker}(g)$ e facciamo vedere che f e g sono proporzionali. Notiamo con H l'iperpiano $\text{Ker}(f) = \text{Ker}(g)$ e sia P un supplementare di H . Inoltre sia p un vettore non nullo di P (quindi una base di P). Abbiamo $f(p) = \lambda$, $g(p) = \mu$ per degli scalari λ, μ . Osserviamo che $\lambda \neq 0, \mu \neq 0$ perché f, g non sono identicamente nulle. Poniamo $\alpha = \mu/\lambda$. Se v è un vettore di $E = H \approx P$ allora $v = h + \beta p$ in modo unico. Abbiamo $g(v) = g(h + \beta p) = g(h) + g(\beta p) = \beta\mu$ (perché $g(h) = 0$); nello stesso modo $f(v) = \beta\lambda$. Vediamo che per ogni vettore v di E : $\alpha f(v) = g(v)$ e l'asserto è dimostrato ♦

10: Corollario: Notiamo con $\text{Gr}(1, E^*)$ l'insieme delle rette (vettoriali) di E^* e con $\text{Gr}(n-1, E)$ l'insieme degli iperpiani (vettoriali) di E . L'applicazione $d: \text{Gr}(1, E^*) \rightarrow \text{Gr}(n-1, E)$: $\langle f \rangle \rightarrow \text{Ker}(f)$ è biettiva.

10.1: Osservazione : Qui f è una forma lineare non nulla su E e $\langle f \rangle = \{\lambda f / \lambda \in k\}$ è il sottospazio (di dimensione uno) generato da f .

Dimostrazione del corollario 10:

L'applicazione d è ben definita (7, 9(ii)) e suriettiva (9). Inoltre se $d(\langle f \rangle) = d(\langle g \rangle)$ allora $\text{Ker}(f) = \text{Ker}(g)$ e per 9(ii) f e g sono proporzionali ossia $\langle f \rangle = \langle g \rangle$. Quindi d è iniettiva ♦

11: Osservazione : L'insieme $\text{Gr}(1, E^*)$ è naturalmente in biiezione con lo spazio proiettivo $P(E^*)$ definito come l'insieme quoziente di $E^* \setminus \{0\}$ per la relazione d'equivalenza: $x R y \Leftrightarrow$ esiste una retta vettoriale contenente x e y $\Leftrightarrow x$ e y sono proporzionali (Es. I, 4.3, 4.4, 4.5).

Il corollario 10 è la generalizzazione attesa (cfr 6) della corrispondenza tra rette del piano ed equazioni cartesiane:

Equazioni degli iperpiani:

Sia E un k -spazio vettoriale, (v_1, \dots, v_n) una base di E e (v^*_1, \dots, v^*_n) la base duale di E^* . Sia H un iper piano di E . Abbiamo visto (9(i)) che esiste una forma lineare non nulla $f : E \rightarrow k$ tale che, se w è un vettore di E : $w \in H \Leftrightarrow f(w) = 0$; f si dice equazione dell'iper piano H . Inoltre sappiamo (9(ii)) che ogni altra equazione è della forma: $\alpha f(w) = 0$, $\alpha \in k$.

Facciamo il caso $E = k^n$. In k^n abbiamo la base canonica (e_1, \dots, e_n) dove $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$. Notiamo con (X_1, \dots, X_n) la base duale (e^*_1, \dots, e^*_n) . Abbiamo quindi $X_k(e_j) = \delta_{kj}$. Se $v = (\alpha_1, \dots, \alpha_n)$ è un vettore di k^n allora $v = \alpha_1 e_1 + \dots + \alpha_n e_n$ e $X_k(v) = \alpha_k$. Quindi la forma lineare $X_k : k^n \rightarrow k$ è l'applicazione che ad ogni $v = (\alpha_1, \dots, \alpha_n)$ associa la sua k -esima coordinata, α_k . Siccome (X_1, \dots, X_n) è una base del duale di k^n ogni forma lineare, f , su k^n è combinazione lineare a coefficienti in k degli X_i : $f = \beta_1 X_1 + \dots + \beta_n X_n$.

Quindi ogni forma lineare su k^n si esprime come un polinomio omogeneo di grado uno, a coefficienti in k , nelle variabili X_i ("funzioni coordinate").

12: Definizione: Un polinomio $P(X_1, \dots, X_n) = \sum a_{i_1 \dots i_n} X_1^{i_1} \dots X_k^{i_k} \dots X_n^{i_n}$, nelle variabili X_1, \dots, X_n è detto omogeneo di grado p se per ogni $a_{i_1 \dots i_n} \neq 0$ si ha $i_1 + \dots + i_n = p$.

Invece di polinomio omogeneo si usa dire "forma di grado p ". Questo è il motivo per cui gli elementi del duale vengono chiamati "forme" lineari.

Sia adesso $H \subseteq k^n$ un iper piano (vettoriale). Da quanto precede H ha un'equazione della forma $f(v) = 0$, f forma lineare. Dopo scelta della base canonica in k^n e della sua base duale possiamo scrivere:

$H = \{(\alpha_1, \dots, \alpha_n) \in k^n / P(\alpha_1, \dots, \alpha_n) = 0\} = \{(\alpha_1, \dots, \alpha_n) \in k^n / \beta_1 \alpha_1 + \dots + \beta_n \alpha_n = 0\}$ (†)

Dove $P(X_1, \dots, X_n)$ è il polinomio omogeneo di grado uno $P(X_1, \dots, X_n) = \beta_1 X_1 + \dots + \beta_n X_n$, espressione di f nella base X_1, \dots, X_n , duale della base canonica.

Inoltre ogni altra equazione di H è della forma $\alpha f = 0$ (con $\alpha \neq 0$) ossia $H = \{(\alpha_1, \dots, \alpha_n) \in k^n / \alpha(\beta_1 \alpha_1 + \dots + \beta_n \alpha_n) = 0\}$. Quindi possiamo dire che, a meno di un fattore moltiplicativo, ogni iper piano, H , possiede un'unica equazione del tipo (†) (detta equazione cartesiana di H).

Sappiamo anche che ogni equazione del tipo (†) determina un iperpiano. A questo punto possiamo riscrivere la corrispondenza tra $\text{Gr}(n-1, k^n)$ e $\text{Gr}(1, k^{n*})$ nel modo seguente:

$$\begin{aligned} \text{Gr}(n-1, k^n) = \{\text{iperpiani vettoriali}\} &\leftrightarrow \text{Gr}(1, k^{n*}) = P(k^{n*}) \\ &= P\{\text{polinomi omogenei di grado} \\ &\text{uno}\} \end{aligned}$$

$H \leftrightarrow \{<P(X_1, \dots, X_n)>\}$, dove $P(X_1, \dots, X_n) = 0$ è un'equazione di H .

Osservare che in quanto precede abbiamo stabilito un isomorfismo tra k^{n*} e $k[X_1, \dots, X_n]_1$ lo spazio vettoriale dei polinomi omogenei di grado uno nelle variabili X_1, \dots, X_n (per definizione: $k[X_1, \dots, X_n]_1 = \{\text{polinomi omogenei di grado uno}\} \cup \{0\}$). Si dimostra (**Es. 1**) che $k[X_1, \dots, X_n]_1$ è uno spazio vettoriale di dimensione n .

Quanto precede si può ripetere per ogni k -spazio vettoriale di dimensione finita, E . Infatti la scelta di una base (v_1, \dots, v_n) di E permette di riportarsi al caso precedente; basta considerare l'applicazione lineare $T: E \rightarrow k^n: v_k \rightarrow e_k, 1 \leq k \leq n$, dove (e_1, \dots, e_n) è la base canonica di k^n . Da (**§4, 16**) T è un isomorfismo; questo permette di identificare E con k^n , in questa identificazione il vettore $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, di E corrisponde al punto $T(v) = (\alpha_1, \dots, \alpha_n)$ di k^n .

L'**isomorfismo canonico** tra E e E^{**} (in dimensione finita).

Il corollario **10** fornisce una corrispondenza "naturale" (può essere enunciata senza scegliere delle basi) tra $\text{Gr}(n-1, E)$ e $\text{Gr}(1, E^*)$. La dualità (i.e. l'**isomorfismo canonico** tra E e E^{**}) permette di estendere questa corrispondenza al caso dei sottospazi di dimensione $< n-1$ di E .

Sia E un k -spazio vettoriale di dimensione finita. Abbiamo visto (**3**) che $\dim(E) = \dim(E^*) = \dim(E^{**})$. Quindi E, E^*, E^{**} sono isomorfi (**§4, 17**). Per stabilire un isomorfismo tra E e E^* bisogna scegliere delle basi (prendendo per esempio la base duale di una base di E , si determina un isomorfismo tra E e E^*). Per quanto riguarda E e il suo biduale E^{**} la situazione è diversa: esiste un isomorfismo "più bello degli altri" (si dice "naturale" o "canonico") tra E e E^{**} , indipendente dalla scelta di basi. La dualità è una conseguenza di questo isomorfismo.

Un elemento, Ψ , di E^{**} è una forma lineare su E^* : $\Psi: E^* \rightarrow k$; ad ogni forma lineare $f: E \rightarrow k$, Ψ associa uno scalare $\Psi(f)$.

Sia x un vettore di E . Per ogni forma lineare, f , su E possiamo considerare la valutazione in x : $f(x) \in k$. In questo modo facciamo corrispondere ad ogni forma lineare, f , lo scalare $f(x)$. Questo definisce un'applicazione: $v_x: E^* \rightarrow k$: $f \mapsto f(x)$.

Si verifica facilmente che l'applicazione $v_x: E^* \rightarrow k$ è lineare:

$$v_x(\lambda f + \mu g) = (\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x) = \lambda v_x(f) + \mu v_x(g).$$

Risulta pertanto che l'applicazione $v_x: E^* \rightarrow k$ è una forma lineare su E^* , in altre parole $v_x \in E^{**}$. Questo ci permette di definire un'applicazione:

$$v: E \rightarrow E^{**}: x \mapsto v_x.$$

Quest'applicazione è lineare:

$v(\lambda x + \mu y) = v_{\lambda x + \mu y}$ è l'applicazione $E^* \rightarrow k$ che ad ogni f di E^* associa $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda v_x(f) + \mu v_y(f)$; quindi le applicazioni $v_{\lambda x + \mu y}$ e $\lambda v_x + \mu v_y$ sono uguali, ossia $v(\lambda x + \mu y) = \lambda v(x) + \mu v(y)$.

13: Teorema: Sia E un k -spazio vettoriale di dimensione finita. L'applicazione lineare $v: E \rightarrow E^{**} : x \mapsto v_x$, è un isomorfismo di k -spazi vettoriali.

Dim: Siccome $\dim(E) = \dim(E^{**})$ (3) basta dimostrare che v è iniettiva (§5, 3). Se $v(x) = 0$ allora per ogni f in E^* : $v_x(f) = f(x) = 0$. Sia (e_1, \dots, e_n) una base di E e (e^*_1, \dots, e^*_n) la base duale. Abbiamo $x = \alpha_1 e_1 + \dots + \alpha_n e_n$. Prendendo $f = e_k^* \in k$, $1 \leq k \leq n$: $e_k^*(x) = \alpha_k = 0$, $1 \leq k \leq n$, quindi $x = 0$ e v è iniettiva ♦

13.1: Osservazioni : (i) L'isomorfismo v è "canonico" perché definito senza usare basi di E e E^{**} .

(ii) Il teorema 13 è falso se E ha dimensione infinita; in questo caso l'applicazione v è sempre iniettiva ma mai suriettiva (cfr Es. 2).

(iii) Un altro modo di formulare il teorema 13: ogni applicazione lineare $F: E^* \rightarrow k$ è del tipo valutazione in x per qualche x in E , i.e. esiste x in E tale che per ogni f in E^* : $F(f) = f(x)$. Teoremi di questo tipo (detti di rappresentazione) sono molto importanti in teoria della misura (generalizzazioni dell'integrale di Riemann).

(iv) Il teorema 13 "è" il principio di dualità che si può enunciare così: se abbiamo dimostrato una proprietà per la coppia (E, E^*) , nello stesso modo dimostriamo questa proprietà per (E^*, E^{**}) adesso siccome E^{**} si identifica in modo canonico (senza scegliere delle basi) a E , otteniamo la proprietà per (E^*, E) . Si stabilisce così una specie di simmetria tra E e E^* . Per esempio se (e_1, \dots, e_n) è una base di E e (e^*_1, \dots, e^*_n) è la base duale, allora (e_1, \dots, e_n) è la base duale di (e^*_1, \dots, e^*_n) (dopo l'identificazione canonica di E^{**} con E).

Spazio ortogonale (o "delle equazioni"):

14: Definizione: Sia F un sottospazio vettoriale di un k -spazio vettoriale, E , di dimensione finita. Definiamo l'ortogonale (o il coniugato) di F come il sottospazio vettoriale, F° , di E^* definito da: $F^\circ = \{f \in E^* / \forall x \in F, f(x) = 0\}$.

Quindi l'ortogonale F° è l'insieme delle formi lineari che si annullano su F . Si verifica facilmente che F° è un sottospazio vettoriale di E^* (Es. 3).

Osservare che l'ortogonale di $F \subseteq E$ non "vive" nello stesso spazio di F (F° è un sottospazio di E^* e non di E).

Prendiamo adesso il caso in cui " $E = E^*$ " nella definizione 14: Se $Y \subseteq E^*$ è un sottospazio allora $Y^\circ = \{\phi \in E^{**} / \forall f \in Y, \phi(f) = 0\}$. Sappiamo però (13) che la forma lineare ϕ su E^* è uguale alla valutazione in x per qualche x di E : $\phi = v(x) = v_x$. Tenuto conto di ciò: $Y^\circ = \{v_x / \forall f \in Y, v_x(f) = 0\}$, dove $v_x(f) = f(x)$. Pertanto nell'isomorfismo canonico tra E^{**} e E l'immagine di Y° è: $\{x \in E / \forall f \in Y, f(x) = 0\}$. D'ora in poi noteremo $Y^\circ = \{x \in E / \forall f \in Y, f(x) = 0\}$ i.e. identifichiamo Y° con la sua immagine, $v^{-1}(Y)$, in E tramite l'isomorfismo canonico.

In conclusione tramite l'ortogonalità:

- ad un sottospazio F di E associamo un sottospazio F° di E^*
- ad un sottospazio Y di E^* associamo un sotto spazio Y° di E .

Osservare che Y° è l'insieme delle soluzioni del "sistema" $f(x) = 0, \forall f \in Y$. In altri termini $Y^\circ = \bigcap_{f \in Y} \{x / f(x) = 0\}$.

In particolare $(F^\circ)^\circ \subseteq E$ e $(Y^\circ)^\circ \subseteq E^*$. Che legame c'è tra F e $F^{\circ\circ} := (F^\circ)^\circ$? (e tra Y e $Y^{\circ\circ}$?).

15: Teorema: Sia E un k -spazio vettoriale di dimensione finita.

(i) Sia $F \subseteq E$ un sotto spazio vettoriale. Allora $\dim(F) + \dim(F^\circ) = \dim(E)$.

(ii) $F^{\circ\circ} = F$. Nello stesso modo se $Y \subseteq E^*$ è un sottospazio allora:

$$\dim(Y) + \dim(Y^\circ) = \dim(E^*) = \dim(E) \quad e \quad Y^{\circ\circ} = Y.$$

Dim: (i) Sia (f_1, \dots, f_p) una base di F . Per il teorema della base incompleta (o prendendo un supplementare) possiamo completarla in una base di E : $(f_1, \dots, f_p, f_{p+1}, \dots, f_n)$. Sia inoltre $(f^*_1, \dots, f^*_{p+1}, \dots, f^*_n)$ la base duale. Sia $g \in F^* \subseteq E^*$. Allora $g = \alpha_1 f^*_1 + \dots + \alpha_n f^*_n$ e $g(x) = 0$ per ogni x in F . Prendendo $x = f_j$, $1 \leq j \leq p$, abbiamo $g(f_j) = (\alpha_1 f^*_1 + \dots + \alpha_n f^*_n)(f_j) = \alpha_j = 0$. Quindi $g = \alpha_{p+1} f^*_{p+1} + \dots + \alpha_n f^*_n$. Vediamo così che ogni elemento g di F^* appartiene al sottospazio generato da f^*_{p+1}, \dots, f^*_n , i.e. $F^* \subseteq \langle f^*_{p+1}, \dots, f^*_n \rangle$. È chiaro d'altra parte che $f^*_j(x) = 0$ per ogni x in F se $j \geq p+1$. Quindi $F^* = \langle f^*_{p+1}, \dots, f^*_n \rangle$ e $\dim(F) + \dim(F^*) = \dim(E)$.

(ii) Applicando (i) con E^* , Y al posto di E , F otteniamo: $\dim(Y) + \dim(Y^\circ) = \dim(E^*) = \dim(E)$. Ponendo $Y = F^*$ viene: $\dim(F^*) + \dim(F^{\circ\circ}) = \dim(E)$. Combinando con (i) risulta $\dim(F) = \dim(F^{\circ\circ})$. Siccome $F^\circ = \{f \in E^* / f(x) = 0, \forall x \in F\}$ e $F^{\circ\circ} = \{x \in E / f(x) = 0, \forall f \in F^\circ\}$ (dopo l'identificazione canonica di E con E^{**}), è chiaro che $F \subseteq F^{\circ\circ}$; dall'uguaglianza delle dimensioni segue $F = F^{\circ\circ}$. Lo stesso ragionamento vale per $Y \subseteq E^*$ ♦

15.1: Osservazione : Si osserverà che, se $F \subseteq E$ è un sottospazio vettoriale, l'uguaglianza $F = F^{\circ\circ}$ ammette la seguente interpretazione geometrica: F è uguale all'intersezione degli iperpiani contenenti F .

16: Definizione: Sia E un k -spazio vettoriale di dimensione n . Per $1 \leq m \leq n$, l'insieme dei sottospazi vettoriali di dimensione m di E viene chiamato varietà di Grassmann (o Grassmanniana) dei sottospazi di dimensione m di E , e si nota $Gr(m, E)$: $Gr(m, E) = \{F / F \text{ è un sottospazio vettoriale di dimensione } m \text{ di } E\}$.

16. 1: Osservazione : (i) Si ha $Gr(1, E) = P(E)$; lo spazio proiettivo è un caso particolare di varietà di Grassmann.

17: Teorema: Sia E un k -spazio vettoriale di dimensione finita, n . Per ogni m , $0 \leq m \leq n$, esiste una biiezione "naturale": $d : Gr(m, E) \rightarrow Gr(n-m, E^*)$ data da: $d(F) = F^\circ$.

Dim: Dal teorema 15 segue che d è ben definita perché $\dim(F^\circ) = \dim(E) - \dim(F) = n-m$. Per mostrare che d è biiettiva basta trovare $i : Gr(n-m, E^*) \rightarrow$

$\text{Gr}(k, E)$ tale che $d \circ i = \text{Id}_{\text{Gr}(n-m, E^*)}$, $i \circ d = \text{Id}_{\text{Gr}(m, E)}$; basta prendere $i(Y) = Y^\circ$ (dopo l'identificazione canonica di E^{**} con E). Dal teorema 15: $(i \circ d)(F) = F^\circ = F$ e $(d \circ i)(Y) = d(Y^\circ) = Y^{\circ\circ} = Y \blacklozenge$

17.1: Osservazione: (i) Per $k = 1$ ritroviamo il corollario 10.

(ii) La dualità inverte le inclusioni: $F \subseteq F' \Leftrightarrow F^\circ \supseteq F'$.

(\Rightarrow) Se $f \in F'$ allora $f|_F = 0$, e quindi, a fortiori, $f|_{F'} = 0$, ossia $f \in F^\circ$.

(\Leftarrow) Per l'implicazione precedente: $F^\circ \subseteq F' \Rightarrow F^{\circ\circ} \subseteq F'^{\circ\circ}$, cioè $F \subseteq F'$.

18: Equazioni dei sottospazi: Sia $F \subseteq E$ un sottospazio vettoriale e $F^\circ \subseteq E^*$. Se (ϕ_1, \dots, ϕ_m) è una base di F° allora $F = \{x \in E / \phi_1(x) = 0, \dots, \phi_m(x) = 0\}$, le $\phi_1, 1 \leq i \leq m$, sono delle equazioni del sottospazio vettoriale F ; il sottospazio F è l'intersezione degli iperpiani H_1, \dots, H_m dove un'equazione di H_i è $\phi_i = 0$. La dualità fa corrispondere ad un sottospazio F il sottospazio delle equazioni che lo definiscono.

In modo analogo a quanto fatto per gli iperpiani vediamo che un sottospazio F di dimensione $n-m$ (i.e. di codimensione $r = m = \dim(E) - \dim(F)$) è definito da un sistema di m polinomi omogenei di grado uno (i.e. da r equazioni lineari), linearmente indipendenti. Tale rappresentazione non è unica: corrisponde alla scelta di una base di F° . Se (ψ_1, \dots, ψ_m) è un'altra base di F° , allora $F = \{x \in E / \psi_1(x) = 0, \dots, \psi_m(x) = 0\}$. E' chiaro che si può rappresentare un sottospazio di dimensione $n-m$ con più di m equazioni: per sempio $F = \{x \in E / \phi_1(x) = 0, \dots, \phi_m(x) = 0, \varphi_1(x) = 0, \dots, \varphi_t(x) = 0\}$ dove $\varphi_i \in F^\circ$. Infatti $\varphi_i = \sum \lambda_p^{(i)} \cdot \phi_p$ (perchè ϕ_1, \dots, ϕ_m è una base di F°) e quindi le condizioni $\{\phi_1(x) = 0, \dots, \phi_m(x) = 0\}$ e $\{\phi_1(x) = 0, \dots, \phi_m(x) = 0, \varphi_1(x) = 0, \dots, \varphi_t(x) = 0\}$ sono equivalenti; in altri termini le equazioni $\varphi_i(x) = 0$, essendo linearmente dipendenti dalle equazioni $\phi_j(x) = 0$, sono superflue. In conclusione: un sistema di r equazioni linearmente indipendenti rappresenta un sottospazio di codimensione r ; viceversa ogni sottospazio di codimensione r può essere rappresentato da r equazioni linearmente indipendenti. Riprenderemo queste considerazioni quando parleremo dei sistemi lineari omogenei.

Per concludere questo paragrafo introduciamo la nozione di applicazione trasposta. Il teorema 20 ci sarà utile nello studio dei determinanti.

Quindi alla forma lineare $\phi: F \rightarrow k$ su F , $\mathbf{t}f$ associa la forma lineare $E \xrightarrow{-f} F$ $-\phi \rightarrow k$, $\phi \circ f$ su E .

20: Teorema: Siano E, F due k -spazi vettoriali di dimensione finita e $f: E \rightarrow F$ un'applicazione lineare.

(i) $\mathbf{t}f: F^* \rightarrow E^*$ è un morfismo lineare

(ii) $\text{Ker}(\mathbf{t}f) = (\text{Im}(f))^\circ$

(iii) $\dim(\text{Im}(\mathbf{t}f)) = \dim(\text{Im}(f))$

Dim: (i) Es. 4.

(ii) Per definizione: $\text{Ker}(\mathbf{t}f) = \{\phi \in F^* / \phi \circ f = 0\} = \{\phi \in F^* / \forall x \in E, \phi(f(x)) = 0\}$
 $= \{\phi \in F^* / \forall y \in \text{Im}(f), \phi(y) = 0\} = (\text{Im}(f))^\circ$.

(iii) Sia $\mathbf{t}f: F^* \rightarrow E^*$. Dal teorema delle dimensioni (§5, 2) abbiamo: $\dim \text{Ker}(\mathbf{t}f) + \dim \text{Im}(\mathbf{t}f) = \dim F^*$. Siccome $\text{Ker}(\mathbf{t}f) = (\text{Im}(f))^\circ$ (cfr (ii)), viene: $\dim(\text{Im}(f))^\circ + \dim \text{Im}(\mathbf{t}f) = \dim F^*$. Adesso $(\text{Im}(f))^\circ$ è un sottospazio di F^* e da 15: $\dim(\text{Im}(f))^\circ + \dim(\text{Im}(f))^{\circ\circ} = \dim F^*$. Confrontando con quanto precede: $\dim(\text{Im}(\mathbf{t}f)) = \dim(\text{Im}(f))^\circ$ e si conclude col teorema 15 ($(\text{Im}(f))^{\circ\circ} = \text{Im}(f)$) ♦

Esercizi:

7.1 Sia k un campo. Si pone $k[X_1, \dots, X_n]_d = \{P(X_1, \dots, X_n) / P \in k[X_1, \dots, X_n], \deg P \leq d\}$.

(i) Dimostrare che $k[X_1, \dots, X_n]_d$ è un k -spazio vettoriale.

(ii) Dimostrare che $\dim(k[X_1, \dots, X_n]_d) = (n-1+d)! / d!(n-1)!$ (suggerimento: procedere per induzione osservando che ci sono due tipi di monomi omogenei di grado d : quelli che non contengono X_1 , ossia del tipo $X_2^{i_2} \cdot X_3^{i_3} \cdots X_n^{i_n}$, con $i_2 + \dots + i_n = d$; e quelli che contengono X_1 , quest'ultimi si possono scrivere $X_1^{i_1} X_2^{i_2} \cdot X_3^{i_3} \cdots X_n^{i_n} = X_1 (X_1^{i_1-1} X_2^{i_2} \cdot X_3^{i_3} \cdots X_n^{i_n})$).

7.2 Sia E un k -spazio vettoriale di dimensione infinita. Si può dimostrare (col lemma di Zorn) che ogni k -spazio vettoriale di dimensione infinita ammette una base. Sia dunque $(e_i)_{i \in I}$ una base di E (l'insieme infinito). Quindi ogni vettore x di E si scrive, in modo unico, come combinazione lineare di un numero finito di e_i . Scopo dell'esercizio è di mostrare che l'applicazione canonica: $v: E \rightarrow E^{**}$ è iniettiva ma mai suriettiva. Si ricorda che v è l'applicazione che a $x \in E$ associa la valutazione in x (i.e. $v(x)$ è la forma lineare su E^* definita da $v(x)(f) = v_x(f) = f(x)$, per ogni f in E^*).

Come in dimensione finita si definisce $e^*_i : E \rightarrow k$ tramite $e^*_i(e_p) = \delta_{ip}$. Visto che un'applicazione lineare è completamente determinata dai suoi valori su una base (vero anche in dimensione infinita), $e^*_i \in E^*$.

(i) Mostrare che gli $(e^*_i)_{i \in I}$ sono linearmente indipendenti.

(ii) E' $(e^*_i)_{i \in I}$ una base di E^* ? (considerare $f : E \rightarrow k$ definita da $f(e_i) = 1$ per ogni i in I).

(iii) Dimostrare che v è iniettiva (riprendere la dimostrazione del teorema 13)

(iv) Si assumerà che il teorema della base incompleta sia ancora vero in dimensione infinita. In particolare $(e^*_i)_{i \in I}$ può essere completato ad una base di E^* ; dedurne che v non è suriettiva.

7.3) Sia E un k -spazio vettoriale e $X \subseteq E$ un sottinsieme. Si pone $X^\circ = \{f \in E^* / f(x) = 0, \forall x \in X\}$. Dimostrare che X° è un sottospazio vettoriale di E^* .

7.4) Sia $f : E \rightarrow F$ un'applicazione lineare tra gli k -spazi vettoriali E, F .

(i) Dimostrare che ${}^t f$ è lineare.

(ii) Sia l'applicazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2 : (x,y,z) \rightarrow (x-y, 2x+z)$. Esprimere ${}^t f$ nelle basi duali delle basi canoniche di $\mathbb{R}^3, \mathbb{R}^2$. Calcolare $\dim(\text{Ker}({}^t f))$.

7.5) Siano E, F degli K -spazi vettoriali e $f : E \rightarrow F$ un'applicazione lineare.

(i) Dimostrare: f non è suriettiva \Leftrightarrow esiste una forma lineare non nulla, φ , tale che $\varphi \circ f = 0$.

(ii) Più precisamente se $\dim(\text{Im}(f)) = \dim(F) - r$, dimostrare che esistono r forme lineari $\varphi_1, \dots, \varphi_r$, tali che: (a) $\varphi_1, \dots, \varphi_r$ siano linearmente indipendenti (b) $\varphi_i \circ f = 0, 1 \leq i \leq r$.

7.6) Consideriamo i seguenti sottospazi vettoriali di \mathbb{R}^3 : $L = \{(x,y,z) / x = 0 \text{ e } y = 0\}$, $L' = \{(x,y,z) / x - 2z = 0 \text{ e } 2x + z = 0\}$, $L'' = \{(x,y,z) / 2x - y + z = 0 \text{ e } y = z\}$.

Osservare che L, L' e L'' sono contenuti nell'iperpiano $\text{Ker}(e_1^*)$ (dove (e_1, e_2, e_3) è la base canonica di \mathbb{R}^3), e determinare $L^\circ \cap L'^\circ \cap L''^\circ$ (non è necessario fare conti).

7.7) Siano E un R -spazio vettoriale di dimensione tre, e L, L', L'' tre sottospazi vettoriali di dimensione uno di E . Si suppone L, L', L'' a due a due distinti e contenuti in un iperpiano H . Determinare $L^\circ \cap L'^\circ \cap L''^\circ$.

7.8) Dire se le forme lineari φ, ψ su \mathbb{R}^3 sono o meno linearmente indipendenti dove:

$\varphi : \mathbb{R}^3 \rightarrow \mathbb{R} : (x,y,z) \rightarrow x+2y+3z$.

$\psi : \mathbb{R}^3 \rightarrow \mathbb{R} : (x,y,z) \rightarrow x-2y+4z$.

7.9) Sia E un k -spazio vettoriale di base (e_1, \dots, e_n) . Sia (e^*_1, \dots, e^*_n) la base duale. Finalmente sia $\psi : E^* \rightarrow k$ la forma lineare su E^* definita da $\psi(e^*_i) = \alpha_i$. Dal teorema di dualità sappiamo che esiste (un unico) x in E tale che ψ sia la valutazione in x : $\psi(f) = f(x)$, per ogni f in E^* . Determinare x .

7.10) Sia E un R -spazio vettoriale di dimensione 3, (e_1, e_2, e_3) una base di E e (e^*_1, e^*_2, e^*_3) la base duale.

- (i) Dimostrare che $B = \{2e_1, 5e_2, -e_3\}$ è una base di E .
(ii) Dimostrare che $(1/2.e^*_1, 1/5.e^*_2, -e^*_3)$ è la base duale di B .

7.11) Sia $E = \{0, u, v, w\}$ un spazio vettoriale di dimensione due su \mathbb{F}_2 (cfr Es. 4.17).

- (i) Dimostrare che $u = v+w$ e dedurne che $v = u+w$, $w = u+v$.
(ii) Per $x \in E$, $x \neq 0$, si definisce $F_x : E \rightarrow \mathbb{F}_2$ tramite $F_x(y) = 1$ se $y \neq 0$, x ; $F_x(x) = F_x(0) = 0$. Dimostrare che F_x è lineare.
(iii) Si pone $F_0 = 0$ e si considera $F : E \rightarrow E^* : x \rightarrow F_x$. Dimostrare che F è un isomorfismo lineare tra E e E^* .

Osservazione: l'isomorfismo F è canonico i.e. non dipende dalla scelta di una base. Questa è l'unica situazione (E di dimensione 2 su $K = \mathbb{F}_2$) in cui esiste un isomorfismo canonico tra E e il suo duale E^* .

7.12) Sia E un k -spazio vettoriale di dimensione $n > 0$, con k campo infinito. Scopo dell'esercizio è di dimostrare che E non è l'unione di un numero finito di sottospazi propri; cioè se F_1, \dots, F_r sono dei sottospazi propri di E allora: $F_1 \cup \dots \cup F_r \neq E$ (*).

- (i) Mostrare che se $F \subseteq E$ è un sottospazio proprio allora F è contenuto in un iperpiano di E .
(ii) Da (i) segue che basta mostrare (*) nel caso in cui ogni F_i sia un iperpiano. Siano f_1, \dots, f_r delle forme lineari tali che $\text{Ker}(f_i) = F_i$. Sia $f : E \rightarrow k : v \rightarrow f_1(v) \dots f_r(v)$. Considerando f mostrare che (*) è conseguenza di: sia $P(X_1, \dots, X_n) \in k[X_1, \dots, X_n]$ un polinomio (omogeneo) di grado r . Se k è infinito esiste un'infinità di punti (x_1, \dots, x_n) di k^n tali che $P(x_1, \dots, x_n) \neq 0$ (†).
(iii) Dimostrare (†) per induzione su n (scrivere $P(X_1, \dots, X_n) = X_1^r \cdot g_0(X_2, \dots, X_n) + \dots + X_1 \cdot g_{r-1}(X_2, \dots, X_n) + g_r(X_2, \dots, X_n)$; per ipotesi di induzione esiste un'infinità di punti (x_2, \dots, x_n) tali che $g_0(x_2, \dots, x_n) \dots g_r(x_2, \dots, x_n) \neq 0$. Per tali punti il polinomio $X_1^r \cdot g_0(x_2, \dots, x_n) + \dots + X_1 \cdot g_{r-1}(x_2, \dots, x_n) + g_r(x_2, \dots, x_n) \in k[X_1]$ è non nullo e ha un numero finito di radici).

8) SISTEMI LINEARI OMOGENEI E DUALITA'.

In questo paragrafo si applica la dualità per ricavare informazioni sulla dimensione dell'insieme delle soluzioni di un sistema lineare omogeneo.

Sia (S) il sistema lineare omogeneo:

$$(S) \quad \left\{ \begin{array}{l} a_1^{(1)} \cdot x_1 + \dots + a_1^{(n)} \cdot x_n = 0 \\ \dots \\ a_p^{(1)} \cdot x_1 + \dots + a_p^{(n)} \cdot x_n = 0 \end{array} \right.$$

Gli elementi $a_i^{(j)}$ sono degli scalari, x_1, \dots, x_n sono le incognite. Abbiamo quindi un sistema di p equazioni in n incognite. Il sistema è lineare perché le equazioni sono lineari nelle incognite (sono polinomi omogeni di grado uno nelle variabili). Il sistema è omogeneo perché tutti i "termini noti" sono nulli.

Una soluzione di (S) è un elemento (b_1, \dots, b_n) di k^n che soddisfa tutte le equazioni di (S) ossia tale che: $a_1^{(1)} \cdot b_1 + \dots + a_1^{(n)} \cdot b_n = 0$

$$\dots \\ a_p^{(1)} \cdot b_1 + \dots + a_p^{(n)} \cdot b_n = 0.$$

Sia Σ l'insieme di tutte le soluzioni del sistema (S).

1: Lemma: L'insieme Σ è un sottospazio vettoriale di k^n .

Dim: E' chiaro che $0 = (0, \dots, 0)$ è soluzione di (S). Inoltre se $b = (b_1, \dots, b_n)$ e $c = (c_1, \dots, c_n)$ sono soluzioni di (S) è facile verificare che per ogni α, β in k , $\alpha b + \beta c$ è ancora soluzione di (S) ♦

Sia adesso E un k -spazio vettoriale di dimensione n (quindi E è isomorfo a k^n). Sia (e_1, \dots, e_n) una base di E (la scelta di questa base ci fissa un isomorfismo con k^n); $e(e^*_1, \dots, e^*_n)$ la base duale (quindi una base di E^*). Consideriamo le seguenti forme lineari su E : $f_1 = a_1^{(1)} \cdot e^*_1 + \dots + a_1^{(n)} \cdot e^*_n, \dots, f_p = a_p^{(1)} \cdot e^*_1 + \dots + a_p^{(n)} \cdot e^*_n$.

Se $v \in E$ allora $v = b_1 \cdot e_1 + \dots + b_n \cdot e_n$; vediamo così che:

$$f_1(v) = a_1^{(1)} \cdot b_1 + \dots + a_1^{(n)} \cdot b_n$$

$$\dots$$

$$f_p(v) = a_p^{(1)} \cdot b_1 + \dots + a_p^{(n)} \cdot b_n$$

Altrimenti detto, (b_1, \dots, b_n) è soluzione di (S) se e solo se $f_1(v) = \dots = f_p(v) = 0$. Diremo allora che v è soluzione di (S) e noteremo ancora con Σ l'insieme dei vettori v soluzioni di (S).

Sia $Y = \langle f_1, \dots, f_p \rangle$ il sottospazio di E^* generato da f_1, \dots, f_p .

2 : Lemma: (i) Si ha $\Sigma = Y^\circ$, in particolare Σ è un sottospazio vettoriale di E .
(ii) $\dim \Sigma + \dim Y = n$.

Dim: E' la definizione di Y° (dopo l'identificazione canonica di E^{**} con E) ♦

Il numero $\dim Y$ si chiama rango del sistema (S): è il "numero" di equazioni linearmente indipendenti.

3: Definizione: Se w_1, \dots, w_k sono dei vettori di un k -spazio vettoriale E , il rango dei vettori w_1, \dots, w_k è la dimensione del sottospazio $\langle w_1, \dots, w_k \rangle$ che generano.

Il lemma dice che l'insieme soluzione del sistema (S) è uno spazio vettoriale di dimensione = il numero di incognite (n) - il numero di equazioni linearmente indipendenti ($\dim Y$).

Osserviamo che $\dim(Y) \leq p$ (p = il numero di equazioni).

In particolare se il numero delle incognite è strettamente maggiore del numero delle equazioni ($n > p$) allora (S) ha soluzioni non banali ($\neq (0, \dots, 0)$).

La dimensione dello spazio delle soluzioni è sempre ≥ 0 (perché $Y \subseteq E^* \Rightarrow \dim Y \leq \dim E^* = n$); d'altra parte uno spazio vettoriale non è mai vuoto ($(0, \dots, 0)$) è sempre soluzione di (S)).

4: Esempio : Sia il sistema di tre equazioni in tre incognite reali: .

$$x - 6y + 4z = 0$$

$$2x + 3y + 3z = 0$$

$$x + 12y - 2z = 0$$

Le tre forme lineari corrispondenti sono: $f_1 = 1.e^*_1 - 6.e^*_2 + 4.e^*_3$, $f_2 = 2.e^*_1 + 3.e^*_2 + 3.e^*_3$, $f_3 = 1.e^*_1 + 12.e^*_2 - 2.e^*_3$. Nella base (e^*_1, \dots, e^*_3) hanno coordinate $(1, -6, 4)$, $(2, 3, 3)$, $(1, 12, -2)$. Queste tre equazioni non sono linearmente indipendenti perché per esempio: $(1, 12, -2) = -7/5.(1, -6, 4) + 6/5.(2, 3, 3)$. Quindi

(con il metodo degli scarti successivi) vediamo che $\dim\langle f_1, \dots, f_3 \rangle = 2$, perciò le soluzioni del nostro sistema formano uno spazio di dimensione $3 - 2 = 1$.

Per sapere chi è questo spazio di soluzioni basta risolvere le prime due equazioni (la terza non darà niente di nuovo perché è combinazione lineare delle altre due).

Esercizi:

8.1) Sia in \mathbb{R}^n il seguente sistema lineare:

$$\begin{array}{ll} a_{11} x_1 + \dots & + a_{1n} x_n = 0 \\ a_{22} x_2 + \dots & + a_{2n} x_n = 0 \\ \dots & \dots \\ a_{pp} x_p + \dots & + a_{pn} x_n = 0 \end{array}$$

dove $p \leq n$ e tutti gli a_{ii} , $1 \leq i \leq p$, sono non nulli. Dire qual è la dimensione dell'insieme delle soluzioni.

8.2) Nell' \mathbb{R} spazio vettoriale $\mathbb{R}[X]$ siano i vettori $P_1(X) = \pi X^3 + 4X^2 + X + 7$, $P_2(X) = X^3 - 7X^2 + X - 6$, $P_3(X) = \sqrt{2}X^3 - X + 27$, $P_4(X) = X^3 - 4X^2 + X - 1$, $P_5(X) = X^2 + X + 1$. Dire se questi vettori sono linearmente indipendenti.

8.3) Per ogni $\alpha \in \mathbb{R}$ si denoti con S_α il sottospazio vettoriale di \mathbb{R}^3 delle soluzioni del sistema:

$$3\alpha x - 4y + \alpha z = 0$$

$$x - 2\alpha y + 3\alpha z = 0$$

Determinare $\dim(S_\alpha)$ per ogni α in \mathbb{R} .

8.4) Siano E un \mathbb{R} -spazio vettoriale di dimensione n e H_1, \dots, H_{n-1} $n-1$ iperpiani vettoriali di E .

(i) Mostrare che $H_1 \cap \dots \cap H_{n-1} \neq \{0\}$.

(ii) Sia H un ulteriore iperpiano di E , che cosa si può dire di $H \cap H_1 \cap \dots \cap H_{n-1}$? Descrivere tutte le possibilità quando $n = 3$.

(iii) Dimostrare che dato un iperpiano $H' \subseteq E$ si possono sempre trovare $n-1$ iperpiani di E , H'_1, \dots, H'_{n-1} tali che $H' \cap H'_1 \cap \dots \cap H'_{n-1} = \{0\}$.

9) SCRITTURA MATRICIALE DELLE APPLICAZIONI LINEARI.

Abbiamo già introdotto le matrici (m,n) a coefficienti in k : una tale matrice, M , è una tabella con m righe e n colonne di elementi in k :

$$\begin{pmatrix} a_{11}, \dots, a_{1n} \\ \cdots \\ a_{m1}, \dots, a_{mn} \end{pmatrix}$$

$a_{ij} \in k$, $1 \leq i \leq m$; $1 \leq j \leq n$.

Si ricorda che l'elemento che sta all'incrocio della i -esima riga con la j -esima colonna si nota a_{ij} (i.e. si mette prima l'indice della riga (i) e poi quello della colonna (j)).

Si indica anche M nel modo seguente: $M = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$.

L'insieme delle matrici (m,n) a coefficienti in k si nota $M_{m,n}(k)$. Su questo insieme abbiamo definito un'addizione: se $M = (a_{ij})$ e $N = (b_{ij})$ allora $M+N = (a_{ij}+b_{ij})$; risulta che $(M_{m,n}(k), +)$ è un gruppo abeliano.

Inoltre abbiamo una moltiplicazione esterna: se $\lambda \in k$ allora $\lambda \cdot M = (\lambda a_{ij})$. Per riassumere:

Proposizione: L'addizione e la moltiplicazione esterna descritte qui sopra conferiscono a $M_{m,n}(k)$ una struttura di k -spazio vettoriale.

L'importanza delle matrici deriva però dal fatto che forniscono (dopo la scelta di basi) un metodo meccanico per rappresentare le applicazioni lineari e le operazioni tra di esse. In particolare la composizione delle applicazioni lineari corrisponde al prodotto "righe per colonne" (RICO) delle matrici, che definiremo tra poco.

Un'avvertenza: non bisogna lasciarsi trascinare troppo dall'entusiasmo per l'aspetto meccanico delle matrici. Molti problemi si risolvono prima, e più elegantemente, ragionando direttamente sulle applicazioni lineari (vedere per esempio gli esercizi 9.3, 9.8). Inoltre molti risultati per gli spazi di dimensione finita sono ancora validi per spazi di dimensione infinita e questo si vede ragionando sulle applicazioni.

Siano E, F due k-spazi vettoriali, (e_1, \dots, e_n) una base di E, (v_1, \dots, v_m) una base di F.

Sia $f : E \rightarrow F$ un'applicazione lineare. Abbiamo già visto (§4, 16) che f è completamente determinata dai vettori $f(e_1), \dots, f(e_n)$. I vettori $f(e_1), \dots, f(e_n)$ di F si possono scrivere in funzione della base (v_1, \dots, v_m) :

$$f(e_1) = a_{11}v_1 + a_{21}v_2 + \dots + a_{m1}v_m$$

$$\dots$$

$$f(e_j) = a_{1j}v_1 + a_{2j}v_2 + \dots + a_{mj}v_m$$

$$\dots$$

$$f(e_n) = a_{1n}v_1 + a_{2n}v_2 + \dots + a_{mn}v_m$$

Vediamo così che **dopo la scelta delle basi** (e_1, \dots, e_n) , (v_1, \dots, v_m) di E, F darsi un'applicazione lineare $f : E \rightarrow F$ è equivalente a darsi gli m.n scalari a_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$. Un altro modo di esprimere questo fatto è, appunto, di usare le matrici: gli scalari a_{ij} si possono scrivere sotto forma matriciale:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Abbiamo così una matrice (m,n) . La prima colonna di M (si dice anche il primo vettore colonna) è :

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

Sono le coordinate del vettore $f(e_1)$ nella base (v_1, \dots, v_m) . In modo analogo il j-esimo vettore colonna di M (i.e. la j-esima colonna di M) è:

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Sono le coordinate del vettore $f(e_j)$ nella base (v_1, \dots, v_m) .

Possiamo visualizzare questo fatto con la scrittura:

$$\begin{array}{cccc}
 f(e_1) & f(e_2) \dots & f(e_j) \dots & f(e_n) \\
 v_1 & a_{11} & a_{12} \dots & a_{1j} \dots & a_{1n} \\
 v_2 & a_{21} & a_{22} \dots & a_{2j} \dots & a_{2n} \\
 \cdot & \cdots & \cdots & \cdots & \cdots \\
 \cdot & \cdots & \cdots & \cdots & \cdots \\
 v_i & a_{i1} & a_{i2} \dots & a_{ij} \dots & a_{in} \\
 \cdot & \cdots & \cdots & \cdots & \cdots \\
 \cdot & \cdots & \cdots & \cdots & \cdots \\
 v_m & a_{m1} & a_{m2} \dots & a_{mj} \dots & a_{mn}
 \end{array} = M \quad (\dagger)$$

1: Definizione: La matrice M si chiama la matrice di f nelle basi $B = (e_1, \dots, e_n)$, $D = (v_1, \dots, v_m)$ di E, F . Si nota $\text{mat}(f; B, D) = M$.

1.1: Osservazione: La matrice di f dipende dalle basi (e_1, \dots, e_n) , (v_1, \dots, v_m) . Se (e'_1, \dots, e'_n) , (v'_1, \dots, v'_m) sono altre basi di E, F , relativamente a queste basi f sarà rappresentata da una matrice M' e, in generale si avrà $M \neq M'$.

Viceversa supponiamo di avere fissato delle basi (e_1, \dots, e_n) , (v_1, \dots, v_m) in E, F e di avere una matrice M :

$$M = \begin{pmatrix} a_{11} \dots a_{1n} \\ \cdots \\ a_{m1} \dots a_{mn} \end{pmatrix}$$

Allora possiamo definire un'applicazione lineare $f : E \rightarrow F$ tramite $f(e_i) = a_{1i}v_1 + \dots + a_{mi}v_m$, $1 \leq i \leq n$. Con queste considerazioni è facile dimostrare:

2: Teorema: Siano E, F due k -spazi vettoriali di basi $B = (e_1, \dots, e_n)$, $D = (v_1, \dots, v_m)$ rispettivamente. L'applicazione: $\text{mat}(\cdot; B, D) : L_k(E, F) \rightarrow M_{m,n}(k)$ che ad ogni f di $L_k(E, F)$ associa la sua matrice, $\text{mat}(f; B, D)$, nelle basi $B = (e_1, \dots, e_n)$, $D = (v_1, \dots, v_m)$ è un isomorfismo di k -spazi vettoriali.

2.1: Osservazione: E' indispensabile precisare le basi scelte in E, F . Infatti non esiste un isomorfismo "canonico" tra $L_k(E, F)$ e $M_{m,n}(k)$; ogni scelta di basi determina un isomorfismo specifico tra questi due spazi. La frase "sia M la matrice associata all'applicazione f " non ha senso: si può associare una matrice

ad un'applicazione lineare **solo dopo** aver scelto delle basi negli spazi di partenza e d'arrivo.

2.2: Osservazione : Supponiamo quindi di aver scelto delle basi in E ed F. Il teorema 2 significa che $\text{mat}(\alpha f + \beta g; B, D) = \alpha \cdot \text{mat}(f; B, D) + \beta \cdot \text{mat}(g; B, D)$.

3: Immagine di un vettore:

Sia $M = \text{mat}(f; B, D)$ la matrice di $f : E \rightarrow F$ nelle basi $B = (e_1, \dots, e_n)$, $D = (v_1, \dots, v_m)$:

$$M = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Sia $a \in E$. Il vettore a si decomponde sulla base (e_1, \dots, e_n) : $a = \alpha_1 e_1 + \dots + \alpha_i e_i + \dots + \alpha_n e_n$.

Abbiamo: $f(a) = f(\alpha_1 e_1 + \dots + \alpha_i e_i + \dots + \alpha_n e_n) = \alpha_1 f(e_1) + \dots + \alpha_i f(e_i) + \dots + \alpha_n f(e_n) = \alpha_1 (a_{11} v_1 + \dots + a_{m1} v_m) + \dots + \alpha_i (a_{1i} v_1 + \dots + a_{mi} v_m) + \dots + \alpha_n (a_{1n} v_1 + \dots + a_{mn} v_m) =$

$$v_1 (\alpha_1 a_{11} + \dots + \alpha_i a_{1i} + \dots + \alpha_n a_{1n}) + \dots + v_m (\alpha_1 a_{m1} + \dots + \alpha_i a_{mi} + \dots + \alpha_n a_{mn}) = \sum_{j=1}^m \left(\sum_{i=1}^n \alpha_i a_{ji} \right) v_j.$$

Grazie al prodotto "righe per colonne", si può fare questo calcolo usando le matrici. Prima di tutto associamo al vettore a la matrice, A , con n righe ed una colonna, delle sue coordinate nella base (e_1, \dots, e_n) :

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Adesso possiamo fare il prodotto righe per colonne della matrice M con questa matrice $(n,1)$. Per fare questo prodotto si procede così: M è una matrice (m,n) , il suo prodotto con una matrice $(n,1)$ sarà una matrice $(m, n, n, 1) = (m, 1)$, quindi qualcosa del tipo:

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Il termine b_i si ottiene "moltiplicando" la i -esima riga di M per l'unica colonna di A :

$$\left(a_{11} \dots a_{ij} \dots a_{in} \right) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_n \end{pmatrix}$$

Questo è il prodotto di una matrice $(1,n)$ per una matrice $(n,1)$; il risultato è una matrice $(1, 1) = (1,1)$, ossia uno scalare. Questo scalare è uguale a: $\alpha_1 a_{11} + \dots + \alpha_j a_{ij} + \dots + \alpha_n a_{in}$; è la somma dei prodotti delle coordinate dei "vettori" $M_{(i)} = (a_{i1}, \dots, a_{in})$, $A = (\alpha_1, \dots, \alpha_n)$. Un altro modo di dire è che b_i è il prodotto scalare ("usuale") della i -esima riga di M , $M_{(i)}$, con A .

Il prodotto scalare di due vettori $(x_1, \dots, x_p), (y_1, \dots, y_p)$ di k^p è per definizione: $x_1 y_1 + \dots + x_p y_p$.

Confrontando con il calcolo precedente (cfr (*)) vediamo che b_1, \dots, b_m non sono altro che le coordinate di $f(a)$ nella base (v_1, \dots, v_m) di F : $f(a) = b_1 v_1 + \dots + b_m v_m$.

In conclusione: se $f : E \rightarrow F$ è rappresentata dalla matrice M nelle basi (e_1, \dots, e_n) , (v_1, \dots, v_m) di E , F e se $a \in E$ ha $(\alpha_1, \dots, \alpha_n)$ per coordinate nella base (e_1, \dots, e_n) di E allora $f(a)$ ha (b_1, \dots, b_m) per coordinate nella base (v_1, \dots, v_m) di F . La matrice $(m,1)$ i cui coefficienti sono b_1, \dots, b_m si ottiene facendo il prodotto righe per colonne di M e A , altrimenti detto b_i è il prodotto scalare della i -esima riga, $M_{(i)}$, di M con l'unica colonna di A :

$$b_i = \alpha_1 a_{i1} + \dots + \alpha_j a_{ij} + \dots + \alpha_n a_{in}$$

$$b_i = \left(a_{11} \dots a_{ij} \dots a_{in} \right) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_n \end{pmatrix}$$

Vediamo così che il prodotto "righe per colonne" si introduce naturalmente nella rappresentazione matriciale delle applicazioni lineari.

3.1: Esempio : Sia $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : (x,y,z) \rightarrow (2x-y, 3x+y-z, y-z)$. Prendiamo come base la base canonica e_1, \dots, e_3 (sia nell' \mathbb{R}^3 di partenza che in quello d'arrivo). Abbiamo $f(e_1) = (2, 3, 0) = 2e_1 + 3e_2$; $f(e_2) = (-1, 1, 1) = -e_1 + e_2 + e_3$; $f(e_3) = (0, -1, -1) = -e_2 - e_3$. La matrice di f nella base canonica (all'arrivo come alla partenza) è:

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

Sia $a = (4, 4, 4)$. Queste sono esattamente le coordinate di a nella base canonica di \mathbb{R}^3 ($a = 4e_1 + 4e_2 + 4e_3$, questo è il vantaggio della base canonica!). Per calcolare $f(a)$ facciamo il prodotto righe per colonne di M con A dove A è la matrice (3,1) i cui coefficienti valgono tutti 4:

$$\begin{pmatrix} 2 & -1 & 0 \\ 3 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \cdot 4 - 4 + 0 \cdot 4 \\ 3 \cdot 4 + 4 - 4 \\ 0 \cdot 4 + 4 - 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 12 \\ 0 \end{pmatrix}$$

In conclusione: $f(a) = (4, 12, 0)$ ossia $f(a) = 4e_1 + 12e_2$.

Composizione delle applicazioni lineari e prodotto righe per colonne delle matrici.

Siano E, F, G tre k -spazi vettoriali e $u: E \rightarrow F, v: F \rightarrow G$ due applicazioni k -lineari. L'applicazione composta $w = v \circ u: E \rightarrow G$ è ancora lineare. Siano $B = (e_1, \dots, e_n), D = (f_1, \dots, f_p), H = (g_1, \dots, g_m)$ delle basi di E, F, G . Dopo la scelta di queste basi le applicazioni u, v possono essere rappresentate da matrici: $U = \text{mat}(u; B, D)$, $V = \text{mat}(v; D, H)$; U è una matrice (p, n) , V è una matrice (m, p) . Adesso anche l'applicazione $w = v \circ u: E \rightarrow G$ può essere rappresentata da una matrice, W , rispetto alle basi $(e_1, \dots, e_n), (g_1, \dots, g_m)$ di E, G . È naturale chiedersi se c'è una relazione tra la matrice W e le matrici U, V .

Vedremo adesso che la matrice W è la matrice prodotto $V \cdot U$ dove il prodotto $V \cdot U$ è inteso come prodotto "righe per colonne" (RICO) delle matrici U, V .

4: Definizione: Siano A una matrice (m, n) a coefficienti in k e B una matrice (t, p) a coefficienti in k .

(i) Il prodotto "righe per colonne" (RICO), $A \cdot B$, è definito se e solo se $n = t$.

(ii) Se $n = t$ il prodotto (RICO), $A \cdot B$, è una matrice $(m, p) = (m, n, n, p)$ il cui termine c_{ij} (all'incrocio della i -esima riga con la j -esima colonna) è il prodotto scalare della i -esima riga di A con la j -esima colonna di B .

Osserviamo che la i -esima riga di A è (a_{i1}, \dots, a_{in}) mentre la j -esima colonna di B è (scritto come "vettore riga"): (b_{1j}, \dots, b_{nj}) ; ha quindi senso parlare di prodotto scalare (sono entrambi vettori di k^n).

Se $A = (a_{kl})$, $1 \leq k \leq m$, $1 \leq l \leq n$; $B = (b_{rs})$, $1 \leq r \leq n$, $1 \leq s \leq p$ allora $A \cdot B = (c_{ij})$, $1 \leq i \leq m$, $1 \leq j \leq p$, dove: $c_{ij} = a_{i1}b_{1j} + \dots + a_{in}b_{nj}$.

4.1: Esempio : Sia A la matrice (3,2):

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e sia B la matrice (2,2):

$$B = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$$

Allora il prodotto RICO $A \cdot B$ è ben definito, è una matrice $(3,2) = (3, 2, 2, 2)$ uguale a:

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 & 2 \cdot 4 + 1 \cdot 3 \\ 1 \cdot 1 + 0 \cdot 2 & 1 \cdot 4 + 0 \cdot 3 \\ 0 \cdot 1 + 1 \cdot 2 & 0 \cdot 4 + 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 4 & 11 \\ 1 & 4 \\ 2 & 3 \end{pmatrix}$$

4.2: Osservazione: Come caso particolare ritroviamo il prodotto "righe per colonne" di una matrice (m,n) e di una matrice $(n,1)$ che abbiamo già usato per calcolare l'immagine di un vettore.

5: Proposizione: Siano E, F, G tre k -spazi vettoriali e $u: E \rightarrow F$, $v: F \rightarrow G$ due applicazioni k -lineari. Sia w l'applicazione composta $w = v \circ u: E \rightarrow G$. Siano $B = (e_1, \dots, e_n)$, $D = (f_1, \dots, f_p)$, $H = (g_1, \dots, g_m)$ delle basi di E, F, G . Dopo la scelta di queste basi le applicazioni u, v sono rappresentate da matrici: $U = \text{mat}(u; B, D)$, $V = \text{mat}(v; D, H)$; U è una matrice (p, n) , V è una matrice (m, p) . Allora l'applicazione $w = v \circ u: E \rightarrow G$ è rappresentata, rispetto alle basi (e_1, \dots, e_n) , (g_1, \dots, g_m) di E, G , da una matrice $W = \text{mat}(v \circ u, B, H)$. Si ha: $W = V \cdot U$ dove $V \cdot U$ è il prodotto RICO delle matrici V, U .

5.1: Osservazione : La matrice U è di tipo (p,n) , la matrice V è di tipo (m,p) quindi il prodotto righe per colonne $V.U$ è di tipo $(m, p, p, n) = (m, n)$. Il fatto che il prodotto RICO, $A.B$, di due matrici A, B sia definito solo se il numero di colonne di A è uguale al numero di righe di B rispecchia il fatto che due applicazioni (lineari) a, b possono essere composte ($a \circ b$) solo se l'insieme d'arrivo di b è uguale all'insieme di partenza di a .

Dimostrazione della proposizione 5: Sia W la matrice associata a w rispetto alle basi $(e_1, \dots, e_n), (g_1, \dots, g_m)$ di E, G . Sappiamo (1, e (t)) che il j -esimo vettore colonna di W è il vettore delle componenti di $w(e_j)$ nella base (g_1, \dots, g_m) : se $W = (w_{rt})$, $1 \leq r \leq m$, $1 \leq t \leq n$, allora: $w(e_j) = w_{1j}g_1 + \dots + w_{mj}g_m$ (*). D'altra parte $w(e_j) = (v_\infty u)(e_j) = v(u(e_j))$. Per definizione della matrice U : $u(e_j) = u_{1j}f_1 + \dots + u_{pj}f_p$. Quindi $v(u(e_j)) = v(u_{1j}f_1 + \dots + u_{pj}f_p) = u_{1j}v(f_1) + \dots + u_{pj}v(f_p)$. Adesso per definizione della matrice V : $v(f_k) = v_{1k}g_1 + \dots + v_{mk}g_m$. Inserendo in quanto precede:

$$\begin{aligned} v(u(e_j)) &= u_{1j}(v_{11}g_1 + \dots + v_{m1}g_m) + \dots + u_{pj}(v_{1p}g_1 + \dots + v_{mp}g_m) \\ &= g_1(u_{1j}v_{11} + \dots + u_{pj}v_{1p}) + \dots + g_m(u_{1j}v_{m1} + \dots + u_{pj}v_{mp}). \end{aligned}$$

Confrontando con (*) viene:
 $w_{1j} = u_{1j}v_{11} + \dots + u_{pj}v_{1p}, \dots, w_{ij} = u_{1j}v_{i1} + \dots + u_{pj}v_{ip}, \dots, w_{mj} = u_{1j}v_{m1} + \dots + u_{pj}v_{mp}$. Vediamo così che w_{ij} è il prodotto scalare della i -esima riga di V con la j -esima colonna di U , questo è esattamente il coefficiente i,j della matrice $V.U$. Pertanto le due matrici $(m,n), V.U$ e W sono uguali♦

Esercizi:

9.1) Siano E, F due k -spazi vettoriali di basi $B = (e_1, \dots, e_n), D = (f_1, \dots, f_p)$. Per ogni (i,j) con $1 \leq i \leq n, 1 \leq j \leq p$, sia $g_{ji} : E \rightarrow F$ l'applicazione lineare definita da: $g_{ji}(e_m) = 0$ se $m \neq i, g_{ji}(e_i) = f_j$.

- (i) Determinare la matrice di g_{ji} nelle basi $(e_1, \dots, e_n), (f_1, \dots, f_p)$.
- (ii) Sia $M_{ji} = \text{mat}(g_{ji}; B, D)$. Dimostrare che le matrici M_{ji} sono linearmente indipendenti in $M_{p,n}(k)$. Concludere che $\dim(L_k(E, F)) = np$ (cfr Es. 6.2).

9.2) Sia $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2 : (x,y,z) \rightarrow (2x-z+y, y-z)$

- (i) Scrivere la matrice di f nelle basi canoniche.
- (ii) Dare un sistema di generatori di $\text{Im}(f)$.
- (iii) Sia $v = (2,1,-1)$, calcolare le coordinate di $f(v)$ nella base canonica di \mathbb{R}^2 .
- (iv) Sia $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x,y) \rightarrow (2x-y, x+y)$. Prendendo la base canonica in ogni \mathbb{R}^n che incontrate, dire qual è la matrice associata a $g \circ f$.
- (v) Siano $w = (1,1), w' = (-1,0)$. Mostrare che (w, w') è una base di \mathbb{R}^2 e dare le coordinate di $(g \circ f)(v)$ nella base (w, w') .

9.3) Trovare una matrice M tale che: $M \in M_2(\mathbb{R})$, $M \neq 0$, $M^2 := M \cdot M = 0$ ($M \cdot M$ è il prodotto RICO di M con se stessa).

Mostrare che per ogni $n \geq 2$, esiste una matrice, M , tale che: $M \in M_n(\mathbb{R})$, $M \neq 0$, $M^2 := M \cdot M = 0$ (cfr Es. I.6.3, Es. 6.4)

9.4) Sia $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x,y) \rightarrow (2x-y, 3y)$. Siano inoltre $u = (1, -2)$, $v = (2, 3)$.

Dimostrare che $B' = (u, v)$ è una base di \mathbb{R}^2 , e determinare $\text{mat}(f; B, B')$.

9.5) Sia $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x,y) \rightarrow (2x+2y, -4x+3y)$. Determinare $M = \text{mat}(g; C, C)$ (C = base canonica). Dire se M è invertibile; in caso di risposta affermativa calcolare M^{-1} .

9.6) Sia l'applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2: (x,y,z) \rightarrow (x-y, 2x+z)$. Determinare la matrice di ${}^t f$ nelle basi duali delle basi canoniche di \mathbb{R}^3 , \mathbb{R}^2 . Usando la matrice di ${}^t f$ determinare $\text{Ker}({}^t f)$ e calcolare la sua dimensione (cfr Es. 7.4).

9.7) Sia α un numero reale non nullo e $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, la matrice (n,n) tale che $a_{ij} = \alpha$ per ogni i, j . Calcolare M^p per ogni naturale $p \geq 1$.

9.8) (i) Descrivere le matrici, $M \in M_2(\mathbb{R})$, che commutano con ogni matrice $(2,2)$ a coefficienti reali: $\forall A \in M_2(\mathbb{R})$, $A \cdot M = M \cdot A$.

(ii) Descrivere le matrici, $M \in M_n(\mathbb{R})$, che commutano con ogni matrice (n,n) a coefficienti reali (suggerimento: sia $f: E \rightarrow E$ un endomorfismo tale che $f(x) = v$ con x e v linearmente indipendenti; definire $g: E \rightarrow E$ tale che $g(v) = g(x) = x$ (perché questo è possibile?) e verificare che f e g non commutano).

9.9) Sia $L = \{(x,y,z) \in \mathbb{R}^3 / 2x - y + z = 0 \text{ e } 3x - 2y + z = 0\}$. Inoltre sia $M = \begin{pmatrix} 1 & 4 & -2 \\ 2 & 1 & 3 \\ 3 & -1 & 7 \end{pmatrix}$.

Esiste $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ lineare, tale che $\text{Ker}(f) = L$ e tale che la matrice di f in basi opportune sia M ?

10) $M_{m,p}(k)$, $M_n(k)$, $GL_n(k)$ ED ALCUNE MATRICI NOTEVOLI.

In questo paragrafo traduciamo in termini di matrici i risultati del §6 mettendo l'accento sugli isomorfismi di k -algebre $\text{mat}(\cdot; B, B) : \text{End}_k(E) \rightarrow M_n(k)$. Infine si introducono alcune matrici notevoli.

Operazioni sulle matrici.

Abbiamo, una volta scelte delle basi (§9, teorema 2, proposizione 5) una corrispondenza tra operazioni fra applicazioni lineari e operazioni fra matrici:

$$\begin{array}{lcl} f+g & \leftrightarrow & A+B \\ \lambda f & \leftrightarrow & \lambda A \\ f \cdot g & \leftrightarrow & AB \text{ (prodotto RICO)} \end{array}$$

Questa corrispondenza permette di dimostrare facilmente (guardando alle applicazioni, §6) le seguenti proprietà del calcolo matriciale:

$$\begin{aligned} A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \\ (AB)C &= A(BC) \text{ (associatività del prodotto RICO).} \\ \alpha(AB) &= (\alpha A)B = A(\alpha B) \quad (\alpha \in k) \end{aligned}$$

E' chiaro che ogni volta assumiamo le necessarie condizioni di compatibilità (AB è definito solo se il numero di colonne di A è uguale al numero di righe di B).

Endomorfismi e matrici quadrate.

Consideriamo gli endomorfismi di un k -spazio vettoriale, E , di dimensione n . Sia $B = (e_1, \dots, e_n)$ una base di E . Nel seguito considereremo B come una base sia dello spazio di partenza sia dello spazio d'arrivo. Osserviamo che questa scelta non è assolutamente obbligatoria, anzi vedremo più avanti (§11) che certe volte conviene scegliere una base diversa nello spazio d'arrivo.

Dopo questa scelta delle basi ad ogni endomorfismo f di E possiamo associare la sua matrice rispetto alla base B (all'arrivo e alla partenza). Sia $M = \text{mat}(f; B, B)$, allora M è una matrice (n,n) . Una tale matrice è detta **quadrata**. L'insieme delle matrici quadrate a coefficienti in k si nota $M_n(k)$. Come già detto prima, dopo la scelta di basi, possiamo stabilire un'applicazione: $\text{mat}(\cdot; B, B) : \text{End}_k(E) \rightarrow M_n(k)$:

$f \rightarrow M = \text{mat}(f; B, B)$. Quest'applicazione stabilisce un **isomorfismo dell'anello** $(\text{End}_k(E), +, \cdot)$ con l'anello $(M_n(k), +, \cdot)$, dove \cdot è il prodotto RICO delle matrici (n,n) . Questo isomorfismo è compatibile con l'isomorfismo di k -spazi vettoriali (§9, 2); in effetti $\text{mat}(-; B, B)$ è un isomorfismo di k -algebre (cfr §6 anello degli endomorfismi):

1: Teorema: Siano E un k -spazio vettoriale di dimensione n e B una base di E . L'applicazione $\text{mat}(-; B, B) : \text{End}_k(E) \rightarrow M_n(k) : f \rightarrow \text{mat}(f; B, B)$ è un isomorfismo di k -algebre.

Dim: Es. 1♦

1.1: Osservazione : Siano f, g due endomorfismi di E e siano $M = \text{mat}(f; B, B)$, $N = \text{mat}(g; B, B)$ le matrici associate dopo scelta di una base B . Consideriamo l'endomorfismo $f \cdot g : E \rightarrow E$. Siccome $\text{mat}(-; B, B)$ è un isomorfismo d'anelli $\text{mat}(f \cdot g; B, B) = MN$. Sia D un'altra base di E e siano $M' = \text{mat}(f; B, D)$, $N' = \text{mat}(g; B, D)$ le matrici associate tramite l'isomorfismo di (solì) k -spazi vettoriali $\text{mat}(-; B, D) : \text{End}(E) \rightarrow M_n(k)$. Allora, in generale, $\text{mat}(f \cdot g; B, D) \neq M'N'$ (cfr Es. 2); questo diventerà più chiaro nel §11.

Il fatto che $\text{mat}(-; B, B)$ sia un isomorfismo anche d'anelli (e non solo di k -spazi vettoriali) è il vantaggio essenziale che si ricava prendendo la stessa base sia alla partenza che all'arrivo.

Sia $\text{Id}_E : E \rightarrow E : x \rightarrow x$, l'applicazione identità. Osserviamo che per ogni endomorfismo f , $\text{Id}_E \cdot f = f \cdot \text{Id}_E = f$, ossia Id_E è il neutro per la moltiplicazione nell'anello $\text{End}(E)$. Se $B = (e_1, \dots, e_n)$ è una base di E , $\text{Id}_E(e_i) = e_i$, $1 \leq i \leq n$. Siccome i vettori colonna di $\text{mat}(\text{Id}_E; B, B)$ sono le componenti nella base B dei vettori $\text{Id}_E(e_i)$, ne deduciamo:

2: Proposizione: Sia E un k -spazio vettoriale di dimensione n . Per ogni base, B , di E $\text{mat}(\text{Id}_E; B, B) = I_n$ dove:

$$I_n = \begin{pmatrix} 100\dots0 \\ 010\dots0 \\ \ddots \\ 0\dots01 \end{pmatrix}$$

2.1: Osservazione : Quindi I_n è la matrice i cui coefficienti sono tutti nulli tranne quelli sulla diagonale che valgono tutti uno. In particolare I_n non

dipende dalla base B . Questo segue dal fatto che, per ogni base B , $\text{mat}(\cdot; B, B)$ è un isomorfismo d'anelli: gli elementi neutri per le moltiplicazioni si corrispondono: $\text{mat}(\text{Id}_E; B, B) = I_n$ (è facile vedere che I_n è il neutro per il prodotto RICO delle matrici (n, n) , cfr Es. I.6.3).

3: Matrici invertibili, automorfismi: Tra gli endomorfismi di E abbiamo quelli invertibili (i.e. f è biiettiva). Questi endomorfismi sono chiamati automorfismi di E . L'insieme degli automorfismi di E viene notato $\text{Aut}_k(E)$ o $\text{GL}_k(E)$; $(\text{Aut}_k(E), \circ)$ è un gruppo (cfr §6). Sia $M = \text{mat}(f; B, B)$, $N = \text{mat}(f^{-1}; B, B)$. Dalle relazioni: $f \circ f^{-1} = f^{-1} \circ f = \text{Id}_E$, deduciamo, considerando le matrici associate: $MN = I_n$. Possiamo scrivere: $N = M^{-1}$.

Viceversa siano M, N due elementi di $M_n(k)$ tali che: $MN = NM = I_n$. Siano f, g gli endomorfismi corrispondenti a M e N dopo scelta di una base B di E : $f = m_B^{-1}(M)$, $g = m_B^{-1}(N)$, dove $m_B := \text{mat}(\cdot; B, B)$. Abbiamo $m_B^{-1}(MN) = f \circ g = m_B^{-1}(NM) = g \circ f = m_B^{-1}(I_n) = \text{Id}_E$. Quindi f è un automorfismo e $g = f^{-1}$. In conclusione abbiamo dimostrato:

4: Proposizione: Una matrice quadrata $M \in M_n(k)$ è invertibile (i.e. esiste $N \in M_n(k)$ tale che $MN = NM = I_n$) se e solo se M può essere associata ad un automorfismo di E .

4.1: Osservazione : Una stessa matrice può essere associata a più automorfismi; questo dipende dalla base scelta in E , ogni scelta determina un isomorfismo specifico di k -algebre: $\text{End}_k(E) \rightarrow M_n(k)$.

Queste considerazioni permettono di dimostrare:

5: Proposizione: Sia $M \in M_n(k)$ una matrice quadrata, sono equivalenti:

- (i) M è invertibile
- (ii) Esiste $N \in M_n(k)$ tale che $MN = I_n$.
- (iii) Esiste $Q \in M_n(k)$ tale che $QM = I_n$.

Se queste condizioni sono verificate allora $N = Q = M^{-1}$.

Dim: (i) \Rightarrow (ii) Questo è ovvio (prendere $N = M^{-1}$).

(ii) \Rightarrow (iii) Fissiamo una base B e un isomorfismo $\text{mat}(\cdot; B, B): \text{End}_k(E) \rightarrow M_n(k)$.

Siano f, g gli endomorfismi corrispondenti a M, N . La relazione matriciale diventa: $f \circ g = \text{Id}_E$. Ne segue (I, §3, 10) che f è suriettiva. Siccome f è un

endomorfismo di uno spazio vettoriale di dimensione finita allora (§5, 3) f è biettiva. Quindi esiste f^{-1} tale che $f \circ f^{-1} = f^{-1} \circ f = \text{Id}_E$. Traducendo in termini di matrici tramite l'isomorfismo $\text{mat}(\cdot; B, B)$, otteniamo l'esistenza di Q (vediamo anche che $Q = N = \text{mat}(f^{-1}; B, B)$).

(iii) \Rightarrow (i) La dimostrazione è simile a quella precedente; questa volta avremo $h \circ f = \text{Id}_E$, quindi f iniettiva (I, §3, 9). Si conclude come sopra ♦

6: Definizione: Le matrici quadrate (n,n) invertibili formano per il prodotto RICO un gruppo isomorfo al gruppo $(\text{Aut}(E), \circ)$. Questo gruppo viene notato $GL_k(n)$ e si chiama il gruppo lineare generale.

6.1: Osservazione : E' chiaro che il problema di sapere se una matrice è invertibile si pone solo se la matrice è quadrata: infatti un'applicazione lineare può essere biettiva solo se è un'applicazione tra due spazi che hanno la stessa dimensione.

7: Alcune matrici notevoli.

7.1: Matrice trasposta: Sia $A = (a_{ij})$, $1 \leq i \leq m$, $1 \leq j \leq p$, una matrice m,p . La matrice trasposta di A si nota ${}^t A$ ed è definita da: ${}^t A = (b_{ji})$, $1 \leq j \leq p$, $1 \leq i \leq m$, con $b_{ji} = a_{ij}$. In altre parole ${}^t A$ è la matrice le cui righe sono le colonne di A (e le cui colonne sono le righe di A).

La trasposizione è un'applicazione biettiva: $M_{m,p}(k) \rightarrow M_{p,m}(k)$. E' chiaro che ${}^t({}^t A) = A$ per ogni matrice A .

7.2: Matrici simmetriche: Una matrice m,p , A , è simmetrica se è uguale alla sua trasposta: $A = {}^t A$. Questo implica in particolare $m = p$ quindi una matrice simmetrica è quadrata.

Sia $A = (a_{ij})$, $1 \leq i \leq m$, $1 \leq j \leq m$, una matrice quadrata. Gli elementi a_{ii} , $1 \leq i \leq m$, vengono chiamati elementi diagonali della matrice A . Il loro insieme forma la diagonale principale di A .

Adesso si può dire che la matrice quadrata A è simmetrica solo se gli elementi di A , simmetrici rispetto alla diagonale principale, sono uguali.

7.3: Matrici antisimmetriche: Una matrice m,p , A , è antisimmetrica se: $-A = {}^t A$. Come prima vediamo che A è quadrata. Se A è antisimmetrica abbiamo: $a_{ij} = -a_{ji}$, per ogni (i,j) . In particolare $2a_{ii} = 0$, $1 \leq i \leq n$. Se $2 \neq 0$ nel campo k (in particolare $k \neq \mathbb{Z}/2\mathbb{Z}$, ma ci sono altri casi!) allora $a_{ii} = 0$, $1 \leq i \leq n$.

7.4: Matrici triangolari: Sia $A = (a_{ij})$, $1 \leq i \leq n$, una matrice quadrata n,n . Si dice che A è triangolare superiore se: $j < i \Rightarrow a_{ij} = 0$. In altri termini A è triangolare superiore se tutti i coefficienti sotto la diagonale principale sono nulli. In modo analogo una matrice quadrata è triangolare inferiore se tutti i coefficienti sopra la diagonale principale sono nulli: $j > i \Rightarrow a_{ij} = 0$.

7.5: Matrici diagonali: Una matrice quadrata è diagonale se è allo stesso tempo triangolare superiore e triangolare inferiore. Altrimenti detto una matrice quadrata è diagonale se i suoi coefficienti al di fuori della diagonale principale sono nulli ($i \neq j \Rightarrow a_{ij} = 0$). La seguente matrice è diagonale:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

7.6: Matrici scalari: Una matrice quadrata, A , è detta scalare se A è diagonale e se tutti i coefficienti della diagonale principale sono uguali (ad uno stesso scalare di k). Una matrice scalare è dunque della forma $\alpha \cdot I_n$.

Esercizi:

10.1) Dimostrare il teorema 1.

10.2) Siano $a = (1, 1)$, $b = (1, 0)$ appartenenti a \mathbb{R}^2 .

(i) Mostrare che $B = (a, b)$ è una base di \mathbb{R}^2 .

(ii) Siano le applicazioni lineari $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2: (x, y) \rightarrow (x+y, x-y)$, $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2: (x, y) \rightarrow (2x-y, x-y)$. Determinare le matrici $M = \text{mat}(f; B, C)$, $N = \text{mat}(g; B, C)$ dove C indica la base canonica di \mathbb{R}^2 .

(iii) Mostrare che $NM \neq \text{mat}(g \circ f; B, C)$.

10.3) Se non li avete ancora fatti, fate gli esercizi 9.2, 9.4, 9.5.

10.4) Una matrice $M = (a_{ij}) \in M_3(\mathbb{R})$ si dice magica se le otto somme dei coefficienti delle tre colonne, delle tre righe e delle due diagonali sono uguali ad uno stesso numero reale.

Altrimenti detto M è magica (di somma $s(M) = \alpha$) se: $\sum_{i=1}^3 a_{ij} = \alpha$, $\sum_{j=1}^3 a_{ij} = \alpha$, $\sum_{i=1}^3 a_{ji} = \alpha$,

$a_{13} + a_{22} + a_{31} = \alpha$. Si nota \mathfrak{M} l'insieme delle matrici magiche: $\mathfrak{M} = \{M \in M_3(\mathbb{R}) / \exists \alpha \in \mathbb{R}, M \text{ è magica con } s(M) = \alpha\}$.

(i) Dimostrare che \mathfrak{M} è un sottospazio vettoriale di $M_3(\mathbb{R})$.

(ii) Se $M \in M_3(\mathbb{R})$, $M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, si pone $\sigma(M) = \begin{pmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{pmatrix}$ (sono state

permutate la prima e la terza colonna). Osservare che se M è magica allora anche $\sigma(M)$ è magica. Abbiamo quindi un'applicazione $\sigma: \mathfrak{M} \rightarrow \mathfrak{M} : M \mapsto \sigma(M)$. Dire se σ è: (a) lineare, (b) iniettiva, (c) suriettiva.

(iii) Si nota \mathfrak{M}_- l'insieme delle matrici magiche che sono antisimmetriche. Se $M \in \mathfrak{M}_-$, quanto vale $s(M)$?

(iv) Determinare \mathfrak{M}_- .

(v) Dimostrare che \mathfrak{M}_- è un sottospazio vettoriale di \mathfrak{M} e determinare la sua dimensione.

(vi) Dedurre da (v) che $\dim(\mathfrak{M}_-) \geq 2$.

11) CAMBIAMENTI DI BASI.

Siano E un k -spazio vettoriale e $B = (e_1, \dots, e_n)$ una base di E ; se $v \in E$ allora v si decompone sulla base B : $v = \alpha_1 e_1 + \dots + \alpha_n e_n$. Gli scalari $\alpha_1, \dots, \alpha_n$, sono le coordinate (o componenti) di v rispetto alla base (e_1, \dots, e_n) . Sia adesso $B' = (e'_1, \dots, e'_n)$ un'altra base di E . Lo stesso vettore v si decompone rispetto a quest'ultima base: $v = \alpha'_1 e'_1 + \dots + \alpha'_n e'_n$. Gli scalari $\alpha'_1, \dots, \alpha'_n$, sono le coordinate di v rispetto alla base B' . In questo paragrafo si dà un procedimento generale per calcolare le coordinate dei vettori di E nella base B' in funzione delle loro coordinate nella base B .

Sia F un altro k -spazio vettoriale e $T = (f_1, \dots, f_p)$ una base di F . Inoltre sia $g : E \rightarrow F$ un morfismo lineare. Rispetto alle basi B e T di E , F , l'applicazione g è associata alla matrice, $M = \text{mat}(g; B, T)$. Sia adesso $T' = (f'_1, \dots, f'_p)$ un'altra base di F e $M' = \text{mat}(g; B', T')$ la matrice associata a g rispetto alle basi B' , T' . Nella seconda parte di questo paragrafo si dà un procedimento per calcolare, per ogni g (i.e. "una volta per tutte"), $\text{mat}(g; B', T')$ in funzione di $\text{mat}(g; B, T)$.

Cambiamento di basi in uno spazio vettoriale.

Sia E un k -spazio vettoriale e $B = (e_1, \dots, e_n)$, $B' = (e'_1, \dots, e'_n)$ due basi di E . Sia $\text{Id}_E : E \rightarrow E$ l'applicazione "identità" ($\forall x \in E$, $\text{Id}_E(x) = x$). Consideriamo la matrice di Id_E rispetto alle seguenti basi: nello spazio di partenza prendiamo la base B , nello spazio d'arrivo la base B' . I vettori colonna della matrice, $P = \text{mat}(\text{Id}_E; B, B')$, associate ad Id_E rispetto a queste basi sono i vettori $\text{Id}_E(e_i)$ nella base B' :

$$P = \text{mat}(\text{Id}_E; B, B') = (\|\text{Id}_E(e_1)\|, \|\text{Id}_E(e_2)\|, \dots, \|\text{Id}_E(e_n)\|)_{B'}$$

Con $\|\text{Id}_E(e_k)\|$ si rappresenta il vettore colonna le cui componenti sono le coordinate del vettore $\text{Id}_E(e_k)$, nella base B' . Il fatto che le coordinate vadano prese nella base B' è indicato con l'indice B' .

Abbiamo $\text{Id}_E(e_k) = e_k = \beta_{1k} e'_1 + \dots + \beta_{ik} e'_i + \dots + \beta_{nk} e'_n$, $1 \leq k \leq n$, per certi scalari β_{ij} . Possiamo scrivere la matrice, P , di Id_E rispetto alle basi B , B' .

$$P = \begin{pmatrix} \text{Id}_E(e_1) & \dots & \text{Id}_E(e_k) & \dots & \text{Id}_E(e_n) \\ e'_1 & \left(\begin{matrix} \beta_{11} & \dots & \beta_{1k} & \dots & \beta_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ \beta_{i1} & \dots & \beta_{ik} & \dots & \beta_{in} \\ \dots & \dots & \dots & \dots & \dots \\ \beta_{n1} & \dots & \beta_{nk} & \dots & \beta_{nn} \end{matrix} \right) \\ e'_n \end{pmatrix}$$

Osserviamo che questa matrice è ben diversa dalla matrice I_n (§10, 2).

Sia $v = \alpha_1 e_1 + \dots + \alpha_n e_n$. Per avere le coordinate $(\alpha'_1, \dots, \alpha'_n)$ di $v = \text{Id}_E(v)$ nella base (e'_1, \dots, e'_n) , basta fare (cfr §9, 3) il prodotto RICO delle matrici P, V dove V è il vettore colonna $(\alpha_1, \dots, \alpha_n)$

$$P \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}$$

Svolgendo il prodotto RICO, con le notazioni precedenti, otteniamo:

$$\alpha'_1 = \beta_{11} \alpha_1 + \dots + \beta_{1k} \alpha_k + \dots + \beta_{1n} \alpha_n.$$

.....

$$(†) \quad \alpha'_i = \beta_{i1} \alpha_1 + \dots + \beta_{ik} \alpha_k + \dots + \beta_{in} \alpha_n$$

.....

$$\alpha'_n = \beta_{n1} \alpha_1 + \dots + \beta_{nk} \alpha_k + \dots + \beta_{nn} \alpha_n.$$

Le formule (†) danno le coordinate di v nella "nuova" base (B') in funzione della matrice P e delle "vecchie" coordinate (nella base B).

1: Teorema : Siano $B = (e_1, \dots, e_n)$, $B' = (e'_1, \dots, e'_n)$ due basi del k -spazio vettoriale E . La matrice (n,n) , P , che ha per i -esima colonna le coordinate di e_i nella base (e'_1, \dots, e'_n) si chiama matrice di passaggio dalla base B alla base B' . La matrice P è invertibile e si ha:

$$P = \text{mat}(\text{Id}_E, B, B'), \quad P^{-1} = \text{mat}(\text{Id}_E, B', B).$$

Sia v un vettore di E . Notiamo con V, V' il vettore colonna delle coordinate di v rispetto alle basi B, B' . Allora: $P.V = V'$; $P^{-1}.V' = V$.

Dim: Sia $\text{Id}_E : E_B \rightarrow E_B$ dove abbiamo indicato in indice le basi scelte. Sia $P = \text{mat}(\text{Id}_E; B, B')$ la matrice associata ad Id_E rispetto alle basi B, B' . L'applicazione "identità" è biettiva e $(\text{Id}_E)^{-1} = \text{Id}_E$. Sia $N = \text{mat}(\text{Id}_E; B', B)$ la matrice associata a $\text{Id}_E : E_B \rightarrow E_B$. Componendo abbiamo: $E_B \rightarrow E_{B'} \rightarrow E_B$ che non è altro che l'identità ("nella base B' "). Trascrivendo in termini di matrici: $NP = I_n$ (osservare che la matrice di Id_E "nella base B' " i.e. con la stessa base B all'arrivo e alla partenza è proprio I_n , §10, 2). Nello stesso modo, guardando la composizione: $E_B \rightarrow E_B \rightarrow E_{B'}$, otteniamo $PN = I_n$. Ne deduciamo che P è invertibile con $P^{-1} = N$ ♦

1.1: Osservazione : Per ottenere la matrice P bisogna risolvere le n equazioni lineari: $e_k = x_1 e'_1 + \dots + x_n e'_n$, $1 \leq k \leq n$. La matrice P è invertibile ed abbiamo un modo per calcolare P^{-1} : "basta" risolvere le n equazioni lineari: $e'_k = x_1 e_1 + \dots + x_n e_n$, $1 \leq k \leq n$.

1.2: Osservazione : Abbiamo chiamato P matrice di passaggio dalla base B alla base B' perché P permette di passare dalle coordinate nella base B a quelle nella base B' ($P \cdot V = V'$). Certi autori chiamano matrice di passaggio dalla base B alla base B' la matrice P^{-1} per il motivo seguente: quando si cambiano le coordinate in un problema, si hanno spesso delle relazioni espresse nelle "vecchie" coordinate, del tipo $F(y_1, \dots, y_n) = 0$, per avere le relazioni corrispondenti nelle "nuove" coordinate è comodo avere delle formule che danno gli y_i in funzione degli y'_i , e non il contrario. L'unico modo per non confondersi è comunque di ragionare sulle applicazioni e di "ritrovare" la formula ogni volta! Per questo basta scrivere le applicazioni composte indicando le basi in indice.

Cambiamenti di basi e applicazioni lineari.

Siano E, F due k -spazi vettoriali e $g : E \rightarrow F$ un morfismo lineare. Siano inoltre $A = (e_1, \dots, e_n)$ una base di E e $B = (f_1, \dots, f_p)$, una base di F . Sia $M = \text{mat}(g; A, B)$. Sia $A' = (e'_1, \dots, e'_n)$ un'altra base di E e $B' = (f'_1, \dots, f'_p)$, un'altra base di F . Qual è la matrice associata a g rispetto alle basi A', B' ?

Consideriamo l'applicazione composta (gli indici indicano le basi scelte):

$$E_{(A')} \xrightarrow{\quad} \text{Id}_E \xrightarrow{\quad} E_{(A)} \xrightarrow{\quad} g \xrightarrow{\quad} F_{(B)} \xrightarrow{\quad} \text{Id}_F \xrightarrow{\quad} F_{(B')}$$

Questa non è altro che $g : E_{(A')} \rightarrow F_{(B')}$ ($\text{Id}_F \circ g \circ \text{Id}_E = g$), la matrice associata è $M' = \text{mat}(g; A', B')$. D'altra parte se P è la matrice (n,n) le cui colonne sono le coordinate di e'_k rispetto alla base (e_1, \dots, e_n) , $1 \leq k \leq n$, allora P è la matrice associata a Id_E rispetto alle basi A' , A ; $P = \text{mat}(\text{Id}_E; A', A)$. Nello stesso modo se Q è la matrice (p,p) le cui colonne sono le coordinate di f_t rispetto alla base (f_1, \dots, f_p) , $1 \leq t \leq p$, allora $Q = \text{mat}(\text{Id}_F, B, B')$.

Trascrivendo l'uguaglianza $\text{Id}_F \circ g \circ \text{Id}_E = g$ in termini di matrici associate otteniamo:

$$M' = Q.M.P.$$

Attenzione a non sbagliare l'ordine! L'ordine da seguire è quello della composizione delle applicazioni: prima si applica Id_E , quindi se applichiamo M' ad un vettore colonna V bisogna per prima cosa moltiplicarlo con $P = \text{mat}(\text{Id}_E; A', A)$ e via di seguito...

Abbiamo dimostrato:

2: Teorema : Siano E, F due k -spazi vettoriali di dimensione finita. Siano A, A' due basi di E e B, B' due basi di F . Siano $P = \text{mat}(\text{Id}_E; A', A)$ e $Q = \text{mat}(\text{Id}_F; B, B')$, le matrici di passaggio. Per ogni morfismo lineare $g : E \rightarrow F$, se $M = \text{mat}(g; A, B)$ e $M' = \text{mat}(g; A', B')$ allora: $M' = Q.M.P.$

3: Corollario : Siano E un k -spazio vettoriale e $B = (e_1, \dots, e_n)$, $B' = (e'_1, \dots, e'_n)$ due basi di E . Per ogni endomorfismo $g : E \rightarrow E$ si ha: $M' = P^{-1}.M.P$ dove $M = \text{mat}(g; B, B)$, $M' = \text{mat}(g; B', B')$ e P è la matrice di passaggio dalla base B' alla base B ($P = \text{mat}(\text{Id}_E; B', B)$).

Matrici equivalenti.

Sull'insieme $M_{n,p}(k)$ delle matrici con n righe e p colonne definiamo la relazione binaria: $M \sim N$ se e solo se esistono delle matrici quadrate, invertibili, A, B , tali che: $M = A.N.B$.

Questa è una relazione d'equivalenza (**Es.4**).

4: Definizione: Siano M, N due elementi di $M_{n,p}(k)$. Le matrici M, N sono dette equivalenti se $M \sim N$.

5: Proposizione: Siano M, N due elementi di $M_{n,p}(k)$. Le matrici M, N sono equivalenti se e solo se possono essere associate ad una stessa applicazione lineare $f : E \rightarrow F$ (E, F due k -spazi vettoriali di dimensioni p, n), i.e. esistono delle basi A, A', B, B' di E, F tali che: $M = \text{mat}(f; A, B)$ e $N = \text{mat}(f; A', B')$.

Dim: (i) Se $M = \text{mat}(f; A, B)$ dove A, B sono delle basi di E, F e se $N = \text{mat}(f; A', B')$ allora facendo il cambiamento di basi (cfr 2): $N = Q.M.P$ e $N \sim M$.

(ii) Viceversa supponiamo $N \sim M$ quindi $N = Q.M.P$ con Q, P matrici invertibili. Siano E, F due k -spazi vettoriali di dimensioni p, n (per esempio $E = k^p, F = k^n$) e A, B delle basi di E, F . Sia $f : E \rightarrow F$ l'applicazione lineare che rispetto alle basi A, B è rappresentata dalla matrice M ($M = \text{mat}(f; A, B)$). Adesso definiamo delle basi A', B' , di E, F nel modo seguente: i vettori di A' , nella base A , sono i vettori colonna della matrice P ; i vettori di B' , nella base B , sono i vettori colonna della matrice Q^{-1} . Nelle basi A', B' la matrice associata ad f è N : $\text{mat}(f; A', B') = N$ ♦

Nello stesso modo sull'insieme $M_n(k)$ delle matrici quadrate con n righe definiamo la relazione binaria: $M \approx N$ se e solo se esiste una matrice quadrata, invertibile, A , tale che: $M = A^{-1}.N.A$. Si verifica che questa è una relazione d'equivalenza (Es.4).

6: Definizione: Due matrici quadrate M, N sono simili se $M \approx N$.

In modo analogo a quanto fatto prima abbiamo:

7: Proposizione: Due matrici quadrate, M, N sono simili se e solo se rappresentano uno stesso endomorfismo, "con la stessa base all'arrivo e alla partenza", i.e. $M, N \in M_n(k)$ sono simili se e solo se esistono delle basi, B, B' , di E (un k -spazio vettoriale di dimensione n), e un endomorfismo $f : E \rightarrow E$ tale che $M = \text{mat}(f; B, B)$, $N = \text{mat}(f; B', B')$.

7.1: Osservazione : Scopo di uno dei prossimi paragrafi ("Diagonalizzazione delle matrici") sarà il problema seguente: data una matrice quadrata M trovare una matrice (quadrata), simile a M , più "semplice" possibile.

Esercizi:

11.1) Siano in \mathbb{R}^3 , $v_1 = (-1, 0, 1)$, $v_2 = (-1, 1, 0)$, $v_3 = (2, 0, 0)$.

(i) Mostrare che $B' = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 .

(ii) Si nota B la base canonica di \mathbb{R}^3 . Sia f l'applicazione lineare da \mathbb{R}^3 in \mathbb{R}^3 tale che $\text{mat}(f; B, B') = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$. Determinare $\text{mat}(f; B', B')$.

11.2) (i) Siano $e'_1 = (1, 0, 0)$, $e'_2 = (0, -1, 0)$, $e'_3 = (1, 0, -1)$ in \mathbb{R}^3 . Dimostrare che $B' = (e'_1, e'_2, e'_3)$ è una base di \mathbb{R}^3 .

(ii) Sia $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, l'applicazione lineare rappresentata, nella base canonica, dalla matrice $M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$. Determinare $\text{mat}(f; B, B')$, $\text{mat}(f; B', B)$.

(iii) Mostrare che M è invertibile e calcolare M^{-1} .

(iv) Determinare $\text{mat}(f^{-1}; B', B')$.

11.3) Mostrare che la matrice $M \in M_2(\mathbb{R})$, $M = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$, è simile ad una matrice diagonale (suggerimento: associare un endomorfismo f a M e cercare due vettori linearmente indipendenti, u, v tali che $f(u) = \lambda u$, $f(v) = \mu v$). Mostrare invece che la matrice $N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ non è mai simile ad una matrice diagonale di $M_2(\mathbb{R})$.

11.4) (i) Mostrare che le relazioni di equivalenza (\sim) e di similitudine (\approx) (cfr 4, 6) sono delle relazioni d'equivalenza.

(i) Descrivere gli insiemi quoziante: $M_1(k)/\sim$ e $M_1(k)/\approx$ (nb: $M_1(k) = k$).

(ii) Mostrare che $M_2(\mathbb{R})/\sim$ e $M_2(\mathbb{R})/\approx$ sono due insiemi di cardinalità diverse.

12) RANGO DI UNA MATRICE.

Prima di tutto ricordiamo (cfr §8) la nozione di rango di un sistema (di un insieme) di vettori:

1: Definizione: Sia E un k -spazio vettoriale di dimensione finita e v_1, \dots, v_k dei vettori di E . Il rango del sistema di vettori $\{v_1, \dots, v_k\}$ è la dimensione del sottospazio $\langle v_1, \dots, v_k \rangle$ generato da questi vettori.

2: Definizione: Siano E, F due k -spazi vettoriali di dimensione finita e $f : E \rightarrow F$ un'applicazione lineare. Il rango di f (si nota $rg(f)$) è uguale a $\dim(\text{Im } f)$. Se (e_1, \dots, e_n) è una base di E , $rg(f)$ è uguale al rango di $\{f(e_1), \dots, f(e_n)\}$ (cfr §5, 1).

3: Definizione: Sia M una matrice (n,p) a coefficienti in k . Il rango di M è il rango del sistema dei vettori colonna di M (visti come vettori di k^n).

3.1: Osservazione : I vettori colonna di una matrice (n,p) a coefficienti in k sono p vettori di k^n quindi $rg(M) \leq \min\{n, p\}$.

Il lemma seguente giustifica la definizione 3 :

4: Lemma: Siano $f : E \rightarrow F$ un'applicazione lineare, B, A delle basi di E, F e $M = \text{mat}(f; B, A)$. In queste condizioni $rg(f) = rg(M)$.

Dim: Infatti, sia $B = (e_1, \dots, e_n)$; i vettori colonna di M non sono altri che $f(e_1), \dots, f(e_n)$ e si conclude con 2♦

Scopo dei due prossimi lemmi è di mostrare che il rango di una matrice M è uguale al rango della sua trasposta ${}^t M$.

5: Lemma: Siano $g : E \rightarrow F$ un'applicazione lineare, $B = (e_1, \dots, e_n)$, $A = (f_1, \dots, f_p)$ delle basi di E, F e $M = \text{mat}(g; B, A)$. Notiamo B^*, A^* le basi duali delle basi B, A . Allora $\text{mat}({}^t g; A^*, B^*) = {}^t M$.

Dim: Ricordiamo che ${}^t g : F^* \rightarrow E^*$ è l'applicazione che alla forma lineare $\Psi : F \rightarrow k$ associa la forma lineare $\Psi \circ g : E \rightarrow k$. La matrice, N , di ${}^t g$ ha per vettori colonne le componenti di ${}^t g(f^*_k)$ nella base (e^*_1, \dots, e^*_n) . Sia ${}^t g(f^*_k) = \alpha_{1k} e^*_1 + \dots + \alpha_{nk} e^*_n$. Se valutiamo in e_i abbiamo ${}^t g(f^*_k)(e_i) = (\alpha_{1k})$

$e^*_1 + \dots + \alpha_{nk} e^*_n)(e_i) = \alpha_{ik}$. D'altra parte ${}^t g(f^*_k)(e_i) = f^*_k(g(e_i))$. Se $M = (m_{kt})$ allora per definizione di matrice associata nelle basi B, A, le componenti di $g(e_i)$ sono quelle dell'i-esimo vettore colonna di M: $g(e_i) = m_{1i} f_1 + \dots + m_{pi} f_p$. Pertanto: $f^*_k(g(e_i)) = f^*_k(m_{1i} f_1 + \dots + m_{pi} f_p) = m_{ki}$. In conclusione se $M = (m_{kt})$ e se $N = (\alpha_{sv})$ allora $m_{ki} = \alpha_{ik}$; questo mostra che $N = {}^t M$ ♦

6: Corollario: Sia M una matrice (n,p) a coefficienti in k. Il rango di M è uguale al rango di ${}^t M$.

Dim: Si può pensare M come la matrice associata ad una certa applicazione lineare $g : E \rightarrow F$, $\dim(E) = p$, $\dim(F) = n$, rispetto a delle basi B, A di E, F. Abbiamo allora $\text{rg}(M) = \dim(\text{Im } g)$ (cfr lemma 4) e $\text{rg}({}^t M) = \dim(\text{Im } {}^t g)$ (cfr lemmi 4, 5). Si conclude con il teorema §7, 20 ♦

6.1: Osservazione: Dal corollario 6 segue che $\text{rg}(M)$ è anche il rango del sistema dei vettori righe di M.

7: Corollario: Sia $g : E \rightarrow F$ un'applicazione lineare tra due k-spazi vettoriali di dimensioni n, p. Allora $\text{rg}(g) = r$ se e solo se esistono delle basi B, A di E, F tali che $\text{mat}(g; B, A) = M_r$, dove:

$$M_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{p-r}^{r \quad n-r}$$

Qui I_r è la matrice identità (r,r) e gli 0 stanno per matrici (di dimensioni opportune) i cui coefficienti sono tutti nulli.

Dim: Supponiamo $\text{rg}(g) = r$. Dal teorema delle dimensioni (§5, 2): $\dim(\text{Ker}(g)) = n-r$. Sia (e_{r+1}, \dots, e_n) una base di $\text{Ker}(g)$. Questa base può essere completata in una base di E: $B = (e_1, \dots, e_n)$. Sappiamo allora che, $(g(e_1), \dots, g(e_r))$ è una base di $\text{Im}(g)$ (cfr con la dimostrazione di §5, 2). Questa base può essere completata in una base, A, di F: $A = (g(e_1), \dots, g(e_r), f_{r+1}, \dots, f_p)$. Si ha $\text{mat}(g; B, A) = M_r$.

Viceversa è chiaro che se $\text{mat}(g; B, A) = M_r$ per delle basi B, A, allora $\text{rg}(g) = r$ ♦

8: Matrici simili e equivalenti:

Sia E un k -spazio vettoriale di dimensione n . Sull'insieme $M_n(k)$ abbiamo due relazioni d'equivalenza: la relazione di equivalenza (\sim) e la relazione di similitudine (\approx) (cfr §11, 4, 6).

Queste due relazioni d'equivalenza sono diverse (cfr Es.11.4). La classe d'equivalenza per \sim dipende solo dal rango: l'insieme quoziente $M_n(k)/\sim$ è in biezione con $\{I_r \mid 0 \leq r \leq n\}$. Questo segue dal corollario 7. Per quanto riguarda l'insieme quoziente $M_n(k)/\approx$, lo studieremo più avanti (cfr diagonalizzazione delle matrici) ma per convincerci che queste relazioni sono diverse osserviamo che l'insieme delle matrici (n,n) equivalenti a I_n è l'insieme di tutte le matrici di rango n (le matrici i cui vettori colonna (o riga) sono linearmente indipendenti) invece l'insieme delle matrici (n,n) simili a I_n è semplicemente $\{I_n\}$. Infatti $P^{-1}I_nP = P^{-1}P = I_n$.

Esercizi:

12.1) Per ogni $\alpha \in \mathbb{R}$, si considerino i vettori di \mathbb{R}^3 : $u_\alpha = (-1, 2\alpha, -1)$, $v_\alpha = (1, \alpha, 0)$, $w_\alpha = (0, 1, 2\alpha)$. Per ogni α in \mathbb{R} determinare il rango di $(u_\alpha, v_\alpha, w_\alpha)$.

12.2) Sia E un k -spazio vettoriale e $f: E \rightarrow E$ un'applicazione lineare. Dimostrare che $rg(f \circ f) = rg(f)$ se e solo se $Ker(f)$ e $Im(f)$ sono in somma diretta.

12.3) Sia $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3 : (x, y, z) \rightarrow (x+2y+z, y+2z, x-y-5z)$. Calcolare $rg(f \circ f)$.

12.4) Sia $M \in M_{n,p}(k)$ tale che $rg(M) = r$. Sia $V = \{B \in M_{p,m}(k) / M \cdot B = 0\}$. Dimostrare che V è un sottospazio vettoriale di $M_{p,m}(k)$ e determinare la sua dimensione.

12.5) Siano E un k -spazio vettoriale di dimensione n e $g: E \rightarrow E$ un endomorfismo di E . Si pone $A(g) = \{f \in \text{End}_k(E) / f \circ g = 0\}$, $B(g) = \{h \in \text{End}_k(E) / g \circ h = 0\}$.

(i) Dimostrare che $A(g)$ e $B(g)$ sono dei sottospazi vettoriali di $\text{End}_k(E)$.

(ii) Calcolare $\dim(A(g))$ in funzione di n e $p := \dim(\text{Im}(g))$.

(iii) Dimostrare che $A(g)$ e $B(g)$ sono in somma diretta se e solo se g è iniettiva.

§13) DETERMINANTI.

In questo paragrafo (un po' tecnico) si definiscono i determinanti: essi forniscono un metodo pratico per calcolare il rango di una matrice (cfr §15).

Applicazioni e forme multilinearì.

Siano E_1, \dots, E_n, F degli insiemi. Sia $E := E_1 \times \dots \times E_n$ il prodotto cartesiano di E_1, \dots, E_n . Sia $f : E \rightarrow F : (x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n)$ un'applicazione. Si dice anche che f è una funzione delle variabili x_1, \dots, x_n .

Per ogni i , $1 \leq i \leq n$, possiamo definire, per ogni valore dell' $(n-1)$ -uplo $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, un'applicazione, f_i , da E_i in F tramite:

$$f_i : E_i \rightarrow F : x \rightarrow f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n).$$

L'applicazione f_i è l'applicazione parziale da E_i in F , associata a f relativamente ai valori, $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ attribuiti alle altre $n-1$ variabili.

Osserviamo che la notazione f_i è incompleta, sarebbe più rigoroso notare $f_{i,(}x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ per tener conto dei valori $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ attribuiti alle altre $n-1$ variabili.

1: Definizione : Siano E_1, \dots, E_n, F , $(n+1)$ k-spazi vettoriali e $f : E_1 \times \dots \times E_n \rightarrow F : (x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n)$ un'applicazione. L'applicazione f è detta n -lineare se ogni applicazione parziale f_i (qualsiasi siano i valori attribuiti alle altre $n-1$ variabili) è lineare. Se $F = k$, si dice che f è una forma n -lineare.

1.1: Osservazione : Le applicazioni e forme n -lineari ($n \geq 2$) si chiamano applicazioni e forme **multilineari**. Se $n = 2$ si dice che l'applicazione o la forma è **bilineare**; se $n = 3$, trilineare.

Se f , come nella definizione 1, è multilineare, per ogni $\lambda, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_i, y_i$, si ha: $f(x_1, \dots, x_{i-1}, x_i + y_i, x_{i+1}, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) + f(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)$

$$f(x_1, \dots, x_{i-1}, \lambda x_i, x_{i+1}, \dots, x_n) = \lambda f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n).$$

1.2: Osservazione : Attenzione a non confondere applicazioni lineari e applicazioni n-lineari!

Se $f : E_1 \times E_2 \rightarrow F$ è bilineare allora:

$$f(x_1 + y_1, x_2 + y_2) = f(x_1, x_2) + f(x_1, y_2) + f(y_1, x_2) + f(y_1, y_2)$$

$$f(\lambda x, y) = f(\lambda x, \lambda y) = \lambda^2 f(x, y).$$

Invece se $g : E_1 \times E_2 \rightarrow F$ è lineare allora:

$$g(x_1 + y_1, x_2 + y_2) = g((x_1, x_2) + (y_1, y_2)) = g(x_1, x_2) + g(y_1, y_2)$$

$$g(\lambda x, y) = \lambda g(x, y)$$

1.3: Esempi : (i) L'applicazione $R \times R \rightarrow R : (x, y) \rightarrow xy$ è bilineare.

(ii) L'applicazione $E \times E^* \rightarrow k : (x, f) \rightarrow f(x)$ è bilineare.

La proposizione seguente è di immediata verifica:

2: Proposizione : Siano E_1, \dots, E_n, F , ($n+1$) k-spazi vettoriali. Sia $L_n(E_1 \times \dots \times E_n, F)$ l'insieme delle applicazioni multilinear da $E_1 \times \dots \times E_n$ in F . Per le usuali operazioni di somma e moltiplicazione per uno scalare, $L_n(E_1 \times \dots \times E_n, F)$ è un k-spazio vettoriale.

Applicazioni e forme n-lineari alternanti.

3: Definizione : Un'applicazione multilineare $f : E^n \rightarrow F$ (E, F due k-spazi vettoriali, $E^n = E \times \dots \times E$) si dice alternante se $f(X) = 0$ per ogni $X = (x_1, \dots, x_n) \in E^n$ avente due coordinate uguali (i.e. $x_i = x_k$ con $i \neq k$). Se $F = k$ si dice anche che f è una forma n-lineare alternante (o anche una n-forma alternante).

3.1: Osservazione : Se f è multilineare alternante allora:

$f(x_1, \dots, x_{i-1}, x_i + y_k, x_{i+1}, \dots, x_{k-1}, x_i + y_k, x_{k+1}, \dots, x_n) = 0$ perché la i-esima e la k-esima coordinata sono uguali ($= x_i + y_k$). D'altra parte per multilinearità:

$$f(x_1, \dots, x_{i-1}, x_i + y_k, x_{i+1}, \dots, x_{k-1}, x_i + y_k, x_{k+1}, \dots, x_n) =$$

$$f(\dots, x_i, \dots, x_i, \dots) + f(\dots, x_i, \dots, y_k, \dots) + f(\dots, y_k, \dots, y_k, \dots) + f(\dots, y_k, \dots, x_i, \dots) =$$

(le variabili non scritte hanno lo stesso valore nei quattro termini). Siccome f è alternante il primo e il terzo termine sono nulli, perciò:

$$f(\dots, x_i, \dots, y_k, \dots) = -f(\dots, y_k, \dots, x_i, \dots).$$

Se f, g sono due applicazioni n-lineari alternanti definite su E^n , è chiaro che $f+g, \lambda f$ sono ancora delle applicazioni n-lineari alternanti:

4: Proposizione: Siano E, F due k -spazi vettoriali. L'insieme delle applicazioni n -lineari alternanti definite su E^n e a valori in F è (per le operazioni usuali) un k -spazio vettoriale (più precisamente è un sottospazio vettoriale di $L_n(E \times \dots \times E, F)$).

5: Lemma: Sia $f : E^n \rightarrow F$ un'applicazione n -lineare alternante. Se i vettori x_1, \dots, x_n di E sono linearmente dipendenti allora $f(x_1, \dots, x_n) = 0$.

Dim: Se i vettori x_1, \dots, x_n sono linearmente dipendenti esistono degli scalari non tutti nulli ξ_1, \dots, ξ_n tali che $\xi_1x_1 + \dots + \xi_nx_n = 0$. A meno di riordinare gli indici possiamo supporre $\xi_1 \neq 0$ e $x_1 = \alpha_2x_2 + \dots + \alpha_nx_n$ ($\alpha_i = \xi_i/\xi_1$). Pertanto $f(x_1, \dots, x_n) = f(\alpha_2x_2 + \dots + \alpha_nx_n, x_2, \dots, x_n)$, per multilinearità $f(\alpha_2x_2 + \dots + \alpha_nx_n, x_2, \dots, x_n) = \alpha_2f(x_2, x_2, \dots, x_n) + \dots + \alpha_nf(x_n, x_2, \dots, x_n)$; in ogni termine del membro di destra la prima variabile è uguale ad una delle altre variabili. Siccome f è alternante concludiamo che $f(x_1, \dots, x_n) = 0$ ♦

6: Corollario: Siano E, F due k -spazi vettoriali. Se $p > \dim(E)$, ogni applicazione p -lineare alternante $f : E^p \rightarrow F$ è identicamente nulla.

Dim: Infatti p vettori di E sono sempre linearmente dipendenti perché $p > \dim(E)$ ♦

Non svilupperemo la teoria generale delle applicazioni n -lineari alternanti ma, d'ora in poi, ci limiteremo alle forme n -lineari alternanti su E dove $\dim(E) = n$ (in un certo senso, cfr Cor.6, è il caso "estremo"). Questo è sufficiente per quello che abbiamo in mente (determinanti). Iniziamo col caso $n = 2$.

Forme n -lineari alternanti su E^n ($n = \dim(E) = 2$).

Sia $B = (e_1, e_2)$ una base di E . Se x, x' sono due vettori di E allora si scrivono nella base B : $x = \alpha e_1 + \beta e_2$, $x' = \alpha' e_1 + \beta' e_2$. Per ogni forma bilineare alternante $f : E^2 \rightarrow k$ abbiamo: $f(x, x') = f(\alpha e_1 + \beta e_2, \alpha' e_1 + \beta' e_2) = \alpha\alpha' f(e_1, e_1) + \alpha\beta' f(e_1, e_2) + \beta\alpha' f(e_2, e_1) + \beta\beta' f(e_2, e_2) = (\alpha\beta' - \alpha'\beta) f(e_1, e_2)$.

(*) Vediamo che f è completamente determinata da $f(e_1, e_2)$.

Adesso, con le notazioni precedenti, l'applicazione $d : E^2 \rightarrow k : (x, x') \mapsto \alpha\beta' - \alpha'\beta$, è chiaramente bilineare alternante ed inoltre $d(e_1, e_2) = 1$. Da quanto precede, se poniamo $\lambda = f(e_1, e_2)$, abbiamo $f = \lambda d$. Abbiamo dimostrato:

7: Proposizione: *Sia E un k -spazio vettoriale di dimensione due. Lo spazio vettoriale delle forme bilineari alternanti su E ha dimensione uno. Inoltre se $B = (e_1, e_2)$ è una base di E esiste un'unica forma bilineare alternante, f , tale che $f(e_1, e_2) = 1$; questa forma si nota \det_B e si chiama il determinante rispetto alla base B .*

Osserviamo un modo pratico per visualizzare \det_B : si fa la matrice (2,2) i cui vettori colonna sono le componenti di x, x' nella base $B = (e_1, e_2)$:

$$\begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$$

poi si fa la differenza dei prodotti in croce: $\alpha\beta' - \alpha'\beta$. Lo scalare così ottenuto è il determinante dei vettori x, x' rispetto alla base B . Il determinante si rappresenta anche sotto forma di tabella ma per distinguerlo dalla matrice si mettono delle sbarre al posto delle parentesi:

$$\det_B(x, x') = \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix} = \alpha\beta' - \alpha'\beta.$$

8: Proposizione: *Sia E un k -spazio vettoriale di dimensione 2. Due vettori x, x' di E sono linearmente indipendenti se e solo se per ogni forma bilineare alternante non nulla $g : E^2 \rightarrow k$, si ha $g(x, x') \neq 0$.*

Dim: (i) Supponiamo x, x' linearmente indipendenti. Allora $B = (x, x')$ è una base di E .

Da 7: $\det_B(x, x') = 1$ e se f è una forma bilineare alternante non nulla allora $f = \lambda \det_B$, con $\lambda \neq 0$. Pertanto $f(x, x') = \lambda \neq 0$.

(ii) Supponiamo che $g(x, x') \neq 0$ per ogni (ne basta una in realtà) forma bilineare alternante non nulla. Dal lemma 5 segue che x, x' non sono linearmente dipendenti.♦

8.1: Osservazione : In particolare x e x' sono linearmente indipendenti se e solo se per una base, B , di E : $\det_B(x, x') \neq 0$.

Abbiamo così un metodo meccanico per vedere se due vettori di un k -spazio vettoriale di dimensione due sono o meno linearmente indipendenti: si

sceglie una base e si calcola il determinante dei due vettori rispetto a questa base. I due vettori sono linearmente indipendenti se e solo se il loro determinante è non nullo.

9: Esempio : Siano $u = (3, 5)$, $v = (-1, 3)$ due vettori di \mathbb{R}^2 . Il loro determinante rispetto alla base canonica di \mathbb{R}^2 è: $\begin{vmatrix} 3 & -1 \\ 5 & 3 \end{vmatrix} = 3 \cdot 3 - (-1) \cdot 5 = 14$, quindi u e v sono linearmente indipendenti.

9.1: Osservazione : Questo criterio del determinante rispecchia perfettamente il fatto che due vettori sono linearmente indipendenti se e solo se non sono proporzionali.

Rango di una matrice (2,2).

Sia M una matrice (2,2) a coefficienti in k , $M = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$, $M \in M_2(k)$. Per conoscere il rango di M basta calcolarne il determinante. Per definizione il determinante di M , $\det(M)$, è il determinante dei vettori u , v le cui coordinate nella base canonica, C , di k^2 sono date rispettivamente dal primo e dal secondo vettore colonna di M : $u = (\alpha, \beta)$, $v = (\alpha', \beta')$.

Se $\det_C(u, v) = \det(M) = \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix} = \alpha\beta' - \alpha'\beta \neq 0$, u e v sono linearmente indipendenti e $\text{rg}(M) = 2$.

Se $\det_C(u, v) = 0$, u e v sono linearmente dipendenti. In questo caso $\dim\langle u, v \rangle = 1$ se e solo se uno dei due vettori u , v non è nullo. Quindi se $\det(M) = 0$, $\text{rg}(M) = 1 \Leftrightarrow (\alpha, \beta, \alpha', \beta') \neq (0, 0, 0, 0)$.

Adesso bisogna ripetere tutto quanto nel caso generale ($\dim(E) = n$). Non ci sono difficoltà essenziali ma il calcolo di un determinante (n, n) necessita di preliminari un po' tecnici.

Forme n-lineari alternanti su E^n ($\dim(E) = n \geq 2$)

Sia E un k -spazio lineare e $B = (e_1, \dots, e_n)$, una base di E . Siano x_1, \dots, x_n , n vettori di E , consideriamo le coordinate di questi vettori nella base B :

$$x_1 = \xi^{(1)}_1 e_1 + \dots + \xi^{(1)}_n e_n.$$

.....

$$x_n = \xi^{(n)}_1 e_1 + \dots + \xi^{(n)}_n e_n.$$

Sia $g : E^n \rightarrow k$ una forma n-lineare alternante e cerchiamo di calcolare $g(x_1, \dots, x_n)$:

$g(x_1, \dots, x_n) = g(\xi^{(1)}_1 e_1 + \dots + \xi^{(1)}_n e_n, \dots, \xi^{(n)}_1 e_1 + \dots + \xi^{(n)}_n e_n)$. Se sviluppiamo il membro di destra usando la multilinearità di g , troviamo una somma di n^n termini della forma:

$$\xi^{(1)}_{i_1} \cdot \xi^{(2)}_{i_2} \cdots \xi^{(n)}_{i_n} g(e_{i_1}, \dots, e_{i_n})$$

dove i_1, \dots, i_n sono degli elementi qualsiasi di $\{1, 2, \dots, n\}$. Osserviamo che darsi i_1, \dots, i_n è equivalente a darsi un'applicazione da $\{1, \dots, n\}$ in se stesso. Se A è l'insieme di tutte le applicazioni da $\{1, \dots, n\}$ in se stesso, allora $\text{card}(A) = n^n$ (cfr Es. I.3.8). Siccome g è alternante un tal termine è nullo tranne se i_1, \dots, i_n sono tutti distinti. Sia $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ l'applicazione definita da $\sigma(k) = i_k$. Gli i_1, \dots, i_n sono tutti distinti se e solo se σ iniettiva, ossia biettiva (Es. I.3.5). In altre parole gli i_1, \dots, i_n sono tutti distinti se e solo se σ appartiene al gruppo, S_n , delle permutazioni di $\{1, 2, \dots, n\}$ (cfr Es. I.5.7).

In conclusione:

$$g(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \xi^{(1)}_{\sigma(1)} \cdot \xi^{(2)}_{\sigma(2)} \cdots \xi^{(n)}_{\sigma(n)} g(e_{\sigma(1)}, \dots, e_{\sigma(n)}). \quad (\dagger)$$

L'idea, per valutare questa somma, è di esprimere tutto in funzione di $g(e_1, \dots, e_n)$. Infatti (cfr 3.1) sappiamo che $g(\dots, e_i, \dots, e_k, \dots) = -g(\dots, e_k, \dots, e_i, \dots)$.

Quindi partendo da $g(e_{i_1}, \dots, e_{i_n})$ e scambiando successivamente due i_j (e lasciando gli altri immutati) si cercherà di arrivare a $g(e_1, \dots, e_n)$. Il problema è di sapere quanti scambi si sono fatti, infatti se il numero di scambi è pari allora $g(e_{i_1}, \dots, e_{i_n}) = g(e_1, \dots, e_n)$, se invece questo numero è dispari allora $g(e_{i_1}, \dots, e_{i_n}) = -g(e_1, \dots, e_n)$.

Facciamo un esempio con $n=3$. Sia $(i_1, i_2, i_3) = (2, 3, 1)$. Bisogna arrivare a $(1, 2, 3)$. Possiamo scambiare 1 e 3: $(2, 1, 3)$ e adesso 1 e 2: $(1, 2, 3)$. Quindi abbiamo finito dopo due scambi, il numero ricercato è +1. Avremo potuto anche fare così: scambiamo 2 e 1: $(1, 3, 2)$ e 3 e 2: $(1, 3, 2)$; il numero ricercato è sempre +1. Si può anche procedere con disordine: scambiamo 2 e 3: $(3, 2, 1)$ poi 2 e 1: $(3, 1, 2)$ poi 3 e 1: $(1, 3, 2)$ e finalmente 3 e 2: $(1, 2, 3)$. Abbiamo fatto 4 scambi, il numero cercato è sempre +1 (per fortuna!).

In altre parole se $f : E^3 \rightarrow k$ è multilineare alternante allora $f(e_2, e_3, e_1) = f(e_1, e_2, e_3)$.

Sia adesso $(i_1, i_2, i_3) = (3, 2, 1)$. Possiamo scambiare 3 e 1 e abbiamo finito. Il numero ricercato è -1. Con le notazioni precedenti: $f(e_3, e_2, e_1) = -f(e_1, e_2, e_3)$.

A questo punto sorgono inevitabilmente alcune domande naturali:

-chi garantisce che partendo da un (i_1, \dots, i_n) qualsiasi si arriverà, usando solo scambi di due indici, a $(1, 2, \dots, n)$ in un numero finito di operazioni?

-chi garantisce che la parità del numero di operazioni non dipende dalle scelte che facciamo?

La risposta alla prima domanda non è difficile: si va a cercare 1, per esempio $i_k = 1$; si scambiano il primo e il k-esimo posto; 1 è al suo posto e non interverrà più d'ora in poi. In seguito si cerca 2: $i_m = 2$. Si scambiano il secondo e l'm-esimo posto; 1 e 2 sono a posto e non interverranno più. Dopo al più $n-1$ scambi, abbiamo finito.

La risposta alla seconda domanda è "teoricamente" facile. Supponiamo di sapere che esiste una forma n -lineare alternante non nulla su un k -spazio vettoriale, E, di dimensione n (per esempio $E = k^n$). Sia $f : E^n \rightarrow k$ questa forma n -lineare alternante. Sappiamo (3.1) che $f(e_{i_1}, \dots, e_{i_n}) = \varepsilon \cdot f(e_1, \dots, e_n)$ con $\varepsilon = +1$ o $\varepsilon = -1$, se $\varepsilon = 1$ il numero di scambi in ogni successione sarà pari altrimenti sarà dispari. Solo che, a priori, non sappiamo che esiste una tale forma n -lineare alternante, non nulla. Risolveremo questo problema considerando, in un caso particolare, applicazioni più semplici, "quasi" alternanti (applicazioni antisimmetriche).

10: Definizione: Sia \mathfrak{S}_n il gruppo delle permutazioni di $\{1, 2, \dots, n\}$ (i.e. $\sigma \in \mathfrak{S}_n$ è una biiezione $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$).

Un elemento, τ , di \mathfrak{S}_n si chiama una **trasposizione** se $\tau(i) = k$, $\tau(k) = i$ e $\tau(t) = t$ se $t \neq i, k$. Una tale trasposizione si può anche notare $\tau = (i, k)$.

11: Definizione: Siano X un insieme, G un gruppo abeliano e $f : X^n \rightarrow G$ un'applicazione. L'applicazione f è **antisimmetrica** se: $f(x_{\tau(1)}, \dots, x_{\tau(n)}) = -f(x_1, \dots, x_n)$ per ogni trasposizione, τ , di \mathfrak{S}_n .

12: Lemma: Siano $f : X^n \rightarrow G$ un'applicazione antisimmetrica e $\sigma \in \mathfrak{S}_n$ una permutazione. Se $\sigma = \tau_1 \circ \dots \circ \tau_r$ è una decomposizione di σ in prodotto di r

trasposizioni τ_1, \dots, τ_r , allora, per ogni (x_1, \dots, x_n) in X^n , $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^r f(x_1, \dots, x_n)$.

Dim: Si ragiona per induzione su r . Se $r = 1$, è la definizione 11. Assumiamo il lemma per $r-1$ e cerchiamo di dimostrarlo per r . Scriviamo $\sigma = \tau \circ \xi$, $\tau = \tau_1, \xi = \tau_2 \circ \dots \circ \tau_r$. Per ipotesi di induzione: $f(y_{\xi(1)}, \dots, y_{\xi(n)}) = (-1)^{r-1} f(y_1, \dots, y_n)$, per ogni (y_1, \dots, y_n) in X^n . Poniamo $y_i = x_{\tau(i)}$, allora $y_{\xi(i)} = x_{\tau(\xi(i))} = x_{\sigma(i)}$. Quindi $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^{r-1} f(x_{\tau(1)}, \dots, x_{\tau(n)})$. Ma τ è una trasposizione quindi $f(x_{\tau(1)}, \dots, x_{\tau(n)}) = -f(x_1, \dots, x_n)$ perché f è antisimmetrica. In conclusione $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^r f(x_1, \dots, x_n)$ e il lemma è dimostrato ♦

13: Lemma: (i) Ogni permutazione $\sigma \in S_n$ si scrive (non in modo unico) come prodotto di $(\leq n-1)$ trasposizioni.

(ii) Se $\sigma = \tau_1 \circ \dots \circ \tau_r = v_1 \circ \dots \circ v_t$ con τ_i, v_j delle trasposizioni allora $r \not\equiv t \pmod{2}$, i.e. r e t hanno la stessa parità.

Dim: (i) E' già stato visto, ricordiamo brevemente la dimostrazione. Se $\sigma(1) = k$ consideriamo la trasposizione, $\tau_1 = (1, k)$ che scambia 1 e k : $\tau_1(1) = k$ e $\tau_1(k) = 1$. Abbiamo $\tau_1(\sigma(1)) = 1$, ossia 1 è a posto. Se $\tau_1(\sigma(2)) = t$, consideriamo la trasposizione, $\tau_2 = (2, t)$ che scambia 2 e t : $\tau_2(2) = t$ e $\tau_2(t) = 2$. Abbiamo $\tau_2(\tau_1(\sigma(1))) = 1$ ($1 \neq t$ perché $\tau_1 \circ \sigma$ è iniettiva) e $\tau_2(\tau_1(\sigma(2))) = 2$; quindi anche 2 è a posto. Vediamo così che dopo al più $n-1$ passaggi otteniamo $\tau_r \circ \dots \circ \tau_1 \circ \sigma = Id$ ($r \leq n-1$, osservare che se 1, 2, ..., $n-1$ sono a posto allora anche n lo è). Pertanto $\sigma = \tau_1^{-1} \circ \dots \circ \tau_r^{-1}$. Si conclude osservando che se τ è una trasposizione allora anche τ^{-1} lo è.

(ii) Sia $f : Z^n \rightarrow Z : (x_1, \dots, x_n) \rightarrow \prod_{1 \leq i < j \leq n} (x_i - x_j)$; quest'ultima notazione significa che si fanno tutti i prodotti $(x_i - x_j)$ dove $1 \leq i < j \leq n$. Per esempio se $n = 3$: $f(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$. Mostriamo che f è un'applicazione antisimmetrica $f : X^n \rightarrow G$ ($X = G = Z$). Sia $\tau = (k, t)$ ($k < t$) una trasposizione. Nel prodotto $\prod_{1 \leq i < j \leq n} (x_i - x_j)$, x_k e x_t compaiono nei seguenti fattori:

$$(i < k): \quad (x_1 - x_k)(x_2 - x_k) \dots (x_i - x_k) \dots (x_{k-1} - x_k) =: P$$

$$(k < j < t): \quad (x_k - x_{k+1})(x_k - x_{k+2}) \dots (x_k - x_j) \dots (x_k - x_{t-1}) =: P'$$

$$(k < j, j > t): \quad (x_k - x_{t+1})(x_k - x_{t+2}) \dots (x_k - x_j) \dots (x_k - x_n) =: P''$$

$$(i = t): \quad (x_k - x_t)$$

Questo per quanto riguarda x_k , adesso vediamo x_t :

$$(i < k): \quad (x_1 - x_t)(x_2 - x_t) \dots (x_i - x_t) \dots (x_{k-1} - x_t) =: A$$

$$(k < j < t): \quad (x_{k+1} - x_t)(x_{k+2} - x_t) \dots (x_j - x_t) \dots (x_{t-1} - x_t) =: B$$

$$(j > t): \quad (x_t - x_{t+1}) \dots (x_t - x_j) \dots (x_t - x_n) =: C$$

Osservare che il fattore $(x_k - x_t)$ è già stato considerato prima. Se scambiamo x_t e x_k , i fattori P e A si scambiano tra di loro, così come P'' e C. Invece ogni fattore di P' si trasforma in un fattore di B moltiplicato per (-1) . Nello stesso modo ogni fattore di B si trasforma in un fattore di P' moltiplicato per (-1) . In conclusione la trasposizione trasforma il prodotto $P.P'.P''.A.B.C$ in $A.((-1)^{t-1-k}B).C.P.((-1)^{t-1-k}P').P'' = P.P'.P''.A.B.C$. Finalmente siccome $(x_k - x_t)$ viene trasformato in $(x_t - x_k) = - (x_k - x_t)$ e visto che i fattori $(x_i - x_j)$ con $i \neq k, t$, $j \neq k, t$ rimangono immutati, deduciamo che $f(x_{\tau(1)}, \dots, x_{\tau(n)}) = (-1)^r f(x_1, \dots, x_n)$ e f è antisimmetrica.

Per il lemma 12 se $\sigma = \tau_1 \circ \dots \circ \tau_r$ allora $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^r f(x_1, \dots, x_n)$, e se $\sigma = v_1 \circ \dots \circ v_t$ allora $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^t f(x_1, \dots, x_n)$. Segue che $(-1)^r f(x_1, \dots, x_n) = (-1)^t f(x_1, \dots, x_n)$. Prendendo (x_1, \dots, x_n) in \mathbb{Z}^n tale che $f(x_1, \dots, x_n) \neq 0$ (i.e. $x_i \neq x_j$ se $i \neq j$), otteniamo $(-1)^r = (-1)^t$ quindi r e t hanno la stessa parità ♦

14: Definizione: Sia $\sigma \in \mathfrak{S}_n$. Se σ ammette una decomposizione come prodotto di r trasposizioni: $\sigma = \tau_1 \circ \dots \circ \tau_r$, si pone $\varepsilon(\sigma) = (-1)^r$ e si dice che $\varepsilon(\sigma)$ è la segnatura della permutazione σ .

Risulta dal lemma 13 che la segnatura, $\varepsilon(\sigma)$, è ben definita.

15: Corollario: Siano E, F due k-spazi vettoriali e $f : E^n \rightarrow F$ un'applicazione multilinear alternante. Allora $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma).f(x_1, \dots, x_n)$ per ogni $\sigma \in \mathfrak{S}_n$ e per ogni (x_1, \dots, x_n) in E^n .

Dim: L'applicazione f è antisimmetrica (cfr 3.1) e il corollario segue da 12, 13 tenuto conto della definizione 14 ♦

16: Teorema: Sia E un k-spazio vettoriale di dimensione n.

(i) Lo spazio vettoriale delle n-forme multilineari alternanti su E ha dimensione uno.

(ii) Per ogni base, $B = (e_1, \dots, e_n)$, di E esiste un'unica forma n -lineare alternante, $\det_B : E^n \rightarrow k$, tale che $\det_B(e_1, \dots, e_n) = 1$. Se $x_i = \sum_j \xi^{(i)}_j e_j$, $1 \leq i \leq n$, sono n vettori di E , allora $\det_B(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \xi^{(1)}_{\sigma(1)} \dots \xi^{(n)}_{\sigma(n)}$. Inoltre se $f : E^n \rightarrow k$ è una forma n -lineare alternante: $f(x_1, \dots, x_n) = \det_B(x_1, \dots, x_n) \cdot f(e_1, \dots, e_n)$.

17: Definizione: Con le notazioni del teorema 16, lo scalare $\sum_{\sigma \in S_n} \epsilon(\sigma) \cdot \xi^{(1)}_{\sigma(1)} \dots \xi^{(n)}_{\sigma(n)}$ si chiama il determinante dei vettori (x_1, \dots, x_n) nella base B e si nota $\det_B(x_1, \dots, x_n)$.

Dimostrazione del teorema 16: Ormai è una facile generalizzazione di quanto fatto nel caso $n = 2$.

(i) Abbiamo visto (cfr (t) all'inizio) che se $f : E^n \rightarrow k$ è multilineare alternante allora (con le notazioni di (t), (ii)): $f(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \xi^{(1)}_{\sigma(1)} \xi^{(2)}_{\sigma(2)} \dots \xi^{(n)}_{\sigma(n)} \cdot f(e_{\sigma(1)}, \dots, e_{\sigma(n)})$. Dal corollario 15: $f(x_1, \dots, x_n) = (\sum_{\sigma \in S_n} \epsilon(\sigma) \cdot \xi^{(1)}_{\sigma(1)} \cdot \xi^{(2)}_{\sigma(2)} \dots \xi^{(n)}_{\sigma(n)}) \cdot f(e_1, \dots, e_n)$. Quindi f è completamente determinata da $f(e_1, \dots, e_n)$. Poniamo $\lambda = f(e_1, \dots, e_n)$ e, con le notazioni precedenti, definiamo $\det_B : E^n \rightarrow k$ tramite:

$$\det_B(x_1, \dots, x_n) = (\sum_{\sigma \in S_n} \epsilon(\sigma) \cdot \xi^{(1)}_{\sigma(1)} \cdot \xi^{(2)}_{\sigma(2)} \dots \xi^{(n)}_{\sigma(n)}) \cdot \lambda.$$

Si verifica che \det_B è multilineare alternante. E' chiaro che $f = \lambda \cdot \det_B$. Vediamo così che ogni forma n -lineare alternante, f , è un multiplo di \det_B . Pertanto lo spazio vettoriale delle forme n -lineari alternanti ha dimensione uno (\det_B per esempio è una base).

(ii) Sempre con le notazioni precedenti abbiamo:

$$\det_B(e_1, \dots, e_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \cdot \alpha^{(1)}_{\sigma(1)} \cdot \alpha^{(2)}_{\sigma(2)} \dots \alpha^{(n)}_{\sigma(n)}.$$

Dove $(\alpha^{(k)}_1, \dots, \alpha^{(k)}_n) = (0, \dots, 0, 1, 0, \dots, 0)$ (1 al k -esimo posto). Il prodotto $\alpha^{(1)}_{\sigma(1)} \cdot \alpha^{(2)}_{\sigma(2)} \dots \alpha^{(n)}_{\sigma(n)}$ è non nullo se e solo se ogni $\alpha^{(i)}_{\sigma(i)}$ è diverso da zero, ossia se e solo se $\sigma(i) = i$ per ogni i . Quindi $\alpha^{(1)}_{\sigma(1)} \cdot \alpha^{(2)}_{\sigma(2)} \dots \alpha^{(n)}_{\sigma(n)} \neq 0$ se e solo se σ è la permutazione identità, Id. Siccome la segnatura di Id vale 1 ricaviamo $\det_B(e_1, \dots, e_n) = 1$.

Questo dimostra l'esistenza, per ogni base $B = (e_1, \dots, e_n)$, di una n -forma lineare alternante, \det_B , tale che $\det_B(e_1, \dots, e_n) = 1$. Mostriamo l'unicità. Se f è una forma n -lineare alternante abbiamo $f = \lambda \cdot \det_B$ per qualche scalare λ (cfr (i)). Quindi $f(e_1, \dots, e_n) = \lambda \cdot \det_B(e_1, \dots, e_n) = \lambda$; e $f(e_1, \dots, e_n) = 1$ se e solo se f

$= \det_B$. L'unicità è dimostrata; inoltre, visto che $\lambda = f(e_1, \dots, e_n)$ osserviamo che, per ogni (x_1, \dots, x_n) in E^n : $f(x_1, \dots, x_n) = \lambda \cdot \det_B(x_1, \dots, x_n) = \det_B(x_1, \dots, x_n) \cdot f(e_1, \dots, e_n)$ ♦

18: Proposizione: Sia E un k -spazio vettoriale di dimensione n e x_1, \dots, x_n , n vettori di E . Allora x_1, \dots, x_n sono linearmente indipendenti se e solo se una delle due condizioni equivalenti seguenti è soddisfatta:

- (i) per ogni forma n -lineare alternante non nulla, f , si ha: $f(x_1, \dots, x_n) \neq 0$.
- (ii) esiste una base, $B = (e_1, \dots, e_n)$, di E tale che $\det_B(x_1, \dots, x_n) \neq 0$.

Dim: Le due condizioni sono equivalenti perché $f(x_1, \dots, x_n) = \det_B(x_1, \dots, x_n) \cdot f(e_1, \dots, e_n)$ (16). Basta quindi dimostrare (ii). Se x_1, \dots, x_n sono liberi formano una base, B' , e da 16(ii) $\det_{B'}(x_1, \dots, x_n) = 1$. Viceversa se $\det_B(x_1, \dots, x_n) \neq 0$ per qualche base B allora x_1, \dots, x_n non sono legati (cfr lemma 5)♦

A questo punto abbiamo dimostrato nel caso generale ($n \geq 2$) tutti i risultati dimostrati prima nel caso $n = 2$. L'unica cosa che ancora non sappiamo fare è calcolare un determinante $(n,n)!$ Per questo servono ancora alcuni preliminari.

IL DETERMINANTE DI UN ENDOMORFISMO.

19: Proposizione: Sia E un k -spazio vettoriale e $u: E \rightarrow E$ un endomorfismo. Esiste uno (unico) scalare, $\det(u)$, tale che per ogni (y_1, \dots, y_n) di E^n e ogni forma multilinear alternante $f: E^n \rightarrow k$ si abbia: $f(u(y_1), \dots, u(y_n)) = \det(u) \cdot f(y_1, \dots, y_n)$.

20: Definizione: Lo scalare $\det(u)$ si chiama il determinante dell'endomorfismo u .

20.1: Osservazione : Il determinante di u , $\det(u)$, è definito indipendentemente da ogni base.

Dimostrazione della Prop. 19: Se f è una n -forma lineare alternante poniamo $f_u: E^n \rightarrow k$: $(x_1, \dots, x_n) \mapsto f(u(x_1), \dots, u(x_n))$. Se si fissano $n-1$ variabili, per esempio $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$, l'applicazione parziale $E \rightarrow k: x \mapsto f(u(x_1), \dots, u(x_{i-1}), u(x), u(x_{i+1}), \dots, u(x_n))$ è lineare (perché è uguale a: $f_i \circ u$ dove f_i è la i -esima applicazione parziale rispetto ai valori $x_j = u(x_j)$, $1 \leq j \leq n$,

$j \neq i$. Per multilinearità f_i è lineare quindi anche $f_i \circ u$ lo è). Inoltre f_u è alternante (perché f lo è). Dal teorema 16 segue che f_u è proporzionale a f : $f_u = \lambda f$. Questo sarà vero per ogni n -forma alternante, f . Facciamo vedere che il coefficiente di proporzionalità è costante al variare di f . Sia g un'altra n -forma alternante. Allora (16) $g = \rho f$ per qualche scalare ρ , e $g_u = (\rho f)_u$. È chiaro che $(\rho f)_u = \rho(f_u)$. In conclusione: $g_u = \rho(f_u) = \rho(\lambda f) = \lambda(\rho f) = \lambda g$. Si conclude ponendo $\det(u) := \lambda$ ♦

21: Osservazione : Prendendo per (y_1, \dots, y_n) una base $B = (e_1, \dots, e_n)$ di E , con $f = \det_B$, otteniamo: $\det_B(u(e_1), \dots, u(e_n)) = \det(u)$.

Se $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, è la matrice associata ad u nella base B : $M = \text{mat}(u; B, B)$; vediamo (17) che $\det(u)$ è il determinante dei "vettori colonna" di M (i.e. degli n vettori di k^n le cui coordinate nella base canonica di k^n sono date dalle colonne di M):

$$\det(u) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Questo ci conduce alla seguente definizione:

DETERMINANTE DI UNA MATRICE.

22: Definizione: Sia $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, una matrice quadrata (n,n) a coefficienti in k . Il determinante di M , $\det(M)$, è lo scalare:

$$\det(M) := \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

22.1: Osservazione: Risulta da quanto precede che $\det(M)$ è uguale a:

-(i) il determinante di un qualsiasi endomorfismo associato ad M tramite un isomorfismo: $\text{mat}(-; B, B) : \text{End}(E) \rightarrow M_n(k)$ (E un k -spazio vettoriale di dimensione n), dato dalla scelta di una base B di E .

-(ii) il determinante dei "vettori colonna" di M .

Queste osservazioni (soprattutto (i)) ci permetteranno di dimostrare alcune proprietà dei determinanti senza (praticamente) fare calcoli.

23: Proposizione: (i) Il determinante dell'applicazione identità, Id_E , è uguale a 1.
(ii) Se u, v sono elementi di $\text{End}(E)$, $\det(u \cdot v) = \det(u) \cdot \det(v)$.

(iii) Un endomorfismo, u , di E è invertibile se e solo se $\det(u) \neq 0$. Inoltre se u è invertibile: $\det(u^{-1}) = (\det(u))^{-1}$.

Dim: (i) Segue dalle definizioni.

(ii) Per ogni (x_1, \dots, x_n) in E^n e ogni forma n-lineare alternante, f :

$f((u \circ v)(x_1), \dots, (u \circ v)(x_n)) = \det(u \circ v) \cdot f(x_1, \dots, x_n)$ (cfr Prop.19). Ma, sempre per la Prop.19 abbiamo anche: $f((u \circ v)(x_1), \dots, (u \circ v)(x_n)) = \det(u) \cdot f(v(x_1), \dots, v(x_n)) = \det(u) \cdot \det(v) \cdot f(x_1, \dots, x_n)$ si conclude prendendo f e (x_1, \dots, x_n) tali che $f(x_1, \dots, x_n) \neq 0$.

(iii) Se u è invertibile esiste u^{-1} tale che $u^{-1} \circ u = \text{Id}_E$. Da (i) e (ii): $\det(u^{-1}) \cdot \det(u) = 1$ quindi $\det(u) \neq 0$. Inoltre $\det(u^{-1}) = (\det(u))^{-1}$. Supponiamo u non invertibile. Sia $B = (e_1, \dots, e_n)$ una base di E . Abbiamo $\text{rg}(u) = \text{rg}(u(e_1), \dots, u(e_n)) < n$. Quindi i vettori $u(e_1), \dots, u(e_n)$ sono linearmente dipendenti. Da 18 (ii): $\det_B(u(e_1), \dots, u(e_n)) = 0$, quindi (22.1(i)) $\det(u) = 0$ ♦

24: Corollario: (i) $\det(I_n) = 1$ dove $I_n \in M_n(k)$ è la matrice unità.

(ii) Se $A, M \in M_n(k)$, $\det(A \cdot M) = \det(A) \cdot \det(M) = \det(M) \cdot \det(A)$

(iii) Una matrice $M \in M_n(k)$ è invertibile se e solo se $\det(M) \neq 0$; inoltre se M è invertibile allora $\det(M^{-1}) = (\det M)^{-1}$.

Dim: (i) Sia B una base di E , un k -spazio vettoriale di dimensione n . Abbiamo $\det(\text{Id}_E) = 1$ (23(i)) e $\det(\text{Id}_E \cdot B, B) = I_n$. Quindi (22.1(i)), $\det(I_n) = 1$.

(ii) Siano u, v gli endomorfismi corrispondenti a A, M nell'isomorfismo $\text{mat}(-; B, B): \text{End}(E) \rightarrow M_n(k)$. Allora $\text{mat}(u \circ v) = AM$ ($\text{mat}(-; B, B)$ è un isomorfismo d'anelli) e, usando 23 (ii): $\det(AM) = \det(u \circ v) = \det(u) \cdot \det(v) = \det(A) \cdot \det(M)$.

(iii) Come sopra sia $M = \text{mat}(w; B, B)$ allora M è invertibile se e solo se w lo è ($\text{mat}(-; B, B)$ è un isomorfismo d'anelli), si conclude con 23 (iii) ♦

Esercizi:

13.1) Nel C -spazio vettoriale C^2 , calcolando un determinante, dire se i vettori u, v sono linearmente indipendenti, dove $u = (2+i, 1-i)$, $v = (4+i, 2+3i)$.

13.2) Nel k -spazio vettoriale k^2 per quali valori di $m \in k$ i due vettori $u = (m, 1)$, $v = (-1, m)$ sono indipendenti? ($k = R, C$).

13.3) La matrice "scacchiera" (n,n) , $S \in M_n(k)$, è definita nel modo seguente: $S = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, e: $a_{ij} = 0$ se $i+j$ è pari, $a_{ij} = 1$ se $i+j$ è dispari. Calcolare $\det(S)$.

13.4) Sia $M \in M_n(k)$. Mostrare che se $M = A \cdot B$ con $A \in M_{n,p}(k)$, $B \in M_{p,n}(k)$ e $n > p$, allora $\det(M) = 0$. E' ancora vero che $\det(M) = 0$ se si suppone invece $n < p$?

13.5) Sia $M_n(\mathbb{R})$ lo spazio vettoriale reale delle matrici $n \times n$ a coefficienti reali. Sia $D = \{A \in M_n(\mathbb{R}) / \det(A) = 0\}$. Dire se D è un sottospazio vettoriale di $M_n(\mathbb{R})$ (fare prima il caso $n = 2$).

§14) CALCOLO DI UN DETERMINANTE.

In questo paragrafo, che è il seguito diretto del precedente, si dà un metodo pratico per calcolare un determinante (n,n) ("sviluppo per riga o per colonna"). Nei casi $n = 2, 3$ si ricavano delle formule esplicite, non troppo complicate.

ALCUNE REGOLE PRATICHE PER IL CALCOLO DI UN DETERMINANTE.

Da quanto fatto finora risulta che il determinante di una matrice quadrata $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, è una funzione multilineare alternante delle colonne di M (dei vettori colonna di M). Pertanto viste le proprietà delle applicazioni multilineari alternanti:

- (1) se si fa una permutazione, σ , delle colonne di M , $\det(M)$ si trasforma in $\varepsilon(\sigma).\det(M)$. In particolare se si scambiano due colonne, il determinante della matrice ottenuta è: $-\det(M)$.
- (2) si ha: $\det(M) = 0$ se e solo se le colonne di M sono linearmente dipendenti.
- (3) $\det(M)$ è lineare rispetto ad una colonna qualsiasi (le altre essendo fissate).
- (4) il determinante non cambia se si aggiunge ad una colonna una combinazione lineare delle altre colonne. Infatti se $[c_1, \dots, c_n]$ denota la matrice (n,n) di vettori colonna c_1, \dots, c_n , abbiamo: $\det[c_1 + \sum_{i \geq 2} \lambda_i c_i, c_2, \dots, c_n] = \det[c_1, \dots, c_n] + \det[\sum_{i \geq 2} \lambda_i c_i, c_2, \dots, c_n]$ e quest'ultimo termine è nullo perché le colonne $\sum_{i \geq 2} \lambda_i c_i, c_2, \dots, c_n$ sono linearmente dipendenti.
- (5) per ogni scalare: $\det(\lambda M) = \lambda^n \cdot \det(M)$; (M quadrata (n,n)).

Inoltre:

6: Proposizione: Se $M \in M_n(k)$ allora $\det(M) = \det(tM)$.

6.1: Osservazione : $\det(tM)$ è una funzione multilineare alternante delle colonne di tM quindi delle righe di M . La proposizione dice che questa

funzione è uguale a $\det(M)$. Quindi $\det(M)$ è una funzione multilineare alternante sia delle colonne che delle righe di M . In particolare (1), ..., (4) qui sopra sono veri con la parola riga al posto di colonna.

Dimostrazione della Prop.6: Se $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, per definizione (cfr §13, 22):

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \quad \text{Invece:}$$

$$\det({}^t M) = \sum_{\rho \in S_n} \varepsilon(\rho) \cdot a_{1\rho(1)} \cdot a_{2\rho(2)} \cdots a_{n\rho(n)}.$$

Possiamo riscrivere la formula per $\det(M)$ nel modo seguente:

$$\det(M) = \sum_{\sigma \in S_n} (\varepsilon(\sigma) \cdot \prod_{i=1}^n a_{\sigma(i)i}). \text{ Poniamo } \sigma(i) = t, \text{ di modo che } i = \sigma^{-1}(t).$$

Visto che σ è una biiezione da $\{1, 2, \dots, n\}$ in $\{1, 2, \dots, n\}$, anche t varia tra 1 e n . Abbiamo quindi:

$$\det(M) = \sum_{\sigma \in S_n} (\varepsilon(\sigma) \cdot \prod_{t=1}^n a_{t\sigma^{-1}(t)}).$$

Osserviamo adesso che per ogni permutazione, σ , di S_n : $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. Infatti se $\sigma = \tau_1 \circ \dots \circ \tau_r$ allora $\sigma^{-1} = \tau_r^{-1} \circ \dots \circ \tau_1^{-1}$ e τ_i^{-1} sono ancora delle trasposizioni.

Quindi $\varepsilon(\sigma^{-1}) = (-1)^r = \varepsilon(\sigma)$. Pertanto: $\det(M) = \sum_{\sigma \in S_n} (\varepsilon(\sigma^{-1}) \cdot \prod_{t=1}^n a_{t\sigma^{-1}(t)})$.

Finalmente, tenuto conto che l'applicazione $S_n \rightarrow S_n : \rho \rightarrow \rho^{-1}$ è una biiezione, fare la somma su σ o su σ^{-1} è la stessa cosa quindi $\det(M) =$

$\sum_{\sigma^{-1} \in S_n} (\varepsilon(\sigma^{-1}) \cdot \prod_{t=1}^n a_{t\sigma^{-1}(t)})$. Finalmente ponendo $\rho = \sigma^{-1}$, $\det(M) =$

$$\sum_{\rho \in S_n} (\varepsilon(\rho) \cdot \prod_{t=1}^n a_{t\rho(t)}) = \det({}^t M) \blacklozenge$$

Sviluppo di un determinante ($n = 2, 3$).

Se $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, per definizione (cfr §13, 22):

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

7) DETERMINANTI (2,2):

Sia $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, una matrice (2,2) a coefficienti in k . Abbiamo $\mathfrak{S}_2 = \{\text{Id}, \tau\}$ dove τ è la trasposizione $\tau(1) = 2$ e $\tau(2) = 1$. Applicando la definizione: $\det(M) = \varepsilon(\text{Id}) a_{11}a_{22} + \varepsilon(\tau) a_{21}a_{12}$. Visto che $\varepsilon(\text{Id}) = 1$, $\varepsilon(\tau) = -1$, ritroviamo la solita formula: $\det(M) = a_{11}a_{22} - a_{21}a_{12}$.

8) DETERMINANTI (3,3):

Sia $M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, una matrice (3,3) a coefficienti in k . Abbiamo $\mathfrak{S}_3 = \{\text{Id}, \tau_{12}, \tau_{13}, \tau_{23}, \sigma, \nu\}$, dove τ_{kt} è la trasposizione $\tau(k) = t$, $\tau(t) = k$; $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ e $\nu = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. Osserviamo che $\sigma = \tau_{31} \circ \tau_{12}$ mentre $\nu = \tau_{23} \circ \tau_{12}$. Abbiamo perciò $\varepsilon(\tau_{kt}) = -1$, $\varepsilon(\sigma) = \varepsilon(\nu) = 1$. Applicando la definizione: $\det(M) = \varepsilon(\text{Id}) a_{11}a_{22}a_{33} + \varepsilon(\tau_{12}) a_{21}a_{12}a_{33} + \varepsilon(\tau_{13}) a_{31}a_{22}a_{13} + \varepsilon(\tau_{23}) a_{11}a_{32}a_{23} + \varepsilon(\sigma) a_{21}a_{32}a_{13} + \varepsilon(\nu) a_{31}a_{12}a_{23} = a_{11}a_{22}a_{33} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$

In conclusione:

$$\det(M) = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23}.$$

Per ricordare questa formula si può usare il trucco seguente (detto regola di Sarrus): si ricopiano le prime due righe di M sotto la terza riga poi si fa la somma dei prodotti delle "diagonali" (da sinistra verso destra) di lunghezza tre meno la somma dei prodotti delle "diagonali" di lunghezza tre (da destra verso sinistra).

Passiamo adesso al caso generale. Iniziamo col seguente risultato, molto utile:

9: Teorema: Sia $M = (a_{ij})$ una matrice (n,n) tale che $a_{ij} = 0$ se $i > p$ e $j < p+1$. Quindi M è della forma: $M = \begin{pmatrix} N & P \\ 0 & Q \end{pmatrix}$ con N quadrata (p,p) , Q è $p \times n-p$

quadrata ($n-p, n-p$), P è *di tipo* ($p, n-p$) e 0 è la *matrice nulla* ($n-p, p$). In queste condizioni:

$$\det(M) = \det(N)\det(Q).$$

Dim: Poniamo $q := n-p$ e $E = k^p$, $F = k^q$. Consideriamo l'applicazione $d: E^p \times F^q \rightarrow k$:

$$(X_1, \dots, X_p; Y_1, \dots, Y_q) \rightarrow \det \begin{pmatrix} X_1 \dots X_p & P \\ 0 & Y_1 \\ & \ddots \\ & Y_q \end{pmatrix}, \text{ in altre parole, se notiamo } X \text{ la matrice } (p,p) \text{ i cui vettori colonna sono } X_1, \dots, X_p \text{ e } Y \text{ la matrice } (q,q) \text{ i cui vettori riga sono } Y_1, \dots, Y_q \text{ allora } d(X_1, \dots, X_p; Y_1, \dots, Y_q) = \det \begin{pmatrix} X & P \\ 0 & Y \end{pmatrix}.$$

Per Y_1, \dots, Y_q fissati otteniamo un'applicazione:

$d_Y: E^p \rightarrow k: (X_1, \dots, X_p) \rightarrow \det \begin{pmatrix} X_1 \dots X_p & P \\ 0 & Y \end{pmatrix}$. Siccome il determinante di una matrice è un'applicazione multilinearare alternante dei vettori colonna della matrice (quindi qui in particolare di X_1, \dots, X_p (gli zeri sotto non disturbano)), vediamo che d_Y è una p -forma alternante su E . Dal teorema §13, 16 (i), segue che d_Y è proporzionale a \det_C dove C è la base canonica di $E = k^p$. Abbiamo quindi $d_Y = \lambda_Y \det_C$; osserviamo infatti che il coefficiente di proporzionalità dipende da Y_1, \dots, Y_q (P è fissata). Sia $\phi: F^q \rightarrow k: (Y_1, \dots, Y_q) \rightarrow \phi(Y_1, \dots, Y_q) := \lambda_Y$. Riassumendo: $d(X_1, \dots, X_p; Y_1, \dots, Y_q) = \phi(Y_1, \dots, Y_q) \det_C(X_1, \dots, X_p)$. Prendendo per (X_1, \dots, X_p) proprio la base canonica $C = (e_1, \dots, e_p)$ otteniamo:

$d(e_1, \dots, e_p; Y_1, \dots, Y_q) = \phi(Y_1, \dots, Y_q) = \det \begin{pmatrix} I_p & P \\ 0 & Y \end{pmatrix}$. Siccome il determinante di una matrice è una funzione multilinearare alternante delle righe della matrice, deduciamo che $\phi: F^q \rightarrow k$ è una q -forma alternante su F . Dal teorema §13, 16(i): $\phi = \alpha \det_{\underline{C}}$ dove \underline{C} è la base canonica di $F = k^q$.

Abbiamo quindi dimostrato che: $d(X_1, \dots, X_p; Y_1, \dots, Y_q) = \alpha \det_{\underline{C}}(Y_1, \dots, Y_q) \det_C(X_1, \dots, X_p)$. Rimane da calcolare α . Per questo osserviamo che se

prendiamo $(X_1, \dots, X_p) = C$, $(Y_1, \dots, Y_q) = \underline{C}$ allora $\alpha = \det \begin{pmatrix} I_p & P \\ 0 & I_q \end{pmatrix}$. Per calcolare questo determinante osserviamo che: $\begin{pmatrix} I_p & P \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix} + \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} =: I + H$. Siano

c_1, \dots, c_q i vettori colonna di P ; sono vettori di $E = k^P$ quindi sono combinazioni lineari dei vettori della base canonica C . Pertanto $\begin{pmatrix} I_p & P \\ 0 & I_q \end{pmatrix}$ è ottenuta da I aggiungendo, ad ognuna delle ultime q colonne, una combinazione lineare delle prime p colonne (quindi di colonne diverse). Da (4) segue che $\alpha = \det \begin{pmatrix} I_p & P \\ 0 & I_q \end{pmatrix} = \det(I)$. Da §13, 24(i): $1 = \det(I) = \alpha$. Finalmente: $d(X_1, \dots, X_p; Y_1, \dots, Y_q) = \det_C(Y_1, \dots, Y_q) \det_C(X_1, \dots, X_p)$. Da 6 e §13, 22.1, possiamo scrivere: $d(X_1, \dots, X_p; Y_1, \dots, Y_q) = \det(X) \cdot \det(Y)$ dove X è la matrice (n,n) le cui colonne sono X_1, \dots, X_p e dove Y è la matrice le cui righe sono Y_1, \dots, Y_q . Prendendo per X_1, \dots, X_p i vettori colonna di N e per Y_1, \dots, Y_q i vettori righe di Q otteniamo $\det(M) = \det(N) \cdot \det(Q)$ e il teorema è dimostrato ♦

9.1: Osservazione : Più generalmente si può dimostrare (cfr Es.10) che se M è

$$\begin{pmatrix} A_1 & \text{termini} \\ 0 & A_2 \text{ qualsiasi} \\ \vdots & \ddots \ddots \\ 0 & \ddots \ddots \\ 0 & \dots \dots 0 & A_n \end{pmatrix}$$

una matrice quadrata del tipo $M = \begin{pmatrix} A_1 & \text{termini} \\ 0 & A_2 \text{ qualsiasi} \\ \vdots & \ddots \ddots \\ 0 & \ddots \ddots \\ 0 & \dots \dots 0 & A_n \end{pmatrix}$, A_i matrici quadrate qualsiasi, allora $\det(M) = \det(A_1) \dots \det(A_n)$.

Passiamo adesso allo sviluppo per riga (o per colonna) di un determinante (n,n) .

10: Definizione: Sia $M = (a_{ij})$ una matrice (n,n) . Sia $M_{k,t}$ la matrice $(n-1,n-1)$ ottenuta da M togliendo la k -esima riga e la t -esima colonna. Il determinante di $M_{k,t}$ si chiama il minore di M relativo a a_{kt} .

Si può visualizzare questa definizione nel modo seguente:

$$\det(M_{k,t}) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ -a_{kt} & \dots & - \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

11: Teorema: Sia M una matrice (n,n) . Per ogni $i, j, 1 \leq i \leq n, 1 \leq j \leq n$, si ha:

(i) $\det(M) = \sum_{r=1}^n (-1)^{r+j} a_{rj} \det(M_{rj})$ (sviluppo di $\det(M)$ secondo la j -esima colonna)

(ii) $\det(M) = \sum_{r=1}^n (-1)^{i+r} a_{ir} \det(M_{ir})$ (sviluppo di $\det(M)$ secondo la i -esima riga).

Dim: Siccome $\det(M) = \det(tM)$ (cfr 6) basta dimostrare (i). La j -esima colonna si può scrivere come somma di n colonne:

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{rj} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ a_{rj} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ \vdots \\ a_{nj} \end{pmatrix}.$$

$$\begin{pmatrix} a_{11} \dots a_{1j-1} & 0 & a_{1j+1} & \dots a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \\ a_{r1} \dots a_{rj-1} & a_{rj} & a_{rj+1} & \dots a_{rn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} \dots a_{nj-1} & 0 & a_{nj+1} & \dots a_{nn} \end{pmatrix}$$

Pertanto $\det(M) = \sum_{r=1}^n \det(M_r)$ dove $M_r =$

$$M'_r = \begin{pmatrix} a_{r1} \dots a_{rj-1} & a_{rj} & a_{rj+1} \dots a_{rn} \\ a_{11} \dots a_{1j-1} & 0 & a_{1j+1} \dots a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} \dots a_{nj-1} & 0 & a_{nj+1} \dots a_{nn} \end{pmatrix}.$$

Da (1) e 6.1 segue che $\det(M'_r) = (-1)^{r-1} \det(M_r)$ Siano C_1, \dots, C_n le colonne di M'_r . Scambiando successivamente C_j con $C_{j-1}, C_{j-2}, \dots, C_1$, otteniamo, dopo $j-1$ scambi la matrice M''_r dove

$$M''_r = \begin{pmatrix} a_{rj} & a_{r1} & a_{r2} & \dots & a_{r,j-1} & a_{r,j+1} & \dots & a_{rn} \\ 0 & a_{11} & a_{12} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots \\ 0 & a_{n1} & a_{n2} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix}$$

Sempre per (1): $\det(M''_r) = (-1)^{j-1} \cdot \det(M'_r)$. In conclusione: $\det(M''_r) = (-1)^{j+r-2} \cdot \det(M_r) = (-1)^{j+r} \cdot \det(M_r)$. Ma dal teorema 9 abbiamo anche $\det(M''_r) = a_{rj} \cdot \det(M_{rj})$. Risulta $\det(M_r) = (-1)^{j+r} \det(M''_r) = (-1)^{j+r} \cdot a_{rj} \cdot \det(M_{rj})$. Perciò $\det(M) = \sum_{r=1}^n (-1)^{r+j} a_{rj} \det(M_{rj})$ e il teorema è dimostrato ♦

12: Definizione: Il termine $(-1)^{r+j} \det(M_{rj})$ si chiama il cofattore di a_{rj} . La matrice, $\text{Co}(M)$, i cui coefficienti sono i cofattori dei coefficienti di M si chiama la matrice dei cofattori di M . Se $\text{Co}(M) = (A_{ij})$ allora $A_{ij} = (-1)^{i+j} \det(M_{ij})$.

13: Definizione: La matrice complementare di M , M^c , è la trasposta della matrice dei cofattori: $M^c = {}^t(\text{Co}(M))$.

14: Teorema: Sia $M \in M_n(k)$ allora $M \cdot M^c = M^c \cdot M = \det(M) \cdot I_n$.

Dim: Sia $M = (a_{ij})$ e $M \cdot M^c = (c_{ij})$. Per definizione c_{ij} è il prodotto scalare della i -esima riga di M con la j -esima colonna di M^c (quindi la j -esima riga di $\text{Co}(M)$): $c_{ij} = \sum_k a_{ik} \cdot A_{jk}$, e tenuto conto della definizione di A_{jk} , $c_{ij} = \sum_k a_{ik} \cdot (-1)^{i+k} \det(M_{jk})$. Se $i = j$ otteniamo: $c_{ii} = \sum_k a_{ik} \cdot (-1)^{i+k} \det(M_{ik})$ e dal teorema 11 segue che $c_{ii} = \det(M)$.

Sia adesso $i \neq j$. Notiamo con R_1, \dots, R_n le righe di M , e sia $M_{i(j)}$ la matrice le cui righe sono R'_t , con $R'_t \doteq R_t$, se $t \neq j$, $R'_j = R_i$. Siccome $M_{i(j)}$ ha due righe uguali ($R'_j = R'_i$), $\det(M_{i(j)}) = 0$ (cfr (2)). Se sviluppamo $\det(M_{i(j)})$ secondo la j -esima riga, viene: $\det(M_{i(j)}) = \sum_k a_{ik} \cdot (-1)^{j+k} \det(M_{jk}) = c_{ij}$. Quindi se $i \neq j$, $c_{ij} = 0$ e $M \cdot M^c = \det(M) \cdot I_n$. Nello stesso modo si dimostra $M^c \cdot M = \det(M) \cdot I_n$ ♦

15: Corollario: Sia $M \in M_n(k)$ allora M è invertibile se e solo se $\det(M)$ è invertibile (i.e., visto che $\det(M)$ appartiene al campo k , se e solo se $\det(M) \neq 0$).

Se M è invertibile allora: $M^{-1} = (\det(M))^{-1} \cdot M^c$.

Dim: Se M è invertibile allora $M \cdot M^{-1} = I_n$, da cui (cfr §13, 24) $\det(M) \cdot \det(M^{-1}) = 1$; nello stesso modo $\det(M^{-1}) \cdot \det(M) = 1$. Quindi $\det(M)$ è invertibile. Viceversa, se $\det(M)$ è invertibile, da 14, $(\det(M))^{-1} \cdot M^c \cdot M = M \cdot (\det(M))^{-1} \cdot M^c = I_n$. Quindi M è invertibile e $M^{-1} = (\det(M))^{-1} \cdot M^c$ ♦

15.1: Osservazione : (i) Il teorema e il corollario forniscono una dimostrazione alternativa a §13, 23(iii), 24(iii).

(ii) Il corollario fornisce un metodo effettivo per calcolare l'inversa di una matrice invertibile (cfr Oss. §11, 1.1).

16: Esempio : Sia $M = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$. Sviluppando secondo la prima colonna:

$\det(M) = \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} - 2 \begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = 4 - (-1) - 2(6-4) = 1$. Quindi M è invertibile. Per determinare M^{-1} iniziamo col calcolare i cofattori:

$$A_{11} = \det M_{11} = \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 4 - (-1) = 5; A_{12} = -\det M_{12} = -\begin{vmatrix} 2 & -1 \\ 0 & 2 \end{vmatrix} = -4.$$

$$A_{13} = \det M_{13} = \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} = 2; A_{21} = -\det M_{21} = -\begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = -(6-4) = -2$$

$$A_{22} = \det M_{22} = \begin{vmatrix} 1 & 4 \\ 0 & 2 \end{vmatrix} = 2; A_{23} = -\det M_{23} = -\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} = -1$$

$$A_{31} = \det M_{31} = \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} = -11; A_{32} = -\det M_{32} = -\begin{vmatrix} 1 & 4 \\ 2 & -1 \end{vmatrix} = 9$$

$$A_{33} = \det M_{33} = \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} = -4.$$

Concludiamo così che $\text{Co}(M) = \begin{pmatrix} 5 & -4 & 2 \\ -2 & 2 & -1 \\ -11 & 9 & -4 \end{pmatrix}$. Visto che $\det(M) = 1$, $M^{-1} =$

$${}^t(\text{Co}(M)) = M^c \text{ ossia: } M^{-1} = \begin{pmatrix} 5 & -2 & -11 \\ -4 & 2 & 9 \\ 2 & -1 & -4 \end{pmatrix}.$$

16.1: Osservazione : Si possono considerare matrici a coefficienti non più in un corpo commutativo ma in un anello commutativo, A. Se $M \in M_n(A)$ si può definire tramite §13, 22, per esempio, il determinante di M. Osservare che $\det(M)$ è un elemento di A. E' facile verificare che 14, 15 qui sopra sono

ancora veri in questo contesto più generale. In particolare nell'esempio 16 possiamo considerare M come un elemento di $M_3(\mathbb{Z})$. Da 14, 15 segue che M è invertibile come elemento di $M_3(\mathbb{Z})$ (osservare che gli unici elementi invertibili dell'anello \mathbb{Z} sono 1 e -1). Infatti M^{-1} è una matrice a coefficienti in \mathbb{Z} . Invece la matrice $N = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$, il cui determinante è 5, non è invertibile in $M_3(\mathbb{Z})$; lo è invece come elemento di $M_3(\mathbb{R})$. In particolare possiamo già dire che $N^{-1} \in M_3(\mathbb{R})$ ha un coefficiente che non è un numero intero.

Esercizi:

14.1) Calcolare i seguenti determinanti:

$$A = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix}, \quad B = \begin{vmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -1 & -2 \\ 0 & 1 & 0 & 3 \\ 3 & 2 & -3 & 2 \end{vmatrix}, \quad C = \begin{vmatrix} 2 & -179 & -4 \\ 0 & 1 & 0 \\ 4 & 350 & 6 \end{vmatrix}.$$

14.2) Dire se le seguenti matrici sono invertibili. In caso di risposta affermativa,

calcolarne la matrice inversa: $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 3 & 0 \\ 0 & 4 & 1 \end{pmatrix}$.

14.3) Se t è un numero reale si pone $M(t) = \begin{pmatrix} t & 0 & -1 \\ 1 & -t & 1 \\ 1 & 1 & t \end{pmatrix}$. Sia $D = \{t \in \mathbb{R} / \det(M(t)) = 0\}$.

Calcolare $\text{card}(D)$.

14.4) Sia $M = (a_{ij}) \in M_n(\mathbb{R})$. Si suppone che esistono tre indici j, j', j'' e un numero reale r tali che: $a_{ij'} = a_{ij} + r$, $a_{ij''} = a_{ij'} + r$, per ogni i , $1 \leq i \leq n$. Dimostrare che $\det(M) = 0$.

14.5) (i) Siano $e'_1 = (1, 0, 0)$, $e'_2 = (0, -1, 0)$, $e'_3 = (1, 0, -1)$ in \mathbb{R}^3 . Calcolando un determinante, dire se $B' = (e'_1, e'_2, e'_3)$ è una base di \mathbb{R}^3 .

(ii) Sia $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, l'applicazione lineare rappresentata, nella base canonica, dalla

matrice $M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$.

Mostrare che M è invertibile e calcolare M^{-1} .

(iii) Determinare $\text{mat}(f^{-1}; B', B')$ dove $B' = (e'_1, e'_2, e'_3)$, cfr (i).

14.6) Sia E un R -spazio vettoriale di dimensione n . Sull'insieme, B , delle basi di E si considera la relazione binaria, R , definita da $B R B' \Leftrightarrow \det_{B'}(B') > 0$. Mostrare che R è una relazione d'equivalenza su B e che l'insieme quoziante ha due elementi.

14.7) Sia $M = (a_{ij}) \in M_n(K)$. Si suppone $a_{ij} = 0$ se $i+j < n+1$. Calcolare $\det(M)$:
 1°) sviluppando per riga o per colonna,
 2°) in un altro modo...

14.8) Sia $n \geq 3$ un numero naturale dispari. Dimostrare che non esiste nessuna matrice $M \in M_n(R)$ tale che $M^2 + I_n = 0$.

14.9) Sia $a \in R$. Per ogni x in R dimostrare che il determinante della matrice quadrata

$$\text{di ordine } n: \begin{pmatrix} x & a & \dots & a \\ ax & a & \dots & a \\ \vdots & & & \\ a & \dots & x \end{pmatrix}, \text{ è } (x + (n-1)a)(x-a)^{n-1}.$$

(Suggerimento: cominciare sommando tutte le colonne alla prima.)

$$\begin{pmatrix} A_1 & \text{termi} \\ 0 & A_2 \text{ qualsiasi} \\ \vdots & \ddots \ddots \\ 0 & \ddots \ddots \ddots \\ 0 & \dots \dots 0 & A_n \end{pmatrix}$$

14.10) Se M è una matrice quadrata del tipo $M = \begin{pmatrix} A_1 & \text{termi} \\ 0 & A_2 \text{ qualsiasi} \\ \vdots & \ddots \ddots \\ 0 & \ddots \ddots \ddots \\ 0 & \dots \dots 0 & A_n \end{pmatrix}$, A_i matrici

quadrate qualsiasi, dimostrare che $\det(M) = \det(A_1)\dots\det(A_n)$.

(ii) Se $D = (d_{ij})$ è una matrice (n,n) triangolare (superiore o inferiore) dimostrare che $\det(D) = d_{11}\dots d_{nn}$.

14.11) Sia $D_n \in M_n(R)$, $D_n = (a_{ij})$ con $a_{11} = 1$, $a_{ij} = 1$ se $i \neq j$ e $a_{ii} = 0$ se $i \neq 1$. Si pone $\Delta_n := \det(D_n)$. Calcolare Δ_n (osservare che $\Delta_n = -\Delta_{n-1}$)

$$\begin{pmatrix} X & a & b & X \\ a & X & X & b \\ b & X & X & a \\ X & b & a & X \end{pmatrix}$$

14.12) Fattorizzare il polinomio $Q(X) = \det \begin{pmatrix} X & a & b & X \\ a & X & X & b \\ b & X & X & a \\ X & b & a & X \end{pmatrix}$ (iniziate sostraendo la quarta colonna alla prima).

§15) RANGO E DETERMINANTI.

In questo paragrafo applichiamo la teoria dei determinanti al calcolo del rango di una matrice.

Iniziamo con la nozione di matrice estratta (o sotto matrice) di una matrice M.

1: Definizione: Sia $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq p$, una matrice (n,p) . Siano $I \subseteq \{1, 2, \dots, n\}$, $J \subseteq \{1, 2, \dots, p\}$ due sottoinsiemi. Con $M_{I,J}$ si nota la matrice ottenuta sopprimendo in M tutte le righe, R_k , con $k \in I$ e tutte le colonne C_t con $t \in J$. Una tale matrice $M_{I,J}$ si dice estratta da M (si dice anche che $M_{I,J}$ è una sottomatrice di M).

1.1: Osservazione : Nello sviluppo per riga (o colonna) di un determinante (cfr §14, 11) abbiamo incontrato la matrice estratta $M_{ij} = M_{I,J}$ con $I = \{i\}$, $J = \{j\}$.

2: Definizione: Un minore d'ordine r di M è il determinante di una matrice (r,r) estratta da M.

Quindi Δ_r è un minore d'ordine r di M se $\Delta_r = \det(M_{I,J})$ per qualche $I \subseteq \{1, 2, \dots, n\}$, $J \subseteq \{1, 2, \dots, p\}$ con $n - \text{card}(I) = r$, $p - \text{card}(J) = r$.

3: Proposizione: Sia $M \in M_{n,p}(k)$, una matrice a coefficienti in k . Si ha $\text{rg}(M) \geq r$ se e solo se esiste un minore d'ordine r di M non nullo.

Dim: Supponiamo che esista un minore d'ordine r , Δ_r , non nullo. Il rango di M è il rango dell'insieme dei suoi vettori colonne e anche il rango dei suoi vettori righe. Perciò una permutazione delle righe o delle colonne di M non cambia il rango di M. Possiamo dunque supporre che i vettori colonne corrispondenti al minore Δ_r siano i primi r . Nello stesso modo possiamo supporre che pure i vettori righe corrispondenti a Δ_r siano i primi r . Chiamiamo c_1, \dots, c_p i p vettori colonne di M. Sono vettori di k^n . Sia $\psi : k^n \rightarrow k^r$: $(x_1, \dots, x_n) \rightarrow (x_1, \dots, x_r)$ ($r \leq n$) la proiezione sulle prime r coordinate. Poniamo $c'_k = \psi(c_k)$. Allora $\Delta_r = \det_{C'}(c'_1, \dots, c'_r)$ (C' è la base canonica) e quindi c'_1, \dots, c'_r sono linearmente indipendenti. Questo implica che c_1, \dots, c_r sono linearmente indipendenti. Infatti se $\lambda_1 c_1 + \dots + \lambda_r c_r = 0$ allora $\psi(\lambda_1 c_1 + \dots + \lambda_r c_r) = 0 = \lambda_1 c'_1 + \dots + \lambda_r c'_r$.

$\dots + \lambda_r c'_r$ e quindi, visto che c'_1, \dots, c'_r sono linearmente indipendenti, $\lambda_1 = \dots = \lambda_r = 0$. Siccome c_1, \dots, c_r sono liberi concludiamo che $\text{rg}(M) \geq r$. Viceversa, se $\text{rg}(M) \geq r$, esistono r vettori colonna di M linearmente indipendenti. Scambiando semmai le colonne possiamo supporre c_1, \dots, c_r linearmente indipendenti. Sia N la matrice le cui colonne sono c_1, \dots, c_r . Abbiamo $\text{rg}(N) = r$. Ci sono quindi r vettori righe di N linearmente indipendenti. A costo di permutare le righe possiamo supporre che le prime r righe, r_1, \dots, r_r , di N sono linearmente indipendenti; quindi $\det_C(r_1, \dots, r_r) \neq 0$ (C è la base canonica di k^r). Pertanto il minore di ordine r di M costruito sulle colonne c_1, \dots, c_r e su le righe corrispondenti a r_1, \dots, r_r , è non nullo.

4: Definizione: Sia $M \in M_{n,p}(k)$ e Δ_r un minore d'ordine r di M . Un orlato di Δ_r è un minore di ordine $r+1$ di M , determinante di una matrice $(r+1, r+1)$ estratta da M e di cui Δ_r è un minore.

4.1: Osservazione : (i) In altre parole se $\Delta_r = \det M_{I,J}$ allora un orlato di Δ_r è $\det M_{I',J}$ con I' (risp. J') contenuto in I (risp. J) e $\text{card}(I') = \text{card}(I) - 1$ (risp. $\text{card}(J') = \text{card}(J) - 1$).

(ii) Se M è una matrice (n,p) ci sono $(n-r)(p-r)$ orlanti di Δ_r . Infatti ci sono $(n-r)$ possibilità di "aggiungere una riga a Δ_r " e per ognuna di queste, $(p-r)$ possibilità di aggiungere una colonna.

4.2: Esempio : Sia $M = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$ e $\Delta = \det M_{I,J}$ dove $I = \{2\}$, $J = \{1,3\}$. Quindi Δ

$= \begin{vmatrix} 2 & 4 \\ 10 & 12 \end{vmatrix}$. Ci sono $(3-2)(4-2) = 2$ orlanti di Δ , sono $\begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 8 \\ 9 & 10 & 12 \end{vmatrix}$ e $\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 10 & 11 & 12 \end{vmatrix}$.

5: Teorema: Sia $M \in M_{n,p}(k)$. Allora $\text{rg}(M) = r$ se e solo se esiste un minore d'ordine r di M , Δ_r , tale che $\Delta_r \neq 0$ e tale che tutti gli orlanti di Δ_r siano nulli.

Dim: Se $\text{rg}(M) = r$ allora esiste un minore di ordine r , Δ_r , di M non nullo (cfr 3), inoltre, sempre per 3, tutti i minori di ordine $r+1$ di M sono nulli, quindi, in particolare, tutti gli orlanti di Δ_r sono nulli.

Sia adesso $\Delta_r \neq 0$. Possiamo supporre che Δ_r sia formato con le righe e le colonne di indice $1, 2, \dots, r$ di M . Siano c_1, \dots, c_n i vettori colonna di M . Siccome $\Delta_r \neq 0$, c_1, \dots, c_r sono linearmente indipendenti (cfr 3). Mostriamo che se $\text{rg}(M) \geq$

$r+1$ allora esiste un orlato di Δ_r non nullo. Se $rg(M) \geq r+1$ esiste $k \geq r+1$ tale che c_1, \dots, c_r, c_k siano linearmente indipendenti ("completare la base dell'immagine"). Sia N la matrice le cui colonne sono c_1, \dots, c_r, c_k . E' chiaro che $rg(N) = r+1$. Siano r_1, \dots, r_n , le righe di N . Siccome $\Delta_r \neq 0$, r_1, \dots, r_r sono linearmente indipendenti (cfr 3). Da $rg(N) = r+1$ segue che esiste $t > r$ tale che r_1, \dots, r_r, r_t siano liberi. Allora il determinante costruito su r_1, \dots, r_r, r_t è non nullo ed è un orlato di Δ_r .

5.1: Esempi : (i) Sia $M = \begin{vmatrix} 0 & 2 & -2 & 2 \\ 1 & 3 & 0 & 5 \\ 2 & 0 & 6 & 4 \end{vmatrix}$. Il minore $\Delta_2 = M_{I,J}$ con $I = \{3\}, J = \{3, 4\}$ vale

$$\begin{vmatrix} 0 & 2 \\ 2 & 0 \end{vmatrix} = -2 \neq 0; \text{ quindi } rg(M) \geq 2 \text{ (cfr 3). Ci sono } (3-2)(4-2) = 2 \text{ orlati di } \Delta_2, \text{ sono:}$$

$$\begin{vmatrix} 0 & 2 & -2 \\ 1 & 3 & 0 \\ 2 & 0 & 6 \end{vmatrix} = - \begin{vmatrix} 2 & -2 \\ 0 & 6 \end{vmatrix} + 2 \begin{vmatrix} 2 & -2 \\ 3 & 0 \end{vmatrix} = -12 + 2 \cdot 6 = 0, \text{ e:}$$

$$\begin{vmatrix} 0 & 2 & 2 \\ 1 & 3 & 5 \\ 2 & 0 & 4 \end{vmatrix} = - \begin{vmatrix} 2 & 2 \\ 0 & 4 \end{vmatrix} + 2 \begin{vmatrix} 2 & 2 \\ 3 & 5 \end{vmatrix} = -8 + 2(10 - 6) = 0. \text{ Si conclude (cfr 5) che } rg(M) = 2.$$

(ii) Siano $u = (x_1, \dots, x_n)$, $v = (y_1, \dots, y_n)$ due vettori di k^n . Questi due vettori sono

linearmente indipendenti se e solo se esiste un minore $\begin{vmatrix} x_i & y_i \\ x_k & y_k \end{vmatrix}$ ($i \neq k$), non nullo.

Esercizi:

15.1) Calcolare il rango delle seguenti matrici:

$$A = \begin{pmatrix} 1 & 7 & 5 & 3 & 2 \\ 0 & 4 & 2 & 2 & 0 \\ 2 & -2 & 4 & 0 & 1 \\ 3 & -1 & 7 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -3 & 4 \\ 3 & 1 & 5 \\ -4 & 0 & -1 \\ 0 & 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 & 0 & -3 & 0 \\ -3 & 0 & 1 & 6 & 1 \\ 2 & 0 & -1 & -4 & 0 \\ -4 & -4 & 0 & 5 & 1 \end{pmatrix}$$

15.2) Usando il metodo dei minori orlati, calcolare il rango della matrice:

$$\begin{pmatrix} 2 & 6 & 2 & -4 \\ 1 & 3 & 1 & -2 \\ -1 & -3 & 1/2 & -1 \end{pmatrix}$$

15.3) Usando il metodo dei minori orlati, calcolare il rango della matrice:

$$\begin{pmatrix} 1 & 7 & -1 & -7 \\ 1 & -1 & 3 & 5 \\ 2 & 2 & 4 & 4 \end{pmatrix}$$

15.4) Sia E un k -spazio vettoriale di dimensione n e x_1, \dots, x_k , k vettori di E . Si pone $V = \langle x_1, \dots, x_k \rangle$ (sottospazio generato da x_1, \dots, x_k). Sia V_l il sotto spazio generato da $x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_k$ con $x'_i = l_1 x_1 + \dots + l_i x_i + \dots + l_k x_k$ (si è rimpiazzato x_i con la combinazione lineare $l_1 x_1 + \dots + l_i x_i + \dots + l_k x_k$). Dimostrare:

- (i) se $l_i \neq 0$ allora $V_l = V$.
- (ii) se x_1, \dots, x_k , sono linearmente indipendenti allora $V_l = V$ se e solo se $l_i \neq 0$.

15.5) RIDUZIONE DELLE MATRICI.

Sia $M \in M_{n,p}(k)$. Una riga nulla di M è una riga i cui coefficienti sono tutti nulli. Si dice che la matrice M è ridotta per righe se in ogni riga non nulla esiste un coefficiente non nullo al di sotto del quale ci sono solo zeri; i.e. se la k -esima riga è non nulla esiste j , $1 \leq j \leq p$, tale che:

$a_{kj} \neq 0$ e $a_{rj} = 0$, $r \geq k+1$. Per esempio $M = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 0 & 5 \\ 0 & 0 & 6 \end{pmatrix}$ è ridotta per righe.

- (i) dimostrare che se M è ridotta per righe allora $\text{rg}(M)$ è uguale al numero di righe non nulle di M .
- (ii) scambiando la parola riga con la parola colonna si ottiene la nozione di matrice ridotta per colonne. Per una tale matrice il rango è uguale al numero di colonne non nulle.

15.6) CALCOLO DEL RANGO DI UNA MATRICE COL METODO DI RIDUZIONE.

Sia $M = (a_{ij})$, $M \in M_{n,p}(k)$. Siano R_1, \dots, R_n le righe di M . Una trasformazione del primo tipo sulle righe di M consiste nel sostituire una riga R_i con $R_i + \lambda R_j$ dove R_j è un'altra riga ($j \neq i$) e $\lambda \in k$. Dopo una tale trasformazione (indicata con $R_i \rightarrow R_i + \lambda R_j$) si ottiene la matrice

$$M' = \begin{pmatrix} R_1 \\ \vdots \\ R_i + \lambda R_j \\ \vdots \\ R_n \end{pmatrix}.$$

- (i) Dimostrare che $\text{rg}(M') = \text{rg}(M)$.
- (ii) Interpretare una trasformazione elementare del primo tipo come un cambiamento di basi.
- (iii) Se $a_{kt} \neq 0$ dimostrare che con trasformazioni elementari del primo tipo si arriva ad una matrice $N = (b_{ij})$ tale che $b_{kt} = a_{kt}$ e $b_{rt} = 0$, $r \geq k+1$. Concludere che con applicazioni ripetute di trasformazioni elementari si può trasformare ogni matrice in una matrice ridotta per righe.

Questo fornisce un procedimento per calcolare il rango di una matrice: data una matrice, M , con trasformazioni elementari la si trasforma in una matrice, M' , ridotta per righe; si ha allora (cfr Es.5) $\text{rg}(M) =$ il numero di righe non nulle di M' .

Siccome $\text{rg}(M) = \text{rg}({}^t M)$, tutto questo è ancora valido scambiando la parola riga con la parola colonna.

(iv) Siano M e M' come sopra (M qualsiasi, M' ridotta per righe (o per colonne), dedotta da M). Sono M , M' matrici equivalenti, simili?

(v) Dedurre che ogni matrice è equivalente ad una matrice del tipo: $S = \begin{pmatrix} \dots & \dots & \dots \\ 0 & \ddots & \vdots \\ 0 & \ddots & \ddots \\ 0 & & 0 \end{pmatrix}$ (gli zeri indicano matrici nulle); se $S = (a_{ij})$ allora $a_{ij} = 0$ se $i \geq r+1$ o se $i > j$; $a_{ii} \neq 0$ se $1 \leq i \leq r$ dove $r = \text{rg}(S)$.

(vi) applicare il metodo di riduzione per righe (o colonne) per calcolare $\text{rg}(M)$ dove $M = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix}$.

(vii) Calcolare $\text{rg}(M)$ usando i determinanti.

15.7) Sia M una matrice quadrata. Notiamo con M' una matrice ridotta per righe ottenuta da M dopo applicazioni di trasformazioni elementari del primo tipo. Dimostrare che $\det(M) = \det(M')$; è ancora vera questa uguaglianza se invece di trasformazioni elementari del primo tipo si fanno trasformazioni qualsiasi? (i.e. $R_i \rightarrow \alpha R_i + \lambda R_j$).

§16) DIAGONALIZZAZIONE.

In questo paragrafo consideriamo il problema seguente: sia u un endomorfismo di un k -spazio vettoriale di dimensione finita, E ; trovare una base, B , di E tale che $\text{mat}(u; B, B)$ sia la più "semplice" possibile. Le matrici più semplici sono, in un certo senso, le matrici diagonali. Questo riflette il fatto che gli endomorfismi più semplici sono le omotetie, ossia le applicazioni del tipo: $f: E \rightarrow E : x \rightarrow \lambda x$; infatti per ogni base B di E : $\text{mat}(f; B, B) = \lambda I_n$.

Se $\text{mat}(u; B, B)$ è diagonale allora in qualche modo abbiamo espresso u come somma di omotetie. Supponiamo $M = \text{mat}(u; B, B)$ diagonale: $M = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$. Poniamo $B = (e_1, \dots, e_n)$. Allora $u(e_i) = \lambda_i e_i$. In particolare gli e_i sono "autovettori" (cfr Def.1) di u ($\lambda_1, \dots, \lambda_n$ sono gli autovalori associati). Vediamo così che $\text{mat}(u; B, B)$ è diagonale se e solo se esiste una base di E fatta da "autovettori" di u . Vedremo che una tale base non esiste sempre e daremo delle condizioni necessarie e sufficienti per la sua esistenza.

La versione "matriciale" del problema è la seguente: sia $M \in M_n(k)$ una matrice quadrata. Dopo scelta di una base, B , in uno k -spazio vettoriale di dimensione n , E (per es. $E = k^n$ e B la base canonica), M corrisponde ad un endomorfismo, u , di E nell'isomorfismo $\text{mat}(-; B, B): \text{End}(E) \rightarrow M_n(k)$. Diagonalizzare M vuol dire diagonalizzare u ossia cercare una base, B' , tale che $\text{mat}(u; B', B') := M'$ sia diagonale. Avremo allora $M' = P^{-1}MP$ dove $P = \text{mat}(\text{Id}_E; B', B)$. In conclusione M è diagonalizzabile se e solo se M è simile ad una matrice diagonale.

1: Definizione: (Autowettore) Sia E un k -spazio vettoriale e u un endomorfismo di E . Un vettore x di E è un autovettore di u se: $x \neq 0$ e se esiste $\lambda \in k$ tale che $u(x) = \lambda x$.

1.1: Osservazione : In queste condizioni si dice che λ è associato a x (nb: ogni λ sarebbe associato a $x = 0$).

2: Definizione: (Autovalore) Con le notazioni precedenti sia $\lambda \in k$. Se esiste un vettore x di E tale che: $x \neq 0$ e $u(x) = \lambda x$ allora si dice che λ è un autovalore di u .

2.1: Osservazione : Se $u(x) = \lambda x$ per qualche $x \neq 0$, il vettore x è un autovettore (associato a λ). Osservare che mentre ad ogni autovettore è associato un unico autovalore (cfr 1.1) ad un autovalore sono associati più autovettori. Più precisamente:

3: Lemma: Sia $u \in \text{End}(E)$; λ è un autovalore di u se e soltanto se $\lambda \cdot \text{Id}_E - u$ non è iniettiva.

Se λ è un autovalore di u , l'insieme degli autovettori associati a λ è $\text{Ker}(\lambda \cdot \text{Id}_E - u) \setminus \{0\}$.

Dim: E' chiaro ♦

4: Definizione : Sia $u \in \text{End}(E)$. Per ogni scalare $\lambda \in k$ si pone $E_u(\lambda) := \text{Ker}(\lambda \cdot \text{Id}_E - u) = \{x \in E / u(x) = \lambda x\}$. Se λ è un autovalore di u , $E_u(\lambda)$ si chiama l'autospazio associato a λ .

4.1: Osservazione : (i) E' chiaro che, per ogni λ in k , $E_u(\lambda)$ è un sottospazio vettoriale di E (è il ker di un morfismo).

(ii) Inoltre λ è un autovalore di u se e solo se $E_u(\lambda) \neq \{0\}$ (cfr 3).

(iii) Osserviamo che se $x \in E_u(\lambda)$ allora $u(x) = y = \lambda x$ e $u(y) = u(\lambda x) = \lambda u(x) = \lambda y$. Quindi $u(E_u(\lambda)) \subseteq E_u(\lambda)$. Si dice che $E_u(\lambda)$ è stabile per u . Gli autospazi di u sono sottospazi stabili per u . La restrizione di u a $E_u(\lambda)$ non è altro che l'omotetia $E_u(\lambda) \rightarrow E_u(\lambda) : x \rightarrow \lambda x = u(x)$.

5: Lemma : Siano $\lambda_1, \dots, \lambda_m$, m autovalori, a due a due distinti ($i \neq j \Rightarrow \lambda_i \neq \lambda_j$), di u e x_1, \dots, x_m degli autovettori associati ($u(x_i) = \lambda_i x_i$, $1 \leq i \leq m$). Allora x_1, \dots, x_m sono linearmente indipendenti.

Dim: Induzione su m . Se $m = 1$, è chiaro. Assumiamo il lemma dimostrato per $m-1$. Sia $\alpha_1 x_1 + \dots + \alpha_m x_m = 0$ (*). Mostriamo che una tale relazione implica $\alpha_1 = \dots = \alpha_m = 0$. Abbiamo $u(\alpha_1 x_1 + \dots + \alpha_m x_m) = \lambda_1 \alpha_1 x_1 + \dots + \lambda_m \alpha_m x_m = 0$. D'altra parte moltiplicando (*) per λ_1 : $\lambda_1 \alpha_1 x_1 + \dots + \lambda_1 \alpha_m x_m = 0$. Per sottrazione: $\alpha_2 x_2 (\lambda_2 - \lambda_1) + \dots + \alpha_m x_m (\lambda_m - \lambda_1) = 0$. Siccome $\lambda_i \neq \lambda_1$ per $i = 2, \dots, m$ (ipotesi), abbiamo $\alpha_2 = \dots = \alpha_m = 0$. Inserendo in (*) viene anche $\alpha_1 = 0$ ♦

6: Corollario: Siano E un k -spazio vettoriale di dimensione n e u un endomorfismo di E . Se u ha n autovalori a due a due distinti allora u è diagonalizzabile.

Dim: Siano $\lambda_1, \dots, \lambda_n$ gli n autovalori a due a due distinti ($i \neq j \Rightarrow \lambda_i \neq \lambda_j$) di u . Siano x_1, \dots, x_n degli autovettori associati. Per 5 , x_1, \dots, x_n sono linearmente indipendenti. Quindi formano una base, B , di E . E' chiaro che $\text{mat}(u; B, B)$ è diagonale ♦

7: Lemma : Siano $\lambda_1, \dots, \lambda_m$ autovalori di u due a due distinti ($i \neq j \Rightarrow \lambda_i \neq \lambda_j$). Gli autospazi corrispondenti, $E_u(\lambda_1), \dots, E_u(\lambda_m)$, sono in somma diretta (i.e. $E_u(\lambda_1) + \dots + E_u(\lambda_m) = E_u(\lambda_1) \oplus \dots \oplus E_u(\lambda_m)$).

Dim: Mostriamo che $E_u(\lambda_k) \cap (E_u(\lambda_1) + \dots + E_u(\lambda_{k-1}) + E_u(\lambda_{k+1}) + \dots + E_u(\lambda_m)) = \{0\}$. Se $x_k = \alpha_1 x_1 + \dots + \alpha_{k-1} x_{k-1} + \alpha_{k+1} x_{k+1} + \dots + \alpha_m x_m$ (con $x_i \in E_u(\lambda_i)$), allora x_1, \dots, x_m sono legati. Il lemma 5 implica $x_1 = \dots = x_m = 0$ ♦

POLINOMIO CARATTERISTICO.

Siano $M \in M_n(k)$ e E un k -spazio vettoriale di dimensione n . La scelta di una base, B , di E stabilisce un isomorfismo (di k -algebre) $\text{mat}(-; B, B) : \text{End}(E) \rightarrow M_n(k)$. Sia u tale che $\text{mat}(u; B, B) = M$.

8: Definizione: Gli autovalori di M sono, per definizione, gli autovalori di u .

8.1: Osservazione : Se x è un autovettore di u allora $u(x) = \lambda \cdot x$. Tenuto conto della scrittura matriciale delle applicazioni lineari questa relazione si può anche scrivere: $M \cdot X = \lambda \cdot X$ dove X è la matrice $(n, 1)$ delle coordinate di x nella base B ; diremo che X è un autovettore di M .

Possiamo riassumere quanto detto in precedenza:

9: Proposizione: Se $M \in M_n(k)$, per ogni λ in k le proprietà seguenti sono equivalenti:

- (i) λ è un autovalore di M
- (ii) $\lambda \cdot I_n - M$ non è invertibile
- (iii) $\det(\lambda \cdot I_n - M) = 0$.

Dim: cfr 4, 4.1 ♦

Per determinare gli autovalori di una matrice, di un endomorfismo, l'idea è di usare 9(iii). Se $M = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$, allora $\lambda \cdot I_n - M = (\lambda \cdot \delta_{ij} - a_{ij})$ dove δ_{ij} è il simbolo di Kronecker ($\delta_{ij} = 1$ se $i = j$, $\delta_{ij} = 0$ se $i \neq j$). Perciò:

$$\det(\lambda \cdot I_n - M) = \sum_{\sigma \in S_n} \epsilon(\sigma) \cdot (\lambda \cdot \delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdot (\lambda \cdot \delta_{\sigma(2)2} - a_{\sigma(2)2}) \cdots (\lambda \cdot \delta_{\sigma(n)n} - a_{\sigma(n)n}).$$

Sviluppando vediamo che $\det(\lambda \cdot I_n - M)$ è un polinomio in λ , $P_M(\lambda)$. Cerchiamo di calcolare il grado di questo polinomio.

Se $\sigma = \text{Id}$, il termine $\epsilon(\sigma) \cdot (\lambda \cdot \delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdot (\lambda \cdot \delta_{\sigma(2)2} - a_{\sigma(2)2}) \cdots (\lambda \cdot \delta_{\sigma(n)n} - a_{\sigma(n)n})$ diventa:

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) = \lambda^n - \lambda^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) + \cdots + (-1)^n a_{11} \cdot a_{22} \cdots a_{nn}.$$

Se $\sigma \neq \text{Id}$, ci sono almeno due indici k, t , $1 \leq k \leq n$, $1 \leq t \leq n$, $k \neq t$ e tali che $\sigma(k) \neq k$ e $\sigma(t) \neq t$. Avremo allora $\delta_{\sigma(k)k} = \delta_{\sigma(t)t} = 0$ quindi il termine:

$\epsilon(\sigma) \cdot (\lambda \cdot \delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdot (\lambda \cdot \delta_{\sigma(2)2} - a_{\sigma(2)2}) \cdots (\lambda \cdot \delta_{\sigma(n)n} - a_{\sigma(n)n})$ ha al più grado $n-2$ in λ .

Finalmente osserviamo che il termine costante del polinomio $P_M(\lambda)$ è $P_M(0) = \det(-M) = (-1)^n \cdot \det(M)$.

Abbiamo dimostrato:

10: Proposizione: Con le notazioni precedenti $P_M(\lambda) = \det(\lambda \cdot I_n - M)$ è un polinomio di grado n in λ ; più precisamente abbiamo: $P_M(\lambda) = \lambda^n - \lambda^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) + \cdots + (-1)^n \cdot \det(M)$.

Sia A una matrice simile a M : $A = P^{-1}MP$. Allora $\lambda \cdot I_n - A = \lambda \cdot (P^{-1}I_n P) - (P^{-1}MP)$ perché ovviamente $P^{-1}I_n P = I_n$. Siccome le matrici scalari permutano con ogni matrice (cfr Es. 9.8) questo si può riscrivere: $P^{-1}(\lambda \cdot I_n)P - (P^{-1}MP) = P^{-1}(\lambda \cdot I_n - M)P$. Quindi $\lambda \cdot I_n - A = P^{-1}(\lambda \cdot I_n - M)P$ e perciò $\det(\lambda \cdot I_n - A) = (\det P)^{-1} \cdot \det(\lambda \cdot I_n - M) \cdot \det(P) = \det(\lambda \cdot I_n - M)$. In altre parole $P_M(\lambda) = P_A(\lambda)$.

Finalmente se $M = \text{mat}(u; B, B)$ allora se B' è un'altra base avremo $\text{mat}(u; B', B') = Q^{-1}MQ = N$ per qualche matrice invertibile Q . Da quanto precede $P_M(\lambda) = P_N(\lambda)$. Possiamo enunciare:

11: Proposizione: Sia u un endomorfismo di E e $M = \text{mat}(u; B, B)$ dove B è una base di E . Il polinomio $P_M(\lambda) = \det(\lambda \cdot I_n - M)$ è invariante quando si rimpiazza M con una matrice simile (i.e. quando si cambia base). Si dice che $P_M(\lambda)$ è il polinomio caratteristico di $M \in M_n(k)$ o il polinomio caratteristico di $u \in \text{End}(E)$ (in questo caso lo si nota anche $P_u(\lambda)$).

11.1: Osservazione : Abbiamo $P_M(\lambda) = \lambda^n - \lambda^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) + \cdots + (-1)^n \cdot \det(M)$. Il termine $a_{11} + a_{22} + \cdots + a_{nn}$ si chiama la traccia della matrice M

e si nota $\text{tr}(M)$. Se N è simile a M , $P_M(\lambda) = P_N(\lambda)$ quindi questi due polinomi hanno gli stessi coefficienti, in particolare $\text{tr}(M) = \text{tr}(N)$ e $\det(M) = \det(N)$.

Prima di enunciare il prossimo risultato ci servono alcuni richiami sui polinomi.

POLINOMI.

Sia k un campo e P un polinomio, a coefficienti in k , in una variabile X : $P \in k[X]$ e $P(X) = a_n X^n + \dots + a_1 X + a_0$. Una radice in k di P è un elemento, $\alpha \in k$, tale che $P(\alpha) = 0$. Si dimostra che α è una radice di P se e solo se $(X-\alpha)$ divide $P(X)$: $P(X) = (X-\alpha)Q(X)$. Se $(X-\alpha)$ divide $Q(X)$ allora $P(X) = (X-\alpha)^2 Q_1(X)$ e si dice che α è radice multipla di P . Più generalmente la molteplicità di una radice α è la più grande potenza di $(X-\alpha)$ che divide $P(X)$. Contando le radici con le loro molteplicità, vediamo che se $\deg(P) = n$ (i.e. con le notazioni precedenti $a_n \neq 0$) allora P ha al più n radici in k . Per esempio $X^4 - X^3 - 3X^2 + 5X - 2$ ha quattro radici in \mathbb{R} infatti $X^4 - X^3 - 3X^2 + 5X - 2 = (X-1)^3(X+2)$; le quattro radici sono 1 (con molteplicità 3) e -2 (con molteplicità 1). Può succedere però che un polinomio P non abbia neanche una radice in k . Per esempio il polinomio a coefficienti reali $P(X) = X^2 + 1$, non ha nessuna radice in $\mathbb{R} = k$. Il fatto che ogni polinomio, $P \in k[X]$, abbia una radice in k è una proprietà del campo k .

12: Definizione: Un campo k è detto algebricamente chiuso se ogni polinomio non costante, $P(X) \in k[X]$ ha una radice in k .

Per esempio \mathbb{R} non è algebricamente chiuso mentre \mathbb{C} lo è (cfr corso di algebra). Si vede facilmente che se k è algebricamente chiuso allora ogni polinomio di grado n a coefficienti in k ha n radici (contate con le loro molteplicità) in k .

Finalmente si può dimostrare che ogni campo, k , è un sottocampo di un campo algebricamente chiuso; inoltre esiste un unico (modulo k -isomorfismi) "più piccolo" campo algebricamente chiuso contenente k (teorema di Steinitz); questo campo si chiama la chiusura algebrica di k . Per esempio la chiusura algebrica di \mathbb{R} è \mathbb{C} ; la chiusura algebrica di \mathbb{Q} è \mathbb{Q}' il campo dei numeri algebrici: $\mathbb{Q}' = \{\alpha \in \mathbb{C} / \alpha \text{ è radice di un polinomio a coefficienti razionali}\}$. Si ha $\mathbb{Q} \subset \mathbb{Q}' \subset \mathbb{C}$, le inclusioni sono tutte strette (\mathbb{Q} è numerabile e quindi anche \mathbb{Q}' lo è, mentre \mathbb{C} non è numerabile; pertanto \mathbb{Q}'

$\neq C$). Un numero complesso appartenente a Q' si dice numero algebrico mentre gli elementi di $C-Q'$ sono i numeri trascendenti. Per esempio $i \in Q'$ mentre $\pi \notin Q'$: π è trascendente. Quindi se $P(X) \in k[X]$ si può considerare $P(X)$ anche come un elemento di $k'[X]$ (k' = chiusura algebrica di k). Visto come elemento di $k'[X]$, $P(X)$ ha $\deg(P)$ radici (contate con molteplicità) in k' . Per esempio $P(X) = X^2 + 1$ ha due radici in C : sono i e $-i$.

Detto ciò torniamo alla diagonalizzazione. Innanzitutto:

13: Proposizione: *Sia u un endomorfismo del k -spazio vettoriale E di dimensione n . Allora gli autovalori di u sono le radici (in k) del polinomio caratteristico $P_u(X)$. In particolare u ha al più n autovalori (contati con molteplicità).*

Dim: Abbiamo visto che λ è un autovalore di u se e solo se $\lambda I - u$ non è iniettiva ossia se e solo se $\det(\lambda I - M) = 0$ (M una matrice associata a u). Da quanto precede questo è equivalente a $P_u(\lambda) = 0$. Finalmente $P_u(X)$ ha grado n quindi ha al più n radici in k (contate con molteplicità)♦

13.1: Osservazione : Nell' enunciato precedente la molteplicità di un autovalore è la sua molteplicità come radice del polinomio caratteristico

14: Teorema : *Sia u un endomorfismo del k -spazio vettoriale E di dimensione n . Allora u è diagonalizzabile se e soltanto se sono verificate le due condizioni seguenti:*

- (i) *il polinomio caratteristico di u ha le sue n radici (contate con molteplicità) in k ;*
- (ii) *per ogni autovalore λ_i , di molteplicità k_i , di u : $\dim(E_u(\lambda_i)) = k_i$.*

Iniziamo con un lemma:

15: Lemma : *Sia u un endomorfismo del k -spazio vettoriale E . Se λ è una radice in k , di molteplicità t , di $P_u(X)$ allora la dimensione dell'autospazio $E_u(\lambda)$ è al più t .*

Dim: Sia $\dim(E_u(\lambda)) = h$ e sia (e_1, \dots, e_h) una base di $E_u(\lambda)$. Possiamo completare (e_1, \dots, e_h) in una base di E : $B = (e_1, \dots, e_n)$. La matrice, M , di u nella base B sarà del tipo: $M = \begin{pmatrix} \lambda I_h & A \\ 0 & A' \end{pmatrix}$; dove A' è quadrata ($n-h, n-h$).

Abbiamo (cfr §14, 9) $P_u(X) = \det(XI_n - M) = \det((X-\lambda)I_h) \cdot \det(XI_{n-h} - A') = (X-\lambda)^h \cdot Q(X)$. Siccome, per ipotesi, λ è una radice di molteplicità t di P_u , $h \geq t$ ♦

Dimostrazione del teorema 14: (i) Supponiamo u diagonalizzabile. Esiste quindi una base $B = (e_1, \dots, e_n)$ di E dove ogni e_i è un autovettore di u . Riordinando semmai i vettori della base possiamo supporre che $M = \text{mat}(u; B, B)$ sia diagonale della forma:

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \lambda_2 & & & \\ & & & & \ddots & & \\ & & & & & \lambda_2 & \\ & & & & & & \ddots \\ & & & & & & & \lambda_m \\ & & & & & & & & \lambda_m \end{pmatrix}$$

Quindi $E = E_u(\lambda_1) \oplus \dots \oplus E_u(\lambda_m)$. Perciò $n = \dim(E) = \dim(E_u(\lambda_1)) + \dots + \dim(E_u(\lambda_m))$. Siccome $\dim(E_u(\lambda_i)) \leq k_i$ (cfr 15), dove k_i è la molteplicità di λ_i come radice di P_u , abbiamo $n \leq k_1 + \dots + k_m$. D'altra parte, siccome P_u ha grado n : $k_1 + \dots + k_m \leq n$. In conclusione $k_1 + \dots + k_m = n$ e quindi $\lambda_1, \dots, \lambda_m$ sono tutte le radici di P_u (con le molteplicità k_1, \dots, k_m). Pertanto P_u ha tutte le sue radici in k . Inoltre siccome $\dim(E_u(\lambda_1)) + \dots + \dim(E_u(\lambda_m)) = n = k_1 + \dots + k_m$ e $\dim(E_u(\lambda_i)) \leq k_i$, $i = 1, \dots, m$, segue che $\dim(E_u(\lambda_i)) = k_i$ per ogni i .

(ii) Supponiamo le condizioni (i) e (ii) soddisfatte. Siano $\lambda_1, \dots, \lambda_m$, tutte le radici distinte di P_u e sia k_i la molteplicità di λ_i . Per ipotesi $\lambda_i \in k$, per ogni i . Siccome queste sono tutte le radici di P_u , che ha grado n , abbiamo $k_1 + \dots + k_m = n$. Dall'ipotesi segue inoltre: $\dim(E_u(\lambda_1)) + \dots + \dim(E_u(\lambda_m)) = k_1 + \dots + k_m = n$. Ma la somma degli autospazi è diretta (cfr 7); quindi $E_u(\lambda_1) \oplus \dots \oplus E_u(\lambda_m) = E$. Se B è una base di E riunione di basi di $E_u(\lambda_1), \dots, E_u(\lambda_m)$, allora $\text{mat}(u; B, B)$ è diagonale ♦

16: Corollario: Sia f un endomorfismo di E , k -spazio vettoriale di dimensione n . Se k è algebricamente chiuso (per es. $k = \mathbb{C}$) e se P_f non ha radici multiple allora f è diagonalizzabile.

Dim: Se k è algebricamente chiuso, $P_f(X) \in k[X]$ ha tutte le sue radici in k . Per ipotesi queste radici sono a due a due distinte. Quindi f ha n autovalori a due a due distinti. Si conclude con 6♦

17: Esempi :

(a) Sia $M \in M_2(\mathbb{R})$, $M = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. Il polinomio caratteristico $P_M(X)$ ha grado due, dunque da (10) segue che $P_M(X) = X^2 - \text{tr}(M)X + \det(M)$. Abbiamo $\text{tr}(M) = 2$ e $\det(M) = 3$ quindi $P_M(X) = X^2 - 2X + 3 = (X-1)^2 + 2$. Questo polinomio non ha nessuna radice reale. Pertanto M non ha autovalori (su \mathbb{R}) e quindi M non è diagonalizzabile.

Se guardiamo adesso M come una matrice a coefficienti complessi ($M \in M_2(\mathbb{C})$) la situazione è diversa. Infatti $P_M(X)$ ha due radici complesse distinte: $X^2 - 2X + 3 = (X-1-i\sqrt{2})(X-1+i\sqrt{2})$. Quindi M ha due autovalori complessi distinti pertanto (cfr 6) M è diagonalizzabile (su \mathbb{C}).

(b) Sia $M \in M_2(\mathbb{R})$, $M = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. Abbiamo $P_M(X) = X^2 - \text{tr}(M)X + \det(M) = (X-2)^2$. Quindi M ha un autovalore (2) di molteplicità due. Da (14), M è diagonalizzabile se e solo se l'autospazio corrispondente, $E_M(2)$, ha dimensione due, ossia se e solo se $E_M(2)$ è tutto E . Quindi per ogni vettore, X , di \mathbb{R}^2 dovremmo avere $MX = 2X$, questo, chiaramente, non è vero, quindi M non è diagonalizzabile.

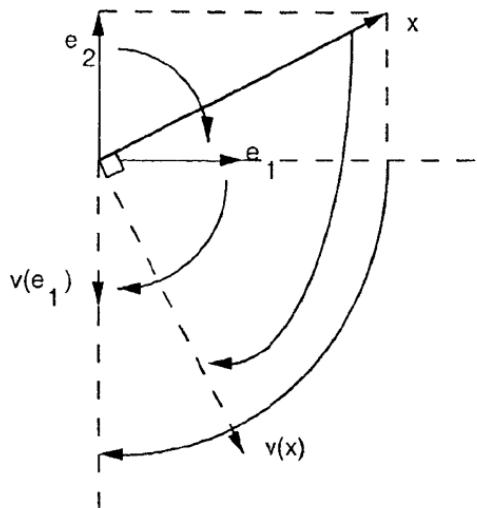
Si può anche ragionare così: se M è diagonalizzabile allora M è simile a $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2I_2$, ma ogni matrice simile a αI è uguale a αI perché: $P^{-1}(\alpha I)P = \alpha I$ (nb: le matrici scalari permangono con ogni matrice, cfr Es. 9.8). Siccome, ovviamente $M \neq 2I$, M non è diagonalizzabile. E' chiaro che M non è diagonalizzabile neanche su \mathbb{C} .

(c) Sia $M \in M_3(\mathbb{R})$, $M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Abbiamo $\det(\lambda I - M) = \begin{vmatrix} \lambda-1 & -1 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda-2 \end{vmatrix}$ e usando

Es. 14.10 viene subito $P_M(X) = (X-1)^2 \cdot (X-2)$. Quindi M ha un autovalore di molteplicità due (1) e uno semplice (2). Da (14) M è diagonalizzabile se e solo se la dimensione dell'autospazio $E_M(1)$ è due. Siccome $E_M(1) = \text{Ker}(I-M)$ per calcolare $\dim(\text{Ker}(I-M))$ basta calcolare il rango della matrice $I-M$ (perché, per il teorema delle dimensioni: $3 = \dim(\text{Ker}(I-M)) + \text{rg}(I-M)$). Abbiamo $I-M =$

$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Il determinante è ovviamente nullo (1 è autovalore di M) e il minore $\det M_{\{2\},\{1\}} = \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} = 1$. Quindi (cfr teorema §15, 5) $\text{rg}(I-M) = 2$ e $\dim(\text{Ker}(I-M)) = 1$; M non è diagonalizzabile (neanche su \mathbb{C}).

(d) Sia $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Sia v l'endomorfismo di \mathbb{R}^2 corrispondente a M nell'isomorfismo $\text{mat}(\cdot; C, C) : \text{End}(\mathbb{R}^2) \rightarrow M_2(\mathbb{R})$ dove $C = (e_1, e_2)$ è la base canonica di \mathbb{R}^2 . Abbiamo $v(e_1) = -e_2$, $v(e_2) = e_1$. Quindi se $x = \alpha e_1 + \beta e_2$, $v(x) = \beta e_1 - \alpha e_2$. Vediamo così che v è una rotazione di angolo $\pi/2$:



In particolare v non ha sottospazi stabili (i.e. non c'è nessuna retta vettoriale, D, tale che $v(D)$ sia contenuto in D). Pertanto v non è diagonalizzabile (cfr 4.1 (iii))

Conclusione.

A parte qualche caso particolare (cfr esempio (e)) il metodo standard per provare a diagonalizzare una matrice (n,n) è il seguente:

- si scrive il polinomio caratteristico, per questo si calcola $\det(\lambda I - M)$ (nel caso $n = 2$ basta ricordarsi che il coefficiente di λ^{n-1} è $\text{tr}(M)$ e il termine costante, $\det(M)$)

- si cercano gli autovalori di M e le loro molteplicità (i.e. le radici in k del polinomio caratteristico e le loro molteplicità).

- se ci sono n radici distinte in k allora M è diagonalizzabile.

- se la somma delle molteplicità delle radici in k è $< n$, M non è diagonalizzabile.

- se la somma delle molteplicità delle radici in k è n ma ci sono radici multiple allora M è diagonalizzabile se e solo se gli autospazi corrispondenti ad autovalori multipli hanno per dimensione la molteplicità dell'autovalore associato. L'autospazio corrispondente a λ è $\text{Ker}(\lambda I - M)$, la dimensione di questo sottospazio è uguale a $n - r$ dove $r = \text{rg}(\lambda I - M)$; r si può calcolare usando la tecnica dei minori. Finalmente per trovare una base fatta da autovettori, bisogna trovare delle basi degli autospazi $\text{Ker}(\lambda_i I - M)$; per questo bisogna risolvere dei sistemi lineari (cfr §17).

Esercizi:

16.1) Sia $M \in M_2(\mathbb{R})$, $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(i) Mostrare che se $\det(M) < 0$ allora M è diagonalizzabile.

(ii) Si suppone $c \neq 0$ o $b \neq 0$ (i.e. M non diagonale). Dimostrare che M è diagonalizzabile se e solo se: $a^2 + d^2 + 4bc > 2ad$.

16.2) Dire se $A \in M_3(\mathbb{R})$ è diagonalizzabile dove $A = \begin{pmatrix} 3 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{pmatrix}$.

In caso di risposta affermativa, dare una base di autovettori.

16.3) Dire se $A \in M_3(\mathbb{C})$ è diagonalizzabile dove $A = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

16.4) Sia E un k -spazio vettoriale dove k è un campo di caratteristica $\neq 2$ (i.e. si può dividere per 2 in k), e $f : E \rightarrow E$ un endomorfismo tale che $f \circ f = \text{Id}_E$. Si pone $E_+ = \{x \in E / f(x) = x\}$ e $E_- = \{x \in E / f(x) = -x\}$.

(i) Mostrare che E_+ e E_- sono due sottospazi supplementari di E .

(ii) Mostrare che f è diagonalizzabile.

16.5) Dimostrare che ogni matrice simmetrica, $M \in M_2(\mathbb{R})$, è diagonalizzabile.

16.6) Sia $M \in M_2(\mathbb{C})$. Dimostrare che M è simile ad una matrice diagonale (a coefficienti in \mathbb{C}) o ad una matrice triangolare del tipo $\begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix}$ ($a, b \in \mathbb{C}$). E' ancora vero questo risultato su \mathbb{R} ?

16.7) Si consideri l'applicazione lineare $f: k^3 \rightarrow k^3$ tale che $\text{mat}(f; B, B) = A$ dove B è la base canonica di k^3 e dove $A = \begin{pmatrix} 5 & -5 & -1 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$

- (i) Si dica in quale dei seguenti casi f è diagonalizzabile: $k = \mathbb{R}$, e $k = \mathbb{C}$.
- (ii) Si determinino gli autovalori e i relativi autospazi di f nei due casi $k = \mathbb{R}$, $k = \mathbb{C}$.
- (iii) Nel caso in cui f sia diagonalizzabile, si dica quante sono le basi di k^3 che rendono la matrice di f diagonale.
- (iv) Sia $k = \mathbb{R}$; si costruisca una base, C , di \mathbb{R}^3 tale che la seconda colonna di $\text{mat}(f; C, C)$ sia $(0, 1, 0)$.

16.8) (i) Sia E un k -spazio vettoriale e $f: E \rightarrow E$ un endomorfismo soddisfacente: (a) f non è iniettivo, (b) esiste un iper piano vettoriale, H , di E tale che la restrizione di f ad H sia l'identità. E' f diagonalizzabile?

(ii) per ogni $n \geq 2$ dare un esempio di un endomorfismo soddisfacente (b) e inoltre tale che:

- (α) f è suriettivo, $f \neq \text{Id}$ e f è diagonalizzabile
- (β) f è suriettivo e f non è diagonalizzabile.

16.9) Sia B la base canonica di \mathbb{R}^3 . Si consideri l'endomorfismo $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ tale che

$$\text{mat}(f; B, B) = \begin{pmatrix} -2 & -6 & -6 \\ 3/2 & 4 & 3 \\ 3/2 & 3 & 4 \end{pmatrix}.$$

(i) Dire se f è diagonalizzabile.

(ii) In caso di risposta affermativa, trovare una base di \mathbb{R}^3 costituita da autovettori di f .

16.10) Sia $A \in M_3(k)$, $A = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Dire se A è diagonalizzabile dove:

(i) $k = \mathbb{R}$, (ii) $k = \mathbb{C}$, (iii) $k = \mathbb{Z}/2\mathbb{Z}$

(iv) Sia E un R -spazio vettoriale di dimensione tre, B una base di E , e $f: E \rightarrow E$ un endomorfismo tale che $\text{mat}(f; B, B) = A$. Dire quante (zero, un numero finito > 0 , infinite)

sono le basi B' di E tali che la seconda colonna di $\text{mat}(f; B', B')$ sia $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

16.11) Siano E un k -spazio vettoriale di dimensione 4, $B = (e_1, e_2, e_3, e_4)$ una base di E ,

$\rightarrow E$ l'endomorfismo tale che $\text{mat}(f; B, B) = M$ dove $M = \begin{pmatrix} 13 & 8 & -12 & -12 \\ -5 & -2 & 5 & 5 \\ 6 & 4 & -5 & -6 \\ 3 & 2 & -3 & -2 \end{pmatrix}$.

- (i) Dire se f è diagonalizzabile dove $k = \mathbf{R}, \mathbf{C}$ (suggerimento: guardare bene la matrice prima di iniziare i conti!).
(ii) Se f è diagonalizzabile dare una base formata da autovettori (si avrà cura di precisare la base nella quale sono espressi gli autovettori).

16.12) Siano E un k -spazio vettoriale di dimensione n , u e v due endomorfismi di E . Se u è biettivo dimostrare che $u \circ v$ e $v \circ u$ hanno lo stesso polinomio caratteristico (suggerimento: considerare $u^{-1} \circ u \circ v \circ u$).

16.13) Sia E un k -spazio vettoriale di dimensione n .

(i) Siano f, g due endomorfismi di E . Dimostrare: $f \circ g$ non iniettivo $\Rightarrow g \circ f$ non iniettivo.

(ii) Sia $\lambda \neq 0, \lambda \in k$, e $h \in \text{End}(E)$, dimostrare: λ è autovalore di $h \Leftrightarrow I - \frac{1}{\lambda} h$ non è biettivo (I è l'identità).

(iii) Siano $A, B \in M_n(k)$. Dimostrare: $I_n - A \cdot B$ invertibile $\Rightarrow I_n - B \cdot A$ invertibile (suggerimento: mostrare che $(I_n - B \cdot A)(I_n + B(I_n - A \cdot B)^{-1}A) = I_n$).

(iv) Concludere che se u, v sono due endomorfismi di E , allora uno scalare è autovalore di $u \circ v$ se e solo se è autovalore di $v \circ u$.

16.14) Sia $A \in M_n(\mathbf{R})$, $n \geq 2$, una matrice tale che $A^2 + I_n = 0$. Dimostrare che A non è diagonalizzabile.

16.15) Siano E un \mathbf{C} -spazio vettoriale di dimensione finita, e f un endomorfismo di E . Sia V un sottospazio vettoriale di E tale che $f(V) \subseteq V$. Dimostrare che V contiene un autovettore di f .

Sia g un altro endomorfismo di E . Si suppone che g commuta con f . Mostrare che ogni autopazio di f è stabile sotto g .

Dedurre che f e g hanno un autovettore in comune.

16.16) Siano A, B in $M_n(\mathbf{R})$. Si suppone A e B diagonalizzabili. E' $A+B$ diagonalizzabile?

16.17) Sia $v \in \text{End}(\mathbf{R}^3)$ tale che $\text{mat}(v; C, C) = N$ dove C è la base canonica e dove

$$N = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

(i) Dire se v è diagonalizzabile.

(ii) Determinare una base B tale che se $M = \text{mat}(v; B, B)$, la prima e la terza colonna di M siano ${}^t(-2, 0, 0)$, ${}^t(0, 0, -2)$.

(iii) Sia $b = \{B / B$ è una base di \mathbf{R}^3 soddisfacente (ii)\}. Dire se b è un insieme finito o infinito. Se $B \in b$ e se $M = (a_{ij})$ è la matrice $\text{mat}(v; B, B)$ quanto vale a_{22} ?

16.18) Sia $H = \{(x, y, z, t) \in \mathbf{R}^4 | 2x-y+z=0, x-y+t=0\}$.

(i) Esiste $f \in \text{End}(\mathbf{R}^4)$ tale che $f|_H = \text{Id}_H$, 2 sia autovalore di f , e f sia diagonalizzabile? Se sì, se ne costruisca una.

(5)(ii) Esiste $g \in \text{End}(\mathbf{R}^4)$ tale che $g|_H = \text{Id}_H$, 2 sia autovalore di g , e g non sia diagonalizzabile? Se una tale g esiste si ponga $\varphi = g - 2\text{Id}_{\mathbf{R}^4}$. Che valori può assumere $\dim(\text{Im } \varphi)$?

16.19) Si considerino le seguenti matrici di $M_3(\mathbf{R})$: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(i) Determinare la dimensione del sotto spazio $\langle A, B, C \rangle$ di $M_3(\mathbf{R})$.

Sia E la base canonica di \mathbf{R}^3 e siano $f, g \in \text{End}(\mathbf{R}^3)$ tali che $A = \text{mat}(f; E, E)$, $B = \text{mat}(g; E, E)$.

(ii) E' possibile determinare un sottospazio U di \mathbf{R}^3 tale che $\dim((f \circ g)(U)) \neq \dim(U)$?

(iii) Gli endomorfismi f, g sono diagonalizzabili?

(iv) Le matrici A e B sono simili?

16.20) Trovare due matrici reali $(2, 2)$, non simili, ma con lo stesso polinomio caratteristico.

16.21) Esiste un'applicazione lineare $f: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tale che $f((0, 1, -1)) = (2, 0, 0)$ e tale che $V = \{(x, y, z) \in \mathbf{R}^3 / y - z = 0\}$ sia l'autospazio di f relativo all'autovalore 2 ?

In caso di risposta affermativa si dica se f è univocamente determinata.

16.22) Siano $A, B \in M_3(\mathbf{R})$: $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & a & 0 \\ 0 & a & 1 \end{pmatrix}$, $B = \begin{pmatrix} a & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

(i) Si dica per quali valori del parametro reale a è A (risp. B) diagonalizzabile.

(ii) Si dica per quali valori del parametro reale a le due matrici A, B sono simili.

16.23) (i) Sia $B \in M_n(\mathbf{R})$ tale che $B^3 + I_n = 0$. Dimostrare che se B è diagonalizzabile (su \mathbf{R}) allora $B = -I_n$.

(ii) Esiste $B \in M_3(\mathbf{C})$, $B \neq -I_3$, B diagonalizzabile e tale che $B^3 + I_3 = 0$?

16.24) Sia E un k -spazio vettoriale, di dimensione ≥ 2 e $f: E \rightarrow E$ un endomorfismo. Si suppone che f non sia un'omotetia (i.e. f non è della forma λId).

Dimostrare che esiste x tale che $f(x)$ e x siano linearmente indipendenti.

16.25) Sia $M \in M_n(k)$ una matrice della forma $M = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{12} & a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$, con

$$a_{11} = \dots = a_{nn}.$$

(i) Trovare una condizione necessaria e sufficiente affinché M sia diagonalizzabile.

(ii) Per ogni campo k e per ogni $n \geq 2$ dare un esempio di matrice $N \in M_k(k)$ non diagonalizzabile.

16.26) Il teorema di Cayley-Hamilton.

Siano k un campo e E un k -spazio vettoriale. Se $Q(X) \in k[X]$ è un polinomio a coefficienti in k , nella variabile X e se $u \in \text{End}(E)$ possiamo definire $Q(u) \in \text{End}(E)$ nel modo seguente: se $Q(X) = a_n X^n + \dots + a_1 X + a_0$, allora $Q(u) = a_n u^n + \dots + a_1 u + a_0 \text{Id}_E$ ($u^k = u \circ \dots \circ u$, k fattori). Osservare che $u^{k+p} = u^k \circ u^p$, e che $(\alpha u)(x) = \alpha \cdot u(x)$ per ogni x in E .

Nello stesso modo se $M \in M_n(k)$ possiamo definire la matrice $Q(M) = a_n M^n + \dots + a_1 M + a_0 I_n$.

Teorema (Cayley-Hamilton): Sia E un k -spazio vettoriale di dimensione n e $u \in \text{End}(E)$. Notiamo $P_u(X)$ il polinomio caratteristico di u . Allora $P_u(u) = 0$.

Sia $A \in M_n(K)$ e notiamo $P_A(X)$ il polinomio caratteristico di A allora $P_A(A) = 0$.

La dimostrazione procede in due tappe. Si suppone dapprima che $P_u(X)$ abbia tutte le sue radici in k : $P_u(X) = (X - \lambda_1) \dots (X - \lambda_n)$; $\lambda_1, \dots, \lambda_n$, sono gli autovalori di u .

(i) Dimostrare (per induzione su n) che esiste una base, B , di E tale che $\text{mat}(u; B, B)$ sia triangolare superiore, di elementi diagonali gli autovalori $\lambda_1, \dots, \lambda_n$ (suggerimento: sia e_1 un autovettore di u associato a λ_1 , completiamo e_1 ad una base $B' = (e_1, e'_2, \dots, e'_n)$ di

$$\begin{pmatrix} \lambda_1 & \dots \\ 0 & \\ \vdots & C \\ 0 & \end{pmatrix}$$

E allora $\text{mat}(u; B', B') = \begin{pmatrix} \lambda_1 & & & \\ 0 & & & \\ \vdots & C & & \\ 0 & & & \end{pmatrix}$, dove C è $(n-1, n-1)$; cercare di applicare l'ipotesi di

induzione a C considerando la proiezione $p: E \rightarrow E'$ dove $E' = \langle e'_2, \dots, e'_n \rangle$.

(ii) Sia $u_i := (\lambda_i \text{Id}_E - u)$. Dimostrare che $u_k \circ u_i = u_i \circ u_k$ per ogni i, k .

(iii) Sia B la base fornita da (i). Se $F_i := u_1 \circ \dots \circ u_i$, dimostrare che $F_i(e_1) = F_i(e_2) = \dots = F_i(e_n) = 0$ e concludere la dimostrazione del teorema sotto queste ipotesi ($P_u(X)$ ha tutte le sue radici in k).

Per dimostrare il caso generale si lavora nella chiusura algebrica, k' , di k . Sia $f \in \text{End}(E)$.

Sia B una base di E e $M = \text{mat}(f; B, B)$. Allora $M \in M_n(k)$ e $P_f(X) = P_M(X)$. Allora $M \in M_n(k) \subseteq M_n(k')$.

(vi) Considerando $P_M(X)$, concludere la dimostrazione del teorema di Cayley-Hamilton.

(vii) Si potrebbe essere tentati di ragionare nel modo seguente: $P_M(X) = \det(XI_n - M)$, siccome $MI_n = M$, ponendo $X=M$ viene $P_M(M) = \det(M-M) = 0$. Spiegare perché questa argomentazione è pericolosa.

16.27) Un endomorfismo $u: E \rightarrow E$ si dice nilpotente se esiste un intero k tale che $u^k = 0$.

(i) Mostrare che se λ è un autovalore di u allora $\lambda = 0$.

(ii) Mostrare che il polinomio caratteristico di u è $P_u(X) = X^n$. (in particolare $u^n = 0$, $n = \dim(E)$).

16.28) Sia $u: E \rightarrow E$ un endomorfismo del k -spazio vettoriale. Se $P(X) \in k[X]$, possiamo considerare l'endomorfismo $P(u)$ (cfr Es.26), otteniamo così un'applicazione $\varphi_u: k[X] \rightarrow \text{End}(E)$, φ_u è un morfismo d'anelli e $\text{Ker}(\varphi_u)$ è un ideale di $k[X]$. Siccome $k[X]$ è un anello principale (divisione euclidea, cfr corso di algebra), esiste un unico polinomio monico, $M_u(X)$, tale che: $P(X) \in \text{Ker}(\varphi_u)$ se e solo se $M_u(X)$ divide $P(X)$. Si ha $M_u(u) = 0$. Il polinomio $M_u(X)$ si chiama il polinomio minimale di u (è il polinomio monico di grado più basso tale che $M(u) = 0$). Il polinomio caratteristico, $P_u(X)$, è un multiplo di $M_u(X)$.

Se u è un endomorfismo nilpotente di che forma è il suo polinomio minimale?

§17) SISTEMI LINEARI.

In questo paragrafo si danno alcuni metodi per risolvere i sistemi lineari. La parte teorica è piuttosto facile; la difficoltà, semmai, sta nella scelta del metodo per calcolare le soluzioni di un dato sistema nel modo più economico possibile. Vedremo che non esiste un metodo "universalmente economico", piuttosto si cercherà, a seconda del sistema, di usare il metodo più adatto (e poi è anche una questione di gusti...).

SISTEMI LINEARI: DEFINIZIONI.

Un'equazione lineare (o di primo grado), a coefficienti in k , nelle incognite X_1, \dots, X_n è un'equazione della forma: $a_1X_1 + \dots + a_nX_n = b$ dove a_1, \dots, a_n, b sono elementi di k . Si dice anche che a_1, \dots, a_n, b sono i coefficienti (tavolta b viene chiamato "termine noto"), e X_1, \dots, X_n le incognite. Una soluzione dell'equazione è un elemento, (x_1, \dots, x_n) , di k^n tale che la relazione $a_1x_1 + \dots + a_nx_n = b$ sia soddisfatta.

Un sistema lineare, a coefficienti in k , nelle incognite X_1, \dots, X_n , è un insieme di equazioni lineari a coefficienti in k , nelle incognite X_1, \dots, X_n :

$$(S) \begin{cases} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n = b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n = b_2 \\ \dots \dots \dots \dots \dots \\ a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pn}X_n = b_p \end{cases}$$

Il sistema (S) è un sistema lineare, a coefficienti in k , di p equazioni nelle n incognite X_1, \dots, X_n . Se $b_1 = 0, \dots, b_p = 0$, si dice che (S) è un sistema lineare **omogeneo** (cfr §8). In ogni caso la matrice dei coefficienti del sistema è la matrice (p,n) :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{pmatrix}.$$

INTERPRETAZIONI VETTORIALE, MATRICIALE.

Siano E, F due k -spazi vettoriali di dimensioni n, p . Scegliendo delle basi $B = (e_1, \dots, e_n), C = (f_1, \dots, f_p)$ di E, F possiamo associare ad A un morfismo lineare $v: E \rightarrow F$, tramite l'isomorfismo $\text{mat}(\cdot; B, C): L_k(E, F) \rightarrow M_{p,n}(k)$. Dopo questa scelta delle basi, possiamo vedere l'insieme delle soluzioni di (S) come un sottinsieme di E : $\underline{S} = \{\underline{x} \in E / \underline{x} = x_1e_1 + \dots + x_ne_n \text{ con } (x_1, \dots, x_n) \text{ soluzione di } (S)\}$. Ovviamente il sottinsieme \underline{S} di E dipende dalle basi, per esempio se $B' = (e'_1, \dots, e'_n)$ è un'altra base di E e se $\underline{S}' = \{\underline{x} \in E / \underline{x} = x_1e'_1 + \dots + x_ne'_n \text{ con } (x_1, \dots, x_n) \text{ soluzione di } (S)\}$, allora in generale $\underline{S} \neq \underline{S}'$.

Sia $B^* = (e^*_1, \dots, e^*_n)$ la base duale di B e poniamo $L_i = a_{i1}e^*_1 + \dots + a_{in}e^*_n$, $1 \leq i \leq p$. Le L_i sono forme lineari su E e se $\underline{x} = x_1e_1 + \dots + x_ne_n$, allora $L_i(\underline{x}) = a_{i1}x_1 + \dots + a_{in}x_n$. Possiamo riscrivere il sistema (S) sotto la forma più compatta: $\{L_1(\underline{x}) = b_1, \dots, L_p(\underline{x}) = b_p\}$. Le L_i corrispondono alle righe della matrice A , per questo chiameremo questa presentazione l'interpretazione per "righe".

Facciamo una parentesi per segnalare che questo approccio permette (partendo dalle forme lineari L_i e dagli scalari b_i) di definire la nozione (intrinseca) di sistema lineare in uno spazio vettoriale E .

Sistemi lineari in uno spazio vettoriale.

Siano L_1, \dots, L_p , delle forme lineari su E e b_1, \dots, b_p degli scalari. Questi dati definiscono il sistema lineare $(S) = \{L_1(\underline{x}) = b_1, \dots, L_p(\underline{x}) = b_p\}$. L'insieme delle soluzioni di (S) è $S = \{\underline{x} \in E / \text{le relazioni } L_1(\underline{x}) = b_1, \dots, L_p(\underline{x}) = b_p \text{ siano soddisfatte}\}$. Osservare che il sottinsieme S di E è definito in modo intrinseco (i.e. senza usare basi). Prendendo una base, B , di E e la base duale, B^* , ci si riporta alla nozione iniziale di sistema lineare "numerico".

Torniamo al nostro morfismo lineare $v: E \rightarrow F$. Per semplificare la scrittura prenderemo $E = k^n, F = k^p$ e B, B' le basi canoniche. Abbiamo quindi:

$$k^n \rightarrow k^p: \underline{x} = (X_1, \dots, X_n) \rightarrow v(\underline{x}) = A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}. \text{ Sia } B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in k^p \text{ e } X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in k^n$$

(noteremo con delle lettere maiuscole i vettori colonna e con delle minuscole sottolineate i vettori riga).

Sotto forma matriciale il sistema (S) si scrive: $A\mathbf{x} = \mathbf{b}$.

Sotto forma vettoriale invece il sistema (S) si scrive: $v(\underline{\mathbf{x}}) = \underline{\mathbf{b}}$,
o anche: $\{L_1(\underline{\mathbf{x}}) = b_1, \dots, L_p(\underline{\mathbf{x}}) = b_p\}$.

STRUTTURA DELL'INSIEME DELLE SOLUZIONI.

Con l'interpretazione vettoriale vediamo che l'insieme delle soluzioni di (S), S , non è altro che la contr'immagine di $\underline{\mathbf{b}}$ tramite v : $S = v^{-1}(\underline{\mathbf{b}}) = \{\underline{\mathbf{x}} = (x_1, \dots, x_n) / v(\underline{\mathbf{x}}) = \underline{\mathbf{b}}\}$. In altre parole S è la fibra di v in $\underline{\mathbf{b}}$.

1: Definizione: Il sistema (S) è detto compatibile se $S \neq \emptyset$, incompatibile se $S = \emptyset$.

1.1: Esempio : Se $a_{ik} = a_{1k}$ per ogni k , $1 \leq k \leq n$, e se $b_i \neq b_1$ allora (S) è incompatibile.

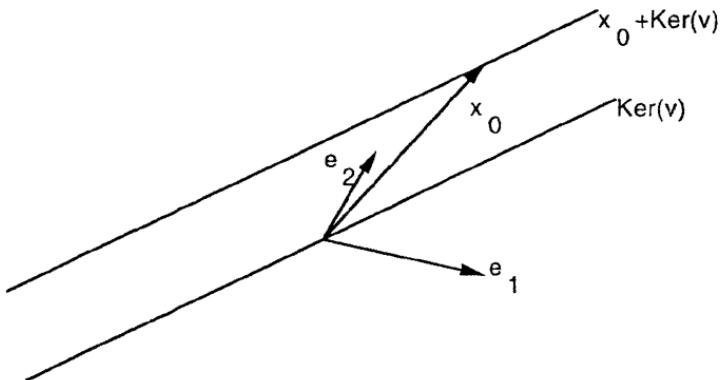
2: Proposizione: L'insieme delle soluzioni di un sistema lineare omogeneo: $A\mathbf{x} = 0$, con $A \in M_{p,n}(k)$, è un sottospazio vettoriale di dimensione $n-r$ di k^n dove $r = rg(A)$.

Dim: cfr §8 ♦

Nel caso non omogeneo (i.e. se $\underline{\mathbf{b}} \neq \underline{\mathbf{0}}$) l'insieme delle soluzioni non è mai un sottospazio vettoriale (osservare tra l'altro che S può essere vuoto): infatti S non è neanche un sottogruppo additivo: $\underline{\mathbf{Q}} \in S$ se e solo se $v(\underline{\mathbf{Q}}) = \underline{\mathbf{b}}$ ma, per linearità, $v(\underline{\mathbf{Q}}) = \underline{\mathbf{0}}$.

Osserviamo però il fatto seguente: se $\underline{\mathbf{x}}_0$ è una soluzione particolare di (S) allora ogni altra soluzione è della forma $\underline{\mathbf{x}} = \underline{\mathbf{x}}_0 + \underline{\mathbf{z}}$ con $\underline{\mathbf{z}} \in \text{Ker}(v)$. Infatti se $\underline{\mathbf{z}} \in \text{Ker}(v)$, allora $v(\underline{\mathbf{x}}_0 + \underline{\mathbf{z}}) = v(\underline{\mathbf{x}}_0) = \underline{\mathbf{b}}$. Viceversa se $v(\underline{\mathbf{x}}) = \underline{\mathbf{b}}$ allora $\underline{\mathbf{x}} = \underline{\mathbf{x}}_0 + (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)$ e $v(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) = \underline{\mathbf{0}}$ quindi $\underline{\mathbf{z}} = \underline{\mathbf{x}} - \underline{\mathbf{x}}_0$ appartiene a $\text{Ker}(v)$.

In conclusione se il sistema è compatibile: $S = \underline{\mathbf{x}}_0 + \text{Ker}(v) = \{\underline{\mathbf{x}}_0 + \underline{\mathbf{z}} / \underline{\mathbf{z}} \in \text{Ker}(v)\}$. Si dice che S è il sottospazio affine di k^n passante per $\underline{\mathbf{x}}_0$ e parallelo a $\text{Ker}(v)$:



Vediamo che S è il traslato, per x_0 , del sottospazio vettoriale $\text{Ker}(v)$. In particolare un sottospazio vettoriale è anche un sottospazio affine ($x_0 = \underline{0}$); un sottospazio vettoriale è un sottospazio affine che "passa per l'origine". La dimensione di un tale sottospazio affine è, per definizione, la dimensione del sottospazio vettoriale associato, nel nostro caso $\text{Ker}(v)$.

3: Proposizione: Se il sistema lineare (S) è compatibile l'insieme delle sue soluzioni è il sottospazio affine di k^n , di dimensione $n-r$ (r è il rango di A), $x_0 + \text{Ker}(v)$ dove x_0 è una soluzione particolare di (S) .

3.1: Osservazione : (i) Come caso particolare ritroviamo **2** perché $\underline{0}$ è sempre soluzione particolare di un sistema omogeneo.

(ii) scriviamo il sistema (S) con l'interpretazione "per righe": $\{L_1(x) = b_1, \dots, L_p(x) = b_p\}$, allora l'insieme delle soluzioni del sistema omogeneo associato è $\langle L_1, \dots, L_p \rangle^*$ (cfr §8); e l'insieme delle soluzioni di (S) è $x_0 + \langle L_1, \dots, L_p \rangle^*$.

RANGO DI UN SISTEMA LINEARE, SISTEMI DI CRAMER.

4: Definizione: Il rango del sistema (S) è il rango della matrice, A , dei coefficienti (= il rango del morfismo v , uguale a $\dim \langle L_1, \dots, L_p \rangle$).

Il rango è un invariante importante del sistema (S) . Abbiamo già visto che se il sistema è compatibile allora l'insieme delle soluzioni è un sottospazio affine di k^n di dimensione $n-r$ (=il numero di incognite meno il rango). In particolare se ci sono più incognite che equazioni ($n > p$) allora, siccome $r \leq \min\{n, p\}$, avremo $n-r > 0$; se compatibile il sistema ammette un'infinità di

soluzioni (ci sono delle "incognite libere" che comportano "un'indeterminazione"). Abbiamo anche:

5: Proposizione: (i) Se $r = p$ il sistema (S) ammette sempre almeno una soluzione (i.e. (S) è compatibile).

(ii) Se $n = r = p$, il sistema (S) ammette una ed un'unica soluzione.

Dim: (i) Se $r = p$ il morfismo $v : k^n \rightarrow k^p$ è suriettivo quindi $v^{-1}(\underline{b}) \neq \emptyset$:

(ii) Se $r = p = n$ il morfismo $v : k^n \rightarrow k^p$ è suriettivo, quindi biiettivo ($n = p$) e la fibra $v^{-1}(\underline{b})$ consta di uno ed un unico elemento♦

6: Definizione: Se $n = r = p$, il sistema (S) si dice sistema di Cramer.

Quindi (S) è di Cramer se ci sono altrettanto incognite che equazioni e se le equazioni sono linearmente indipendenti. Questo è ancora equivalente a dire che la matrice, A, dei coefficienti è quadrata (n,n) ed invertibile (le righe di A sono le equazioni, cfr §8).

RISOLUZIONE DI UN SISTEMA DI CRAMER.

Se il sistema (S) $AX = B$ è di Cramer, abbiamo appena detto che la matrice A è invertibile. Quindi l'unica soluzione del sistema è: $X = A^{-1}B$.

Per calcolare effettivamente la soluzione di un sistema di Cramer ci sono diversi metodi, indichiamone uno che usa i determinanti.

Siano $\underline{c}_1, \dots, \underline{c}_n$ i vettori colonna della matrice A. Il sistema (S) si può riscrivere:

$$(S) \quad x_1 \cdot \underline{c}_1 + \dots + x_n \cdot \underline{c}_n = \underline{b}.$$

Si tratta quindi di trovare le coordinate x_1, \dots, x_n , del vettore \underline{b} di k^n nella base $(\underline{c}_1, \dots, \underline{c}_n)$ di k^n (nb: $(\underline{c}_1, \dots, \underline{c}_n)$ è una base di k^n perché $rg(A) = n$). Sia quindi $x_1 \cdot \underline{c}_1 + \dots + x_n \cdot \underline{c}_n = \underline{b}$ (*).

Consideriamo la matrice A_i , ottenuta da A rimpiazzando la i-esima colonna, \underline{c}_i , con \underline{b} : $A_i = [\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{b}, \underline{c}_{i+1}, \dots, \underline{c}_n]$. Adesso calcoliamo $\det(A_i)$ sostituendo \underline{b} con la sua espressione (*):

$\det(A_i) = \det(\underline{c}_1, \dots, \underline{c}_{i-1}, x_1 \cdot \underline{c}_1 + \dots + x_n \cdot \underline{c}_n, \underline{c}_{i+1}, \dots, \underline{c}_n)$, per linearità nella i-esima variabile:

$$\det(A_i) = x_1 \det(\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{c}_i, \underline{c}_{i+1}, \dots, \underline{c}_n) + x_2 \det(\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{c}_2, \underline{c}_{i+1}, \dots, \underline{c}_n) + \dots + x_i \det(\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{c}_i, \underline{c}_{i+1}, \dots, \underline{c}_n) + \dots + x_n \det(\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{c}_n, \underline{c}_{i+1}, \dots, \underline{c}_n).$$

Tutti i determinanti che compaiono in questa somma sono nulli tranne $\det(\underline{c}_1, \dots, \underline{c}_{i-1}, \underline{c}_i, \underline{c}_{i+1}, \dots, \underline{c}_n) = \det(A)$; infatti gli altri determinanti sono determinanti di matrici con due colonne uguali.

Perciò $\det(A_i) = x_i \cdot \det(A)$ e siccome $\det(A) \neq 0$ (A è invertibile) si ricava: $x_i = \det(A_i)/\det(A)$. Abbiamo dimostrato:

7: Proposizione: *Un sistema di Cramer: $AX = B$ ($n = r = p$) ha una ed un'unica soluzione $X = (x_1, \dots, x_n)$ data da: $x_i = \det(A_i)/\det(A)$, $1 \leq i \leq n$, dove A_i è la matrice ottenuta sostituendo la i -esima colonna di A con B .*

7.1: Esempio : Sia (S) il sistema:

$$x_1 + 2x_2 = 3$$

$$-4x_1 + x_2 = -2.$$

Abbiamo $A = \begin{pmatrix} 1 & 2 \\ -4 & 1 \end{pmatrix}$ e $\det(A) = 9$. Quindi il sistema è di Cramer e la soluzione è: $x_1 = \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} / 9 = 7/9$, $x_2 = \begin{vmatrix} 1 & 3 \\ -4 & -2 \end{vmatrix} / 9 = 10/9$.

IL CASO GENERALE.

L'idea per risolvere un sistema generale di p equazioni a n incognite, è di cercare di ricondursi ad un sistema di Cramer. Prima si eliminano le equazioni superflue (i.e. che sono combinazioni lineari di altre equazioni); in altre parole si cerca una base dello "spazio delle righe di A ". A questo punto intervengono le condizioni di compatibilità. Per esempio nel sistema (S): $\{L_1(\underline{x}) = b_1, L_2(\underline{x}) = b_2, \lambda L_1(\underline{x}) + \eta L_2(\underline{x}) = b_3\}$, con $L_i(\underline{x}) = a_{i1}x_1 + \dots + a_{in}x_n$, possiamo eliminare la terza equazione purchè $\lambda b_1 + \eta b_2 = b_3$. Infatti se \underline{x} è soluzione di (S) dalle prime due equazioni si ricava $\lambda L_1(\underline{x}) + \eta L_2(\underline{x}) = \lambda b_1 + \eta b_2 = b_3$. Se invece $\lambda b_1 + \eta b_2 \neq b_3$ il sistema è chiaramente incompatibile.

Una volta eliminate le equazioni superflue e verificate le condizioni di compatibilità si è ricondotti a risolvere un sistema (S') in n incognite, di r equazioni e di rango r . Inoltre l'insieme delle soluzioni di (S') è uguale a quello di (S). Il sistema (S') è compatibile (cfr 5 (i)). Per risolverlo si prende una base "dello spazio delle colonne" di A' , la matrice (r, n) dei coefficienti di (S'). Le incognite che non figurano nelle r colonne prescelte sono le incognite libere; le si fanno passare nel secondo membro ottenendo così per

ogni valore delle incognite libere un sistema di Cramer che si risolve, per esempio col metodo precedente. Vediamo i dettagli.

(i) Riduzione delle equazioni (o righe).

Consideriamo l'interpretazione per righe. Il sistema (S) si può riscrivere: (S) $\{L_1(\underline{x}) = b_1, \dots, L_p(\underline{x}) = b_p\}$. Sia r il rango del sistema, il sottospazio $\langle L_1, \dots, L_p \rangle$ di $(k^n)^*$ ha dimensione r e, riordinando eventualmente gli indici, possiamo supporre che (L_1, \dots, L_r) sia una base di $\langle L_1, \dots, L_p \rangle$. Esistono quindi degli scalari $\lambda^{(j)}_k$ tali che: $L_j = \lambda^{(j)}_1.L_1 + \dots + \lambda^{(j)}_r.L_r, j = r+1, \dots, p$.

Sia (S') il sistema: (S') $\{L_1(\underline{x}) = b_1, \dots, L_r(\underline{x}) = b_r\}$.

Si dice che (S') è un sistema principale associato a (S). Osserviamo che (S') dipende dalla scelta di una base di $\langle L_1, \dots, L_p \rangle$; se per esempio L_2, \dots, L_{r+1} , sono liberi otteniamo un altro sistema principale associato a (S).

8: Definizione: Con le notazioni precedenti siano le relazioni:

$$(t) b_j = \lambda^{(j)}_1.b_1 + \dots + \lambda^{(j)}_r.b_r, j = r+1, \dots, p.$$

Le relazioni (t) sono dette relazioni di compatibilità (di (S) rispetto a (S')).

9: Proposizione: Con le notazioni precedenti:

(i) Se le condizioni di compatibilità (t) non sono soddisfatte, il sistema (S) non è compatibile.

(ii) Se le condizioni di compatibilità (t) sono soddisfatte, l'insieme delle soluzioni di (S) è uguale a quello di (S') (si dice che (S) e (S') sono equivalenti).

Dim: (i) Se esiste k , $r+1 \leq k \leq n$, tale che $b_k \neq \lambda^{(k)}_1.b_1 + \dots + \lambda^{(k)}_r.b_r$, e se \underline{x} è una soluzione di (S) allora siccome $L_k = \lambda^{(k)}_1.L_1 + \dots + \lambda^{(k)}_r.L_r$, si ottiene $b_k = L_k(\underline{x}) = \lambda^{(k)}_1.L_1(\underline{x}) + \dots + \lambda^{(k)}_r.L_r(\underline{x}) = \lambda^{(k)}_1.b_1 + \dots + \lambda^{(k)}_r.b_r$, ma questo è assurdo.

(ii) Supponiamo (t) soddisfatte. E' chiaro che se \underline{x} è soluzione di (S) allora \underline{x} è pure soluzione di (S'). Viceversa se \underline{x} è soluzione di (S') allora \underline{x} verifica le prime r equazioni di (S) e le condizioni (t) permettono di verificare le rimanenti equazioni♦

Questa proposizione ci porta a risolvere sistemi di r equazioni in n incognite e di rango r . Abbiamo già visto (cfr 5 (i)) che tali sistemi sono sempre compatibili.

(ii) Riduzione delle incognite (o delle colonne).

Sia $A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rn} \end{pmatrix}$, la matrice dei coefficienti del sistema (S') . Per ipotesi $\text{rg}(A') = r$. Ci sono quindi r colonne di A' linearmente indipendenti. Riordinando gli indici possiamo supporre che le r prime colonne c_1, \dots, c_r , siano linearmente indipendenti. A questo punto riscriviamo (S') nella forma seguente:

$$\begin{aligned} a_{11}X_1 + \dots + a_{1r}X_r &= b_1 - a_{1,r+1}X_{r+1} - \dots - a_{1n}X_n \\ \dots & \\ a_{r1}X_1 + \dots + a_{rr}X_r &= b_r - a_{r,r+1}X_{r+1} - \dots - a_{rn}X_n \end{aligned}$$

In forma matriciale $A'' \cdot X' = B(X_{r+1}, \dots, X_n)$ dove $X' = (X_1, \dots, X_r)$, A'' è la matrice (r, r) costruita sulle prime r righe e colonne di A' , $B(X_{r+1}, \dots, X_n)$ è il vettore colonna $(r, 1)$ di componenti $b_k - a_{k,r+1}X_{r+1} - \dots - a_{kn}X_n$, $1 \leq k \leq r$.

Ogni volta che diamo dei valori $\alpha_{r+1}, \dots, \alpha_n$, a X_{r+1}, \dots, X_n , otteniamo un sistema di Cramer la cui unica soluzione è $(x_1, \dots, x_r) = (A'')^{-1} \cdot B(\alpha_{r+1}, \dots, \alpha_n)$, o ancora, con notazioni analoghe a quelle di 7: $x_i = \det(A''_{-i}) / \det(A'')$, $1 \leq i \leq r$. Osserviamo che $\det(A''_{-i})$ è una funzione di $\alpha_{r+1}, \dots, \alpha_n$. Infatti sviluppando $\det(A''_{-i})$ secondo la i -esima colonna (uguale a $B(\alpha_{r+1}, \dots, \alpha_n)$) vediamo che x_i si esprime nella forma: $x_i = f^{(i)}_{r+1} \cdot \alpha_{r+1} + \dots + f^{(i)}_n \cdot \alpha_n + \beta_i$, dove $f^{(i)}_j, \beta_i$, sono elementi di k . Poniamo $l_i(\underline{\alpha}) := f^{(i)}_{r+1} \cdot \alpha_{r+1} + \dots + f^{(i)}_n \cdot \alpha_n$, di modo che $x_i = l_i(\underline{\alpha}) + \beta_i$, $1 \leq i \leq r$.

L'insieme, S' , delle soluzioni di (S') è quindi:

$$S' = \{(l_1(\underline{\alpha}) + \beta_1, \dots, l_r(\underline{\alpha}) + \beta_r; \underline{\alpha}) / \underline{\alpha} = (\alpha_{r+1}, \dots, \alpha_n) \in k^{n-r}\}.$$

Siccome $S' = S$, l'insieme delle soluzioni di (S) (cfr 9 (ii)), concludiamo che:

$$\begin{aligned} S &= \{(l_1(\underline{\alpha}) + \beta_1, \dots, l_r(\underline{\alpha}) + \beta_r; \underline{\alpha}) / \underline{\alpha} = (\alpha_{r+1}, \dots, \alpha_n) \in k^{n-r}\} \\ &= \{(\beta_1, \dots, \beta_r, 0, \dots, 0) + (l_1(\underline{\alpha}), \dots, l_r(\underline{\alpha}), \alpha_{r+1}, \dots, \alpha_n) / (\alpha_{r+1}, \dots, \alpha_n) \in k^{n-r}\} \\ &= \{(\beta_1, \dots, \beta_r, 0, \dots, 0) + W\} \end{aligned}$$

dove $W = \text{Im}(f)$ con $f : k^{n-r} \rightarrow k^n$: $(\alpha_{r+1}, \dots, \alpha_n) \rightarrow \begin{pmatrix} \dots & 1 & \dots \\ \dots & \dots & \dots \\ \dots & 1_r & \dots \\ I_{n-r} & \dots \end{pmatrix} \begin{pmatrix} \alpha_{r+1} \\ \vdots \\ \vdots \\ \alpha_n \end{pmatrix} = (l_1(\underline{\alpha}), \dots, l_r(\underline{\alpha}), \alpha_{r+1}, \dots, \alpha_n)$.

Siccome il rango di $\begin{pmatrix} \dots & 1 & \dots \\ \dots & \dots & \dots \\ \dots & 1_r & \dots \\ I_{n-r} & \dots \end{pmatrix}$ è $(n-r)$, $\dim(W) = n-r$, e ritroviamo (cfr 3) il

fatto che $S = S'$ è un sottospazio affine di dimensione $n-r$ di k^n .

(III) MODO OPERATIVO (riassunto):

Sia (S) un sistema lineare qualsiasi; per risolverlo si può procedere così:

(1) si calcola, con la tecnica dei minori orlati (cfr §15), il rango della matrice, A, dei coefficienti. Sia $\Delta_r \neq 0$ un minore di A, di ordine r, e non nullo, dove $r = \text{rg}(A)$. Si riordinano gli indici in modo che Δ_r sia costruito sulle prime r righe e colonne di A:

$$\Delta_r = \det(A') = \det \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}.$$

(2) nel caso non omogeneo, si guarda se le condizioni di compatibilità sono verificate. Osserviamo un modo semplice di dare queste condizioni: il sistema (S) si può anche scrivere: $X_1\underline{c}_1 + \dots + X_n\underline{c}_n = \underline{b}$ dove \underline{c}_i sono i vettori colonna di A e dove \underline{b} è il vettore colonna dei termini noti. Il sistema è compatibile se e solo se \underline{b} appartiene al sottospazio generato da $\underline{c}_1, \dots, \underline{c}_n$. Siccome questo sottospazio ha per base $\underline{c}_1, \dots, \underline{c}_r$ (cfr (1)), il sistema è compatibile se e solo se il rango di:

$\begin{pmatrix} a_{11} & \dots & a_{1r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r1} & \dots & a_{rr} & b_r \\ \vdots & & \vdots & \vdots \\ a_{p1} & \dots & a_{pr} & b_p \end{pmatrix}$ è uguale al rango di $\begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pr} \end{pmatrix}$, che vale r (cfr(1)).

Questo è equivalente a dire che i $p-r$ orlati di Δ_r sono nulli (cfr §15, 5): \det

$$\begin{pmatrix} & \vdots & b_1 \\ \Delta_r & \vdots & \vdots \\ & \vdots & b_r \\ a_{j1} \dots a_{jr} & b_j \end{pmatrix} = 0, \quad j = r+1, \dots, p \quad (\text{condizioni di compatibilità, cfr 8, 9}).$$

Osservare che questo è equivalente a dire che il sistema è compatibile se e solo se: $\text{rango}(A) = \text{rango}(A|B)$, dove A è la matrice dei coefficienti e dove $(A|B)$ è la matrice completa del sistema:

$$(A|B) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pn} & b_p \end{pmatrix}.$$

(3) se le condizioni di compatibilità (2) sono verificate il sistema ammette come insieme delle soluzioni un sottospazio affine di dimensione $n-r$ di k^n . Per calcolare le soluzioni ci si riporta al sistema di Cramer:

$$\begin{aligned} a_{11}X_1 + \dots + a_{1r}X_r &= b_1 - a_{1,r+1}X_{r+1} - \dots - a_{1n}X_n \\ \dots & \\ a_{r1}X_1 + \dots + a_{rr}X_r &= b_r - a_{r,r+1}X_{r+1} - \dots - a_{rn}X_n \end{aligned}$$

che si risolve formalmente col metodo della Prop.7, calcolando ancora determinanti.

Abbiamo dimostrato:

10: Teorema: (Rouché-Capelli) *Con le notazioni precedenti, il sistema lineare (S) è compatibile se e solo se: $\text{rango}(A) = \text{rango}(A|B)$. Se questa condizione è soddisfatta l'insieme delle soluzioni di (S) è un sottospazio affine di dimensione $n-r$ dove $r = \text{rango}(A)$.*

11: Osservazione : Ben inteso, nella pratica, il metodo appena esposto non è necessariamente il più rapido; può però essere utile in questioni teoriche in cui i coefficienti non sono completamente specificati e dove si tratta solo di determinare la dimensione dell'insieme delle soluzioni. Un altro interesse è che questo metodo fornisce un algoritmo per calcolare le soluzioni; in altre parole, questo metodo può essere programmato su un computer (se il computer sa calcolare i determinanti). Finalmente, nel caso di un sistema

con un gran numero di equazioni ed incognite, e con alcune proprietà di simmetria, i determinanti possono rispecchiare le simmetrie e quindi essere di gran aiuto. Per concludere diciamo che, nella pratica, si avrà tendenza a preferire il buon vecchio metodo di "combinazioni e sostituzioni" che illustreremo adesso usando la riduzione delle matrici.

$$\text{Sia } (S) \text{ il sistema lineare: } (S) \left\{ \begin{array}{l} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n = b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n = b_2 \\ \dots \dots \dots \dots \dots \dots \\ a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pn}X_n = b_p \end{array} \right.$$

in forma più compatta: $(S) = \{L_1(\underline{x}) = b_1, \dots, L_p(\underline{x}) = b_p\}$ dove $\underline{x} = (x_1, \dots, x_n)$, $L_k(\underline{x}) = a_{k1}x_1 + \dots + a_{kn}x_n$. Sia (S') il sistema ottenuto da (S) rimpiazzando la i -esima equazione $(L_i(\underline{x}) = b_i)$ con una combinazione lineare di tutte le equazioni: $T_i(\underline{x}) = c_i$, dove: $T_i = \lambda_1 L_1 + \dots + \lambda_i L_i + \dots + \lambda_p L_p$, $c_i = \lambda_1 b_1 + \dots + \lambda_i b_i + \dots + \lambda_p b_p$.

12: Lemma: Se $\lambda_i \neq 0$, (S) e (S') sono due sistemi lineari equivalenti (i.e. hanno lo stesso insieme delle soluzioni, eventualmente vuoto).

Dim: E' chiaro che se \underline{x} è soluzione di (S) allora è anche soluzione di (S') . Supponiamo \underline{x} soluzione di (S') allora \underline{x} soddisfa ovviamente tutte le equazioni di (S) tranne, forse, la i -esima, ma $T_i(\underline{x}) = c_i$, ossia: $\lambda_1 L_1(\underline{x}) + \dots + \lambda_i L_i(\underline{x}) + \dots + \lambda_p L_p(\underline{x}) = \lambda_1 b_1 + \dots + \lambda_i b_i + \dots + \lambda_p b_p$, da cui, siccome $\lambda_i \neq 0$: $\lambda'_1 L_1(\underline{x}) + \dots + \lambda'_i L_i(\underline{x}) + \dots + \lambda'_p L_p(\underline{x}) = \lambda'_1 b_1 + \dots + b_i + \dots + \lambda'_p b_p$ ($\lambda'_k = \lambda_k / \lambda_i$). Per ipotesi $L_k(\underline{x}) = b_k$ se $k = i$, pertanto dalla relazione appena vista: $L_i(\underline{x}) = b_i$ ♦

RIDUZIONE PER RIGHE DELLE MATRICI E SISTEMI LINEARI.

Si può sistematizzare il procedimento di risoluzione per combinazioni e sostituzioni nel modo seguente (per la riduzione per righe delle matrici vedere Es. 15.4, 15.5, 15.6).

1) Si riduce la matrice A per righe riportando le operazioni eseguite sulle righe di $(A|B)$. In altre parole si comincia a ridurre $(A|B)$ per righe fino ad arrivare ad una matrice $(A'|B')$ con A' ridotta per righe.

2) Poi si riordinano le colonne in modo da ottenere da A' una matrice del

tipo $A'' = \begin{pmatrix} \dots & \dots & \dots \\ 0 & \ddots & \vdots \\ & \ddots & \ddots \\ 0 & 0 & 0 \end{pmatrix}$ (cfr Es. 15.5). A questo punto si scrive il sistema (S') di

matrice dei coefficienti A'' e di matrice dei termini noti B' , facendo ATTENZIONE che, se si sono scambiate delle colonne in A' per ottenere A'' , bisogna anche scambiare le relative incognite (cfr esempio 13).

Il sistema (S') è equivalente ad (S) in virtù del lemma 12. Il sistema (S') è un sistema a "gradini", ossia della forma:

$$\begin{array}{lcl} a'_{11}y_1 + a'_{12}y_2 + \dots + a'_{1n}y_n & = & b'_1 \\ a'_{22}y_2 + \dots + a'_{2n}y_n & = & b'_2 \\ \dots & & \\ a'_{rr}y_r + \dots + a'_{rn}y_n & = & b'_r \\ 0 & = & b'_{r+1} \\ \dots & & \\ 0 & = & b'_p \end{array}$$

con $a'_{11} \neq 0, \dots, a'_{rr} \neq 0$ dove r è il rango di (S) . Le condizioni di compatibilità sono: $b'_{r+1} = 0, \dots, b'_p = 0$. Il sistema è compatibile se e soltanto se sono soddisfatte. Se il sistema è compatibile, per risolverlo si procede per sostituzioni successive, partendo dal basso. L'ultima equazione fornisce: $y_r = (b'_r - a'_{r,r+1}y_{r+1} - \dots - a'_{r,n}y_n) / a'_{rr}$. L'equazione precedente è: $a'_{r-1,r-1}y_{r-1} + \dots + a'_{r-1,n}y_n = b'_{r-1}$. Da cui $y_{r-1} = (b'_{r-1} - a'_{r-1,r}y_r - \dots - a'_{r-1,n}y_n) / a'_{r-1,r-1}$. Rimpiazzando y_r con l'espressione precedente, si esprime y_{r-1} in funzione delle $n-r$ incognite "libere" y_{r+1}, \dots, y_n . Procedendo man mano in questo modo si ricavano y_1, \dots, y_r come funzioni (affini) delle incognite libere y_{r+1}, \dots, y_n ; ritrovando così il fatto che l'insieme delle soluzioni di (S) è un sottospazio affine di dimensione $n-r$ di k^n .

13: Esempio : Sia (S) il sistema:
$$\begin{array}{lcl} x+2y+3z & = & -1 \\ -x+y & = & 2 \\ 2x+3y-4z & = & 1 \end{array}$$

Abbiamo $(A|B) = \begin{pmatrix} 1 & 2 & 3 & -1 \\ -1 & 1 & 0 & 2 \\ 2 & 3 & -4 & 1 \end{pmatrix}$. Si fa la trasformazione elementare $R_3 \rightarrow R_3 + 4R_1 / 3$ e si ottiene:

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ -1 & 1 & 0 & 2 \\ \frac{10}{3} & \frac{17}{3} & 0 & -\frac{1}{3} \end{pmatrix}$$

$$\text{Poi si fa } R_3 \rightarrow R_3 - 17R_2 / 3 \text{ e si ottiene: } \begin{pmatrix} 1 & 2 & 3 & -1 \\ -1 & 1 & 0 & 2 \\ 9 & 0 & 0 & -\frac{35}{3} \end{pmatrix}$$

Si riordinano le colonne per avere la forma a gradini facendo seguire gli scambi alle incognite:

$$\begin{pmatrix} x & y & z \\ 1 & 2 & 3 & -1 \\ -1 & 1 & 0 & 2 \\ 9 & 0 & 0 & -\frac{35}{3} \end{pmatrix} \rightarrow \begin{pmatrix} z & y & x \\ 3 & 2 & 1 & -1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 9 & -\frac{35}{3} \end{pmatrix}$$

Il sistema (S') è:

$$3x + 2y + z = -1$$

$$y - x = 2$$

$$9x = -35/3$$

Il sistema (S') (e quindi anche (S)) ha un'unica soluzione: $(x, y, z) = (-35/27, 19/27, -10/27)$. E' altresì chiaro che si può risolvere il sistema (S) per combinazioni e sostituzioni senza riportarsi alla forma a gradini.

Esercizi:

17.1) Risolvere, in \mathbb{R} , il seguente sistema lineare: $x - y + z = 27$
 $x - 2y + z = -7$
 $x + y - z = 33$.

17.2) Risolvere, a seconda del parametro reale λ , il sistema lineare:
 $\lambda x - 2y + 4 = 0$
 $x + (\lambda - 3)y + 6 - 2\lambda = 0$

17.3) Risolvere, a seconda del parametro reale λ , il sistema lineare:
 $(\lambda + 3)x - 4y + 4z = \lambda$

$$4x + (\lambda - 5)y + 4z = \lambda$$

$$4x - 4y + (\lambda + 3)z = -\lambda.$$

17.4) (i) Discutere, a seconda del parametro reale λ , il seguente sistema in \mathbb{R}^3 :

$$\begin{array}{l|l} & \lambda x + 6y + 5z = \lambda \\ S(\lambda) & \left\{ \begin{array}{l} \lambda x + (6+\lambda)y + 6z = 1 \\ x + 2\lambda y + \lambda z = 1 \end{array} \right. \end{array}$$

(ii) Per ogni λ in \mathbb{R} sia $S(\lambda)$ l'insieme delle soluzioni di $S(\lambda)$. Senza risolvere il sistema si può dire che: (a) esistono al più tre valori di λ tali che $\text{card}(S(\lambda)) \neq 1$; (b) per ogni λ in \mathbb{R} , $S(\lambda)$ è un sottospazio affine di dimensione ≤ 1 . Perché? (giustificare!)

17.5) Come costruire un sistema lineare non omogeneo: $f_i(x, y, z) = a_i$, $1 \leq i \leq 3$, con i coefficienti di x, y, z , non nulli in ogni f_i , la cui unica soluzione sia $(1, 2, 3)$?

17.6) (Messaggio in codice) Si può adoperare la teoria dei sistemi lineari per scrivere messaggi in codice. Ecco un esempio. Per scrivere si useranno p simboli dove p è un numero primo. Se si vogliono usare le lettere dell'alfabeto inglese (ABCDEFGHIJKLMNPQRSTUVWXYZ), che sono 26, basterà aggiungere tre simboli ($\$, \&, ?$) per arrivare a $p = 29$. Si stabilisce quindi una biiezione, f , tra i 29 simboli e gli elementi del campo $F_{29} = \mathbb{Z}/29\mathbb{Z}$, per esempio le lettere A, B...Z corrispondono a 1, 2, ..., 26 (mod.29) e $\$, \&, ?$ a 27, 28, 0. Poi si sceglie un intero n , per esempio $n = 3$. Il messaggio da criptare sarà diviso in gruppi di 3 ($=n$) lettere. Per esempio se vogliamo comunicare che il giorno prescelto è LUNEDI scriviamo (LUN). (EDI), tramite f questi due gruppi corrispondono a $(12, 21, 14), (5, 4, 9)$. Poi scegliamo una matrice, E , invertibile 3×3 a coefficienti in F_{29} : $E = (a_{ij})$. Per cifrare i due gruppi $(x, y, z) = (12, 21, 14), (5, 4, 9)$, prendiamo il vettore (u, v, w) immagine di ${}^t(x, y, z)$ tramite E . Per

esempio se $E = \begin{pmatrix} 102 \\ 010 \\ 002 \end{pmatrix}$ abbiamo $E.{}^t(12, 21, 14) = {}^t(40, 21, 28) = {}^t(11, 21, 28)$ (perché 40 è congruo a 11 mod.29).

Usando di nuovo la biiezione f , $(11, 21, 28)$ corrisponde a (KU&). Nello stesso modo $E.{}^t(5, 4, 9) = (23, 4, 18)$ e (EDI) corrisponde a (WDR). Il nostro messaggio in codice sarà quindi: KU&WDR. Per decifrare questo messaggio il ricevente deve conoscere: p , la biiezione f e la matrice E (o E^{-1} , e quindi n). Infatti per decifrare il primo gruppo (KU&) che corrisponde tramite f a $(11, 21, 28)$ dovrà risolvere, in F_{29} , il sistema lineare (di Cramer): $11 = x + 2z, 21 = y, 28 = 2z$, ecc... Questo procedimento è abbastanza sicuro perché, anche se il "nemico" conosce p ed f , avrà difficoltà nel trovare la matrice E (ma se il messaggio è lungo è probabile che un bravo decriptatore sia in grado di decifrarlo). Esistono metodi più sicuri (basati sul fatto che è praticamente impossibile fattorizzare numeri molto grandi) per costruire codici.

- (i) Perché lavorare in un campo finito (e non in \mathbb{R})? Perché E deve essere invertibile?
- (ii) Trovare una matrice E' tale che, con le notazioni precedenti, il deciframento di KU& sia DOM.
- (iii) E' possibile trovare E'' tale che il deciframento di KU&WDR sia DOMANI? Se si determinare una tale matrice.

17.7) Si consideri il sistema reale (S_λ) :

$$\begin{cases} \lambda x - y + (1-\lambda)z = 1 \\ 2x + \lambda y - z = \lambda \\ -x + (1+\lambda)y + z = 5 - \lambda \end{cases}$$

- (i) Determinare $\Lambda := \{\lambda \in \mathbf{R} / (S_\lambda) \text{ è compatibile}\}$.
(ii) Risolvere S_2, S_3 .

17.8) Sia il sistema $S(\lambda)$:

$$\begin{cases} \lambda x - y + (1-\lambda)z = 1 \\ x + (2+\lambda)y + z = \lambda \\ 2\lambda x - 2y + 3z = 2 \end{cases}$$

- (i) Determinare $S := \{\lambda \in \mathbf{R} / S(\lambda) \text{ è compatibile}\}$
(ii) Discutere i sistemi $S(-1), S(1)$.

17.9) Si consideri il sistema lineare, S , in \mathbf{C} :

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

(i) Senza risolvere il sistema dimostrare che S è compatibile per ogni λ in \mathbf{C} .

(ii) Si aggiunge una terza equazione al sistema S :

$$\begin{cases} \lambda x + y - 3\lambda z = 1 \\ x - \lambda y + 2\lambda z = 2 \\ ax + by + cz = d \end{cases}$$

dove a, b, c, d sono numeri reali (indipendenti da λ). È possibile scegliere a, b, c, d in modo che il nuovo sistema sia: (a) compatibile per ogni λ in \mathbf{C} , (b) compatibile per ogni λ in \mathbf{R} ?

17.10) Si consideri il sistema lineare, con parametro reale λ , $S(\lambda)$:

$$\begin{cases} \lambda x + y + \lambda z + t = 1 \\ -2x + \lambda y + 2z + 2t = 0 \\ x - 2y + z - \lambda t = 0 \end{cases}$$

- (i) Determinare i valori di λ per i quali il sistema è compatibile.
(ii) Determinare, per ogni λ tale che $S(\lambda)$ sia compatibile, la dimensione dell'insieme delle soluzioni corrispondente.
(iii) Risolvere $S(1)$.

1) Sottospazi affini di uno spazio vettoriale.

L'algebra lineare comprende la geometria dei sottospazi vettoriali di un k-spazio vettoriale (per esempio k^n) ovvero la geometria delle "sottovarietà lineari che passano per l'origine". La geometria delle sottovarietà lineari qualsiasi (non passanti necessariamente per l'origine) è l'oggetto della geometria affine. Una sottovarietà lineare qualsiasi si ottiene, per traslazione, da una sottovarietà lineare passante per l'origine.

1: Definizione: Siano E un k-spazio vettoriale e a un elemento di E. La traslazione di vettore a è l'applicazione: $t_a: E \rightarrow E : x \rightarrow x+a$.

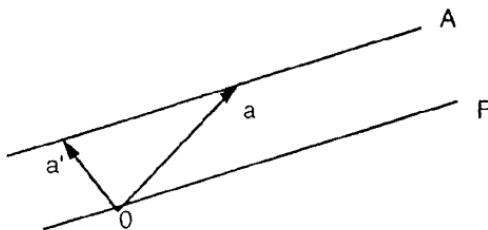
1.1: Osservazione : La traslazione non è un'applicazione lineare se $a \neq 0$. Infatti $t_a(0) = a$. Osserviamo che t_a è biiettiva ($t_a^{-1} = t_{-a}$).

2: Definizione : Sia E un k-spazio vettoriale. Si chiama sottospazio affine di E ogni sottoinsieme di E dedotto per traslazione da un sottospazio vettoriale di E. In altri termini $A \subseteq E$ è un sottospazio affine di E se esistono $a \in E$ e $F \subseteq E$, F sottospazio vettoriale tali che: $A = a + F := \{x \in E / \exists v \in F, x = a + v\}$.

1.1: Esempio : Sia $E = \mathbb{R}^2$ con la sua struttura abituale di R-spazio vettoriale. Un sottospazio vettoriale di dimensione uno di E è una retta, F, passante per il vettore nullo ($= (0,0)$). Un sottospazio affine $A = a + F$ non è altro che una retta, parallela a F e passante per il punto a.

Osserviamo che F ha un'equazione della forma: $\lambda x + \mu y = 0$ (cfr II, §7). Infatti F è un iperpiano di E e quindi esiste una forma lineare $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $F = \text{Ker}(f)$. Nelle basi canoniche f sarà rappresentata dalla matrice (λ, μ) . Quindi $F = \{(x,y) \in \mathbb{R}^2 / \lambda x + \mu y = 0\}$. Sia adesso $a = (\alpha, \beta)$. Il sottospazio affine $A = a + F$ è uguale a $\{(\alpha+x, \beta+y) / (x,y) \in \mathbb{R}^2 \text{ e } \lambda x + \mu y = 0\}$. Abbiamo $\lambda(x+\alpha) + \mu(y+\beta) = \lambda\alpha + \mu\beta$. Poniamo: $\lambda\alpha + \mu\beta = \delta$. Allora $A = \{(x,y) \in \mathbb{R}^2 / \lambda x + \mu y = \delta\}$. Si ritrova così che $\lambda x + \mu y = \delta$ è l'equazione della retta parallela a F e passante per il punto a.

E' chiaro che A non determina univocamente il punto a. Per esempio abbiamo anche $A = a' + F$.



D'altra parte si ha però:

3: Lemma : Sia E un k -spazio vettoriale, F, F' due sotto spazi vettoriali di E ; a, a' due elementi di E . Se $a+F = a'+F'$ allora $F = F'$.

Dim: E' chiaro ($a = a+0$) che $a \in a+F$. Quindi $a \in a'+F'$, perciò $a = a'+v'$, $v' \in F'$. Sia $v \in F$, abbiamo $a+v = a'+v'+v$. Ma $a+v \in a+F = a'+F'$. Pertanto $a'+v'+v = a'+w'$ per qualche $w' \in F'$. Da quest'ultima relazione segue $v = w'-v'$ e quindi $v \in F'$. Questo dimostra $F \subseteq F'$. Nello stesso modo si dimostra $F' \subseteq F$ ♦

Se riprendiamo l'uguaglianza $A = a+F$ vediamo che il sottospazio vettoriale F è determinato univocamente da A . Quindi ad ogni sottospazio affine, A , è associato un unico e ben determinato sottospazio vettoriale, F , di E .

4: Definizione: Con le notazioni precedenti, F si chiama la direzione (o giacitura) di A . La dimensione del sottospazio affine A è, per definizione, la dimensione del k -spazio vettoriale F . Noteremo $F = \text{Dir}(A)$.

4.1: Osservazione : (i) Il fatto che la direzione di un sottospazio affine sia unica riflette bene il fatto che i sottospazi affini sono i traslati dei sottospazi vettoriali.
(ii) Se F è una retta (risp. un piano, iperpiano) vettoriale, si dice che A è una retta (risp. piano, iperpiano) affine. I sottospazi affini di dimensione zero vengono chiamati punti di E (sono i vettori, elementi di E).

Tornando all'equazione $A = a+F$, abbiamo già detto che il punto a non è determinato da A , più precisamente abbiamo:

5: Lemma: Sia A un sottospazio affine del k -spazio vettoriale E . Se F è la direzione di A , per ogni punto $b \in A$ abbiamo: $A = b+F$.

Dim: Si tratta di dimostrare che se a, b sono due punti di A allora $a+F = b+F$. Siccome $b \in A$ abbiamo $b = a+v$ per qualche v in F , in particolare $a-b = -v$

appartiene ad F . Sia $x = a+w$, $w \in F$. Allora $x = b + (a-b) + w = b+v'$ con $v' = (a-b)+w \in F$. Questo dimostra $a+F \subseteq b+F$. Nello stesso modo si dimostra $b+F \subseteq a+F$ ♦

5.1: Osservazione : Si può riformulare il lemma precedente nel modo seguente: se A è un sottospazio affine di direzione F allora $A = t_a(F)$ per ogni $a \in A$.

6: Proposizione : Sia $(A_i)_{i \in I}$ una famiglia di sottospazi affini di E e $A = \cap_{i \in I} A_i$. Se $A \neq \emptyset$ allora A è un sottospazio affine di direzione $\cap_{i \in I} F_i$ dove, per ogni i in I , F_i è la direzione di A_i .

Dim: Supponiamo $A \neq \emptyset$ e sia $a \in A$. Allora $a \in A_i$ e $A_i = t_a(F_i)$, per ogni i in I . Quindi $A = \cap_{i \in I} t_a(F_i)$ e siccome t_a è iniettiva (cfr 1.1), $\cap_{i \in I} t_a(F_i) = t_a(\cap_{i \in I} F_i)$. Finalmente $\cap_{i \in I} F_i$ è un sottospazio vettoriale di E (II.§1, 7.3), e la proposizione è dimostrata♦

7: Definizione: Sia E un k -spazio vettoriale e siano v_1, \dots, v_k degli elementi di E . Il sottospazio affine generato da v_1, \dots, v_k , si nota $[v_1, \dots, v_k]$ ed è, per definizione, l'intersezione di tutti i sottospazi affini contenenti v_1, \dots, v_k .

Più generalmente se A è un sottinsieme di E il sottospazio affine generato da A , $[A]$, è il più piccolo sottospazio affine di E contenente A .

7.1: Osservazione : (i) In virtù della Prop. 6, $[v_1, \dots, v_k]$ è ben definito; è il più piccolo (per l'inclusione) sottospazio affine contenente v_1, \dots, v_k .

(ii) Non bisogna confondere $[v_1, \dots, v_k]$ con $\langle v_1, \dots, v_k \rangle$, il sottospazio vettoriale generato dai vettori v_1, \dots, v_k . Per esempio, se $k = 1$: $[v_1] = \{v_1\}$ mentre $\langle v_1 \rangle = \{\lambda v_1 / \lambda \in k\}$; quindi se $v_1 \neq 0$, $[v_1] \neq \langle v_1 \rangle$ ($0 \in \langle v_1 \rangle$ ma $0 \notin [v_1]$).

8: Lemma: Con le notazioni precedenti: per ogni i , $1 \leq i \leq k$, $[v_1, \dots, v_k] = v_i + \langle v_1 - v_i, \dots, v_{i-1} - v_i, v_{i+1} - v_i, \dots, v_k - v_i \rangle$.

Dim: Poniamo $A_i := v_i + \langle v_1 - v_i, \dots, v_{i-1} - v_i, v_{i+1} - v_i, \dots, v_k - v_i \rangle$. Se $j \neq i$, $v_j = v_i + (v_j - v_i)$ mentre $v_i = v_i + 0$ quindi A_i contiene v_1, \dots, v_k . Se A è un sottospazio affine contenente v_1, \dots, v_k allora A è della forma $v_i + F$ e $v_j = v_i + f_j$, $f_j \in F$. Quindi $v_j - v_i = f_j \in F$ e $\langle v_1 - v_i, \dots, v_{i-1} - v_i, v_{i+1} - v_i, \dots, v_k - v_i \rangle \subseteq F$, perciò $A_i \subseteq A$ ♦

8.1: Osservazione : Risulta in particolare che $A_1 = A_2 = \dots = A_k = [v_1, \dots, v_k]$.

9: Definizione: I punti v_1, \dots, v_k si dicono (affinamente) indipendenti se $\dim[v_1, \dots, v_k] = k-1$.

9.1: Osservazione : I punti v_1, \dots, v_k sono affinamente indipendenti se e solo se i $k-1$ vettori $v_1 - v_2, \dots, v_1 - v_k$ sono linearmente indipendenti.

10: Definizione: I punti v_1, \dots, v_k si dicono allineati (risp. complanari) se $\dim[v_1, \dots, v_k] \leq 1$ (risp. ≤ 2).

10.1: Esempi : Due punti v_0, v_1 sono indipendenti se e solo se sono distinti: in questo caso $[v_0, v_1]$ è una retta (l'unica retta passante per v_0, v_1 , cfr Es.1). Tre punti v_0, v_1, v_2 , sono indipendenti se e solo se generano un piano ossia se e solo se non sono allineati. Quattro punti sono indipendenti se e solo se non sono complanari ecc....

APPLICAZIONI AFFINI.

11: Definizione: Siano E, F due k -spazi vettoriali. Un'applicazione $u : E \rightarrow F$ è un'applicazione affine se esistono:

- un'applicazione lineare $v : E \rightarrow F$
- e una traslazione $t_a : F \rightarrow F$ ($a \in F$)

tali che $u = t_a \circ v$.

11.1: Esempio : Un'applicazione affine $u : k \rightarrow k$ è quindi della forma $u(x) = bx + a$ ($a, b \in k$).

11.2: Osservazione : Un'applicazione affine è lineare se e solo se $a = 0$. Infatti $u(0) = a$ quindi se $a \neq 0$, u non è lineare. D'altra parte se $a = 0$ allora $u = v$ è lineare; in particolare ogni applicazione lineare è un'applicazione affine.
Una traslazione $t_a : E \rightarrow E$ è un'applicazione affine; basta prendere $F = E$ e $v = \text{Id}_E$ nella definizione 11.

12: Lemma: La decomposizione di u in una traslazione e in un'applicazione lineare è unica.

Dim: Supponiamo $u(x) = v'(x) + a' = v(x) + a$. Abbiamo $u(0) = a' = a$ e $v'(x) = u(x) - a = v(x)$ per ogni x in E . Quindi $v = v'$ ♦

Il lemma 12 permette di dare la seguente:

13: Definizione: Con le notazioni precedenti l'applicazione lineare v si dice associata all'applicazione affine u e si noterà $v = L(u)$.

14: Proposizione: Siano E, F, G dei k -spazi vettoriali e $u : E \rightarrow F, u' : F \rightarrow G$ delle applicazioni affini.

(i) $u' \circ u$ è un'applicazione affine

(ii) Se v, v' sono le applicazioni lineari associate a u, u' allora $v' \circ v$ è l'applicazione lineare associata a $u' \circ u$.

Dim: Abbiamo $u : E \rightarrow F : x \rightarrow v(x)+a$ e $u' : F \rightarrow G : y \rightarrow v'(y)+a'$. Pertanto: $(u' \circ u)(x) = u'(v(x)+a) = v'(v(x)+a)+a' = v'(v(x))+v'(a)+a' = v' \circ v(x)+b$ dove $b = v'(a)+a' \in G$. Quindi $u' \circ u : E \rightarrow G$ è un'applicazione affine di applicazione lineare associata $v' \circ v$ ♦

15: Proposizione: Siano E, F due k -spazi vettoriali e $u : E \rightarrow F$ un'applicazione affine, di applicazione lineare associata $L(u) = v$. Se $A \subseteq E$ è un sottospazio affine di direzione D allora $u(A) \subseteq F$ è un sottospazio affine di direzione $v(D)$.

Dim: Abbiamo $A = a+D$ per qualche a in A . Siccome v è lineare $v(A) = v(a+D) = v(a)+v(D)$ e inoltre $v(D)$ è un sottospazio vettoriale di F . Pertanto $v(A)$ è un sottospazio affine, di direzione $v(D)$ di F . Siccome $u(A)$ si deduce da $v(A)$ per traslazione, anche $u(A)$ è un sottospazio affine di direzione $v(D)$ di F ♦

16: Proposizione: Siano E, F due k -spazi vettoriali e $u : E \rightarrow F$ un'applicazione affine. Allora u è iniettiva (risp. suriettiva, biiettiva) se e solo se $L(u)$ è iniettiva (risp. suriettiva, biiettiva).

Dim: Abbiamo $u = t \circ L(u)$ dove t è una traslazione. La proposizione segue dal fatto che t è una biiezione (cfr 1.1) ♦

17: Teorema: Sia E un k -spazio vettoriale. L'insieme delle applicazioni affini biiettive da E in E è un gruppo (per la composizione delle applicazioni).

Dim: Sia $G_A(E)$ l'insieme delle applicazioni affini biiettive da E in E . Se u, u' sono elementi di $G_A(E)$ allora anche $u \circ u'$ appartiene a $G_A(E)$ (cfr 14, 16). E' chiaro che $Id_E \in G_A(E)$ (è lineare; cfr 11.2). La composizione delle applicazioni è associativa, quindi rimane da dimostrare che se $u \in G_A(E)$ allora anche u^{-1} appartiene a $G_A(E)$. Per ipotesi, per ogni x in E , $u(x) = v(x)+a$ dove $v : E \rightarrow E$ è

lineare e dove $a \in E$. Da (16), v è biiettiva. Abbiamo $x = u^{-1}(y) \Leftrightarrow u(x) = y \Leftrightarrow v(x) + a = y$. Ma v essendo un isomorfismo lineare, da $v(x) = y - a$ si ricava: $x = v^{-1}(y - a) = v^{-1}(y) - v^{-1}(a)$. In conclusione $u^{-1}(y) = v^{-1}(y) - v^{-1}(a)$, quindi u^{-1} è un'applicazione affine♦

18: Definizione: Il gruppo delle applicazioni affini biiettive da E in E si nota $G_A(E)$ (o $\text{aff}(E)$) e si chiama il gruppo affine di E . Gli elementi di $G_A(E)$ si chiamano affinità di E .

18.1: Osservazione : La geometria affine è lo studio delle proprietà geometriche invarianti sotto $G_A(E)$, più precisamente:

- due sottinsiemi X, X' di E sono affinamente equivalenti se esiste un'affinità di E che trasforma X in X' : $\exists u \in G_A(E)$ tale che $u(X) = X'$.
- una proprietà affine di un sottinsieme, X , di E è una proprietà comune a tutti i sottinsiemi affinamente equivalenti a X (i.e. comune a tutti i sottinsiemi di E della forma $u(X)$ dove u percorre $G_A(E)$).

Per quanto riguarda i sottinsiemi di E che sono sotto spazi affini abbiamo un criterio "dimensionale" per stabilire se sono o meno affinamente equivalenti:

19: Proposizione: Siano A, A' due sottospazi affini di E ; A e A' sono affinamente equivalenti se e solo se $\dim(A) = \dim(A')$.

Dim: Abbiamo $A = a + F$ e $A' = a' + F'$.

(i) supponiamo $\dim(A) = \dim(A')$, cioè (cfr 4) $\dim(F) = \dim(F')$. Esiste un isomorfismo lineare $v: E \rightarrow E$ tale che $v(F) = F'$. Infatti siano $B = (e_1, \dots, e_r)$ e $B' = (e'_1, \dots, e'_r)$ delle basi di F, F' ($r := \dim(F) = \dim(F')$). Completiamo B, B' a basi B, B' di E : $B = (e_1, \dots, e_r, \dots, e_n)$, $B' = (e'_1, \dots, e'_r, \dots, e'_n)$, e definiamo un'applicazione lineare $v: E \rightarrow E$ tramite $v(e_i) = e'_i$, $1 \leq i \leq n$. Siccome $(v(e_1), \dots, v(e_n)) = B'$ è una base di E , v è un isomorfismo (cfr II.§4, 16). Da (11.2, 15) risulta che $v(A)$ è un sottospazio affine di E di direzione $F' = v(F)$; quindi $v(A) = b + F'$. Sia $t_{a'-b}$ la traslazione di vettore $a' - b$. Allora $t_{a'-b}(v(A)) = a' + F' = A'$. Finalmente, siccome v è un isomorfismo, $u = t_{a'-b} \circ v$ è un'affinità.

(ii) supponiamo che esista un'affinità $u: E \rightarrow E$ tale che $u(A) = A'$. Abbiamo $u = t_{a'-b} \circ v$ dove $v: E \rightarrow E$ è un isomorfismo lineare (11, 18). Inoltre $v(F) = F'$ (ipotesi e 15). Gli (sotto) spazi vettoriali F, F' sono isomorfi quindi $\dim(F) = \dim(F')$; ossia $\dim(A) = \dim(A')$ ♦

20: Osservazione : (i) La dimensione è quindi una proprietà affine dei sottospazi affini di E.

(ii) Possiamo già individuare due sotto gruppi notevoli di $G_A(E)$:

- il gruppo lineare $GL(E)$ degli automorfismi lineari di E.

- il gruppo, T, delle traslazioni di E. E' chiaro infatti che l'insieme, T, delle traslazioni di E è un gruppo per la composizione delle applicazioni ($t_a \circ t_b = t_{a+b}$, $Id_E = t_0$, $(t_a)^{-1} = t_{-a}$).

Osserviamo inoltre che l'applicazione $L: G_A(E) \rightarrow GL(E)$ che all'affinità $u = t \circ v$ associa l'isomorfismo lineare $v = L(u)$, è un morfismo di gruppi (cfr 14 (ii)).

Vediamo adesso come dare le "equazioni" di un'affinità o, più generalmente di un'applicazione affine $u: E \rightarrow F$.

21: Proposizione: Siano E, F due k-spazi vettoriali di basi $B = (e_1, \dots, e_n)$, $C = (f_1, \dots, f_p)$ e $u: E \rightarrow F$ un'applicazione affine, di applicazione lineare associata v. Sia $M = \text{mat}(v; B, C)$. Inoltre siano $(\alpha_1, \dots, \alpha_p)$ le coordinate di $u(0)$ nella base C. Se $v = x_1e_1 + \dots + x_ne_n$, sia X il vettore colonna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e Y il vettore colonna fatto sulle coordinate di $u(v)$. Allora: $Y = MX + T$ dove T è il vettore colonna $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}$.

Dim: Abbiamo $u = t_f \circ v$ per qualche f in F. Quindi $u(0) = v(0) + f = f$ e pertanto $u(x) = v(x) + f = v(x) + u(0)$. La proposizione risulta allora dalla scrittura matriciale delle applicazioni lineari♦

21.1: Osservazione : Se $u(x) = y_1f_1 + \dots + y_pf_p$ e se $M = (\alpha_{ij})$, $1 \leq i \leq p$, $1 \leq j \leq n$, allora la relazione $Y = MX + T$ è equivalente alle relazioni:

$$y_1 = x_1\alpha_{11} + \dots + x_n\alpha_{1n} + \alpha_1$$

.....

$$y_p = x_1\alpha_{p1} + \dots + x_n\alpha_{pn} + \alpha_p.$$

Viceversa è chiaro che, dopo la scelta di basi, relazioni di questo tipo definiscono un'applicazione affine da E in F.

Esercizi:

1.1) Siano E un k-spazio vettoriale e v, u due elementi di E ($v \neq u$). Dimostrare che esiste una ed un'unica retta affine passante per u e v.

1.2) Siano E, F due k -spazi vettoriali, $A \subseteq F$ un sottospazio affine, e $u : E \rightarrow F$ un'applicazione affine. Dimostrare che $u^{-1}(A)$ è un sottospazio affine di E .

1.3) Siano E, F due k -spazi vettoriali. Si ricorda che $j : E \rightarrow F$ è un'applicazione affine se $j = t_f \circ u$ dove $u : E \rightarrow F$ è un'applicazione lineare e $t_f : F \rightarrow F$ è la traslazione di vettore $f \in F$.

(i) Sia $u : E \rightarrow F$ un'applicazione lineare e $t_e : E \rightarrow E$ la traslazione di vettore $e \in E$. Mostrare che $f = u \circ t_e$ è un'applicazione affine da E in F . Sono u ed e univocamente determinati da f ?

(ii) Mostrare che ogni applicazione della forma $t_f \circ u \circ t_e : E \rightarrow F$ è affine e che ogni applicazione affine $j : E \rightarrow F$ si può scrivere nella forma $t_f \circ u \circ t_e$ con opportuni $f \in F$, $e \in E$, $u \in L(E, F)$. E' questa scrittura unica?

(iii) Mostrare che esistono delle applicazioni affini $j : E \rightarrow F$ che non si scrivono nella forma $j = u \circ t_e$.

(iv) Mostrare che ogni affinità $j : E \rightarrow E$ si scrive in modo unico $j = u \circ t_e$.

1.4) ("Il teorema fondamentale sulle affinità del piano"). Siano P, P', P'', Q, Q', Q'' dei punti di \mathbb{R}^2 . Si suppone che P, P', P'' non siano allineati e che pure Q, Q', Q'' non siano allineati. Dimostrare che esiste una ed un'unica affinità $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tale che $f(P) = Q$, $f(P') = Q'$, $f(Q'') = P''$.

Che cosa si può dire se P, P', P'' (risp. Q, Q', Q'') sono allineati?

1.5) Sia E un \mathbf{R} -spazio vettoriale e $g : E \rightarrow E$ un'applicazione affine. Un punto $p \in E$ si dirà fisso per g se $g(p) = p$.

(i) Sia $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3 : (x, y, z) \rightarrow (x+y-z+2, -x+y+z-2, x-y+z)$. Dimostrare che f è un'affinità e determinare l'insieme dei suoi punti fissi.

(ii) Sia $g : E \rightarrow E$ un'applicazione affine. Dimostrare che l'insieme dei suoi punti fissi è un sottospazio affine di E .

(iii) Dimostrare che g ha un unico punto fisso se e solo se 1 non è un autovalore dell'applicazione lineare associata a g .

1.6) Sia E un k -spazio vettoriale e $G_A(E)$ il gruppo delle affinità di E . Sia $T \subseteq G_A(E)$ il sottogruppo delle traslazioni; inoltre se $x \in E$ si nota $F_x = \{f \in G_A(E) / f(x) = x\}$.

(i) Dimostrare che F_x è un sottogruppo di $G_A(E)$.

(ii) Mostrare $F_x \cap T = \{\text{Id}\}$, e che: $\forall t \in T, \forall g \in G_A(E), g^{-1} \circ t \circ g \in T$.

(iii) E' vero che: $\forall f \in F_x, \forall g \in G_A(E), g^{-1} \circ f \circ g \in F_x$?

(iv) Mostrare che: $\forall f \in F_x, \forall g \in G_A(E), g^{-1} \circ f \circ g$ ha un punto fisso.

1.7) Siano E un \mathbf{R} -spazio vettoriale di dimensione tre, $\alpha \in \mathbf{R}$, e $v \in E, v \neq 0$. Si pone $v(\alpha) = \{f \in \text{End}(E) / f(v) = \alpha v\}$. Dimostrare che $v(\alpha)$ è un sottospazio affine di $\text{End}(E)$. Determinare $\dim(v(\alpha))$. (sugg: scegliere una base (e_1, e_2, e_3) di E con $e_1 = v$).

1.8) Sia $f : E \rightarrow F$ un'applicazione lineare suriettiva tra due k -spazi vettoriali di dimensione finita con $\dim_k E > \dim_k F$. Siano $x, y, x \neq y$, due vettori non nulli di F , linearmente dipendenti.

(i) E' vero che se u e v sono vettori di E tali che $f(u) = x, f(v) = y$ allora u e v sono linearmente dipendenti?

(ii) Mostrare che esistono sempre u_0 e v_0 in E tali che u_0 e v_0 siano linearmente indipendenti, e $f(u_0) = x, f(v_0) = y$.

1.9) Sia $k = \mathbb{Z}/3\mathbb{Z}$ il campo con tre elementi e $E = k^2$. Quante sono le rette affini di E ?

1.10) In \mathbb{R}^2 siano i tre punti $M = (x_1, y_1), M' = (x_2, y_2), M'' = (x_3, y_3)$. Dimostrare che

$$M, M', M'' \text{ sono allineati se e solo se } D = 0 \text{ dove } D = \det \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix}.$$

2) Equazioni dei sottospazi affini di uno spazio vettoriale.

La proposizione 1 di questo paragrafo mette in evidenza la corrispondenza tra la teoria "geometrica" dei sottospazi affini e la teoria "algebrica" dei sistemi lineari (cfr anche Oss. 1.1(iv)). Questa corrispondenza diventerà più evidente nel paragrafo successivo (parallelismo e incidenze).

1: Proposizione: Siano E un k -spazio vettoriale di dimensione n e $A \subseteq E$ un sottospazio affine di dimensione p . Allora esiste un sistema lineare in E , di $n-p$ equazioni e di rango $n-p$, il cui insieme delle soluzioni è A .

Dim: Sia $A = a+F$ e consideriamo l'ortogonale di F : $F^\circ \subseteq E^*$ (cfr II.§7, Def.14). Abbiamo $\dim(F^\circ) = n-p$ (II.§7, 15). Sia $(\varphi_1, \dots, \varphi_{n-p})$ una base di F° e poniamo $\varphi_1(a) = \beta_1, \dots, \varphi_{n-p}(a) = \beta_{n-p}$. Sia (S) il sistema lineare: $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}\}$. Siccome $(\varphi_1, \dots, \varphi_{n-p})$ è una base di F° , il sistema ha rango $n-p$ (rango delle righe della matrice dei coefficienti). Una soluzione particolare del sistema è ovviamente $v = a$. Le soluzioni di (S) si ottengono sommando ad a le soluzioni del sistema omogeneo associato: (S_0) : $\{\varphi_1(v) = 0, \dots, \varphi_{n-p}(v) = 0\}$ (cfr II.§17, 3). L'insieme delle soluzioni di (S_0) non è altro che $\langle \varphi_1, \dots, \varphi_{n-p} \rangle^\circ = (F^\circ)^\circ = F$ (cfr II.§8, 2). In conclusione l'insieme delle soluzioni di (S) è: $a+F = A \blacklozenge$

1.1: Osservazione : (i) Con le notazioni della dimostrazione precedente si dice che $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}\}$, è un sistema di equazioni (cartesiane) di A .

(ii) Possiamo riassumere quanto fatto nel modo seguente: per trovare delle equazioni (cartesiane) di $A = a+F$ basta prendere una base $(\varphi_1, \dots, \varphi_{n-p})$ di F° e considerare il sistema lineare $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}\}$ dove $\varphi_i(a) = \beta_i, 1 \leq i \leq n-p$.

(iii) E' chiaro che uno stesso sottospazio affine, A , può essere definito da sistemi lineari diversi (ma equivalenti!): per esempio basta considerare un'altra base di F° . Inoltre si possono anche aggiungere altre equazioni; per esempio, con le notazioni precedenti, sia $\phi = \alpha_1\varphi_1 + \dots + \alpha_{n-p}\varphi_{n-p}, \alpha_i \in k$, una combinazione lineare di $\varphi_1, \dots, \varphi_{n-p}$. Poniamo $\beta = \alpha_1\beta_1 + \dots + \alpha_{n-p}\beta_{n-p}$, allora l'insieme delle soluzioni del sistema $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}, \phi(v) = \beta\}$ è sempre A .

(iv) Viceversa, se prendiamo un sistema lineare in E: $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}\}$, allora, se (S) è compatibile, l'insieme delle soluzioni di (S) è un sottospazio affine di E di dimensione $n-r$, dove r è il rango del sistema (S).

EQUAZIONI CARTESIANE (IN COORDINATE) DI UN SOTTOSPAZIO.

Sia $B = (e_1, \dots, e_n)$ una base di E, e $B^* = (e_1^*, \dots, e_n^*)$ la base duale. Se $\varphi_t \in E^*$, possiamo scrivere φ_t nella base B^* : $\varphi_t = \lambda_{t1}e_1^* + \dots + \lambda_{tn}e_n^*$, $1 \leq t \leq n-p$. Se $v = x_1e_1 + \dots + x_ne_n$ è un vettore di E allora $\varphi_t(v) = \lambda_{t1}x_1 + \dots + \lambda_{tn}x_n$, e possiamo riscrivere il sistema (S): $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-p}(v) = \beta_{n-p}\}$ in coordinate:

$$\lambda_{11}x_1 + \dots + \lambda_{1n}x_n = \beta_1$$

$$\lambda_{21}x_1 + \dots + \lambda_{2n}x_n = \beta_2$$

$$\dots$$

$$\lambda_{n-p,1}x_1 + \dots + \lambda_{n-p,n}x_n = \beta_{n-p}$$

Abbiamo $A = \{v = x_1e_1 + \dots + x_ne_n / (x_1, \dots, x_n) \text{ è soluzione del sistema (S)}\}$; il sistema (S) è un sistema di equazioni cartesiane di A (rispetto alle basi B, B^*).

2: Esempio : L'insieme dei vettori $v = x_1e_1 + \dots + x_ne_n$ le cui coordinate (x_1, \dots, x_n) nella base B soddisfano l'equazione $\alpha_1x_1 + \dots + \alpha_nx_n = \beta$ ($\alpha_1, \dots, \alpha_n$ scalari non tutti nulli) formano un iperpiano affine, H, di E. Si dice che $\alpha_1X_1 + \dots + \alpha_nX_n = \beta$ è un'equazione cartesiana di H (qui X_1, \dots, X_n sono variabili). Nelle stesse condizioni sia $\alpha'_1X_1 + \dots + \alpha'_nX_n = \beta'$ un'equazione dell'iperpiano affine H'. Allora $H = H'$ se e solo se esiste $\lambda \in k$, $\lambda \neq 0$, tale che: $\lambda\alpha_i = \alpha'_i$, e $\lambda\beta = \beta'$. E' chiaro che se questa condizione è soddisfatta allora $H = H'$. Viceversa se $H = H'$ e se F è la direzione di H allora $\dim(F^\circ) = 1$. Le forme lineari $f = \alpha_1e_1^* + \dots + \alpha_ne_n^*$, $f' = \alpha'_1e_1^* + \dots + \alpha'_ne_n^*$ appartengono a F° e quindi sono proporzionali: $\lambda f = f'$ per qualche $\lambda \in k$ non nullo. Inoltre se a è un punto di H: $\beta' = f'(a) = \lambda f(a) = \lambda\beta$.

Useremo spesso in seguito il seguente risultato:

3: Lemma: Siano (S): $\{\varphi_1(v) = \beta_1, \dots, \varphi_r(v) = \beta_r\}$ un sistema di equazioni cartesiane del sottospazio affine A e (S'): $\{\psi_1(v) = \delta_1, \dots, \psi_u(v) = \delta_u\}$ un sistema di equazioni cartesiane del sottospazio affine A'. Allora $A \cap A'$ è l'insieme (eventualmente vuoto) delle soluzioni del sistema lineare (S)+(S'): $\{\varphi_1(v) = \beta_1, \dots, \varphi_r(v) = \beta_r; \psi_1(v) = \delta_1, \dots, \psi_u(v) = \delta_u\}$.

Dim: Se $w \in A \cap A'$ allora w è soluzione di (S) ($w \in A$) e di (S') ($w \in A'$), quindi w è anche soluzione di $(S)+(S')$. Viceversa se w è soluzione di $(S)+(S')$ allora w è soluzione di (S) (quindi $w \in A$) e w è soluzione di (S') (quindi $w \in A'$); pertanto $w \in A \cap A'$ ♦

RAPPRESENTAZIONE PARAMETRICA DI UN SOTTOSPAZIO AFFINE.

Sia E un k -spazio vettoriale di base $B = (e_1, \dots, e_n)$. Sia $A = a + F$ un sottospazio affine di E , di dimensione p . Prendiamo una base (f_1, \dots, f_p) di F . Il sottospazio affine A è l'insieme dei vettori, v , di E della forma: $v = a + \lambda_1 f_1 + \dots + \lambda_p f_p$; $\lambda_1, \dots, \lambda_p$ degli scalari qualsiasi.

Esprimendo quanto fatto finora nelle coordinate rispetto alla base B si ottengono delle equazioni parametriche di A .

Se $a = \alpha_1 e_1 + \dots + \alpha_n e_n$, $f_1 = \beta_{11} e_1 + \dots + \beta_{n1} e_n$, $f_2 = \beta_{12} e_1 + \dots + \beta_{n2} e_n$, ..., $f_p = \beta_{1p} e_1 + \dots + \beta_{np} e_n$, allora A è l'insieme dei vettori le cui coordinate (x_1, \dots, x_n) nella base B sono della forma:

$$\begin{aligned} x_1 &= \alpha_1 + \lambda_1 \beta_{11} + \lambda_2 \beta_{12} + \dots + \lambda_p \beta_{1p}, \\ x_2 &= \alpha_2 + \lambda_1 \beta_{21} + \lambda_2 \beta_{22} + \dots + \lambda_p \beta_{2p}, \quad (\dagger) \\ &\dots \\ x_n &= \alpha_n + \lambda_1 \beta_{n1} + \lambda_2 \beta_{n2} + \dots + \lambda_p \beta_{np}. \end{aligned}$$

dove $\lambda_1, \dots, \lambda_p$ sono degli scalari qualsiasi.

4: Definizione: Con le notazioni precedenti le relazioni (\dagger) sono delle equazioni parametriche del sottospazio affine A ($\lambda_1, \dots, \lambda_p$ sono i "parametri").

4.1: Osservazione : E' chiaro che equazioni del tipo (\dagger) dipendono dalla scelta del punto a in A e dalla scelta di una base (f_1, \dots, f_p) di F (e anche dalla scelta della base B di E). Quindi uno stesso sottospazio affine può avere più rappresentazioni parametriche.

4.2: Esempio : Sia $E = \mathbb{R}^3$ e B la base canonica. Le equazioni:

$$x = x_0 + \lambda \alpha$$

$$y = y_0 + \lambda \beta$$

$$z = z_0 + \lambda \delta$$

sono delle equazioni parametriche della retta affine, A, che passa per il punto a = (x₀, y₀, z₀) e di direzione la retta vettoriale, D, generata dal vettore (α, β, δ): A = a + D.

Esercizi:

2.1) Determinare, secondo il parametro reale α, l'intersezione, P ∩ Q ∩ R, dei tre piani, P, Q, R, di \mathbb{R}^3 di equazioni: (P) $\alpha x - 2y + z = 1$; (Q) $x - 2\alpha y + z = -2$; (R) $x - 2y + \alpha z = 1$.

2.2) In \mathbb{R}^3 dare delle equazioni del sottospazio affine generato da P, Q, R dove.

- (i) P = (0, 0, 2), Q = (0, 1, 2), R = (0, 2, 2)
- (ii) P = (0, 0, 2), Q = (-1, 0, 2), R = (0, 2, 2).

2.3) In \mathbb{R}^3 sia D la retta di equazioni: $x+y+1=0$, $x-y+2z=0$, e R la retta di equazioni $x-5y+6z=0$, $5x-y+6z+1=0$.

- (i) Mostrare che $R \cap D = \emptyset$.

(ii) Dare delle equazioni parametriche di D e R.

(iii) Dare delle equazioni cartesiane e parametriche del sottospazio affine generato (cfr §1, Def.7) da R e D.

2.4) Si pone $S = \{M \in M_3(\mathbb{R}) / M \text{ è simmetrica}\}$, $A = \{M \in M_3(\mathbb{R}) / M \text{ è antisimmetrica}\}$.

(i) Mostrare che S e A sono dei sottospazi vettoriali di $M_3(\mathbb{R})$, calcolare la loro dimensione, e mostrare che $M_3(\mathbb{R}) = S \oplus A$ (cfr Es. II.5.3).

Sia $\mathfrak{M} \subseteq M_3(\mathbb{R})$ il sottospazio vettoriale delle matrici magiche (si rimanda all'esercizio II.10.4 per le notazioni e definizioni). Si definisce $\mathfrak{M}_+ = S \cap \mathfrak{M}$, $\mathfrak{M}_- = A \cap \mathfrak{M}$.

(ii) Dimostrare che \mathfrak{M}_+ , \mathfrak{M}_- sono dei sottospazi vettoriali di \mathfrak{M} , e che $\mathfrak{M} = \mathfrak{M}_- \oplus \mathfrak{M}_+$.

Risulta da Es. II.10.4 che $\mathfrak{M}_- = \{\alpha \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} / \alpha \in \mathbb{R}\}$.

Se $\alpha \in \mathbb{R}$ si pone $\mathfrak{M}_+(\alpha) := \{M \in \mathfrak{M}_+ / s(M) = \alpha\}$.

(iii) Determinare $\mathfrak{M}_+(0)$ (scrivere un sistema lineare, calcolarne il rango e osservare che

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \text{ appartiene a } \mathfrak{M}_+(0))$$

(iv) Mostrare che $\mathfrak{M}_+(\alpha)$ è un sottospazio affine di \mathfrak{M}_+ e determinare la sua direzione (scrivere un sistema lineare e guardarlo...).

(v) Concludere che \mathfrak{M} è uno spazio vettoriale di dimensione tre, di base $\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$,

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

(vi) Sia $M = (a_{ij})$ una matrice 3×3 . Notiamo R_i (risp. C_i, D_i) la somma degli elementi esima riga (risp. colonna, diagonale). La matrice M è magica se $R_i = R_1, 2 \leq i \leq 3; D_j = .$ $j = 1, 2; C_k = R_1, 1 \leq k \leq 3$. Si ottiene così un sistema lineare omogeneo di 7 equazioni in 9 incognite. Dedurne che $\dim(\mathfrak{M}) \geq 2$ (cfr Es.II.10.4). Calcolare il rango del sistema (si potrà usare il metodo della riduzione per righe, cfr Es.II.15.5) e ritrovare che $\dim(\mathfrak{M}) = 3$.

3) Parallelismo e incidenze

In questo paragrafo si prosegue lo studio, iniziato nel paragrafo precedente, della corrispondenza, tra la teoria, algebrica, dei sistemi lineari e quella, geometrica, dei sottospazi affini. A posteriori si osserverà come il linguaggio sintetico e "intuitivo" della geometria abbia le basi solide e "rigorose" dell'algebra.

1: Definizione: Due sottospazi affini, di dimensione > 0 , A e A' di un k -spazio vettoriale, E , sono paralleli se $D \subseteq D'$ o $D' \subseteq D$ dove D, D' sono le direzioni di A, A' . Se A e A' sono paralleli si scrive $A // A'$.

1.1: Osservazione : (i) La relazione $A // A'$ non è una relazione d'equivalenza sull'insieme dei sottospazi affini di dimensione > 0 di E . Infatti $A // A'$ e $A' // A''$ non implica $A // A''$ come si vede prendendo per A e A'' due rette incidenti di un piano A' . Vedremo però che il parallelismo è una relazione d'equivalenza sull'insieme dei sottospazi affini di dimensione fissata, $d > 0$.

(ii) Nel seguito quando si parlerà di sottospazi paralleli si assumerà sempre tacitamente che questi sottospazi abbiano dimensione > 0 .

2: Lemma: Siano A, A' due sottospazi affini di E . Se $A \subseteq A'$ allora $A // A'$.

Dim: Sia $A = a + F$ allora $a \in A'$ e possiamo scrivere (cfr §1, lemma 5) $A' = a + F'$. Sia v in F quindi $a + v \in A \subseteq A'$, pertanto $a + v = a + v'$ con $v' \in F'$. Quindi $v = v'$ appartiene a F' ; questo dimostra $F \subseteq F'$ ossia $A // A'$ ♦

3: Proposizione: Siano A, A' due sottospazi affini paralleli di E con $\dim(A) \leq \dim(A')$.

(i) Se A e A' hanno un punto in comune allora $A \subseteq A'$.

(ii) Se $\dim(A) = \dim(A')$ e se A ed A' hanno un punto in comune allora $A = A'$.

Dim: (i) Sia $a \in A \cap A'$, per (§1, 6) $A \cap A'$ è un sottospazio affine di direzione $F \cap F'$ (F, F' le direzioni di A, A'). Per ipotesi $F \subseteq F'$ (perchè $A // A'$ e $\dim(F) \leq \dim(F')$). Quindi $F \cap F' = F$ e $A \cap A' = a + F = A$ ossia $A \subseteq A'$.

(ii) Da (i) abbiamo $A \subseteq A'$ e $A' \subseteq A$ ♦

3.1: Osservazione : Si può dimostrare adesso che $//$ è una relazione d'equivalenza sull'insieme dei sottospazi affini di dimensione fissata $d > 0$.

4: Corollario: Sia $A \subseteq E$ un sottospazio affine e $x \in E$. Esiste uno ed un unico sottospazio affine $A' \subseteq E$ tale che: $A // A'$, $\dim(A) = \dim(A')$, $x \in A'$.

Dim: Sia $A = a+F$. Per l'esistenza basta considerare $A'=x+F$. L'unicità segue da 3(ii)♦

4.1: Osservazione : Identifichiamo $E = \mathbf{R}^2$ col "piano della geometria elementare". Il corollario 4 afferma che data una retta R nel piano e un punto x del piano, esiste una ed un'unica retta del piano passante per x e parallela a R . Questo non è altro che l'assioma delle parallele di Euclide. Gli assiomi usati finora (teoria degli insiemi, definizioni di k -spazi vettoriali, sottospazi affini) e l'identificazione del "piano della geometria elementare" con \mathbf{R}^2 (e delle "rette" con i sottospazi affini di dimensione uno) permettono quindi di dimostrare l'assioma delle parallele.

5: Lemma: Sia E un k -spazio vettoriale di dimensione n e siano A, A' due sottospazi affini di E di dimensioni rispettive t, s . Se $A \cap A' \neq \emptyset$, allora $\dim(A \cap A') \geq t+s-n$.

Dim: Sia $a \in A \cap A'$, allora (§1, 6): $A \cap A' = a + (F \cap F')$ dove F, F' sono le direzioni di A, A' . Per ipotesi $\dim(F) = t$, $\dim(F') = s$. Dalla relazione di Grassmann (cfr II.§5, 4): $\dim(F+F') = \dim(F) + \dim(F') - \dim(F \cap F')$, altrimenti detto: $\dim(F \cap F') = \dim(F) + \dim(F') - \dim(F+F') = t+s - \dim(F+F')$. Siccome $\dim(F+F') \leq n$ ($F+F' \subseteq E$), si deduce $\dim(F \cap F') \geq t+s-n$ ♦

5.1: Osservazione : (i) Se A e A' sono due sottospazi vettoriali allora $A \cap A' \neq \emptyset$ (contiene il vettore nullo) quindi $\dim(A \cap A') \geq \dim(A) + \dim(A') - n$ (un sottospazio vettoriale è anche un sottospazio affine).

(ii) Siano $H = a+F$, $H' = F'$ due sottospazi affini paralleli. Se $a \notin F$ allora $H \cap H' = \emptyset$; questo esempio mostra che l'ipotesi $A \cap A' \neq \emptyset$ nel lemma 5 è necessaria. Si può, aggiungendo ipotesi sulla posizione relativa delle direzioni, garantire che l'intersezione di due sottospazi affini sia non vuota.

6: Proposizione: Sia E un k -spazio vettoriale di dimensione n . Siano inoltre $A = a+F$, $A' = a'+F'$, due sottospazi affini di dimensioni t , s . Sono equivalenti:

- (i) $F+F' = E$
- (ii) $A \cap A' \neq \emptyset$ e $\dim(A \cap A') = s+t-n$.

Dim: (i) \Rightarrow (ii) Siccome $F+F' = E$, $\exists f \in F$ e $\exists f' \in F'$ tali che $a-a' = f+f'$. Quindi $a-f = a'+f' \in A \cap A'$. Se $b \in A \cap A'$ possiamo scrivere $A = b+F$, $A' = b+F'$ e $A \cap A' = b+(F \cap F')$ e si conclude con la relazione di Grassmann.

(ii) \Rightarrow (i) Come prima: se $b \in A \cap A'$ possiamo scrivere $A = b+F$, $A' = b+F'$ e $A \cap A' = b+(F \cap F')$, si conclude con la relazione di Grassmann visto che $\dim(F \cap F') = s+t-n$ per ipotesi \blacklozenge

7: Osservazione : L'ipotesi $F+F' = E$ implica $s+t-n \geq 0$.

8: Corollario: Sia H un iperpiano affine del k -spazio vettoriale E . Se $A \subseteq E$ è un sottospazio affine (di dimensione ≥ 1) non parallelo a H allora $H \cap A \neq \emptyset$ e $\dim(H \cap A) = \dim(A)-1$.

Dim: Sia $A = a+F$ e $H = h+F$. Siccome A e H non sono paralleli esiste $v \in F$ tale che $v \notin F'$. Pertanto la retta vettoriale $D := \langle v \rangle$ è un supplementare di F' : $E = F' \oplus D$. Dal lemma precedente: $\dim(H \cap A) = \dim(A) + (n-1)-n = \dim(A)-1$ \blacklozenge

Per concludere questo paragrafo consideriamo (dal punto di vista delle equazioni cartesiane e parametriche) le possibili incidenze tra due iperpiani; un iperpiano e una retta.

INTERSEZIONE DI DUE IPERPIANI.

Siano $H = h+F$, $H' = h'+F'$, due iperpiani affini di un k -spazio vettoriale, E , di dimensione n .

9: Lemma: Se $\{\varphi(v) = \beta\}, \{\varphi'(v) = \beta'\}$ sono delle equazioni cartesiane di H , H' allora:

- (i) $H = H'$ se e solo se esiste λ in k , $\lambda \neq 0$, tale che $\varphi = \lambda\varphi'$, $\beta = \lambda\beta'$.
- (ii) H e H' sono paralleli se e solo se esiste λ in k , $\lambda \neq 0$, tale che $\varphi = \lambda\varphi'$.
- (iii) H e H' non sono paralleli se e soltanto se φ e φ' sono linearmente indipendenti. Se φ e φ' sono linearmente indipendenti $H \cap H'$ è un

sottospazio affine di dimensione $n-2$, un sistema di equazioni cartesiane di $H \cap H'$ è $\{\varphi(v) = \beta, \varphi'(v) = \beta'\}$.

Dim: cfrr §2, esempio 2 e il corollario 8 qui sopra ♦

Dal punto di vista parametrico: sia (v_1, \dots, v_{n-1}) una base di F e (v'_1, \dots, v'_{n-1}) una base di F' allora:

10: Lemma: Sono equivalenti:

- (i) $H // H'$
- (ii) per ogni i , $1 \leq i \leq n-1$, $\det_B(v_1, \dots, v_{n-1}, v'_i) = 0$ (B una base di E).
- (iii) per ogni i , $1 \leq i \leq n-1$, $\det_B(v'_1, \dots, v'_{n-1}, v_i) = 0$ (B una base di E).

Dim: $H // H'$ se e solo se $F = \langle v_1, \dots, v_{n-1} \rangle = \langle v'_1, \dots, v'_{n-1} \rangle = F'$, quindi se e solo se ogni v'_i è combinazione lineare di v_1, \dots, v_{n-1} (o ogni v_i di v'_1, \dots, v'_{n-1}) ♦

Mostriamo adesso come si fa a passare da equazioni cartesiane ad equazioni parametriche e vice versa.

Equazioni degli iperpiani:
dal cartesiano al parametrico.

Sia $B = (e_1, \dots, e_n)$ una base di E , $f = \alpha_1 e_1^* + \dots + \alpha_n e_n^*$ e $H = \{v = x_1 e_1 + \dots + x_n e_n / \alpha_1 x_1 + \dots + \alpha_n x_n = \beta\}$. Per ottenere delle equazioni parametriche di H bisogna individuare una base di $F = \langle f \rangle$ e una soluzione particolare del sistema $f(v) = \beta$. Osserviamo che esiste k , $1 \leq k \leq n$, tale che $\alpha_k \neq 0$. Per ogni i , $i \neq k$, $1 \leq i \leq n$, definiamo $v_i = x^{(i)}_1 e_1 + \dots + x^{(i)}_n e_n$ nel modo seguente: $x^{(i)}_i = 1$, $x^{(i)}_k = (\beta - \alpha_i)/\alpha_k$, $x^{(i)}_j = 0$ se $j \neq i, k$. Se $\beta = 0$ questi $n-1$ vettori (linearmente indipendenti) formano una base di $\langle f \rangle$. Se $\beta \neq 0$ sia $v_k = \beta/\alpha_k e_k$ e $w_i = v_i - v_k$ ($i \neq k$). E' facile verificare che i w_i formano una base di $\langle f \rangle$. Osservare la differenza: se $\beta \neq 0$, gli n vettori v_i , $1 \leq i \leq n$, sono linearmente indipendenti ed appartengono ad H mentre se $\beta = 0$ non si possono trovare n vettori liberi in H (se $\beta = 0$, H è un sottospazio vettoriale di dimensione $n-1$).

Equazioni degli iperpiani:
dal parametrico al cartesiano.

Se $H = a + F$ e se (v_1, \dots, v_{n-1}) è una base di F per ottenere un'equazione cartesiana si scrive che $x \in H$ se e solo se $x-a \in F$ ossia se e solo se $x-a$ è combinazione lineare di v_1, \dots, v_{n-1} , o ancora se e solo se $\det_B(x-a, v_1, \dots, v_{n-1}) = 0$. In coordinate nella base B : $\det_B(x-a, v_1, \dots, v_{n-1}) =$

$$\begin{vmatrix} x_1 - a_1 & \alpha_1^{(1)} & \dots & \alpha_1^{(n-1)} \\ \vdots & \vdots & & \vdots \\ x_n - a_n & \alpha_n^{(1)} & \dots & \alpha_n^{(n-1)} \end{vmatrix} = 0$$

dove (x_1, \dots, x_n) sono le coordinate di x , (a_1, \dots, a_n) quelle di a e $(\alpha^{(i)}_1, \dots, \alpha^{(i)}_{n-1})$ quelle di v_i . Sviluppando secondo la prima colonna otteniamo un'equazione cartesiana di H : $(x_1 - \alpha_1)_1 \Delta_1 + \dots + (x_n - \alpha_n) \Delta_n = 0$.

INTERSEZIONE DI UNA RETTA E DI UN IPERPIANO.

Per iniziare facciamo vedere come si passa da equazioni cartesiane di una retta ad equazioni parametriche e viceversa.

Equazioni delle rette:
dal parametrico al cartesiano.

Sia $R = a + D$ una retta affine di E . Sia $u \neq 0$ un vettore di D (si dice che u è un vettore direttore di R), allora $R = \{x \in E / \exists \lambda \in k, x = a + \lambda u\}$. Sia $B = (e_1, \dots, e_n)$ una base di E . Per ottenere delle equazioni parametriche di R basta scrivere in coordinate la relazione $x = a + \lambda u$:

$x_1 = \alpha_1 + \lambda v_1, \dots, x_n = \alpha_n + \lambda v_n$, sono delle equazioni parametriche di R dove $(\alpha_1, \dots, \alpha_n)$, (v_1, \dots, v_n) sono le coordinate di a , u nella base B . Per ottenere delle equazioni cartesiane si scrive che $x \in R$ se e solo se i vettori $x-a$ e u sono legati; questo è equivalente a dire che la matrice $M(x) = \begin{pmatrix} x_1 - \alpha_1 & x_2 - \alpha_2 & \dots & x_n - \alpha_n \\ v_1 & v_2 & \dots & v_n \end{pmatrix}$ ha rango uno. Questo è ancora equivalente a

dire che tutti i minori $(2,2)$ di $M(x)$ sono nulli. Osserviamo che l'ipotesi $u \neq 0$ implica che esiste k tale che $v_k \neq 0$. Quindi v_k è un minore $(1,1)$ non nullo di $M(x)$ e $M(x)$ ha rango uno se e solo se gli $n-1$ orlati di v_k sono nulli: $(x_i - \alpha_i)v_k$

- $v_i(x_k - \alpha_k) = 0$, $1 \leq i \leq n$, $i \neq k$. Queste $n-1$ equazioni sono delle equazioni cartesiane della retta R .

Nella pratica si cercherà più semplicemente di eliminare il parametro. Per esempio se R è la retta di \mathbb{R}^3 di equazioni:

$x = 1 + 2t$, $y = 3 - t$, $z = 2 + 3t$, allora $t = 3+y$ da cui: $x = 7 + 2y$, $z = 11 + 3y$ sono delle equazioni cartesiane di R .

Equazioni delle rette:
dal cartesiano al parametrico.

Sia $f_1(v) = \beta_1, \dots, f_{n-1}(v) = \beta_{n-1}$, un sistema di equazioni cartesiane che definisce la retta R (quindi f_1, \dots, f_{n-1} linearmente indipendenti). Per ottenere delle equazioni parametriche dobbiamo innanzitutto determinare un vettore direttore di R ossia una base di $\langle f_1, \dots, f_{n-1} \rangle^\circ$. Questo è ancora equivalente a trovare una soluzione non nulla del sistema lineare omogeneo: $f_1(v) = 0, \dots, f_{n-1}(v) = 0$. Per questo useremo il seguente risultato generale sui sistemi omogenei di $n-1$ equazioni in n incognite:

11: Lemma: Sia S il sistema omogeneo in $n-1$ equazioni e n incognite:

$$\alpha_{11}x_1 + \dots + \alpha_{1n}x_n = 0$$

.....

$$\alpha_{n-1,1}x_1 + \dots + \alpha_{n-1,n}x_n = 0$$

Sia $A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n-1,1} & \dots & \alpha_{n-1,n} \end{pmatrix}$ la matrice $(n-1, n)$ dei coefficienti e A_i la matrice ottenuta da A togliendo la i -esima colonna. Finalmente sia $\Delta_i = (-1)^{i+1} \cdot \det(A_i)$. Se il sistema S ha rango $n-1$, $(\Delta_1, \dots, \Delta_n)$ è una soluzione non banale di S .

Dim: Per k , $1 \leq k \leq n$, consideriamo il determinante $\left| \begin{array}{ccc} \alpha_{k1} & \dots & \alpha_{kn} \\ \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n-1,1} & \dots & \alpha_{n-1,n} \end{array} \right|$. E'

il determinante della matrice (n, n) ottenuta aggiungendo ad A la sua k -esima riga. Questo determinante è nullo perché ha due righe uguali. Sviluppando secondo la prima riga: $\alpha_{k1}\Delta_1 + \alpha_{k2}\Delta_2 + \dots + \alpha_{kn}\Delta_n = 0$, $1 \leq k \leq n-1$. Quindi $(\Delta_1, \dots, \Delta_n)$ è soluzione di S . Inoltre siccome S ha rango $n-1$, la

matrice A ha un minore di ordine $n-1$ non nullo. Quindi esiste j , $1 \leq j \leq n$, tale che $\Delta_j \neq 0$ ♦

Tornando alla nostra retta affine, il lemma precedente fornisce un metodo per ricavare un vettore direttore a partire da un sistema di equazioni cartesiane. Rimane poi da trovare una soluzione particolare del sistema non omogeneo, per questo si daranno valori opportuni alle incognite libere.

Intersezione di un iperpiano e di una retta.

Siano H , R , un iperpiano e una retta affini di E . Sia $\{\varphi(v) = \beta\}$ un'equazione cartesiana di H e $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-1}(v) = \beta_{n-1}\}$ delle equazioni cartesiane di R .

12: Lemma: Con le notazioni precedenti, sono equivalenti:

- (i) $R // H$
- (ii) $\det_B(\varphi_1, \dots, \varphi_{n-1}, \varphi) = 0$ (B^* una base di E^*).

Dim: Se $H = h + F$, $R = a + D$ allora $R // H \Leftrightarrow D \subseteq F$. Ma (cfr §2, 1.1 (ii)) $D = \langle \varphi_1, \dots, \varphi_{n-1} \rangle^\circ$, $F = \langle \varphi \rangle^\circ$. Quindi: $R // H \Leftrightarrow \langle \varphi_1, \dots, \varphi_{n-1} \rangle^\circ \subseteq \langle \varphi \rangle^\circ \Leftrightarrow \langle \varphi \rangle \subseteq \langle \varphi_1, \dots, \varphi_{n-1} \rangle$ ("la dualità inverte le inclusioni") e quest'ultima condizione è equivalente a dire che φ è combinazione lineare di $\varphi_1, \dots, \varphi_{n-1}$; visto che $\varphi_1, \dots, \varphi_{n-1}$ sono liberi, questo è equivalente a dire che $\varphi_1, \dots, \varphi_{n-1}, \varphi$ sono linearmente dipendenti quindi che il loro determinante in ogni base è nullo ♦

12.1: Osservazione: Sia $B = (e_1, \dots, e_n)$ una base di E e B^* la base duale. Se $\varphi = \alpha_1 e_1^* + \dots + \alpha_n e_n^*$, $\varphi_i = \alpha_{i1} e_1^* + \dots + \alpha_{in} e_n^*$, $1 \leq i \leq n-1$, in modo che un'equazione in coordinate di H sia $\alpha_1 x_1 + \dots + \alpha_n x_n = \beta$; mentre equazioni in coordinate per R siano: $\alpha_{i1} x_1 + \dots + \alpha_{in} x_n = \beta_i$, $1 \leq i \leq n-1$, abbiamo: $R // H \Leftrightarrow \begin{vmatrix} \alpha_1 & \alpha_{11} & \dots & \alpha_{1,n-1} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_n & \alpha_{n1} & \dots & \alpha_{n,n-1} \end{vmatrix} = 0$.

Se $R // H$ abbiamo $R \cap H = \emptyset$ o $R \subseteq H$ (cfr Prop.3). Siamo nel secondo caso se e solo se il sistema: $\{\varphi_1(v) = \beta_1, \dots, \varphi_{n-1}(v) = \beta_{n-1}, \varphi(v) = \beta\}$ è compatibile.

13: Lemma: Se R e H non sono paralleli, R e H s'intersecano in punto, v_0 , soluzione del sistema di Cramer $S = \{\varphi_1(v) = \beta_1, \dots, \varphi_{n-1}(v) = \beta_{n-1}, \varphi(v) = \beta\}$.

Dim: Dal lemma 8, R e H s'intersecano in un punto, da (§2, lemma 3) questo punto è soluzione di S; perciò S è compatibile e quindi di Cramer♦

Passiamo adesso al punto di vista parametrico. Sia u un vettore direttore di R e (u_1, \dots, u_{n-1}) una base di F, la direzione di H. Abbiamo $R // H \Leftrightarrow u \in F \Leftrightarrow u, u_1, \dots, u_{n-1}$ sono legati $\Leftrightarrow \det_B(u, u_1, \dots, u_{n-1}) = 0$ (B una base di E).

Sia B = (e_1, \dots, e_n) una base di E, e $(v_1, \dots, v_n), (v_{1k}, \dots, v_{nk})$ le coordinate di u, v_k nella base B. Se $\Delta := \det_B(u, u_1, \dots, u_{n-1}) \neq 0$ sappiamo che R e H s'intersecano in un punto $v_0 = x_1 e_1 + \dots + x_n e_n$. Quindi esistono $\lambda, \lambda_1, \dots, \lambda_{n-1}$ in k tali che:

$$x_1 = \alpha_1 + \lambda v_1 = \beta_1 + \lambda_1 v_{11} + \dots + \lambda_{n-1} v_{1,n-1}$$

.....

$$x_n = \alpha_n + \lambda v_n = \beta_n + \lambda_1 v_{n1} + \dots + \lambda_{n-1} v_{n,n-1}$$

$((\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n)$ sono le coordinate di un punto di R, H). Otteniamo il sistema:

$$-\lambda v_1 + \lambda_1 v_{11} + \dots + \lambda_{n-1} v_{1,n-1} = \alpha_1 - \beta_1$$

.....

$$-\lambda v_n + \lambda_1 v_{n1} + \dots + \lambda_{n-1} v_{n,n-1} = \alpha_n - \beta_n$$

Le incognite sono $-\lambda, \lambda_1, \dots, \lambda_n$; il determinante della matrice dei coefficienti è $\Delta \neq 0$. Quindi il sistema è di Cramer e l'unica soluzione fornisce i valori dei parametri che determinano il punto $R \cap H$.

Esercizi:

3.1) In \mathbb{R}^2 dare delle equazioni cartesiane e parametriche della retta passante per $P = (1,1)$ e parallela alla retta vettoriale $R = \{(X, Y) \in \mathbb{R}^2 / 2X - Y = 0\}$.

3.2) Sia E un k-spazio vettoriale e R, D due rette affini distinte di E. Se R e D sono complanari dimostrare che esiste un unico piano affine contenente R e D (questo piano si chiama il piano generato da R e D).

3.3) (i) Risolvere a seconda del parametro reale λ il sistema lineare:

$$\lambda x - 2y + 4 = 0$$

$$x + (\lambda - 3)y + 6 - 2\lambda = 0$$

(ii) per ogni λ si nota $S(\lambda)$ l'insieme soluzione del sistema. Determinare il più piccolo sottospazio affine, F , di \mathbf{R}^2 , tale che: $\forall \lambda \in \mathbf{R}, S(\lambda) \subseteq F$.

3.4) Sia E un k -spazio vettoriale di dimensione n e H' , H'' due iperpiani affini di E , non paralleli, di equazioni $f(v) = b'$, $f''(v) = b''$. Si pone $T := H' \cap H''$.

(i) Qual è la dimensione di T ?

(ii) Dimostrare che un iper piano contiene T se e solo se ha un'equazione della forma $\lambda f(v) + \mu f''(v) = \lambda b' + \mu b''$. L'insieme degli iperpiani contenente T si chiama fascio di iperpiani di asse (o base) T .

3.5) Siano R , D due rette parallele distinte di \mathbf{R}^3 , descrivere le affinità $f: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tali che $f(R) = D$ (e' f necessariamente una traslazione?).

3.6) Esistono due piani affini sghembi in \mathbf{R}^4 ? (un piano affine $P \subseteq \mathbf{R}^4$ è un sottospazio affine di dimensione due; due piani P , P' sono sghembi se non sono paralleli e se non s'intersecano).

Più generalmente esistono due sottospazi affini, di codimensione due (i.e. di dimensione $n-2$), sghembi in \mathbf{R}^n ?

3.7) Si considerino le applicazioni $f_i: \mathbf{R}^3 \rightarrow \mathbf{R}$, $1 \leq i \leq 3$, così definite:

$$f_1(x, y, z) = x - 2y + z + 1, f_2(x, y, z) = -x + 2y + 3z + 2, f_3(x, y, z) = -2x + 4y + 18z - 1.$$

Sia $f: \mathbf{R}^3 \rightarrow \mathbf{R}^3$: $(x, y, z) \rightarrow (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z))$.

(i) Dimostrare che f è affine e trovare il rango di f .

Si consideri la retta L di equazioni: $f_1(x, y, z) = 0, f_2(x, y, z) = 0$.

(ii) Determinare i piani passanti per L e \parallel al piano H di equazione $f_3(x, y, z) = 0$.

(iii) Determinare i piani passanti per L ed incidenti il piano H .

(iv) Sia $X = \{(a, b, c) \in \mathbf{R}^3 / \text{il sistema: } f_1(x, y, z) = a, f_2(x, y, z) = b, f_3(x, y, z) = c, \text{ è compatibile}\}$. Dimostrare che X è affine e determinare la sua dimensione.

3.8) Siano H_1, H_2, H_3, H_4 quattro piani di \mathbf{R}^3 , a due a due non paralleli. Per $i \neq j$ si denoti L_{ij} la retta $H_i \cap H_j$. Dimostrare che sono equivalenti:

(i) L_{13} e L_{24} sono incidenti (risp. sghembe, parallele)

(ii) L_{12} e L_{34} sono incidenti (risp. sghembe, parallele)

(iii) L_{14} e L_{23} sono incidenti (risp. sghembe, parallele)

(suggerimento: considerare $H_1 \cap H_2 \cap H_3 \cap H_4$).

4) Incidenze in un piano ed in uno spazio tridimensionale affini.

PIANO AFFINE.

Sia E un k -spazio vettoriale di dimensione due. Diremo che E è un piano affine (su k). I sottospazi affini propri (i.e. diversi da E) di E sono i punti (dimensione zero) e le rette. Per quanto riguarda la posizione reciproca di due rette del piano affine basta riportarsi al §3 ("intersezione di due iperpiani" o "intersezione di un iperpiano e di una retta"). In particolare se R, D sono due rette affini di E , di equazioni cartesiane rispettive $\varphi(v) = \beta$, $\varphi'(v) = \beta'$ e se S è il sistema lineare $\{\varphi(v) = \beta, \varphi'(v) = \beta'\}$, ci sono solo tre casi possibili: $R // D$ (φ e φ' sono proporzionali, i.e. S non è di Cramer), e abbiamo due possibilità: $R = D$ (S è compatibile), $R \cap D = \emptyset$ (S è incompatibile); oppure R e D s'intersecano in un unico punto (S è di Cramer).

SPAZIO TRIDIMENSIONALE.

Sia adesso E un k -spazio vettoriale di dimensione tre. Diremo che E è un spazio affine tridimensionale (su k). I sottospazi affini propri, di dimensione > 0 , di E sono le rette ed i piani (in questo caso un piano è un iperpiano). Ci rimane da studiare l'intersezione di due rette.

1: Definizione: Due rette R, R' affini di E si dicono complanari se esiste un piano affine che le contiene entrambe. Le rette R, R' si dicono sghembe se R e R' non sono parallele e se $R \cap R' = \emptyset$.

1.1: Osservazione : Nel piano due rette non sono mai sghembe.

2: Lemma: Siano E un k -spazio vettoriale di dimensione n , R e R' due rette affini di E .

- (i) Se R e R' sono complanari e non parallele la loro intersezione consta di un ed un solo punto.
- (ii) Se R e R' s'incontrano in un punto allora sono complanari.

Dim: (i) Sia $R = a + D$. Siccome per ipotesi R è contenuta nel piano affine H , a $\in H$ e possiamo scrivere $H = a + F$. Sia $R' = a' + D'$. L'ipotesi $R' \subseteq H$ implica $D' \subseteq F$ (§3, lemma 2). Quindi $D + D' \subseteq F$, inoltre non essendo R e R' parallele si

ha: $D \oplus D' = F$. Abbiamo $a' \in R' \subseteq H$ quindi $a' = a + f$, ossia: $a' - a \in F = D \oplus D'$. Pertanto $a' - a = d + d'$, $d \in D$, $d' \in D'$. Segue che $p := a' - d' = a + d$ appartiene a $R \cap R'$. Se $q \neq p$ è un altro punto di $R \cap R'$ allora $R = R'$ perché per due punti distinti passa una ed un'unica retta (Es.1.1); ma $R \neq R'$ perché R e R' non sono parallele.

(ii) Se $p \in R \cap R'$ allora $R = p + D$, $R' = p + D'$; pertanto R ed R' sono contenute in $p + \langle D + D' \rangle$ che è un sottospazio affine di dimensione al più due♦

3: Lemma: Siano E un k -spazio vettoriale di dimensione n , R e R' due rette affini di E . Se R e R' sono parallele allora sono complanari.

Dim: Le rette R , R' essendo parallele hanno la stessa direzione: $R = a + D$, $R' = a' + D$. Sia u un vettore non nullo di D . Se $\alpha(a-a') + \lambda u = 0$ allora $\alpha \neq 0$ e $a' = a + \beta u$; quindi $a' \in R$. In questo caso $R = a' + D = R'$. Supponiamo u e $a-a'$ linearmente indipendenti e consideriamo il piano affine $H = a + \langle a-a', u \rangle$. E' chiaro che $R \subseteq H$. Se $a'+\lambda u$ è un punto di R' allora $a'+\lambda u = a-(a-a')+\lambda u$ appartiene ad H , quindi $R' \subseteq H$ ♦

3.1: Osservazione : Risulta dai lemmi precedenti che due rette sono sghembe se e solo se non sono complanari.

Consideriamo adesso due rette R , R' in uno spazio, E , di dimensione tre.

Sia $\{f_1(v) = \beta_1, f_2(v) = \beta_2\}$ un sistema di equazioni cartesiane di R , $\{f'_1(v) = \beta'_1, f'_2(v) = \beta'_2\}$, un sistema di equazioni di R' . Sia $B = (e_1, e_2, e_3)$ una base di E , $B^* = (e^*_1, e^*_2, e^*_3)$ la base duale, e $f_k = \alpha_{k1}e^*_1 + \alpha_{k2}e^*_2 + \alpha_{k3}e^*_3$, $f'_k = \alpha'_{k1}e^*_1 + \alpha'_{k2}e^*_2 + \alpha'_{k3}e^*_3$, $1 \leq k \leq 2$, le coordinate di f_k , f'_k . Finalmente sia

$$M = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \beta_2 \\ \alpha'_{11} & \alpha'_{12} & \alpha'_{13} & \beta'_1 \\ \alpha'_{21} & \alpha'_{22} & \alpha'_{23} & \beta'_2 \end{pmatrix}.$$

Osservare che M è la matrice completa del sistema, S , che determina $R \cap R'$.

4: Proposizione: Con le notazioni precedenti, sono equivalenti:

(i) $\det(M) = 0$

- (ii) $R // R'$ o $R \in R'$ s'intersecano in un punto
 (iii) $R \in R'$ sono complanari.

Dim: (i) \Rightarrow (ii): Sia A la matrice $(4,3)$ ottenuta da M sopprimendo l'ultima colonna; A è la matrice dei coefficienti del sistema che determina $R \cap R'$. Abbiamo $\text{rg}(A) \geq 2$ perché, per esempio f_1 e f_2 sono linearmente indipendenti per ipotesi. Inoltre $\text{rg}(A) = 2$ se e solo se $\langle f_1, f_2 \rangle = \langle f'_1, f'_2 \rangle$ ossia se e solo se $\langle f_1, f_2 \rangle^\circ = \langle f'_1, f'_2 \rangle^\circ$ quindi se e solo se R e R' sono parallele. Possiamo quindi supporre $\text{rg}(A) = 3$. L'ipotesi $\det(M) = 0$ implica $\text{rg}(A) = \text{rg}(M)$, pertanto il sistema S è compatibile (cfr II.§17, "modo operativo"(2), teorema 10) ed equivalente ad un sistema di Cramer; l'unica soluzione fornisce le coordinate del punto $R \cap R'$.

(ii) \Rightarrow (i): Se $R // R'$, come già visto, $\text{rg}(A) \leq 2$, e questo implica $\text{rg}(M) \leq 3$ e quindi $\det(M) = 0$. Se $R \cap R' = \{p\}$, il sistema S è compatibile e perciò $\det(M) = 0$ (condizione di compatibilità, (cfr II.§17 loc. cit.), osservare che in ogni caso $\text{rg}(A) \leq 3$).

(ii) \Leftrightarrow (iii): Segue dai lemmi 2, 3 ♦

Esercizi:

4.1) In \mathbf{R}^3 sia D la retta di equazioni: $x+y+1=0$, $x-y+2z=0$, e R la retta di equazioni $x-5y+6z=0$, $5x-y+6z+1=0$. Mostrare che $R // D$ (cfr Es.2.3).

4.2) In \mathbf{R}^3 sia P il piano d'equazione $x + y + z = 0$ e D la retta di equazioni parametriche:

$$x = 1 + 2\lambda, \quad y = 2 + \lambda, \quad z = 1 - 2\lambda.$$

Dimostrare che esiste un'unica retta, L , passante per $p = (1,1,1)$, parallela a P , e con $L \cap D \neq \emptyset$. Determinare $L \cap D$.

4.3) Siano nello spazio affine \mathbf{R}^3 , due rette D, L non parallele e complanari. Si noti H il piano contenente D e L . Sia R una retta non contenuta in H e che incontra D in un punto. Mostrare che R e L sono sghembe (suggerimento: non fare conti ma un disegno)

4.4) Siano nello spazio affine \mathbf{R}^3 , due rette sghembe, R, S , e p un punto non appartenente a $R \cup S$. Dimostrare che esiste un'unica retta, L , passante per p e complanare sia con R che con S . Mostrare inoltre che L incontra almeno una delle due rette R, S (suggerimento: non fare conti).

4.5) Siano in \mathbf{R}^3 i piani $H(\lambda)$, $H'(\lambda)$, $H''(\lambda)$, di equazioni:

$$H(\lambda): (\lambda+3)x - 4y + 4z = \lambda$$

$$H'(\lambda): 4x + (\lambda - 5)y + 4z = \lambda$$

$$H''(\lambda): 4x - 4y + (\lambda + 3)z = -\lambda.$$

(i) Determinare, secondo il parametro reale λ , l'intersezione $H(\lambda) \cap H'(\lambda) \cap H''(\lambda) =: I(\lambda)$.

(ii) Determinare il più piccolo sottospazio affine, $E \subseteq \mathbf{R}^3$, tale che: $\forall \lambda \in \mathbf{R}, I(\lambda) \subseteq E$.

(iii) Per $\lambda = 3$ dare l'equazione di un piano passante per $I(\lambda)$ e parallelo a F dove:

(a) $F = \{(x, y, z) \in \mathbf{R}^3 / z+x = 1\}$, (b) $F = \{(x, y, z) \in \mathbf{R}^3 / x+y = 1\}$.

4.6 Siano in \mathbf{R}^3 le rette affini D, D', D'' :

$D = \{(x, y, z) \in \mathbf{R}^3 / y = 0 \text{ e } z = 0\}; D'' = \{(x, y, z) \in \mathbf{R}^3 / x = 0 \text{ e } z = 1\}$ e D' data dalle equazioni parametriche: $x = 1, y = 1, z = \lambda; \lambda$ parametro reale.

(i) Dimostrare che queste tre rette sono due a due sghembe (i.e. D e D' (risp. D e D'' , D' e D'') sono sghembe).

Siano a', b' due punti distinti di D' . Per ogni $p \in D$ si pone $P_p := [p, a', b']$; P_p è il piano affine generato da p e da D' .

(ii) Sia $S = \{p \in D / P_p \cap D'' = \emptyset\}$. Determinare S .

(iii) Se $p \in (D \setminus S)$ (quindi $P_p \cap D'' \neq \emptyset$), e se $p \neq (0, 0, 0)$, dimostrare che esiste un'unica retta, L_p , passante per p e che incontra D' e D'' .

(iv) Siano $p \neq q$ due punti distinti di $(D \setminus S)$ e diversi dall'origine. Con le notazioni di (iii), dimostrare che $L_p \cap L_q = \emptyset$.

4.7 Sia P il piano di \mathbf{R}^3 , di equazione $y = 0$.

(i) Mostrare che esistono infinite rette $(L_i)_{i \in I}$, tali che: (a) $L_i \parallel P$, (b) se $i \neq j$, L_i e L_j sono sghembe, (c) per ogni i , L_i incontra la retta $x = 0, z = 0$.

(ii) Siano H un piano affine di \mathbf{R}^3 , R una retta affine incidente H in un unico punto. E' possibile trovare infinite rette $(D_i)_{i \in I}$, tali che: (a) $D_i \parallel H$, (b) se $i \neq j$, D_i e D_j sono sghembe, (c) per ogni i , D_i incontra la retta R ? (giustificare!).

(iii) Siano L_1, L_2, L_3 , tre rette come in (i). Dimostrare che per ogni punto $p \in L_1, L_3$ non è parallela al sottospazio affine generato da p e L_2 .

4.8 Si consideri in \mathbf{R}^3 il fascio di piani, F , di asse la retta $A \cap B$ dove A è il piano passante per $p = (0, 0, 1)$, di direzione $\langle(1, 2, 0), (0, 1, 1)\rangle$, e B è il piano passante per $q = (-3, 0, 0)$, di direzione $\langle(0, 1, 0), (-2, 0, 1)\rangle$.

(i) Si determini i piani di F paralleli alla retta S che passa per il punto $(0, 0, 2)$ con vettore direttore $(-2, -3, 1)$.

(ii) Sia H un piano non parallelo ad S . Esiste un piano $H' \in F$ tale che $\dim(H \cap H') \neq 1$? (suggerimento: non fare conti ma ragionare sulle posizioni reciproche).

4.9 Siano in \mathbf{R}^3 il piano H di equazione $x + y - z = 1$ e la retta $R = v + \langle v \rangle$ dove $v = (1, 0, 1)$. Infine sia D la retta di equazioni $y = x + 1, z = y + 1$.

(i) Dare delle equazioni cartesiane di R .

(ii) Determinare le posizioni reciproche di H, R, D (sono paralleli, incidenti, sghembi? in particolare determinare le eventuali intersezioni $q = H \cap D, H \cap R, D \cap R$).

(iii) Sia $p = (3, 2, -4) \in H$ e sia $P = \langle p, R \rangle$. Senza fare conti mostrare che P passa per l'origine. Dare un'equazione cartesiana di P .

(iv) Determinare $t := P \cap D$.

4.10) Siano R_1, R_2, R_3 tre rette distinte di \mathbb{R}^3 .

(i) Si assume R_1, R_2, R_3 due a due sghembe. Mostrare che esiste sempre almeno una retta (risp. un'infinità di rette) L t.c. $L \cap R_i \neq \emptyset, \forall i$.

(ii) Si assume ancora le rette R_i due a due sghembe. Sia $p \in R_1$. Mostrare con un esempio che non si può sempre trovare una retta L tale che: $p \in L, L \cap R_i \neq \emptyset, i = 2, 3$.

(iii) E' ancora vero l) se le rette R_i non sono due a due sghembe?

4.11) Siano in \mathbb{R}^4 i due sottospazi affini H, H' dove:

$$H = \{(x, y, z, t) \in \mathbb{R}^4 / x+y-z = 1 \text{ e } x-y+t = 1\},$$

$$H' = \{(x, y, z, t) \in \mathbb{R}^4 / x-z+t = -1 \text{ e } y+z-t = -1\}.$$

(i) Determinare la dimensione di H e H' . Dare delle equazioni parametriche di H .

(ii) Determinare l'intersezione e la posizione reciproca di H e H' .

(iii) Determinare il sottospazio affine generato da H e H' .

(iv) Esiste un sottospazio affine di dimensione due di \mathbb{R}^4 , V , tale che:

$V // H$ e $V \cap H' = \emptyset$? (senza fare tanti conti, ma ragionando)

5) Riferimenti affini ed affinità.

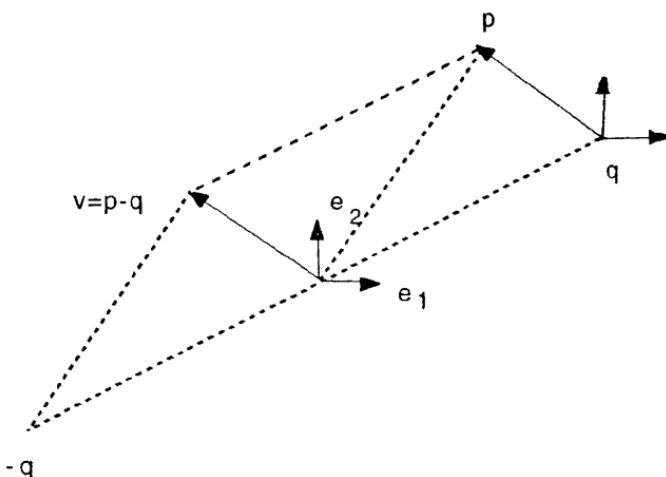
Sia E un k -spazio vettoriale. Possiamo considerare E come uno (sotto)spazio affine in tanti modi diversi: $\forall q \in E: E = q+E$. Dal punto di vista vettoriale la scelta naturale consiste nel prendere $q = 0$; però, se guardiamo E come uno spazio affine (per esempio "il piano della lavagna"), ci accorgiamo che tutti i punti sono "uguali" non c'è nessun punto "più bello degli altri" (un spazio affine è omogeneo). Vogliamo però poter reperire ogni punto di E tramite delle coordinate. In conclusione, $\forall q \in E$, vogliamo definire la nozione di sistema di coordinate di origine q . Se $q = 0$ sappiamo già come fare: basta prendere una base $B = (e_1, \dots, e_n)$ dello spazio vettoriale E . Nel caso generale, basta traslare.

Sia $q \in E$ un punto qualsiasi di E . Se $p \in E$ possiamo considerare il vettore $p-q = v$, lo si può pensare come il vettore ("libero") di origine q ed estremità p ("estremità meno l'origine").

E' chiaro che $p = q+v$ (i.e. $p = q+(p-q)$); in questa uguaglianza p e q vanno pensati come punti e v come vettore. Una volta fissato q (come origine dello spazio affine E) il punto p è completamente determinato dal vettore v .

Se $B = (e_1, \dots, e_n)$ è una base dello spazio vettoriale E allora possiamo scomporre v su B : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$.

Quindi una volta fissati il punto q (come origine dello spazio affine E) e la base B , il punto p è completamente determinato dagli scalari $(\lambda_1, \dots, \lambda_n)$.



1: Definizione: Un riferimento affine dello spazio affine E è una coppia $(q; B)$ dove q è un punto di E e B è una base dello spazio vettoriale E . Il punto q viene chiamato origine (del riferimento).

1.1: Osservazione : Come risulta dal disegno qui sopra, il riferimento $(q; B)$, $B = (e_1, e_2)$, non è altro che la base B traslata nel punto q .

1.2: Esempio : (i) Sia $E = k^n$. Come origine prendiamo $q = (0, \dots, 0)$ e come base B prendiamo la base canonica, il riferimento ottenuto si chiama riferimento standard (o canonico). Lo spazio affine k^n con il riferimento standard viene notato $A^n(k)$ e chiamato l' n -spazio affine numerico su k .

Nel seguito (e come già fatto in precedenza) scriveremo indifferentemente k^n o $A^n(k)$ per indicare lo spazio affine k^n con il riferimento canonico.

(ii) Il vettore $p-q$ si nota anche \vec{qp} . La relazione di Chasles: $\vec{qp} + \vec{pr} = \vec{qr}$, non è altro che: $(p-q) + (r-p) = r-q$.

2: Definizione: Siano (q, B) un riferimento affine di E e p un punto di E . Le coordinate $(\lambda_1, \dots, \lambda_n)$ del vettore $p-q$ nella base B sono le coordinate (affini) del punto p nel riferimento (q, B) .

2.1: Osservazione : Le coordinate dell'origine q sono $(0, \dots, 0)$.

Cambiamento di riferimento:

Siano $(q, B), (q', B')$ due riferimenti affini e $p \in E$. Il problema consiste nel determinare le coordinate di p nel riferimento (q', B') a partire da quelle nel riferimento (q, B) . Per questo basta conoscere le coordinate di q' nel riferimento (q, B) .

Infatti sia $p = q + v$, allora $p = q' + ((q - q') + v) = q' + (v - (q' - q))$. Le coordinate di p rispetto a (q, B) sono le coordinate $(\lambda_1, \dots, \lambda_n)$ del vettore v nella base B , le coordinate di p rispetto a (q', B') sono le coordinate $(\lambda'_1, \dots, \lambda'_n)$, del vettore $v - (q' - q)$ nella base B' . Siano $(\beta_1, \dots, \beta_n)$ le coordinate del vettore $q' - q$ nella base B (i.e. le coordinate di q' rispetto a (q, B)). Se $M = \text{mat}(\text{Id}; B, B')$ allora: $M \cdot {}^t(\lambda_1 - \beta_1, \dots, \lambda_n - \beta_n) = {}^t(\lambda'_1, \dots, \lambda'_n)$. Abbiamo quindi:

3: Proposizione: Se $(\lambda_1, \dots, \lambda_n)$ sono le coordinate di p nel riferimento (q, B) allora le coordinate $(\lambda'_1, \dots, \lambda'_n)$ di p nel riferimento (q', B') sono date da:

$$M \begin{pmatrix} \lambda_1 - \beta_1 \\ \vdots \\ \lambda_n - \beta_n \end{pmatrix} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix}$$

dove $(\beta_1, \dots, \beta_n)$ sono le coordinate di q' rispetto a (q, B) e dove $M = \text{mat}(\text{Id}; B, B')$.

3.1: Osservazione : La formula precedente $M(\underline{\lambda} - \underline{\beta}) = \underline{\lambda}'$ ($\underline{\lambda} = {}^t(\lambda_1, \dots, \lambda_n)$) si può anche riscrivere $\underline{\lambda}' = M\underline{\lambda} + \underline{\delta}$ dove $\underline{\delta} = -M\underline{\beta}$ sono le coordinate di q nel riferimento (q', B') .

Affinità e riferimenti:

4: Lemma: Siano E un k -spazio vettoriale e $f: E \rightarrow E$ un'applicazione. L'applicazione f è un'affinità se e solo se esiste un isomorfismo lineare $v: E \rightarrow E$ tale che: $\forall (p, q) \in E^2, f(p) - f(q) = v(p-q)$. Si ha allora $L(f) = v$.

Dim: Supponiamo la condizione verificata. Fissiamo un punto q . Da $f(x) - f(q) = v(x-q)$ segue $f(x) = (t_a \cdot v)(x)$ dove $a = f(q) - v(q)$. Siccome v è invertibile $t_a \cdot v$ è un'affinità. Viceversa se $f = t_a \cdot v$ è un'affinità è chiaro che $f(p) - f(q) = v(p-q)$ per ogni coppia di punti (p, q) ♦

5: Lemma: Sia E un k -spazio vettoriale. Fissiamo un punto q in E . Allora: $\forall q' \in E$ e $\forall v \in GL(E)$, esiste un'unica affinità $u: E \rightarrow E$ tale che $u(q) = q'$ e $L(u) = v$.

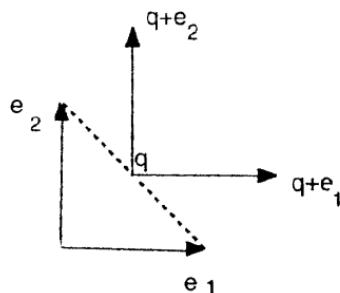
Dim: Poniamo $u(x) = v(x-q)+q'$ allora u è un'affinità perché $u = t_b \circ v$ ($b = q'-v(q)$) e $u(q) = q'$. Se $f = t_c \circ w$ è un'affinità che soddisfa le condizioni del lemma allora $w = v$ e $f(q) = c+v(q) = b+v(q)$ da cui $c = b$. Pertanto (§1, 12), $f = u$ ♦

5.1: Osservazione : Il lemma 5 mostra che un'affinità è completamente determinata dalla sua applicazione lineare associata e dall'immagine di un punto qualsiasi.

6: Teorema: ("Teorema fondamentale della geometria affine") Sia E un k -spazio vettoriale di dimensione n . Siano q_0, q_1, \dots, q_n (risp. p_0, p_1, \dots, p_n) $n+1$ punti affinamente indipendenti. Esiste una ed un'unica affinità $u: E \rightarrow E$ tale che $u(q_i) = p_i$, $0 \leq i \leq n$.

Dim: Siccome i punti q_i (risp. p_i) sono affinamente indipendenti i vettori $(q_i - q_0)$ (risp. $(p_i - p_0)$), $1 \leq i \leq n$, sono linearmente indipendenti (§1, 9), e quindi formano una base, C , (risp. C') di E . Sia $v: E \rightarrow E$ l'applicazione lineare definita da $v(q_i - q_0) = p_i - p_0$; v è un isomorfismo lineare (cfr II. §4, 16). Sia u l'unica affinità tale che $u(q_0) = p_0$ e $L(u) = v$ (lemma 5). Allora $u(q_i) - u(q_0) = v(q_i - q_0) = p_i - p_0$, ossia $u(q_i) = p_i$ ♦

6.1: Osservazione : Se (q, B) , $B = (e_1, \dots, e_n)$, è un riferimento, i punti q, e_1, \dots, e_n non sono necessariamente affinamente indipendenti; è invece vero che i punti $q, e_1 + q, \dots, e_n + q$ sono affinamente indipendenti:



Si osserverà che se p_1, \dots, p_{n+1} sono $n+1$ punti affinamente indipendenti, allora per ogni i , $1 \leq i \leq n+1$, $(p_i; B_i)$ dove $B_i = (p_1 - p_i, \dots, p_{i-1} - p_i, p_{i+1} - p_i, \dots, p_{n+1} - p_i)$, è un riferimento.

In realtà questa osservazione conduce a una definizione alternativa di riferimento: $(p_0; p_1-p_0, \dots, p_n-p_0)$ è un riferimento se e solo se i punti p_0, p_1, \dots, p_n sono affinamente indipendenti.

Esercizi:

5.1) Siano E un k -spazio vettoriale e $R = (q, B)$, $B = (e_1, \dots, e_n)$, un riferimento di E .

(i) Quanto può valere la dimensione del sottospazio affine generato dai punti q, e_1, \dots, e_n ?

(ii) Sia $u: E \rightarrow E$ un'affinità. Si pone $q' = u(q)$, $e'_i = q(e_i)$, $B' = (e'_i)$. E' $R' = (q', B')$ un riferimento di E ?

5.2) Siano in $A^3(\mathbb{R})$, $q = (1, 1, 1)$, $v_1 = (-1, 0, 0)$, $v_2 = (1, 2, 0)$, $v_3 = (2, 0, -3)$, e $b = (4, -1, 0)$.

(i) Mostrare che $B = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 .

(ii) Determinare le coordinate del punto b nel riferimento (q, B) .

(iii) Sia c il punto di coordinate $(1, 1, 1)$ nel riferimento (q, B) . E' possibile determinare un riferimento $R' = (q', C)$ (C la base canonica) tale che le coordinate di c in R' siano $(1, -1, 1)$? In caso di risposta affermativa determinare esplicitamente un tale riferimento R' (R' unicamente determinato?).

5.3) Sia in $A^2(\mathbb{R})$ la retta R di equazione $x+y=1$. Dare l'equazione di R nel riferimento (q, B) , $q = (1, -1)$, $B = (u, v)$, $u = (0, -1)$, $v = (2, 1)$ (si verificherà che si tratta effettivamente di un riferimento).

5.4) Si considerino in \mathbb{R}^3 le rette R, S di equazioni: $R = \{(x, y, z) / x = 2t, y = t+1, z = t-1, t \in \mathbb{R}\}$, $S = \{(x, y, z) / x-y-z+2 = 0, y-z+1 = 0\}$. Sia F il fascio di piani di asse la retta R .

(i) Determinare i piani di F incidenti il piano di equazione $x+y-3z=0$.

(ii) Determinare la posizione reciproca di R e S .

(iii) Sia $A = \begin{pmatrix} -3 & 0 & 1 \\ -5 & 0 & -1 \\ 5 & 1 & 2 \end{pmatrix} \in M_3(\mathbb{R})$ e si consideri l'applicazione $g \in \text{End}()$ tale che $\text{mat}(g; B, B) = A$. Determinare una base, (u, v, w) , di autovettori di g .

(iv) Sia $h \in \text{Aff}(\mathbb{R}^3)$ tale che $h(O) = P$, dove $O = (0, 0, 0)$ e $P = (1, -1, 0)$, $L(h) = g$. Si provi che le uniche rette per O mandate da h in rette parallele (passanti necessariamente per P) sono le rette di direzioni u, v, w .

5.5) Sia in \mathbb{R}^4 il sottospazio affine $P = \{(x, y, z, t) \in \mathbb{R}^4 / x+y-2z=1, y-z+t=1\}$.

(i) Determinare delle equazioni parametriche di P e la dimensione di P .

- (ii) Dare delle equazioni per una generica retta, $D \subseteq \mathbb{R}^4$, passante per il punto $q := (1,1,1,1)$ e parallela a P . Sia A il più piccolo sottospazio affine di \mathbb{R}^4 contenente tutte le rette passanti per q e parallele a P . Determinare delle equazioni di A .
- (iii) E' possibile trovare un'affinità, f , di \mathbb{R}^4 tale che $f(P) \subseteq A$? Se sì, determinare una tale affinità dandone delle equazioni.
- (iv) Sia B il più piccolo sottospazio affine contenente P e il punto q . Determinare B e $A \cap B$.

5.6) Siano E un k -spazio vettoriale di dimensione n , $\lambda_1, \dots, \lambda_p$ degli scalari tali che $\lambda = \lambda_1 + \dots + \lambda_p \neq 0$. Se v_1, \dots, v_p sono degli elementi di E , il baricentro dei punti v_1, \dots, v_p con i pesi $\lambda_1, \dots, \lambda_p$ è l'unico punto v tale che $\lambda v = \lambda_1 v_1 + \dots + \lambda_p v_p$.

(i) Mostrare che si possono dare a v_1, \dots, v_p dei pesi $\alpha_1, \dots, \alpha_p$ tali che $\sum \alpha_i = 1$, e tali che v sia il baricentro dei v_i con i pesi α_i .

(i) Con le notazioni precedenti supponiamo $\{1, \dots, p\}$ unione disgiunta di I_1, \dots, I_r tali che $\sum_{i \in I_k} \lambda_i \neq 0$, $1 \leq k \leq r$. Sia x_k il baricentro di $(v_i)_{i \in I_k}$ coi pesi $(\lambda_i)_{i \in I_k}$. Mostrare che v è il baricentro di x_1, \dots, x_r coi pesi $\alpha_k = \sum_{i \in I_k} \lambda_i$, $1 \leq k \leq r$.

(ii) Sia $u: E \rightarrow F$ un'applicazione affine. Mostrare che $u(v)$ è il baricentro dei punti $u(v_i)$ con i pesi $\lambda_1, \dots, \lambda_p$.

(iii) Dimostrare che un sottinsieme A di E è un sottospazio affine se e soltanto se ogni baricentro di punti di A appartiene a A .

5.7) Sia E uno spazio vettoriale reale. Siano u, v due punti di E . Il segmento chiuso di estremità u, v è l'insieme, $[u, v]$, degli elementi del tipo $(\lambda u + \mu v)/(\lambda + \mu)$ dove $\lambda \geq 0, \mu \geq 0, \lambda + \mu \neq 0$. E' anche il sottinsieme $\{tu + (1-t)v / 0 \leq t \leq 1\}$.

(i) Mostrare che $[u, v]$ è l'insieme dei baricentri dei punti u, v con pesi positivi.

(ii) Un sottinsieme B di E è convesso se per ogni coppia (u, v) di punti in B il segmento $[u, v]$ è tutto contenuto in B . Dimostrare che un'intersezione qualsiasi di sottinsiemi convessi è convessa. E' vero che l'unione di due sottinsiemi convessi è convessa?

(iii) Sia $f: E \rightarrow \mathbb{R}$ una forma lineare e $\lambda \in \mathbb{R}$. Mostrare che $\{v \in E / f(v) \geq \lambda\}$ è convesso.

(iv) L'involucro convesso di un sottinsieme X di E è il più piccolo (per l'inclusione) sottinsieme convesso di E contenente X . Spiegare perché questa nozione è ben definita. Mostrare che l'involucro convesso di X è l'insieme dei baricentri di punti di X con pesi positivi.

(v) Se $C \subseteq \mathbb{R}^k$ è un sottinsieme convesso e se $f: \mathbb{R}^k \rightarrow \mathbb{R}^n$ è un'applicazione affine, allora $f(C)$ è convesso.

(vi) Un k -simplex è l'involucro convesso di $X = \{P_0, \dots, P_k\}$ dove P_0, \dots, P_k sono $k+1$ punti affinamente indipendenti di \mathbb{R}^k . Rappresentare graficamente i k -simplessi, $0 \leq k \leq 3$.

5) Spazi affini: teoria generale.

Si può definire su un insieme X una struttura di spazio affine (su un k-spazio vettoriale E) in modo assiomatico ("operazione semplicemente transitiva del gruppo $(E,+)$ su X " cfr Es.2). Questo punto di vista, molto utile in certe questioni, si riconduce comunque (cfr Prop. 13) a quello, studiato finora, dei sottospazi affini di uno spazio vettoriale.

1: Definizione: Sia E un k-spazio vettoriale di dimensione n . Uno spazio affine su E è una terna (X, E, ϕ) dove:

- X è un insieme
- $\phi : X \times E \rightarrow X$ è un'applicazione soddisfacente le seguenti condizioni:

$$(A1) \forall x \in X, \forall (v,w) \in E^2: \phi(\phi(x,v), w) = \phi(x, v+w)$$

(A2) per ogni (x,x') in X^2 , esiste uno ed un'unico v in E tale che $\phi(x, v) = x'$.

1.1: Osservazione : (i) Se lo spazio vettoriale è precisato dal contesto si usa dire che X è uno spazio affine.

1.2: Convenzioni: Per capire meglio questa definizione facciamo le seguenti convenzioni di scrittura e di terminologia:

- gli elementi di X (i "punti") saranno denotati con lettere maiuscole (P, Q, \dots)

- gli elementi di E (i "vettori") saranno denotati con lettere minuscole (v, u, \dots)

- se $(P, v) \in X \times E$ si pone $\phi(P, v) =: P+v$. **Attenzione:** questo segno + non ha niente a che fare con l'addizione in E ; è solo una convenzione per scrivere ϕ ; vedremo però, nell'esempio 2, la motivazione per questa scrittura.

Con queste convenzioni abbiamo:

(A1) $\forall P \in X, \forall (v,w) \in E^2: (P+v)+w = P + (v+w)$ (nel membro di sinistra i due + hanno significati diversi)

(A2) per ogni (P, P') in X^2 , esiste uno ed un unico v in E tale che: $P+v = P'$.

Finalmente se P, P' sono due punti di X , per (A2) esiste un unico vettore, v , tale che $P+v = P'$, si usa notare questo vettore $P'-P$ o anche $\vec{PP'}$.

Vedremo più avanti (cfr 10.1) il perché di questa notazione.

2: Esempio : Prendiamo $X = E$ e $\phi : E \times E \rightarrow E : (x, v) \mapsto x+v$, l'addizione del gruppo abeliano $(E, +)$. Le condizioni (A1), (A2) sono verificate e (E, E, ϕ) è uno spazio affine; è la struttura che abbiamo studiato finora, viene chiamata struttura naturale di spazio affine su E . E' per ricalcare questo esempio "naturale" che sono state introdotte le convenzioni fatte in 1.2.

D'ORA IN POI MANTERREMO LE CONVENZIONI FATTE IN 1.2.

3: Lemma: Per ogni (P, v) in $X \times E$: $P+v = P$ se e solo se $v = 0$ (0 = il vettore nullo di E).

Dim: Mostriamo che $P+0 = P$, l'equivalenza risulterà da l'unicità nella condizione (A2). Sia $v \in E$ tale che $P+v = P$ (cfr (A2)). Quindi: $P+0 = (P+v)+0$, e per via di (A1) questo è uguale a: $P+(v+0) = P+v$; in conclusione $P+0 = P$ ♦

4: Definizione: Per ogni $v \in E$ l'applicazione $t_v : X \rightarrow X : P \mapsto P+v$ si chiama la traslazione di vettore v .

5: Lemma: Per ogni $v \in E$, la traslazione $t_v : X \rightarrow X$ è una biiezione.

Dim: Come nel caso vettoriale si verifica (usando il lemma 3) che $(t_v)^{-1} = t_{-v}$ ♦

5.1: Osservazione : Nel caso $X = E$ con la struttura naturale di spazio affine ritroviamo le traslazioni studiate in precedenza (cfr §1).

6: Definizione: Un sottoinsieme X' di X è un sottospazio affine di X se esiste un sottospazio vettoriale, W , di E e un punto Q di X tali che $X' = \{P \in X / \exists v \in W, Q+v = P\}$.

6.1: Osservazione : Con le convenzioni precedenti: $X' = Q+W$. Come nel caso vettoriale si dimostra: $Q \in X'$ e se $Q+W = Q'+W'$ allora $W = W'$ (cfr §1, Lemma 3); il sottospazio vettoriale W è la direzione (o giacitura) del sottospazio affine X' . La direzione di X' è univocamente determinata da X' . E' chiaro inoltre che X' è uno spazio affine su la sua direzione. La direzione di X è E .

7: Definizione: Sia X uno spazio affine di direzione E . La dimensione di X è, per definizione, la dimensione di E .

8: Definizione: Siano X, Y due spazi affini sugli spazi vettoriali E, E' . Un'applicazione $f : X \rightarrow Y$ è un'applicazione affine se esiste un morfismo lineare $v : E \rightarrow E'$ tale che: $\forall (P, Q) \in X^2: f(P) - f(Q) = v(P-Q)$.

8.1: Osservazione : Si ricorda che $f(P)-f(Q)$ è l'unico vettore, v' , di E' tale che $f(P) = f(Q)+v'$. Come nel caso vettoriale, si dimostra che v è univocamente determinata da f . Osservare però che non ha senso richiedere che f sia la composizione di un morfismo lineare con una traslazione: un'applicazione lineare $E \rightarrow E'$ e una traslazione $Y \rightarrow Y$ non possono essere composte (tranne nel caso vettoriale in cui $Y = E'$).

ORIGINE, RIFERIMENTI AFFINI.

La differenza essenziale tra la struttura naturale di spazio affine su uno spazio vettoriale, E , e la struttura "generale" di spazio affine, consiste nell'esistenza, nel primo caso, di un punto "privilegiato": il vettore nullo; invece, nel caso generale, non esiste nessun punto privilegiato dell'insieme X (non esiste, a priori, nessuna struttura canonica di k -spazio vettoriale su X). Queste due situazioni, apparentemente contraddittorie, si conciliano tramite la nozione di riferimento affine. Abbiamo già visto, nel caso vettoriale, come la scelta di un'origine faccia perdere al vettore nullo il suo statuto privilegiato (traslando la struttura di spazio vettoriale nella nuova origine). In modo analogo, nel caso generale, la scelta di un'origine in X permette di dare a X una struttura di k -spazio vettoriale (isomorfo a E) e quindi di ricondurci alla situazione precedente.

Sia X uno spazio affine su E e sia Q un punto di X . Sia inoltre l'applicazione parziale: $\varphi_Q : E \rightarrow X: v \rightarrow Q+v$.

9: Lemma: Per ogni punto Q di X l'applicazione φ_Q è una biiezione.

Dim: E' una conseguenza di (A2): per ogni P in X , esiste uno ed un unico v in E tale che $\varphi_Q(v) = P$ ♦

La biiezione φ_Q permette di trasportare la struttura di k -spazio vettoriale di E ad X (cfr Es.II.2.5): si definisce un'addizione su X tramite: $P+P' := \varphi_Q(\varphi_Q^{-1}(P)$

+ $\varphi_Q^{-1}(P')$) e una moltiplicazione esterna tramite: $\lambda P := \varphi_Q(\lambda \varphi_Q^{-1}(P))$. In particolare $\varphi_Q^{-1}(Q)$ è il vettore, w , tale che: $Q+w = \varphi_Q(w) = Q$, dal lemma 3, $w = 0$.

10: Definizione: La struttura di k-spazio vettoriale su X ottenuta tramite φ_Q si nota X_Q . Si dice che Q è l'origine di X_Q (in pratica, per abuso di linguaggio, si dice che Q è (stato scelto come) l'origine di X).

10.1: Osservazione : (i) L'applicazione $\varphi_Q: E \rightarrow X_Q$ è un isomorfismo lineare per la struttura di k-spazio vettoriale appena definita su X_Q .

(ii) Siano P, P' due punti di X , per (A2) esiste un unico vettore v di E tale che $P+v = P'$. Sia Q un'origine di X ; nello spazio vettoriale X_Q possiamo calcolare $P'-P$: se $Q+w = P'$, $Q+u = P$ allora $P'-P$ è il vettore $w-u$; abbiamo $P+(w-u) = (Q+u)+(w-u) = Q+w = P'$ quindi $P'-P$ è l'unico vettore, v , tale che $P+v = P'$. Si trova così giustificata la convenzione fatta in 1.2.

11: Definizione: Sia X uno spazio affine su E . Un sistema di coordinate affini (o un riferimento affine) in X consiste nella scelta di un'origine, Q , in X e di una base (e_1, \dots, e_n) di E . Un tale riferimento viene denotato Qe_1, \dots, e_n .

Sia P un punto di X . Per (A2) esiste un unico vettore v in E tale che $P = Q+v$. Scriviamo v nella base (e_1, \dots, e_n) : $v = \lambda_1 e_1 + \dots + \lambda_n e_n$:

12: Definizione: Con le notazioni precedenti, l' n -upla $(\lambda_1, \dots, \lambda_n)$ si dice l' n -upla delle coordinate affini del punto P nel riferimento Qe_1, \dots, e_n .

12.1: Osservazione : L'origine Q ha coordinate $(0, \dots, 0)$.

Usando i riferimenti affini possiamo ricondursi alla struttura naturale di spazio affine su uno spazio vettoriale:

13: Proposizione: Siano E un k-spazio vettoriale di dimensione n e X uno spazio affine su E . Allora esiste un isomorfismo affine $f: X \rightarrow \mathbb{A}^n(k)$ (i.e. f è un'applicazione affine biiettiva tale che f^{-1} sia affine).

Dim: Sia Qe_1, \dots, e_n un riferimento affine di X . Se $P \in X$ ha per coordinate $(\lambda_1, \dots, \lambda_n)$, si definisce $f(P) = (\lambda_1, \dots, \lambda_n)$. L'applicazione f è chiaramente biiettiva ($f^{-1}(\beta_1, \dots, \beta_n)$ è l'unico punto P' tale che $P' = Q+w$ dove $w = \beta_1 e_1 + \dots + \beta_n e_n$). L'applicazione lineare associata è $u: E \rightarrow k^n$: $w \mapsto \beta_1 c_1 + \dots + \beta_n c_n$ dove $(c_1, \dots,$

$c_n)$ è la base canonica di k^n ; u è ovviamente un isomorfismo lineare. Verifichiamo che f è affine: se $(\lambda_1, \dots, \lambda_n), (\lambda'_1, \dots, \lambda'_n)$ sono le coordinate di P, P' allora $f(P) - f(P') = (\lambda_1 - \lambda'_1, \dots, \lambda_n - \lambda'_n)$. Abbiamo $Q+v = P, Q+v' = P'$ quindi $P-P' = v-v' = (\lambda_1 - \lambda'_1)e_1 + \dots + (\lambda_n - \lambda'_n)e_n$, e la condizione $f(P)-f(P') = u(P-P')$ è verificata. Nello stesso modo (col riferimento standard in $A^n(k)$) si verifica che f^{-1} è affine♦

14: Applicazioni affini:

Sempre usando i riferimenti possiamo ritrovare la nozione precedente di applicazione affine.

Siano X, Y due spazi affini su due k -spazi vettoriali E, E' . Siano O, O' delle origini in X, Y (da non confondere con i vettori nulli di E, E'). Sia $f : X \rightarrow Y$ un'applicazione allora f è affine se e solo se esiste un'applicazione affine (nel senso di §1, Def.7) $\underline{f} : E \rightarrow E'$ tale che: $\varphi_{O'} \circ \underline{f} \circ \varphi_O^{-1} = f$. Quest'ultima condizione si esprime dicendo che il seguente diagramma è commutativo:

$$\begin{array}{ccc} E & \xrightarrow{\quad f \quad} & E' \\ \downarrow \varphi_O^{-1} & & \downarrow \varphi_{O'} \\ X & \xrightarrow{\quad \underline{f} \quad} & Y \end{array}$$

Se f esiste allora: $f(P) = \varphi_{O'} \circ \underline{f} \circ \varphi_O^{-1}(P) = (\varphi_{O'} \circ \underline{f})(P-O) = \varphi_{O'}(v(P-O)+t)$ dove v è l'applicazione lineare associata a \underline{f} e dove t indica il vettore della traslazione. Finalmente: $f(P) = O'+v(P-O)+t$. In modo analogo: $f(P') = O'+v(P-O)+t$. Quindi $f(P) - f(P') = v(P-O) - v(P'-O) = v(P-O-P'+O) = v(P-P')$, e v è l'applicazione lineare associata a f .

Viceversa se f è un'applicazione affine nel senso della definizione 8 allora esiste un morfismo lineare $v : E \rightarrow E'$ tale che $f(P) - f(P') = v(P-P')$. In particolare $f(P) - f(O) = v(P-O)$. Sia $t = f(O)-O'$ e $\underline{f} : E \rightarrow E' : x \rightarrow v(x)+t$ allora \underline{f} è l'applicazione cercata: $\varphi_{O'} \circ \underline{f} \circ \varphi_O^{-1}(P) = (\varphi_{O'} \circ \underline{f})(P-O) = \varphi_{O'}(v(P-O)+f(O)-O') = v(P-O)+f(O) = f(P)$.

Esercizi:

6.1) Siano X un insieme e $(G, *)$ un gruppo. Denoteremo con $\mathfrak{S}(X)$ il gruppo delle permutazioni di X (cfr Es.1.5.7). Si dice che G opera su X se esiste un morfismo di gruppi: $\varphi : G \rightarrow \mathfrak{S}(X)$. In questo caso si nota φ_g (o anche $g(x)$) invece di $\varphi(g)$. Si ha quindi $\varphi_{g * g'} = \varphi_g \circ \varphi_{g'}$, $\varphi_{g^{-1}} = (\varphi_g)^{-1}$, $\varphi_e = \text{Id}_X$ dove e è l'elemento neutro di G . L'orbita del punto x è $O_x = \{g(x) / g \in G\}$.

(i) Mostrare che due orbite distinte sono disgiunte e che X è l'unione delle orbite. In altri termini le orbite sono le classi d'equivalenza per la relazione d'equivalenza: $x R y \Leftrightarrow \exists g \in G$ t.c. $g(x) = y$. Lo spazio quoziante si nota X/G .

(ii) Sia $S^1 = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 = 1\}$. Mostrare che il gruppo $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$ opera su S^1 tramite: $\varphi : \mathbb{Z}_2 \rightarrow \mathfrak{S}(S^1)$, $\varphi(0) = \text{Id}$, $\varphi(1) = -\text{Id}$. Descrivere lo spazio quoziante S^1/\mathbb{Z}_2 (cfr Es.1.4.5).

6.2) Sia G un gruppo che opera su un insieme X (cfr Es.1). Se: $\forall (x, x') \in X^2, \exists g \in G$ tale che $g(x) = x'$, si dice che G opera transitivamente. Se G opera transitivamente su X allora X viene chiamato spazio omogeneo sotto G .

(i) Se $x \in X$ si pone $G_x = \{g \in G / g(x) = x\}$. Dimostrare che G_x è un sottogruppo di G ; G_x si chiama il sottogruppo di isotropia di x (o stabilizzatore di x). Dimostrare che $\text{Ker}(\varphi) = \cap_{x \in X} G_x$.

(ii) Se G è abeliano, φ è iniettivo e se G opera transitivamente, dimostrare che: $G_x = \{e\}$ per ogni x in X . Dedurne che: $\forall (x, x') \in X^2, \exists! g \in G$ tale che $\varphi_g(x) = x'$; se (\dagger) è verificata si dice che l'operazione di G è semplicemente transitiva. Sono le operazioni di Es.1, (ii), (iii), (iv) transitive, semplicemente transitive?

(iii) Sia E un k -spazio vettoriale. Giustificare la seguente definizione: uno spazio affine su E è una terna (X, E, φ) dove X è un insieme sul quale il gruppo $(E, +)$ opera in modo semplicemente transitivo tramite $\varphi : E \rightarrow \mathfrak{S}(X)$ (i.e. $(E, +)$ opera in modo semplicemente transitivo tramite "traslazioni" su X).

1) Forme bilineari.

Le forme bilineari sono un caso particolare di applicazioni multilinear (cfr II. §13). In questo paragrafo si presentano le prime proprietà generali delle forme bilineari (interpretazione matriciale, congruenza, rango, ...). Le forme bilineari simmetriche sono particolarmente importanti per la geometria perché permettono di definire le nozioni di distanza, ortogonalità, angolo, ...

1: Definizione: Sia E un k -spazio vettoriale. Un'applicazione $f : E \times E \rightarrow k$ si dice forma bilineare se per ogni x in E le applicazioni parziali $f_x : E \rightarrow k$: $y \mapsto f(x,y)$; $f_{,x} : E \rightarrow k$: $y \mapsto f(y,x)$ sono lineari.

1.1: Osservazione : Confrontare con II. §13, Def.1.

2: Definizione: Una forma bilineare $f : E \times E \rightarrow k$ è:

- (i) simmetrica se: $\forall (x, y) \in E^2 : f(x,y) = f(y,x)$.
- (ii) antisimmetrica se: $\forall (x, y) \in E^2 : f(x,y) = -f(y,x)$
- (iii) alternante se: $\forall x \in E : f(x,x) = 0$.

2.1: Osservazione : (i) Abbiamo già visto che una forma bilineare alternante è antisimmetrica (cfr II. §13, 3.1). Se $2 (=1+1)$ è diverso da zero in k (in questo caso si dice che la caratteristica di k è diversa da due e si nota $ch(k) \neq 2$), una forma antisimmetrica è alternante. Infatti da $f(x,x) = -f(x,x)$ viene $2.f(x,x) = 0$ e $2 \neq 0$ implica 2 invertibile in k e quindi si ricava $f(x,x) = 0$. In particolare se $k = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, una forma bilineare è antisimmetrica se e solo se è alternante. Invece se $ch(k) = 2$ allora $1 = -1$ e una forma antisimmetrica è simmetrica.

(ii) L'insieme $L_2(E; k)$ (o $Bil(E)$) delle forme bilineari su E è un k -spazio vettoriale (cfr II. §13, Prop.2). È facile verificare che l'insieme $S^2(E)$, delle forme bilineari simmetriche è un sottospazio vettoriale di $Bil(E)$. Lo stesso vale per $A^2(E)$, l'insieme delle forme bilineari antisimmetriche.

INTERPRETAZIONE MATRICIALE DELLE FORME BILINEARI.

Sia $B = (e_1, \dots, e_n)$ una base di E , $f : E \times E \rightarrow k$ una forma bilineare e $x = x_1e_1 + \dots + x_ne_n$, $y = y_1e_1 + \dots + y_ne_n$ due vettori di E . Usando la bilinearità di f abbiamo: $f(x,y) = f(x_1e_1 + \dots + x_ne_n, y) = x_1f(e_1,y) + \dots + x_nf(e_n,y)$.

Inoltre $f(e_k,y) = f(e_k, y_1e_1 + \dots + y_ne_n) = y_1f(e_k,e_1) + \dots + y_nf(e_k,e_n)$.

In conclusione:
$$f(x,y) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} x_i y_j f(e_i, e_j) \quad (\dagger)$$

Risulta che f è completamente determinata dagli n^2 scalari $f(e_i, e_j)$, $1 \leq i \leq n$, $1 \leq j \leq n$.

3: Definizione: La matrice associata alla forma bilineare f , rispetto alla base B di E , è la matrice (n,n) , $\text{mat}_B(f) := (f(e_i, e_j))$, $1 \leq i \leq n$, $1 \leq j \leq n$.

3.1: Osservazione : L'elemento all'incrocio della riga i e della colonna j in $\text{mat}_B(f)$ è $f(e_i, e_j)$.

Possiamo adesso esprimere la relazione (\dagger) sotto forma matriciale. La scelta della base B ci permette di identificare E con k^n e quindi i vettori di E con i vettori, diciamo colonna, di k^n : al vettore $x = x_1e_1 + \dots + x_ne_n$ corrisponde il vettore colonna $X = {}^t(x_1, \dots, x_n)$ delle coordinate di x nella base B (useremo lettere maiuscole per indicare i corrispondenti vettori colonna di k^n).

4: Lemma: Con le notazioni precedenti:

- (i) $f(x, y) = {}^tX \cdot A \cdot Y$ dove $A := \text{mat}_B(f)$
- (ii) se $M \in M_n(k)$ l'applicazione $E \times E \rightarrow k: (x, y) \rightarrow {}^tX \cdot M \cdot Y$, è una forma bilineare, g , su E tale che $\text{mat}_B(g) = M$.

Dim: (i) Il prodotto delle matrici è associativo quindi $({}^tX \cdot A) \cdot Y = {}^tX \cdot (A \cdot Y)$. Per definizione $A \cdot Y$ è la matrice $(n, 1)$ di coefficienti (α_i) , $1 \leq i \leq n$, con $\alpha_i = y_1f(e_i, e_1) + \dots + y_nf(e_i, e_n)$ ("prodotto scalare" della i -esima riga di A con la colonna di Y ; cfr II. §9). Nello stesso modo ${}^tX \cdot (A \cdot Y)$ è il "prodotto scalare" della riga tX con la colonna $(A \cdot Y)$:

$${}^tX \cdot (A \cdot Y) = x_1[y_1f(e_1, e_1) + \dots + y_nf(e_1, e_n)] + \dots + x_n[y_1f(e_n, e_1) + \dots + y_nf(e_n, e_n)] =$$

$$\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} x_i y_j f(e_i, e_j) \text{ e per } (\dagger) \text{ questo è uguale a } f(x, y).$$

(ii) Segue dalle proprietà del prodotto delle matrici. Mostriamo per esempio che g , y è lineare: $g(\lambda x + \mu x', y) = {}^t(\lambda X + \mu X') \cdot M \cdot Y = [(\lambda {}^tX + \mu {}^tX') \cdot M] \cdot Y = [\lambda ({}^tX \cdot M) + \mu ({}^tX' \cdot M)] \cdot Y = \lambda g(x, y) + \mu g(x', y)$. Inoltre è chiaro che $g(e_i, e_j) = m_{ij}$ ♦

4.1: Esempio : Se in 4(ii) prendiamo $M = I_n$ otteniamo $f(x, y) = x_1y_1 + \dots + x_ny_n$. Se $E = k^n$ si dice che f è la forma bilineare simmetrica standard; vedremo più avanti il perché di questa terminologia.

5: Proposizione: Sia B una base del k -spazio vettoriale E . L'applicazione $\text{mat}_B: \text{Bil}(E) \rightarrow M_n(k)$: $f \mapsto \text{mat}_B(f)$ è un isomorfismo di k -spazi vettoriali.

L'immagine di $S^2(E)$ è il sottospazio vettoriale, $S_n(k)$, delle matrici simmetriche e l'immagine di $A^2(E)$ è il sottospazio vettoriale, $A_n(k)$, delle matrici antisimmetriche.

Dim: E' facile verificare che l'applicazione è k -lineare. Da 4(ii) mat_B è suriettiva. Dalla relazione (†) una forma bilineare f è completamente determinata dagli scalari $f(e_i, e_j)$, quindi mat_B è iniettiva. Sempre da (†) segue che f è simmetrica (risp. antisimmetrica) se e solo se $f(e_i, e_j) = f(e_j, e_i)$ (risp. $f(e_i, e_j) = -f(e_j, e_i)$); quindi se e solo se $\text{mat}_B(f)$ è simmetrica (risp. antisimmetrica)♦

5.1: Osservazione : Confrontare con Es. III.2.4, II.5.3.

Ovviamente l'isomorfismo mat_B dipende dalla base B , e le matrici che rappresentano una stessa forma bilineare in basi distinte sono generalmente diverse. Vediamo che legame c'è tra due tali matrici.

Siano $B = (e_1, \dots, e_n)$, $B' = (e'_1, \dots, e'_n)$ due basi di E e x, y due vettori di E :
 $x = x_1e_1 + \dots + x_ne_n = x'_1e'_1 + \dots + x'_ne'_n$; $y = y_1e_1 + \dots + y_ne_n = y'_1e'_1 + \dots + y'_ne'_n$.
Sia $P = \text{mat}(\text{Id}_E; B, B')$ (i vettori colonna di P sono le coordinate di e_1, \dots, e_n nella base B'). Quindi (cfr II.§11): $PX = X'$, $PY = Y'$.

Sia $f : E \times E \rightarrow k$ una forma bilineare e $A = \text{mat}_B(f)$, $A' = \text{mat}_{B'}(f)$. Abbiamo $f(x, y) = {}^t X \cdot A \cdot Y = {}^t(PX) \cdot A \cdot (PY) = {}^t X \cdot ({}^t P \cdot A' \cdot P) \cdot Y = {}^t X \cdot A' \cdot Y$ pertanto (cfr Prop.5): ${}^t P \cdot A' \cdot P = A$. Osserviamo che siccome P è invertibile abbiamo anche:
 $A' = {}^t(P^{-1}) \cdot A \cdot P^{-1}$ (si è usato $({}^t P)^{-1} = {}^t(P^{-1})$; Es.2). Abbiamo dimostrato:

6: Proposizione: Due matrici (n,n) , A, A' , rappresentano la stessa forma bilineare se e solo se esiste una matrice (n,n) invertibile, P , tale che: ${}^t P \cdot A' \cdot P = A$.

7: Definizione: Due matrici $A, A' \in M_n(k)$ sono congruenti se esiste una matrice invertibile (n,n) , P , tale che: ${}^t P \cdot A' \cdot P = A$.

7.1: Osservazione : Altrimenti detto due matrici A, A' rappresentano la stessa forma bilineare se e solo se sono congruenti.

8: Lemma: La relazione \equiv su $M_n(k)$ definita da: $A \equiv A' \Leftrightarrow A$ e A' sono congruenti, è una relazione d'equivalenza.

Dim: Es.3♦

9: Lemma: Due matrici congruenti hanno lo stesso rango.

Dim: Infatti due matrici congruenti sono equivalenti (II. §11, Def.4) e due matrici equivalenti hanno lo stesso rango (II. §12, Cor.9, 10) •

Risulta dal lemma 9 che il rango, r , della matrice che rappresenta una data forma bilineare f rispetto ad una base qualsiasi non dipende dalla base ma solo da f ; questo giustifica la seguente:

10: Definizione: Sia $f : E \times E \rightarrow k$ una forma bilineare. Il rango di f è il rango della matrice $\text{mat}_B(f)$ che rappresenta f rispetto ad una base, B , qualsiasi di E .

11: Definizione: Una forma bilineare $f : E \times E \rightarrow k$ si dice non degenere se ha rango massimo ($= \dim(E)$). Se $\text{rg}(f) < \dim(E)$, f si dice degenere.

Esercizi:

1.1) Sia E un k -spazio vettoriale di dimensione n con $\text{ch}(k) \neq 2$.

- (i) Dimostrare che $\text{Bil}(E) = S^2(E) \oplus A^2(E)$ e dare le dimensioni di $\text{Bil}(E)$, $S^2(E)$, $A^2(E)$ (cfr Es. III.2.4)
- (iii) Trascrivere (i) e (ii) in termini di matrici.
- (iv) E' ancora vero (i) se $\text{ch}(k) = 2$?

1.2) (i) Se $M, N \in M_{n,p}(K)$ dimostrare ${}^t(M+N) = {}^tM + {}^tN$.

(ii) Se $M \in M_{n,p}(K)$, $N \in M_{p,t}(K)$ dimostrare che ${}^t(M \cdot N) = {}^tN \cdot {}^tM$.

(iii) Se $P \in M_n(K)$ è invertibile dimostrare che ${}^t(P^{-1}) = ({}^tP)^{-1}$.

(iv) Se $M \in M_n(K)$ mostrare che ${}^tM \cdot M$ è una matrice simmetrica.

1.3) Dimostrare che la relazione di congruenza è una relazione d'equivalenza su $M_n(k)$.

1.4) Dimostrare che esistono delle matrici M, N in $M_2(\mathbf{R})$ tali che: $\text{rango}(M) = \text{rango}(N) = 2$; M e N non sono congruenti (suggerimento: considerare le matrici congruenti a I_2).

2) Forme bilineari simmetriche e forme quadratiche.

Sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica. Per definizione $f(x,y) = f(y,x)$ per ogni (x,y) in E^2 , e la matrice $\text{mat}_B(f)$ che rappresenta f rispetto ad una base B qualsiasi è simmetrica (§1, Prop.5). Definiamo adesso un'applicazione: $q : E \rightarrow k$ tramite $q(x) = f(x,x)$.

1: Definizione: L'applicazione q è la forma quadratica associata (o determinata) da f .

1.1: Osservazione : La forma quadratica q soddisfa le seguenti proprietà:

$$(Q1) \forall x \in E, \forall \lambda \in k: q(\lambda x) = \lambda^2 \cdot q(x).$$

$$(Q2) \forall (x,y) \in E^2: q(x+y) - q(x) - q(y) = 2 \cdot f(x,y).$$

In particolare (Q'2): l'applicazione $E \times E \rightarrow k: (x,y) \rightarrow q(x+y) - q(x) - q(y)$ è una forma bilineare.

La verifica delle proprietà (Q1), (Q2) sono lasciate al lettore.

2: Definizione: Sia E un k -spazio vettoriale. Una forma quadratica su E è un'applicazione $q : E \rightarrow k$ che soddisfa (Q1) e (Q'2).

3: Lemma: Sia E un k -spazio vettoriale. Se la caratteristica di k è diversa da due esiste una corrispondenza biunivoca tra $S^2(E)$ e $Q(E)$, l'insieme delle forme quadratiche su E .

Dim: La corrispondenza $\varphi : S^2(E) \rightarrow Q(E)$ è data da $\varphi(f) = q$ con $q(x) = f(x,x)$ e $\varphi^{-1}(q) = f$ con $f(x,y) = [q(x+y) - q(x) - q(y)]/2$ ♦

3.1: Osservazione : Il lemma 3 è falso se $\text{ch}(k) = 2$: non si può dividere per 2! (dare un controesempio...)

Come avremo occasione di vedere anche nel seguito, la teoria delle forme bilineari e delle forme quadratiche è molto diversa secondo che $\text{ch}(k) = 2$ o $\text{ch}(k) \neq 2$. D'ora in poi supporremo sempre $\text{ch}(k) \neq 2$ (e ogni tanto, in certi enunciati dove questa ipotesi è essenziale, lo ricorderemo).

Nell'ipotesi $\text{ch}(k) \neq 2$ abbiamo quindi (cfr lemma 3) una corrispondenza perfetta tra forme bilineari simmetriche e forme quadratiche.

4: Definizione: Sia q una forma quadratica su E (k -spazio vettoriale di dimensione finita, $\text{ch}(k) \neq 2$). La forma bilineare simmetrica, f , associata a q si

chiama la forma bilineare polare di q . Il rango di q è, per definizione, il rango della sua forma bilineare polare.

4.1: Osservazione : Sia f la forma polare di q , B una base di E e $A = \text{mat}_B(f)$. Il discriminante di q (nella base B) è $\det(A)$. Se $A' = \text{mat}_B(f)$ allora ${}^t P A' P = A$ e $\det(A) = \det(A')(\det P)^2$, quindi il discriminante di q è definito a meno di un quadrato non nullo in k .

FORME QUADRATICHE E POLINOMI OMOGENEI DI GRADO DUE.

Ricordiamo (cfr II.§7, 12) la definizione di polinomio omogeneo:

5: Definizione: Sia $A = K[X_1, \dots, X_n]$ l'anello dei polinomi a coefficienti in k , nelle variabili X_1, \dots, X_n . Il polinomio $P(X_1, \dots, X_n) = \sum_{(i_1, \dots, i_n) \in \mathbb{N}^n} a_{i_1 \dots i_n} X_1^{i_1} \dots X_n^{i_n}$ (gli $a_{i_1 \dots i_n}$ tutti nulli tranne un numero finito) è omogeneo di grado (totale) k se $i_1 + \dots + i_n = k$ per ogni (i_1, \dots, i_n) tale che $a_{i_1 \dots i_n} \neq 0$.

5.1: Osservazione : Un polinomio del tipo $a_{i_1 \dots i_n} X_1^{i_1} \dots X_n^{i_n}$ ($a_{i_1 \dots i_n} \neq 0$) si chiama monomio. Il suo grado totale è $i_1 + \dots + i_n$. Ogni polinomio è somma di monomi. Un polinomio è omogeneo di grado (totale) k se e solo se è somma di monomi di grado totale k . In particolare un polinomio omogeneo non costante verifica $P(0, \dots, 0) = 0$.

(ii) In particolare un polinomio omogeneo di grado due nelle variabili X_1, \dots, X_n , si scrive: $P(X_1, \dots, X_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} X_i X_j$. Se $i = 1$ ci sono n valori possibili

per j . Se $i = 2$ ci sono $n-1$ valori possibili per j . In generale ci sono $n-(i-1)$ valori possibili di j ; quindi facendo variare i da 1 a n abbiamo: $n + (n-1) + (n-2) + \dots + 1 = n(n+1)/2$ valori possibili per (i,j) ovvero P è determinato dagli $n(n+1)/2$ coefficienti a_{ij} . Vediamo così che $K[X_1, \dots, X_n]_2 := \{\text{polinomi omogenei di grado due}\} \cup \{\text{polinomio nullo}\}$ è un k -spazio vettoriale di dimensione $n(n+1)/2$ (cfr Es. 1). Paragonare con la dimensione di $S_n(k)$, il sottospazio vettoriale delle matrici (n,n) simmetriche.

Torniamo alle nostre forme quadratiche.

Sia $B = (e_1, \dots, e_n)$ una base di E e $v = x_1 e_1 + \dots + x_n e_n$ un vettore di E . Per definizione $q(v) = f(v,v) = {}^t X A X = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{ij} x_i x_j$; quindi "q(v) è un polinomio

omogeneo di grado due nelle coordinate di v ". Cerchiamo di essere più precisi.

Siano p, t tali che $1 \leq p < t \leq n$. Nella sommatoria precedente abbiamo un termine $a_{pt} \cdot x_p \cdot x_t$ ($(i,j) = (p,t)$) e un termine $a_{tp} \cdot x_t \cdot x_p$ ($(i,j) = (t,p)$). Essendo la matrice A simmetrica $a_{pt} = a_{tp}$, quindi possiamo riscrivere:

$$q(v) = \sum_{1 \leq i \leq n} a_{ii} \cdot x_i^2 + \sum_{1 \leq i < j \leq n} 2 \cdot a_{ij} \cdot x_i \cdot x_j.$$

Adesso sia $Q(X_1, \dots, X_n) \in k[X_1, \dots, X_n]$

$$\text{il polinomio omogeneo di grado due: } Q(X_1, \dots, X_n) = \sum_{1 \leq i \leq j \leq n} q_{ij} \cdot X_i \cdot X_j \text{ dove } q_{ii} =$$

a_{ii} , $q_{ij} = 2 \cdot a_{ij}$ se $i \neq j$. Allora $q(v)$ è uguale allo scalare $Q(x_1, \dots, x_n)$ dove x_1, \dots, x_n

sono le coordinate di v nella base B . In queste condizioni si dice che *il polinomio Q rappresenta la forma quadratica q nella base B*.

Viceversa sia $Q(X_1, \dots, X_n) = \sum_{1 \leq i \leq j \leq n} b_{ij} \cdot X_i \cdot X_j$ un polinomio omogeneo di grado due. Poniamo $a_{ii} =$

b_{ii} , $1 \leq i \leq n$, e $a_{ji} = a_{ij} = b_{ij}/2$ se $i < j$ ($\text{ch}(k) \neq 2!$). La matrice $A = (a_{ij})$ è simmetrica e $Q(X) = {}^t X \cdot A \cdot X$ dove ${}^t X = (X_1, \dots, X_n)$.

Pertanto, una volta scelta una base B di E, abbiamo una biiezione tra $k[X_1, \dots, X_n]_2$ e $Q(E)$; quindi $k[X_1, \dots, X_n]_2$, $Q(E)$, $S^2(E)$ e $S_n(k)$ sono in biiezione.

Finalmente osserviamo che due polinomi omogeni di grado due $Q(X) = {}^t X \cdot A \cdot X$, $Q'(X) = {}^t X \cdot B \cdot X$ nelle indeterminate $(X_1, \dots, X_n) = {}^t X$, rappresentano la stessa forma quadratica in due basi di E se e solo se le matrici simmetriche A e B sono congruenti.

Esercizi:

2.1) Dimostrare che $k[X_1, \dots, X_n]_t := \{\text{polinomi omogeni di grado } t\} \cup \{\text{polinomio nullo}\}$ è un k -spazio vettoriale di dimensione $(n-1+t)! / [(n-1)!t!]$ (ragionare per induzione).

2.2) (i) Sia C la base canonica di \mathbb{R}^3 . Sia $f: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ la forma bilineare simmetrica tale

che $\text{mat}_C(f) = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$. Scrivere la forma quadratica associata (i.e. il corrispondente polinomio omogeneo di grado due) ad f rispetto alla base C .

(ii) Siano $u = (1, 1, 0)$, $v = (0, -1, 0)$, $w = (0, 1, 1)$. Mostrare che $B = (u, v, w)$ è una base di \mathbb{R}^3 . Determinare $\text{mat}_B(f)$ e la forma quadratica associata ad f rispetto alla base B .

2.3) Una forma quadratica q su E rappresenta uno scalare $\alpha \in k$ se esiste v in E tale che $q(v) = \alpha$. Scegliendo una base di E la relazione $q(v) = \alpha$ si scrive $Q(x_1, \dots, x_n) = \alpha$ dove (x_1, \dots, x_n) sono le coordinate di v nella base scelta. In particolare se possiamo trovare una base B tale che $\text{mat}_B(f) = I_n$, dove f è la forma polare di q , allora rispetto a questa base la relazione $q(v) = \alpha$ diventa $z_1^2 + \dots + z_n^2 = \alpha$ e abbiamo espresso α come somma di n quadrati in k . Ovviamente questo non è sempre possibile ($k = \mathbb{R}$ e $\alpha < 0$).

Dimostrare che se q è una forma quadratica non degenere su C allora q rappresenta ogni $\alpha \in C$. E' ancora vero questo se q è degenere?

2.4) Siano f, g due forme lineari sul k -spazio vettoriale E .

- (i) Dimostrare che $F: E^2 \rightarrow k: (u, v) \mapsto f(u)g(v) + g(u)f(v)$ è una forma bilineare simmetrica.
- (ii) Sia $E = \mathbb{R}^2$ e C la base canonica. Siano f, g le forme lineari determinate dalle seguenti condizioni: $f((1, 1)) = 1$, $\text{Ker}(f) = \{(x, y) / x+2y = 0\}$, $g((1, 1)) = 2$, $\text{Ker}(g) = \{(x, y) / x+y = 0\}$. Determinare $\text{mat}_C(F)$ (F definita come in (i)). Dire se F è degenera. Determinare la forma quadratica associata ad F nella base C .
- 2.5** (i) Sia k un campo infinito. Se $P(X), Q(X)$ sono due polinomi a coefficienti in k tali che $P(\alpha) = Q(\alpha)$ per ogni α in k tranne al più un numero finito, allora $P(X) = Q(X)$. È ancora vero questo risultato se k è un campo finito?
- (ii) Se $A \in M_n(k)$ si noterà con $P_A(X)$ il polinomio caratteristico di A . Se A è invertibile dimostrare che, per ogni B in $M_n(k)$, $P_{AB}(X) = P_{BA}(X)$ (cfr Es.II.16.12)
- (iii) Sia k un campo infinito. Usando (ii) mostrare che: $\forall (A, B) \in M_n(k)^2$, $P_{AB}(X) = P_{BA}(X)$ (se A non è invertibile considerare $A+\lambda I_n$ ed usare (i)). Dedurne che $\text{Tr}(AB) = \text{Tr}(BA)$.
- (iv) Dimostrare direttamente che (k campo qualsiasi): $\forall (A, B) \in M_n(k)^2$, $\text{Tr}(AB) = \text{Tr}(BA)$.
- (v) Dimostrare che $\text{Tr}: M_n(k) \times M_n(k) \rightarrow k: (A, B) \mapsto \text{Tr}(AB)$ è una forma bilineare simmetrica non degenera.

8: Corollario: Sia $v \in E$, allora v^\perp è un iperpiano se e solo se $v \notin \text{Ker}(f^*)$. In particolare se v non è isotropo allora v^\perp è un iperpiano e $E = \langle v \rangle \oplus v^\perp$.

Dim: Dalla proposizione precedente $v^\perp = f^*(v)^\circ$ quindi $v^\perp = E$ se e solo se $f^*(v) = 0$; altrimenti v^\perp è un iperpiano. Se v non è isotropo allora $f(v,v) \neq 0$, quindi $f_v \neq 0$ e v non appartiene a $\text{Ker}(f^*)$, quindi v^\perp è un iperpiano, inoltre $v \notin v^\perp$ (perché v non è isotropo) e quindi $E = \langle v \rangle \oplus v^\perp$ ♦

9: Proposizione: Sia $f : ExE \rightarrow k$ una forma bilineare simmetrica. Si ha $\text{rango}(f) = \text{rango}(f^*)$ (per la definizione di $\text{rg}(f)$, cfr §1, 10). In particolare f è non degenere se e solo se f^* è un isomorfismo.

Dim: È una conseguenza immediata delle definizioni e del lemma 5♦

9.1: Osservazione : Se f è una forma bilineare simmetrica non degenere, l'isomorfismo $f^* : E \rightarrow E^*$ permette di identificare in modo "naturale" rispetto ad f^* il suo duale; in questo caso l'ortogonalità rispetto ad f si identifica, tramite f^* , all'ortogonalità nel senso di II.§7 (cfr Prop.7), e possiamo usare i risultati di dualità. Per esempio:

10: Proposizione: Sia $f : ExE \rightarrow k$ una forma bilineare simmetrica non degenere e sia $U \subseteq E$ un sottospazio vettoriale.

- (i) $\dim(U^\perp) = \dim(E) - \dim(U)$
- (ii) $(U^\perp)^\perp = U$.

Dim: (i) Dalla Prop.7: $\dim(U^\perp) = \dim(f^*(U)^\circ)$. Per II.§7, Teo.13: $\dim(f^*(U)^\circ) = \dim(E^*) - \dim(f^*(U))$ ma $\dim(E^*) = \dim(E)$ perché E ha dimensione finita e $\dim(f^*(U)) = \dim(U)$ perché f^* è un isomorfismo.

(ii) $U^\perp = \{x \in E / f(x,u) = 0, \forall u \in U\}$ e $(U^\perp)^\perp = \{y \in E / f(y,x) = 0, \forall x \in U^\perp\}$; è chiaro che $U \subseteq (U^\perp)^\perp$ ma da (i) questi due sottospazi vettoriali hanno la stessa dimensione; quindi sono uguali ♦

Esercizi:

3.1) Dimostrare che l'ortogonale di un sottinsieme, X , di E , X^\perp , è un sottospazio vettoriale di E .

3.2) Sia E un k -spazio vettoriale e $f : ExE \rightarrow k$ una forma bilineare simmetrica non degenere. Sia V un sottospazio vettoriale di E e $f|_V$ la restrizione di f a $V \times V$. È $f|_V$ una forma bilineare simmetrica non degenere?

3.3) Sia E un k -spazio vettoriale e $f : ExE \rightarrow k$ una forma bilineare simmetrica. Dimostrare che f è non degenere se e solo se il suo radicale è ridotto al vettore nullo.

3.4) Sia $f: E \times E \rightarrow k$ una forma bilineare simmetrica e V un sottospazio vettoriale di E tale che V non contiene nessun vettore isotropo non nullo. Mostrare che $E = V^\perp \oplus V$.

3.5) Sia $f: E \times E \rightarrow k$ una forma bilineare simmetrica non degenere. Se U è un sottospazio vettoriale di E è sempre vero che $E = U^\perp \oplus U$?

3.6) Sia E un k -spazio vettoriale e $f: E \times E \rightarrow k$ una forma bilineare simmetrica. La forma f si dice anisotropa se non possiede vettori isotropi non nulli.

(i) Mostrare che se f è anisotropa allora f è non degenere. È vero il viceversa?

(ii) Se f è anisotropa e se U è un sottospazio vettoriale di E mostrare che $E = U^\perp \oplus U$.

3.7) Sia E un k -spazio vettoriale, $\text{ch}(k) \neq 2$, e $f: E \times E \rightarrow k$ una forma bilineare simmetrica non degenere. Un sottospazio U di E si dice (totalmente) isotropo se per ogni x in U , $f(x,x) = 0$. Dimostrare che se U è isotropo allora $\dim(U) \leq \dim(E) / 2$. Dare un esempio in cui l'uguaglianza è raggiunta.

3.8) Sia la forma bilineare simmetrica non degenere $\text{Tr}: M_n(\mathbb{R}) \times M_n(\mathbb{R}) \rightarrow \mathbb{R}$ (cfr Es.2.5). Determinare S^\perp dove S indica il sottospazio vettoriale delle matrici simmetriche.

4) Basi ortogonali (diagonalizzazione delle forme quadratiche).

In questo paragrafo si mostra che, su un campo qualsiasi di caratteristica diversa da due, ogni matrice simmetrica è congruente ad una matrice diagonale; in altri termini, ogni forma quadratica è equivalente ad una forma "senza termini misti".

1: Definizione: Sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica sul k -spazio vettoriale E . Una base $B = (e_1, \dots, e_n)$ di E si dice f -ortogonale (o ortogonale) se $f(e_i, e_j) = 0$ se $i \neq j$.

2: Lemma: La base B è f -ortogonale se e solo se $\text{mat}_B(f)$ è diagonale.

Dim: E' chiaro perché $\text{mat}_B(f) = (f(e_i, e_j))$, cfr §1, Def.3 ♦

3: Osservazione : Sia q la forma quadratica associata ad f e $v = x_1e_1 + \dots + x_ne_n$ un vettore di E . Se la base B è f -ortogonale $q(v) = \alpha_1x_1^2 + \dots + \alpha_nx_n^2$ dove $\alpha_i = f(e_i, e_i)$, quindi il polinomio omogeneo di grado due che rappresenta q nella base B è privo di termini misti $x_i x_j$ ($i \neq j$).

Il risultato principale sull'argomento è

4: Teorema: Sia E un k -spazio vettoriale di dimensione finita su k , con $\text{ch}(k) \neq 2$, e sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica, allora esiste una base di E , f -ortogonale.

Dim: Se f è identicamente nulla il teorema è evidente. Sia dunque $f \neq 0$ e ragioniamo per induzione su $n = \dim(E)$. Se $n = 1$, il teorema è vero (non c'è niente da dimostrare). Sia $n > 1$ e supponiamo il teorema dimostrato per $n-1$. Mostriamo che esiste un vettore non isotropo. Siccome $f \neq 0$, esistono v, w tali che $f(v, w) \neq 0$. Abbiamo $f(v+w, v+w) = f(v, v) + f(w, w) + 2f(v, w)$. Siccome $\text{ch}(k) \neq 2$ e $f(v, w) \neq 0$, $2f(v, w) \neq 0$. Pertanto uno dei tre termini $f(v+w, v+w)$, $f(v, v)$, $f(w, w)$ è non nullo. Sia dunque u un vettore non isotropo. Da §3, Cor.8, $H = u^\perp$ è un iperpiano e $E = \langle u \rangle \oplus u^\perp$. Sia $f' : H \times H \rightarrow k$ la restrizione di f a $H \times H$ ($f'(x, y) = f(x, y), \forall (x, y) \in H^2$). L'applicazione f' è una forma bilineare simmetrica sul k -spazio vettoriale H di dimensione $n-1$. Per ipotesi di induzione esiste una base di H , $B' = (e_1, \dots, e_{n-1})$, f' -ortogonale. Visto che $E = \langle u \rangle \oplus H$, $B = (e_1, \dots, e_{n-1}, u)$ è f -ortogonale.

$e_n = u$) è una base di E (Es.II.4.13). La base B è f -ortogonale. Infatti se $i \neq n$ e $k \neq n$ allora $f(e_i, e_k) = f(e_i, e_n) = 0$ perché B' è f -ortogonale per ipotesi di induzione; inoltre $f(e_i, e_n) = f(e_i, u) = 0$ se $i \neq n$ perché $H = u^\perp$ ♦

4.1: Osservazione : L'ipotesi $\text{ch}(k) \neq 2$ è essenziale: siano $k = \mathbb{Z}/2\mathbb{Z}$ e E un k -spazio vettoriale di dimensione due. Se $B = (e_1, e_2)$ è una base di E definiamo una forma bilineare simmetrica $f : E \times E \rightarrow k$ tramite $f(e_1, e_1) = f(e_2, e_2) = 0$, $f(e_1, e_2) = f(e_2, e_1) = 1$; quindi $\text{mat}_B(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In virtù di §1, (f) (o Prop.5), f è ben definita. Sia $v = \alpha e_1 + \beta e_2$ un vettore qualsiasi di E . Abbiamo $f(v, v) = 2\alpha\beta = 0$ (quindi la forma quadratica associata ad f è identicamente nulla). Se $B' = (v_1, v_2)$ è una base f -ortogonale allora $f(v_1, v_2) = f(v_2, v_1) = 0$ (per definizione) e, per quanto appena visto $f(v_i, v_i) = 0$, $1 \leq i \leq 2$. Quindi $\text{mat}_{B'}(f) = 0$ e per §1, Prop.5, $f = 0$; in contraddizione con $f(e_1, e_2) = 1$. Quindi non esiste nessuna base f -ortogonale.

Altri modi di formulare il teorema 4:

5: Corollario: Sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica ($\text{ch}(k) \neq 2$) allora esiste una base B di E tale che $\text{mat}_B(f)$ sia diagonale.

O ancora:

6: Corollario: Ogni matrice simmetrica $A \in M_n(k)$ ($\text{ch}(k) \neq 2$) è congruente ad una matrice diagonale.

6.1: Osservazione : Il corollario 6 non implica che ogni matrice simmetrica sia diagonalizzabile. Infatti A è diagonalizzabile se esiste una matrice, N , invertibile tale che $N^{-1}A.N = D$ dove D è diagonale. Invece A è congruente a D se esiste P invertibile tale che ${}^t P.A.P = D$. Per poter dedurre dal corollario 6 che A è diagonalizzabile bisognerebbe dimostrare che si può prendere P tale che: ${}^t P = P^{-1}$ (questo è equivalente a: ${}^t P.P = I_n$); questo non risulta affatto dalla dimostrazione del corollario 6! Osserviamo inoltre che se ${}^t P = P^{-1}$ allora $\det(P) = 1/\det(P)$ e pertanto $\det(P) = \pm 1$.

7: Definizione: Una matrice $P \in M_n(k)$ tale che ${}^t P.P = I_n$ si dice ortogonale. L'insieme delle matrici ortogonali si nota $O_n(k)$. Per la moltiplicazione delle matrici $O_n(k)$ è un gruppo, chiamato il gruppo ortogonale di ordine n . Il sottinsieme di $O_n(k)$ costituito dalle matrici con determinante uguale a 1 è un

sottogruppo di $O_n(k)$ chiamato gruppo ortogonale speciale di ordine n , e denotato $SO_n(k)$ o $O^+_n(k)$. Finalmente si definisce $O^-_n(k) := O_n(k) \setminus O^+_n(k)$.

7.1: Osservazione : (i) $O^-_n(k)$ non è un gruppo per la moltiplicazione delle matrici (perché?).

(ii) Vedremo più avanti che, in effetti, ogni matrice simmetrica reale è diagonalizzabile.

Esercizi:

4.1) Siano E un \mathbb{R} -spazio vettoriale di dimensione tre, $B = (e_1, e_2, e_3)$ una base di E e

$$f : E \times E \rightarrow \mathbb{R} \text{ la forma bilineare simmetrica tale che } \text{mat}_B(f) = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}.$$

(i) Mostrare che e_1 non è isotropo per f .

(ii) Determinare e_1^\perp .

(iii) Sia $a = e_1 + e_3$. Osservare che $a \in e_1^\perp$. Mostrare che a non è isotropo per f e determinare a^\perp .

(iv) Se b è un vettore non nullo in $e_1^\perp \cap a^\perp$, dimostrare che $B' = (e_1, a, b)$ è una base ortogonale di E ; scrivere la matrice di f in questa base.

4.2) Diagonalizzare la forma quadratica reale $q(x, y, z) = x^2 + 4xy + 2yz - z^2$.

4.3) Sia $V = \{A \in M_2(\mathbb{R}) / A \text{ è simmetrica}\}$ e $q : V \rightarrow \mathbb{R} : A \mapsto \det(A)$.

(i) Dimostrare che q è una forma quadratica su V . Dare la matrice della forma bilineare associata nella base "standard" di V .

(ii) Sia $W = \{A \in V / \text{tr}(A) = 0\}$. Dimostrare che W è un sottospazio di V e determinare la sua dimensione. Dimostrare che la forma bilineare associata a q è definita negativa su W .

(iii) Determinare W^\perp .

(iv) Determinare una base ortogonale.

5) Basi ortonormali; teorema di Sylvester.

Il risultato principale di questo paragrafo è il teorema di Sylvester che mostra l'esistenza di una forma canonica (o ridotta) per ogni forma quadratica reale. Praticamente si userà il metodo di Gauss (cfr Es. 1, 2, 3) per determinare tale forma ridotta.

1: Definizione: Sia E un k -spazio vettoriale e $f : E \times E \rightarrow k$ una forma bilineare simmetrica. Si dice che k vettori distinti x_1, \dots, x_k di E formano una famiglia f -ortonormale se sono due a due ortogonali ($f(x_i, x_j) = 0$ se $i \neq j$) e se $f(x_i, x_i) = 1$, $1 \leq i \leq k$.

Una base ortonormale di E è una base formata da vettori di una famiglia f -ortonormale.

2: Lemma: Se B è una base f -ortonormale allora $\text{mat}_B(f) = I_n$. In particolare f è non degenera.

Dim: Segue immediatamente dalle definizioni ♦

3: Lemma: Se x_1, \dots, x_k formano una famiglia f -ortonormale allora x_1, \dots, x_k sono linearmente indipendenti.

Dim: Sia $\lambda_1 x_1 + \dots + \lambda_k x_k = z = 0$ una relazione di dipendenza lineare. Abbiamo $f(z, x_i) = f(0, x_i) = 0$ ma, d'altra parte $f(z, x_i) = f(\lambda_1 x_1 + \dots + \lambda_k x_k, x_i) = \lambda_i f(x_i, x_i)$ (per ortogonalità) e $\lambda_i f(x_i, x_i) = \lambda_i$ (per ortonormalità). In conclusione $\lambda_i = 0$, $1 \leq i \leq k$, e x_1, \dots, x_k sono linearmente indipendenti ♦

4: Corollario: Se x_1, \dots, x_k formano una famiglia f -ortonormale allora $\text{rg}(f) \geq k$.

Dim: Dal lemma 3, x_1, \dots, x_k sono liberi; si può completarli ad una base $B = (x_1, \dots, x_k, x_{k+1}, \dots, x_n)$ di E . Sia $A = \text{mat}_B(f)$. La sottomatrice di A costruita sulle prime k righe e sulle prime k colonne è I_k . Dunque A ha un minore non nullo di ordine k e pertanto $\text{rg}(A) = k$ ♦

Quindi, rispetto all'esistenza di basi ortonormali, bisogna tenere conto del rango di f , perciò per avere un risultato generale (valido per ogni forma bilineare simmetrica) bisogna chiedere un po' meno. Se k è algebricamente

chiuso (per es. $k = \mathbb{C}$), di caratteristica diversa da 2, il seguente risultato è ottimale:

5: Teorema: *Sia E un k -spazio vettoriale di dimensione n con k algebricamente chiuso, di caratteristica diversa da 2. Sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica. Esiste una base, B , di E tale che $\text{mat}_B(f)$ sia della forma: $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, dove $r = \text{rg}(f)$ e dove gli zeri indicano matrici nulle di ordini opportuni.*

Dim: Da §4, Teo.4, esiste una base, $B' = (e'_1, \dots, e'_n)$, f -ortogonale. Quindi $\text{mat}_{B'}(f) = (a_{ij})$ è diagonale e $f(e'_i, e'_j) = a_{ii}$. Salvo riordinare gli indici possiamo supporre $a_{ii} \neq 0$, $1 \leq i \leq r$, $a_{ii} = 0$ se $i > r$, dove $r = \text{rg}(f)$. Siano $\alpha_1, \dots, \alpha_r$ degli scalari tali che $\alpha_i^2 = a_{ii}$, $1 \leq i \leq r$. Gli α_i esistono perché k è algebricamente chiuso (i polinomi $X^2 - a_{ii}$ hanno le loro radici in k). Poniamo $e_i = e'_i / \alpha_i$ se $1 \leq i \leq r$, e $e_i = e'_i$ se $i > r$. E' chiaro che $B = (e_1, \dots, e_n)$ è ancora una base di E . Per $1 \leq i \leq r$, abbiamo $f(e_i, e_i) = f(e'_i / \alpha_i, e'_i / \alpha_i) = f(e'_i, e'_i) / \alpha_i^2 = 1$ e $\text{mat}_B(f)$ è della forma annunciata♦

6: Corollario: *Se k è algebricamente chiuso, $\text{ch}(k) \neq 2$, ogni matrice simmetrica di rango r , a coefficienti in k , è congruente a $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.*

6.1: Osservazione : L'ipotesi $\text{ch}(k) \neq 2$ è necessaria per usare il teorema 4 del §4. L'ipotesi k algebricamente chiuso serve per affermare che gli elementi a_{ii} hanno una radice quadrata in k ; quindi il teorema 5 è vero sotto l'ipotesi più debole che ogni elemento di k è un quadrato in k (i.e. se $a \in k$, $X^2 - a$ ha una radice in k).

Se $k = \mathbb{R}$ si ha una versione più debole del teorema 5:

7: Teorema: (Sylvester): *Sia E un \mathbb{R} -spazio vettoriale di dimensione n e $f : E \times E \rightarrow \mathbb{R}$ una forma bilineare simmetrica. Esistono un numero naturale p , dipendente solo da f e una base, B , di E tali che: $0 \leq p \leq r$ ($r := \text{rango}(f)$) e $\text{mat}_B(f) = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$, dove gli zeri indicano matrici nulle di ordini opportuni.*

Dim: Da §4, Teo.4, esiste una base $B' = (e'_1, \dots, e'_n)$, f-ortogonale. Quindi $\text{mat}_{B'}(f) = (a_{ij})$ è diagonale e $f(e'_i, e'_j) = a_{ii}$. Riordinando semmai gli indici possiamo supporre: $a_{ii} > 0$ se $1 \leq i \leq p$, $a_{ii} < 0$ se $p+1 \leq i \leq r$ e $a_{ii} = 0$ se $i > r$. Se a_{ii} è positivo allora ammette una radice quadrata α_i : $\alpha_i^2 = a_{ii}$. Se $a_{ii} < 0$ allora esiste α_i tale che $\alpha_i^2 = -a_{ii}$. Come nella dimostrazione del teorema 4, §4 si pone: $e_i = e'_i / \alpha_i$ se $1 \leq i \leq r$; $e_i = e'_i$ se $i > r$. Allora $B = (e_1, \dots, e_n)$ è una base di E e $\text{mat}_B(f)$ ha la forma annunciata. Rimane da mostrare che il numero p di elementi positivi sulla diagonale non dipende dalle basi scelte. Per questo consideriamo la forma quadratica, q , associata ad f . Se $v = x_1e_1 + \dots + x_ne_n$ allora $q(v) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2$ (*). Supponiamo che in un'altra base f-ortogonale $B'' = (f_1, \dots, f_n)$ si abbia $q(v) = y_1^2 + \dots + y_t^2 - y_{t+1}^2 - \dots - y_r^2$ (†) per ogni $v = y_1f_1 + \dots + y_nf_n$. Osserviamo che il numero di termini non nulli è sempre uguale al rango, in entrambe le basi. Se $p \neq t$ possiamo supporre, per esempio, $t < p$. Consideriamo i seguenti sottospazi vettoriali di E: $F := \langle e_1, \dots, e_p \rangle$, $G := \langle f_{t+1}, \dots, f_n \rangle$. Abbiamo $\dim(F) + \dim(G) = p + (n-t) = n + (p-t) > n$; quindi $F \cap G \neq \{0\}$ (usare la relazione di Grassmann, II.§5). Sia $w \neq 0$ un vettore di $F \cap G$. Siccome $w \in F$, $w = x_1e_1 + \dots + x_pe_p$ e da (*) segue che $q(w) = x_1^2 + \dots + x_p^2 > 0$ (c'è un x_i , $1 \leq i \leq p$, diverso da zero perché $w \neq 0$). D'altra parte $w \in G$ quindi $w = y_{t+1}f_{t+1} + \dots + y_nf_n$ e da (†): $q(w) = -y_{t+1}^2 - \dots - y_r^2 \leq 0$; otteniamo una contraddizione. Questo conclude la dimostrazione ♦

8: Definizione: Con le notazioni del teorema 7, p e $r-p$ sono gli indici di positività, negatività di f (o di q). La coppia $(p, r-p)$ è la **segnatura** di f (risp. q). Se $B = (e_1, \dots, e_n)$ è la base del teorema 7 e se $v = x_1e_1 + \dots + x_ne_n$ allora $q(v) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2$. In queste condizioni si dice che q è in **forma canonica**.

9: Osservazione : Il teorema di Sylvester viene anche chiamato "legge dell'inerzia".

10: Corollario: Ogni matrice simmetrica reale, A , è congruente ad una matrice del tipo $\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$ dove $r = \text{rg}(A)$ e dove p dipende solo dalla classe di congruenza di A .

11: Definizione: Sia E un R-spazio vettoriale. Una forma bilineare simmetrica f : $E \times E \rightarrow \mathbb{R}$ si dice:

- positiva se $q(v) = f(v,v) \geq 0$ per ogni v in E .
- definita positiva se $q(v) = f(v,v) > 0$ per ogni $v \neq 0$ in E .
- negativa se $q(v) = f(v,v) \leq 0$ per ogni v in E .
- definita negativa se $q(v) = f(v,v) < 0$ per ogni $v \neq 0$ in E .
- indefinita se f non è né positiva né negativa.

11.1: Osservazione : (i) Dal teorema di Sylvester segue che data una forma bilineare simmetrica esiste sempre una base rispetto alla quale la forma quadratica associata ha un'espressione del tipo (detta forma "ridotta"): $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2$. Si vedrà in esercizi (Es.1, 2) un metodo pratico ("metodo di Gauss") per ricondursi a tale forma.

(ii) Dal teorema di Sylvester segue che se f è definita positiva o definita negativa allora f è non degenere. Ci sono però delle forme indefinite non degeneri.

(iii) Questa terminologia ha senso perché $k = \mathbb{R}$; su un corpo qualsiasi non abbiamo a priori la nozione di positivo, negativo; per questo bisognerebbe disporre di un "ordine compatibile" su k . Un "ordine compatibile" su k è un sottoinsieme, P , di k tale che k è la riunione disgiunta di P , $\{0\}$ e $-P$ in altre parole se x appartiene a k allora una ed una sola delle tre condizioni seguenti è soddisfatta: $x \in P$, $x = 0$, $-x \in P$. Inoltre la seguente condizione di compatibilità con le operazioni dev'essere soddisfatta: **(O2)** per ogni x, y in P , $x+y$ e xy appartengono ancora a P .

Se $y - x \in P$ si scrive $x < y$ (o $y > x$). Se $y < 0$ si dice che y è negativo; si ha: $x \in P \Leftrightarrow x > 0$ (si dice che x è positivo). Si scrive anche $x \geq y$ per indicare: $x > y$ o $x = y$. Osserviamo che se k è compatibilmente ordinato allora $1 > 0$ infatti $1 \neq 0$, quindi $1 \in P$, e abbiamo finito, o $-1 \in P$ e da **(O2)**: $(-1)^2 = 1 \in P$ appartiene a P . Segue anche che -1 è negativo (altrimenti da **(O2)**: $1 + (-1) = 0 \in P$, assurdo). Lo stesso ragionamento mostra che per ogni $x \neq 0$ in k , x^2 è positivo. In particolare non esiste nessun ordine compatibile su C (c'è un quadrato negativo: $i^2 = -1$).

Quindi per estendere le nozioni di forma bilineare simmetrica positiva ecc... a C bisogna seguire un'altra strada, per esempio "imporre" che $f(v,v)$ sia reale per ogni v in E ; questo punto di vista è sviluppato nella teoria delle forme hermitiane.

(iv) Le forme bilineari simmetriche definite positive su uno spazio vettoriale reale sono fondamentali per la geometria euclidea (sono i prodotti scalari), ma

anche altri tipi di forme bilineari simmetriche rivestono una grande importanza; per esempio lo "spazio tempo" della relatività ristretta è lo spazio affine \mathbb{R}^4 con la forma quadratica indefinita $q(v) = x_1^2 + \dots + x_3^2 - c \cdot x_4^2$ (detta forma di Minkowski) dove c è la velocità della luce. Se $q(v) > 0$ (risp. $q(v) < 0$, $q(v) = 0$), v si dice vettore spazio (risp. vettore tempo, vettore luce).

Esercizi:

5.1) Sia $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ una matrice, 2×2 , simmetrica, a coefficienti in k ($\text{ch}(k) \neq 2$). Sia E un k -spazio vettoriale di dimensione due, e $B = (e_1, e_2)$ una base di E . Siano f la forma bilineare simmetrica tale che $\text{mat}_B(f) = M$, e q la forma quadratica associata. Se $v = xe_1 + ye_2$, allora $q(v) = {}^t X M X = ax^2 + cy^2 + 2bxy$ (${}^t X = (x, y)$). Sappiamo che $q(f)$ può essere diagonalizzata e quindi messa in forma ridotta (senza termine misto xy). Scopo dell'esercizio è dare un metodo effettivo ("metodo di Gauss") per ottenere una forma ridotta.

(i) Si suppone $a \neq 0$. L'idea è di considerare $x^2 + 2bxy$ come l'inizio dello sviluppo di un quadrato: $x^2 + 2bxy = (x + by)^2 - b^2y^2$. Verificare che $q(v) = (ax + by)^2/a + (ac - b^2)y^2/a$.

(ii) Trovare una matrice invertibile, P , tale che ${}^t P M P = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & c - \frac{b^2}{a} \end{pmatrix}$.

(iii) Sia $D = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & c - \frac{b^2}{a} \end{pmatrix}$. Mostrare che, in generale, D non è simile a M (dare una condizione necessaria affinché D e M siano simili).

(iv) Si suppone $a = 0$ e $b \neq 0$. Scrivere $q(v)$ nella forma $\alpha L(v) \cdot L'(v)$ dove L, L' sono due forme lineari su E . Osservare che L e L' sono linearmente indipendenti. Tenuto conto che $4L \cdot L' = (L + L')^2 - (L - L')^2$, trovare una matrice invertibile, Q , tale che ${}^t Q M Q$ sia diagonale.

5.2) Si passa adesso al caso generale. Sia q una forma quadratica su E , k -spazio vettoriale di dimensione n ($\text{ch}(k) \neq 2$): $q(v) = \sum_{i=1}^n a_{ii}(x_i)^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j = P(x_1, \dots, x_n)$, P omogeneo di grado due in n variabili. Si tratta di dimostrare, per induzione su n , che si può sempre scrivere q nella forma $\sum_{1 \leq i \leq n} \lambda_i L_i^2$, dove L_1, \dots, L_r sono delle forme lineari (polinomi omogenei di grado uno) linearmente indipendenti. Facendo il cambiamento di variabili $y_i = L_i(x_1, \dots, x_n)$ si avrà $q(v) = \sum \lambda_i y_i^2$; in particolare se $k = \mathbb{R}$ la segnatura di q sarà $(p, r-p)$ dove il rango r è uguale a $\#\{\lambda_i / \lambda_i \neq 0\}$ e dove $p = \#\{\lambda_i / \lambda_i > 0\}$.

Più precisamente sia f la forma polare di q e sia $M = \text{mat}_B(f)$, $B = (e_1, \dots, e_n)$. Sia $R = (a_{ij})$ dove $L_i = \sum a_{ij} e_j^*$.

(i) Dimostrare che R è invertibile e che $M = {}^t R D R$ dove $D = (d_{ij})$ è la matrice diagonale con $d_{ii} = \lambda_i$.

Procediamo adesso nella dimostrazione dell'esistenza della forma ridotta. La dimostrazione è per induzione e sfrutta essenzialmente le identità algebriche: $x^2 + 2bxy = (x+by)^2 - b^2y^2$, $4xy = (x+y)^2 - (x-y)^2$.

Se $n = 1$ non c'è niente da dimostrare (per $n = 2$ cfr Es.1).

Supponiamo l'asserto vero per $n-1$ e sia $q(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii}(x_i)^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j$.

Distinguiamo due casi:

(a) Compare il termine x_1^2 (cioè $a_{11} \neq 0$):

scriviamo q nella forma seguente: $q = ax_1^2 + 2Rx_1 + S$, dove $a = a_{11}$, R è una forma lineare nelle variabili x_2, \dots, x_n e dove S è una forma quadratica nelle variabili x_2, \dots, x_n . Adesso possiamo riscrivere q nella forma: $q = a(x_1 + R/a)^2 - (R/a)^2 + S$. Per ipotesi di induzione $S - (R/a)^2 = \sum_{2 \leq i \leq n} \lambda_i L_i^2$, dove L_2, \dots, L_n sono forme lineari, linearmente indipendenti, nelle variabili x_2, \dots, x_n .

(ii) Sia $L = x_1 + R/a$. Dimostrare che L_1, L_2, \dots, L_n sono linearmente indipendenti e concludere.

(b) Non compare il termine x_1^2 (più generalmente non compaiono termini quadrati):

Possiamo assumere che compaia un termine misto $x_1 x_i$ (altrimenti q è una forma nelle variabili x_2, \dots, x_n). A meno di riordinare gli indici possiamo supporre $i = 2$ e scrivere q nella forma: $q = ax_1 x_2 + Rx_1 + x_2 S + T$, dove $a \neq 0$ e dove R, S sono forme lineari nelle variabili x_3, \dots, x_n e T è una forma quadratica nelle variabili x_3, \dots, x_n . Adesso: $q = a[(x_1 + S/a)(x_2 + R/a)] + T - RS/a$. Si pone $L = x_1 + S/a$, $L' = x_2 + R/a$ e si ha: $q = a[(L+L')^2 - (L-L')^2]/4 + T - RS/a$.

Per ipotesi di induzione: $T - RS/a = \sum_{3 \leq i \leq n} \lambda_i L_i^2$, dove L_3, \dots, L_n sono forme lineari, linearmente indipendenti, nelle variabili x_3, \dots, x_n .

(iii) Sia $L_1 = L+L'$, $L_2 = L-L'$. Mostrare che $L_1, L_2, L_3, \dots, L_n$ sono linearmente indipendenti e concludere.

5.3) Si prende $k = \mathbb{R}$, $E = \mathbb{R}^3$ con la base canonica. Determinare la segnatura delle forme bilineari simmetriche di forma quadratica associata q dove:

$$(a) q(v) = x_1^2 + x_2^2 + x_3^2 - 4(x_2 x_3 + x_1 x_3 + x_2 x_1);$$

$$(b) q(v) = x_1^2 + x_2^2 + x_3^2 - (x_2 x_3 + x_1 x_3 + x_2 x_1)$$

$$(c) q(v) = x_1^2 + 6x_2^2 - 4x_2 x_1 + 8x_1 x_3$$

5.4) Riprendere l'esercizio 4.3 e usando il metodo di Gauss determinare la segnatura di q .

5.5) Mostrare che non si possono trovare due numeri interi ($\in \mathbb{Z}$) non entrambi nulli, x, y , tali che $3x^2 + 5y^2 - 10xy = 0$.

5.6) Si consideri l'equazione $ax^2 + bx + c = 0$ (\dagger), con $a, b, c \in \mathbb{R}$, $a \neq 0$. Dimostrare, con la teoria delle forme quadratiche, che se $D = b^2 - 4ac < 0$ allora (\dagger) non ha nessuna soluzione in \mathbb{R} . Si osserverà che si può supporre $a > 0$.

5.7) Sia su \mathbb{R}^3 la forma quadratica $f(x, y, z) = px^2 + (p-q)y^2 - 2pxy - qz^2 + 2qyz$, dove $p \neq q$ sono due numeri reali strettamente positivi.

(i) Trovare il rango e la segnatura di f .

- (ii) Determinare $(\mathbf{R}^3)^\perp$.
 (iii) Dimostrare che, se p e q sono due numeri primi, l'equazione $f(x, y, 0) = 0$ non ha soluzioni non banali in \mathbf{Z}^2 .

5.8) Si consideri su \mathbf{R}^n la forma quadratica $q(x_1, \dots, x_n) = \sum a_{ii}x_i^2 + \sum_{i < j} 2a_{ij}x_i x_j$, e $A = (a_{ij})$ la matrice simmetrica associata. Sia inoltre $q'(x_1, \dots, x_{n-1}) = q(x_1, \dots, x_{n-1}, 0)$; q' è una forma quadratica su \mathbf{R}^{n-1} .

- (i) Dimostrare che se q è definita positiva allora q' è definita positiva e $\det(A) > 0$.
 (ii) Dimostrare il viceversa: se q' è definita positiva e $\det(A) > 0$ allora q è definita positiva.
 (iii) Per ogni r , $1 \leq r \leq n$, sia A_r la sottomatrice di A costruita sulle prime r righe e le prime r colonne di A . Dimostrare che q è definita positiva se e solo se $\det(A_r) > 0$ per ogni r , $1 \leq r \leq n$ (suggerimento: in una delle due implicazioni, usare (ii) e procedere per induzione).

6) Spazi metrici.

Questo paragrafo, di natura puramente culturale, può essere omesso in prima lettura. Si introducono le nozioni di spazio metrico e di spazio topologico (Es. 2) e si mostra come la nozione di applicazione continua si estenda a queste situazioni più generali.

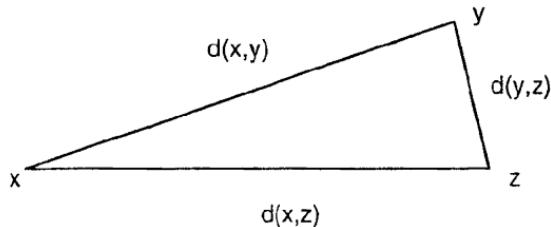
1: Definizione: Sia X un insieme. Una distanza su X è un'applicazione $d : X \times X \rightarrow \mathbb{R}$ soddisfacente le proprietà seguenti:

(I) $d(x,y) \geq 0$ per ogni x, y in X , inoltre: $d(x,y) = 0 \Leftrightarrow x = y$

(II) per ogni x, y : $d(x,y) = d(y,x)$

(III) $d(x,z) \leq d(x,y) + d(y,z)$ per ogni terna (x,y,z) di elementi di X ("*disuguaglianza triangolare*").

Gli assiomi (I), ..., (II) sono molto naturali; (III) esprime che la "strada più corta tra due punti è la retta che li congiunge" (anche se non sappiamo che cos'è una retta in X):



1.1: Esempi: (i) Se $X = \mathbb{R}$ l'applicazione $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$: $(x,y) \mapsto |x-y|$ è una distanza su \mathbb{R} (è la distanza usuale).

(ii) Se $X = \mathbb{R}^3$ l'applicazione $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$: $(x,y) \mapsto [(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2]^{1/2}$ è una distanza, chiamata distanza euclidea "usuale"; la ritroveremo dopo.

(iii) Sia A un insieme e $X = \{f / f : A \rightarrow \mathbb{R}\}$ è un'applicazione limitata}. Osservare che se f, g sono due elementi di X allora anche $f-g$ appartiene a X . Si pone $d(f,g) = \sup_{t \in A} |f(t) - g(t)|$, allora d è una distanza su X .

(iv) Sia X un insieme qualsiasi e, per ogni x, y in X , poniamo $d(x,y) = 1$ se $x \neq y$, $d(x,x) = 0$. Si verifica facilmente che d è una distanza (chiamata distanza banale o discreta). Le verifiche di queste asserzioni sono lasciate al lettore.

2: Definizione: Uno spazio metrico è una coppia (X, d) dove X è un insieme e d una distanza su X .

3: Definizione: Siano (X, d) , (X', d') due spazi metrici. Un'applicazione $f: X \rightarrow X'$ è un'isometria se conserva le distanze: $d'(f(x), f(y)) = d(x, y)$ per ogni $(x, y) \in X^2$.

3.1: Osservazione : (i) Se f è un'isometria biiettiva allora anche f^{-1} è un'isometria (Es.1).

(ii) Sia (X, d) uno spazio metrico, Y un insieme qualsiasi e $f: X \rightarrow Y$ una biiezione. Si può definire una distanza, d' , su Y tramite: $d'(y, y') := d(f^{-1}(y), f^{-1}(y'))$. La distanza d' (verificare che è una distanza!) si dice ottenuta da d per trasporto tramite f . Per le distanze d, d' , f è un'isometria.

4: Definizione: Sia (X, d) uno spazio metrico e $x \in X$. Il disco aperto (risp. chiuso) di centro x e raggio r è il sottinsieme $D(x, r)$ (risp. $D'(x, r)$) definito da $D(x, r) = \{y \in X / d(x, y) < r\}$ (risp. $D'(x, r) = \{y \in X / d(x, y) \leq r\}$).

La sfera di centro x e raggio r è il sottinsieme $S(x, r) = \{y \in X / d(x, y) = r\}$.

5: Definizione: Sia (X, d) uno spazio metrico. Un sottinsieme A di X si dice aperto (per la distanza d) se vale la seguente proprietà: $\forall x \in A, \exists r > 0$ tale che il disco aperto di centro x raggio r , $D(x, r)$, sia contenuto in A .

5.1: Osservazione : (i) L'insieme vuoto è aperto (sempre!); pure X è aperto.

(ii) Un disco aperto è un sottinsieme aperto (c'è qualcosa da dimostrare!).

(iii) Si può definire in un contesto più generale (spazi topologici) la nozione di sottinsieme aperto ecc...

6: Definizione: Sia (X, d) uno spazio metrico e $x \in X$. Un intorno aperto di x è un sottinsieme aperto di X contenente x .

7: Definizione: Siano (X, d) , (X', d') due spazi metrici, $f: X \rightarrow X'$ un'applicazione e $x \in X$. L'applicazione f è continua nel punto x se per ogni intorno aperto, V' , di $f(x)$ in X' esiste un intorno aperto, V , di x in X tale che $f(V) \subseteq V'$.

L'applicazione f è continua se è continua in ogni punto x di X .

7.1: Osservazione : (i) Si consiglia vivamente di mostrare che nel caso $X = X' = \mathbb{R}$, con la distanza usuale, la definizione 7 è equivalente alla definizione classica "con gli ε "(Es. 3).

(ii) La definizione 7 usa solo il concetto di sottinsieme aperto, quindi la nozione di continuità è definita nel contesto generale degli spazi topologici (Es. 2). Lo studio degli spazi topologici è l'oggetto della topologia.

Esercizi:

6.1) Sia (X,d) uno spazio metrico e A, B due sottinsiemi di X . La distanza da A a B , $d(A,B)$, è definita tramite: $d(A,B) := \inf_{x \in A, y \in B} d(x,y)$.

(i) Se $A \cap B \neq \emptyset$ allora $d(A,B) = 0$.

(ii) Sia $A = N^*$ (N^* è l'insieme dei naturali ≥ 1) e sia $B = \{n - 1/n \mid n \in N, n \geq 2\}$. Osservare che $A \cap B = \emptyset$. Se d indica la distanza usuale su \mathbf{R} (cfr 1.1(i)), mostrare che $d(A,B) = 0$.

(iii) Un sottinsieme A di X si dice limitato se è contenuto in un disco. Mostrare che A è limitato se e solo se $\sup_{x \in A, y \in A} d(x,y) =: d(A)$ è finito.

6.2) Sia τ la famiglia dei sottinsiemi aperti di uno spazio metrico, X . Mostrare che:

(T1) X e \emptyset appartengono a τ .

(T2) Se $U_i \in \tau, \forall i, i \in I$, allora $\cup_{i \in I} U_i \in \tau$ (un'unione qualsiasi di elementi di τ è un elemento di τ)

(T3) Se $U_i \in \tau, 1 \leq i \leq n$, allora $(U_1 \cap U_2 \cap \dots \cap U_n) \in \tau$ (l'intersezione di un numero finito di elementi di τ è un elemento di τ).

Siano Y un insieme qualsiasi e τ una famiglia di sottinsiemi di Y (i.e. τ è un sottinsieme dell'insieme delle parti di Y) soddisfacente (T1), ..., (T3), allora τ definisce una struttura di spazio topologico su Y ; gli elementi di τ sono gli aperti di questa topologia.

6.3) Mostrare che nel caso $X = X' = \mathbf{R}$, con la distanza usuale, la definizione 7 è equivalente alla definizione classica "con gli ε ".

6.4) Sia $f: (X,d) \rightarrow (X',d')$ un'applicazione tra due spazi metrici. Dimostrare che f è continua se e solo se per ogni sottinsieme aperto, V' , di X' , $f^{-1}(V')$ è aperto in X .

7) Spazi vettoriali normati.

Si introduce la nozione di norma su uno spazio vettoriale reale (o complesso) e si mostra come la norma permetta di definire una distanza.

In tutto questo paragrafo E indica uno spazio vettoriale sul campo dei numeri reali o sul campo dei numeri complessi.

Si ricorda che se $z = a+ib$ è un numero complesso, il coniugato di z è il numero complesso $z^* = a-ib$. Il modulo di z è $|z| = \sqrt{z \cdot z^*} = \sqrt{a^2+b^2}$. Osserviamo in particolare che $|z|$ è un numero reale; inoltre se $z \in \mathbb{R}$ allora $|z| = \sqrt{z^2}$ è il valore assoluto di z .

1: Definizione: Sia E un k-spazio vettoriale ($k = \mathbb{R}, \mathbb{C}$). Una norma, $\|\cdot\|$, su E è un'applicazione $E \rightarrow \mathbb{R}$ soddisfacente le seguenti proprietà:

- (I) $\|x\| \geq 0$ per ogni x in E
- (II) $\|x\| = 0$ se e solo se $x = 0$.
- (III) $\|\lambda x\| = |\lambda| \|x\|$, per ogni x in E e per ogni scalare λ in k.
- (IV) $\|x+y\| \leq \|x\| + \|y\|$.

Un k-spazio vettoriale munito di una norma si dice spazio vettoriale normato.

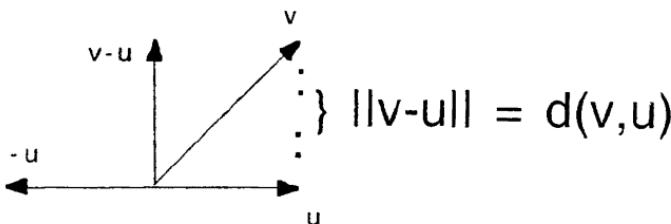
1.1: Osservazione : (i) La norma di un vettore è sostanzialmente la sua "lunghezza".

Per poter definire una norma su un k-spazio vettoriale (k qualsiasi) è necessario disporre di una nozione di modulo (o "valore assoluto") a valori in \mathbb{R} (cfr (III)) su k; ossia bisogna potere definire (in modo coerente) $|\lambda|$ per ogni scalare λ in k. Questo è il motivo per cui ci limitiamo ai casi $k = \mathbb{R}, \mathbb{C}$.

2: Esempio: Consideriamo \mathbb{C} come \mathbb{C} -spazio vettoriale e prendiamo come norma il modulo: $\|z\| = |z|$; la verifica delle proprietà (I), ..., (IV) è immediata.

3: Proposizione: Sia E un k-spazio vettoriale normato. L'applicazione d: $E \times E \rightarrow \mathbb{R}$: $(x,y) \mapsto d(x,y) := \|x-y\|$ è una distanza su E tale che $d(x+z,y+z) = d(x,y)$ e $d(\lambda x, \lambda y) = |\lambda| \cdot d(x,y)$, per ogni x, y, z in E e ogni scalare λ .

Dim: Verifichiamo gli assiomi di distanza: $d(x,y) \geq 0$ segue da 1(I); $d(x,y) = 0$ è equivalente a $\|x-y\| = 0$ e da 1(II), questo è equivalente a: $x-y = 0$ ossia $x = y$. Abbiamo $d(x,y) = \|x-y\| = \|(-1)(y-x)\|$, da 1(III) quest'ultima espressione è uguale a $\|-1\| \cdot \|y-x\| = d(y,x)$. Finalmente rimane da verificare la diseguaglianza triangolare: $d(x,z) \leq d(x,y) + d(y,z)$ ossia $\|x-z\| \leq \|x-y\| + \|y-z\|$; questo segue da 1(IV) visto che $x-z = (x-y) + (y-z)$. Le altre asserzioni sono immediate ♦



3.1: Osservazione : (i) Abbiamo appena visto che ogni spazio vettoriale normato è uno spazio metrico. Come fanno presupporre le proprietà supplementari ($d(x+z,y+z) = d(x,y)$ e $d(\lambda x,\lambda y) = |\lambda| \cdot d(x,y)$), il viceversa non è vero. Per esempio consideriamo la distanza banale (cfr §6, 1.1(iv)), su \mathbb{R} : per definizione $d(4,2) = 1$; se $d(4,2) = \|4-2\|$ allora $\|4-2\| = \|2(2-1)\| = 2 \cdot \|2-1\| = 2d(2,1) = 2$, contraddizione.

(ii) Essendo, uno spazio vettoriale normato, uno spazio metrico, tutte le nozioni sviluppate nel contesto degli spazi metrici (sottoinsiemi aperti, topologia, continuità,...) hanno senso su uno spazio vettoriale normato. Gli spazi vettoriali normati di dimensione finita (o più generalmente "completi", spazi di Banach) sono il quadro naturale per sviluppare, almeno in un primo tempo, il calcolo differenziale.

8) Spazi vettoriali euclidei.

Un prodotto scalare su uno spazio vettoriale reale è una forma bilineare simmetrica definita positiva. Un prodotto scalare permette di definire una distanza (disuguaglianza di Schwarz, Prop. 2, 3). In una base ortonormale (cfr teorema di Sylvester) il prodotto scalare è il prodotto scalare standard e la distanza è la distanza "usuale" (euclidea). Il prodotto scalare e la distanza permettono di misurare le lunghezze e gli angoli e quindi naturale studiare quelle applicazioni che conservano la distanza (isometrie).

In tutto questo paragrafo considereremo spazi vettoriali reali.

FORME BILINEARI SIMMETRICHE DEFINITE POSITIVE.

1: Lemma: *Sia E un R-spazio vettoriale e $f : E \times E \rightarrow \mathbb{R}$ una forma bilineare simmetrica, definita positiva. Due vettori v, w di E sono linearmente dipendenti se e solo se $f(w,w).v = f(v,w).w$.*

Dim: Supponiamo v e w linearmente dipendenti. Se $v = 0$ o $w = 0$, il lemma è chiaramente vero. Sia dunque $v \neq 0$ e $w \neq 0$; quindi $v = \alpha w$, $\alpha \neq 0$. Abbiamo $f(v,w) = f(\alpha w, w) = \alpha f(w,w)$. Siccome $w \neq 0$, $f(w,w) \neq 0$ perché f è definita positiva. Quindi $\alpha = f(v,w) / f(w,w)$ e pertanto $f(w,w).v = f(v,w).w$. Viceversa se $f(w,w).v - f(v,w).w = 0$ allora v e w sono linearmente dipendenti tranne forse se $f(w,w) = f(v,w) = 0$. La prima condizione implica $w = 0$ perché f è definita positiva e, in ogni caso, w e v sono linearmente dipendenti ♦

La proposizione seguente è fondamentale:

2: Proposizione: (disuguaglianza di Schwarz): *Sia E un R-spazio vettoriale e $f : E \times E \rightarrow \mathbb{R}$ una forma bilineare simmetrica, definita positiva. Per ogni v, w in E si ha:*

$$f(v,w)^2 \leq f(v,v).f(w,w)$$

con uguaglianza se e solo se v e w sono linearmente dipendenti (in questo caso necessariamente $f(w,w).v = f(v,w).w$).

Dim: Se $w = 0$, la proposizione è chiara; supponiamo quindi $w \neq 0$. Per ogni α, β in \mathbb{R} abbiamo $f(\alpha v + \beta w, \alpha v + \beta w) \geq 0$ perché f è positiva. Sviluppando: $\alpha^2 f(v, v) + \beta^2 f(w, w) + 2\alpha\beta f(v, w) \geq 0$ con uguaglianza se e solo se $\alpha v + \beta w = 0$ perché f è definita positiva. Ponendo $\alpha = f(w, w)$, $\beta = -f(v, w)$ viene: $f(w, w)^2 \cdot f(v, v) + f(v, w)^2 \cdot f(w, w) - 2f(w, w) \cdot f(v, w) \geq 0$ con uguaglianza se e solo se $f(w, w) \cdot v = f(v, w) \cdot w$; ossia se e solo se v e w sono linearmente dipendenti (lemma 1). Siccome $w \neq 0$, $f(w, w) > 0$ perché f è definita positiva, e possiamo dividere per $f(w, w)$ senza cambiare il senso della diseguaglianza: $f(w, w) \cdot f(v, v) + f(v, w)^2 - 2f(v, w) \geq 0$ ossia $f(w, w) \cdot f(v, v) \geq f(v, w)^2$ con uguaglianza se e solo se v e w sono legati. ♦

3: Corollario: *Nelle condizioni della proposizione 2, l'applicazione :*

$\| \cdot \|_f : E \rightarrow \mathbb{R} : v \mapsto \| v \|_f := \sqrt{f(v, v)}$ è una norma su E tale che: $\| v+w \|_f = \| v \|_f + \| w \|_f$ se e solo se v e w sono legati.

Dim: Verifichiamo gli assiomi di norma (cfr §7, Def.1). Scriviamo $\| \cdot \|$ al posto di $\| \cdot \|_f$. La verifica di (I), ..., (III) è immediata, per (IV) bisogna mostrare: $\| v+w \| = (f(v+w, v+w))^{1/2} \leq (f(v, v))^{1/2} + (f(w, w))^{1/2} = \| v \| + \| w \|$. Abbiamo $\| v+w \|^2 = (\sqrt{f(v+w, v+w)})^2 = f(v+w, v+w) = \| v \|^2 + \| w \|^2 + 2f(v, w)$. Per la diseguaglianza di Schwarz: $f(v, w)^2 \leq \| v \|^2 \cdot \| w \|^2$ con uguaglianza se e solo v e w sono legati. Quindi $f(v, w) \leq \| v \| \cdot \| w \|$ con uguaglianza se e solo se v e w sono legati. Perciò $\| v+w \|^2 = \| v \|^2 + \| w \|^2 + 2f(v, w) \leq \| v \|^2 + \| w \|^2 + 2\| v \| \cdot \| w \| = (\| v \| + \| w \|)^2$, ossia $\| v+w \| \leq \| v \| + \| w \|$ con uguaglianza se e solo se v e w sono legati. ♦

4: Definizione: Un prodotto scalare è una forma bilineare simmetrica, definita positiva su un \mathbb{R} -spazio vettoriale, E .

Un spazio vettoriale euclideo è una coppia (E, f) dove E è un \mathbb{R} -spazio vettoriale di dimensione finita e dove f è un prodotto scalare su E .

4.1: Osservazione: Da 3 segue che uno spazio vettoriale euclideo è uno spazio vettoriale normato; non è vero il viceversa: per esempio si può munire \mathbb{R}^2 di una norma che non proviene da un prodotto scalare (cfr Es.1).

In uno spazio vettoriale euclideo vale il teorema di Pitagora:

5: Proposizione: (teorema di Pitagora): *Sia (E, f) uno spazio vettoriale euclideo. Se v e w sono due vettori ortogonali allora $\|v+w\|^2 = \|v\|^2 + \|w\|^2$.*

Dim: Abbiamo $\|v+w\|^2 = f(v+w, v+w) = f(v, v) + f(w, w) = \|v\|^2 + \|w\|^2$ ($f(v, w) = 0$ perché v e w sono ortogonali)♦

Inoltre:

6: Lemma: *Sia (E, f) uno spazio vettoriale euclideo. Per ogni sottospazio vettoriale $F \subseteq E$ si ha: $f|_F$ è un prodotto scalare e: $E = F^\perp \oplus F$.*

Dim: La restrizione di f è ancora definita positiva. La forma f è non degenere quindi $\dim F + \dim F^\perp = \dim E$ (§3, Prop.10) inoltre F non contiene vettori isotropi non nulli (perché f è definita positiva), pertanto $F \cap F^\perp = \{0\}$ ♦

6.1: Osservazione : Con le notazioni del lemma 6, ogni vettore v di E si scrive in modo unico $v = f + f'$ con $f \in F$, $f' \in F^\perp$. L'applicazione lineare $\pi_F: E \rightarrow F$: $v \rightarrow f$, si chiama la proiezione ortogonale su F . Chiaramente $\text{Ker}(\pi_F) = F^\perp$.

PRODOTTO SCALARE.

Sia (E, f) uno spazio vettoriale euclideo. Dal teorema di Sylvester segue che esiste una base, $B = (e_1, \dots, e_n)$, f -ortonormale. Quindi $\text{mat}_B(f) = I_n$. Se $v = x_1e_1 + \dots + x_ne_n$ e $w = y_1e_1 + \dots + y_ne_n$, allora $f(v, w) = x_1y_1 + \dots + x_ny_n$, quindi in ogni base f -ortonormale f si riduce al prodotto scalare standard ("usuale"). Lavorando in basi f -ortonormali ci si riconduce a considerare solo il prodotto scalare standard che d'ora in poi denoteremo con (.1.): con le notazioni precedenti: $(v | w) := x_1y_1 + \dots + x_ny_n$.

Salvo avviso contrario lavoreremo in basi f -ortonormali, quindi col prodotto scalare standard e diremo ortonormale, ortogonale, senza più menzionare f . Sarà però indispensabile in certe situazioni tornare alla notazione generale (per esempio se si lavora con due prodotti scalari, f , g , simultaneamente: una base f -ortonormale non è mai g -ortonormale se $f \neq g$; perché?).

Sia E uno spazio vettoriale euclideo e $B = (e_1, \dots, e_n)$ una base ortonormale.

Se $v = x_1e_1 + \dots + x_ne_n$ e $w = y_1e_1 + \dots + y_ne_n$ allora:

$$(1) \|v\| = \sqrt{\langle v, v \rangle} = \sqrt{x_1^2 + \dots + x_n^2} \text{ (norma euclidea)}$$

$$(2) d(v, w) = \|v - w\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \text{ (distanza euclidea)}$$

$$(3) x_i = \langle v, e_i \rangle, 1 \leq i \leq n.$$

$$(4) (x_1y_1 + \dots + x_ny_n)^2 \leq (x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2)$$

(diseguaglianza di Schwarz); si può anche formulare così: $\|v \wedge w\| \leq \|v\| \cdot \|w\|$.

Per trovare una base ortonormale si può usare il metodo di ortonormalizzazione di Gram-Schmidt:

7: Proposizione: Sia (E, f) uno spazio euclideo e (v_1, \dots, v_n) una base di E . Esiste una base ortonormale di E , (e_1, \dots, e_n) , tale che, per ogni p : $f(e_p, v_p) > 0$, e $\langle e_1, \dots, e_p \rangle = \langle v_1, \dots, v_p \rangle$.

Dim: Si dimostra, per induzione su p , l'esistenza di una famiglia ortonormale e_1, \dots, e_p tale che $f(e_i, v_i) > 0$ e $\langle e_1, \dots, e_p \rangle = \langle v_1, \dots, v_p \rangle$. Per $p = 1$ si prende $e_1 = v_1 / \|v_1\|$. Supponiamo di avere costruito e_1, \dots, e_{p-1} con le proprietà richieste e cerchiamo di costruire e_p . Siccome deve essere $\langle e_1, \dots, e_p \rangle = \langle v_1, \dots, v_p \rangle$, cerchiamo e_p della forma $e_p = \lambda v_p + \sum_{1 \leq i \leq p-1} \alpha_i e_i$. La condizione di ortonormalità implica $0 = f(e_p, e_k) = \lambda f(v_p, e_k) + \alpha_k$; cioè $\alpha_k = -\lambda f(v_p, e_k)$. Quindi e_p sarà della forma $e_p = \lambda w$ con $w = v_p - \sum_{1 \leq i \leq p-1} f(v_p, e_i) e_i$. Abbiamo $\|e_p\| = \|\lambda w\| = |\lambda| \cdot \|w\|$, per avere $\|e_p\| = 1$ basta porre $|\lambda| = 1 / \|w\|$. Inoltre $f(e_p, v_p) = f(e_p, w)$ (perché $v_p = w + \sum f(v_p, e_i) e_i$) ma dalla relazione $f(e_p, e_p) = f(\lambda w, e_p)$ segue che $f(e_p, v_p) = f(e_p, w) = 1 / \lambda$. Pertanto la condizione $f(e_p, v_p) > 0$ implica $\lambda = 1 / \|w\|$. Si vede facilmente che il vettore e_p così definito verifica le proprietà richieste♦

7.1: Osservazione : Con le notazioni della dimostrazione precedente e se $F = \langle e_1, \dots, e_{p-1} \rangle$, il vettore w non è altro che $\pi_F(v_p)$, la proiezione ortogonale di v_p su F .

Vedere gli esercizi per un'interpretazione matriciale del procedimento di Gram-Schmidt.

8: Corollario: In uno spazio euclideo ogni famiglia di vettori ortonormali può essere completata ad una base ortonormale.

Dim: Es. 2♦

La grande novità degli spazi vettoriali euclidei è l'introduzione della distanza; è quindi naturale studiare le applicazioni che conservano la distanza:

ISOMETRIE VETTORIALI:

9: Definizione: Sia E uno spazio vettoriale euclideo. Un'isometria è un'applicazione $\phi: E \rightarrow E$ tale che $d(\phi(v), \phi(w)) = d(v, w)$ ($o \parallel \phi(v) - \phi(w) \parallel = \|v - w\|$) per ogni v, w in E .

10: Lemma: Sia ϕ un'isometria dello spazio vettoriale euclideo E . Se $\phi(0) = 0$ allora $(\phi(v) \mid \phi(w)) = (v \mid w)$ per ogni v, w in E .

Dim: Abbiamo $\|\phi(v) - \phi(0)\| = \|v - 0\|$, dall'ipotesi $\phi(0) = 0$: $\|o(v)\| = \|v\|$ (*). Per ogni v, w abbiamo $\|\phi(v) - \phi(w)\|^2 = \|v - w\|^2$ ossia $(\phi(v) - \phi(w) \mid \phi(v) - \phi(w)) = (v - w \mid v - w)$. Sviluppando: $\|\phi(v)\|^2 + \|\phi(w)\|^2 - 2(o(v) \mid \phi(w)) = \|v\|^2 + \|w\|^2 - 2(v \mid w)$. Usando (*) si ottiene l'asserto ♦

11: Lemma: Sia ϕ un'isometria dello spazio vettoriale euclideo, E , tale che $\phi(0) = 0$. Sia $B = (e_1, \dots, e_n)$ una base ortonormale, allora $(\phi(e_1), \dots, \phi(e_n))$ è una base ortonormale.

Dim: Basta mostrare che $\phi(e_1), \dots, \phi(e_n)$ formano una famiglia ortonormale (cfr §5, Lemma 3), ma dal lemma precedente $(\phi(e_i) \mid \phi(e_k)) = (e_i \mid e_k) = \delta_{ik}$ ♦

12: Lemma: Sia ϕ un'isometria dello spazio vettoriale euclideo E . Se $\phi(0) = 0$ allora ϕ è lineare.

Dim: Sia $B = (e_1, \dots, e_n)$ una base ortonormale, dal lemma precedente $(\phi(e_1), \dots, \phi(e_n))$ è una base ortonormale; quindi (cfr Prodotti scalari (3)): $\phi(v) = \sum_i (\phi(v) \mid \phi(e_i)).\phi(e_i)$. Usando il lemma 10: $\phi(v) = \sum_i (v \mid e_i).\phi(e_i) = \sum_i x_i.\phi(e_i)$ dove $v = \sum_i x_i.e_i$. In conclusione $\phi(\sum_i x_i.e_i) = \sum_i x_i.\phi(e_i)$, quindi ϕ è lineare ♦

Quanto precede giustifica la:

13: Definizione: Sia E uno spazio vettoriale euclideo e $\phi: E \rightarrow E$ un'isometria tale che $\phi(0) = 0$ allora ϕ viene chiamata isometria vettoriale (o ancora: operatore unitario).

14: Proposizione: Sia E uno spazio vettoriale euclideo e $g: E \rightarrow E$ un endomorfismo. Sono equivalenti:

- (i) g è un'isometria vettoriale
- (ii) per ogni v in E , $\|g(v)\| = \|v\|$
- (iii) per ogni base ortonormale di E , $B = (e_1, \dots, e_n)$, $g(B) = (g(e_1), \dots, g(e_n))$ è una base ortonormale di E .
- (iv) esiste una base ortonormale di E , $B = (e_1, \dots, e_n)$, tale che $g(B) = (g(e_1), \dots, g(e_n))$ sia una base ortonormale di E .
- (v) se $B = (e_1, \dots, e_n)$ è una base ortonormale di E , $\text{mat}(g; B, B)$ è una matrice ortogonale.

Dim: (i) \Leftrightarrow (ii): l'implicazione \Rightarrow è chiara. Per (ii) \Rightarrow (i) basta osservare che g essendo lineare, da (ii), abbiamo $\|v-w\| = \|g(v-w)\| = \|g(v)-g(w)\|$ quindi g è un'isometria.

(i) \Rightarrow (iii): risulta dal lemma 11.

(iii) \Rightarrow (iv): ovvio.

(iv) \Rightarrow (ii): se $v = x_1e_1 + \dots + x_ne_n$ allora $g(v) = x_1g(e_1) + \dots + x_ng(e_n)$ e si ha subito $\|v\| = \|g(v)\|$

(iii) \Rightarrow (v): sia $A = \text{mat}(g; B, B)$. Vogliamo mostrare ${}^t A \cdot A = I_n$. Sia ${}^t A \cdot A = (m_{ij})$ allora m_{ij} è il prodotto scalare "usuale" tra la i-esima riga di (ossia la i-esima colonna di A) e la j-esima colonna di A quindi $m_{ij} = (g(e_i) \mid g(e_j)) = \delta_{ij}$ perché, da (iii), $(g(e_1), \dots, g(e_n))$ è una base ortonormale.

(v) \Rightarrow (iv): sia A la matrice ortogonale, per ipotesi, di g in una certa base ortonormale. Riprendendo la dimostrazione precedente si ha che la condizione ${}^t A \cdot A = I_n$ implica che $g(e_1), \dots, g(e_n)$ formano una famiglia (quindi una base, cfr §5 lemma 3) ortonormale♦

15: Corollario: Sia E uno spazio vettoriale euclideo e $g: E \rightarrow E$ un'isometria vettoriale. Se $\lambda \in \mathbb{R}$ è un autovalore di g allora $\lambda = \pm 1$.

Dim: Sia x un autovettore associato a λ . Abbiamo $\|g(x)\| = \|\lambda x\| = |\lambda| \|x\|$ ma g essendo un'isometria vettoriale: $\|g(x)\| = \|x\|$ (cfr 14 (ii)). Segue che $|\lambda| = 1$ ♦

16: Corollario: Sia E un spazio vettoriale euclideo e B una base ortonormale di E . Denotiamo con $O(E)$ l'insieme delle isometrie vettoriali di E . L'isomorfismo di \mathbb{R} -algebre $\text{mat}(\cdot; B, B) : \text{End}(E) \rightarrow M_n(\mathbb{R})$ induce un isomorfismo tra $O(E)$ e $O_n(\mathbb{R})$. In particolare $O(E)$ è un gruppo per la composizione delle applicazioni.

Dim: Per semplificare noteremo m invece di $\text{mat}(\cdot; B, B)$. Da 14 (v) segue che m induce una suriezione da $O(E)$ su $O_n(\mathbb{R})$; siccome m è una biezione, m induce una biezione da $O(E)$ in $O_n(\mathbb{R})$. Finalmente m essendo un morfismo d'anelli tra $(\text{End}(E), \cdot)$ e $(M_n(\mathbb{R}), \cdot)$ e $O_n(\mathbb{R})$ essendo un gruppo, la biezione $m|_{O(E)} : O(E) \rightarrow O_n(\mathbb{R})$ è un isomorfismo di gruppi ♦

16.1: Osservazione : In particolare se g è un'isometria vettoriale allora g è invertibile e g^{-1} è un'isometria; se g e h sono due isometrie vettoriali allora $g \cdot h$ è ancora un'isometria vettoriale.

17: Definizione: Noteremo $SO(E)$ o $O^+(E)$ l'immagine reciproca di $SO_n(\mathbb{R}) = O_n^+(\mathbb{R})$ tramite l'isomorfismo $\text{mat}(\cdot; B, B)$ e definiamo $O^-(E) := O(E) \setminus O^+(E)$.

17.1: Osservazione : (i) A priori la definizione precedente dipende dall'isomorfismo $\text{mat}(\cdot; B, B)$ e quindi dalla base B ; in realtà non è così: sia g un'isometria vettoriale e $A = \text{mat}(g; B, B)$, sia $A' = \text{mat}(g; B', B')$ dove B' è un'altra base ortonormale. Allora $A' = P^{-1}AP$ e quindi $\det(A') = \det(A)$ (che non è altro che $\det(g)$) quindi $O^+(E)$ è ben definito.

(ii) $O^+(E)$ è un sottogruppo di $O(E)$ (perché $SO_n(\mathbb{R})$ è un sottogruppo di $O_n(\mathbb{R})$)

Esercizi:

8.1) Se $u = (x, y) \in \mathbb{R}^2$ si pone $\|u\| = \max\{|x|, |y|\}$. Dimostrare che $\|\cdot\|$ è una norma e che questa norma non proviene da un prodotto scalare.

8.2) Dimostrare il corollario 8: (i) usando la Prop.7, (ii) senza usare la Prop.7.

8.3) Dimostrare direttamente che $O(E)$ è un gruppo per la composizione delle applicazioni.

8.4) Nello spazio euclideo \mathbb{R}^3 con il prodotto scalare usuale, applicare il procedimento di Gram-Schmidt alla base $v = (1, 1, 1)$, $v' = (1, 1, 0)$, $v'' = (1, 0, 0)$.

8.5) Siano v, w , due vettori non nulli di un piano euclideo. Si determini un numero reale α tale che $v+\alpha w$ abbia lunghezza (i.e. norma) minima fra tutti i vettori della forma $v+\beta w$, $\beta \in \mathbb{R}$. Il numero α è univocamente determinato?

8.6) Sia E un spazio vettoriale euclideo, e v_0, w_0 , due vettori non nulli di E . Si consideri l'applicazione $f: E \rightarrow E: v \mapsto (v_0|v).w_0$.

(i) E è un'applicazione lineare?

(ii) Determinare la matrice di f rispetto ad una base ortonormale. Determinare la dimensione del nucleo di f .

8.7) Sia E uno spazio vettoriale euclideo. Un endomorfismo, f , di E si dice positivo se per ogni v in E si ha $(f(v)|v) \geq 0$.

(i) Sia $W \subseteq E$ un sottospazio vettoriale. Dimostrare che $E = W \oplus W^\perp$.

(ii) Da (i) segue che ogni v in E si scrive in modo unico $v = w + w'$, con w in W e w' in W^\perp . Si definisce $p: E \rightarrow E: v \mapsto w$. Dimostrare che p è positivo.

(iii) Dimostrare che per ogni v, u , si ha: $(p(v)|u) = (v|p(u))$.

(iv) Mostrare con un esempio che esistono degli endomorfismi non nulli, g , tali che $(g(v)|v) = 0$ per ogni v in E (sugg: $\dim E = 2$).

8.8) Sia $T_+(3) := \{M \in M_3(\mathbb{R}) / M \text{ è triangolare superiore con gli elementi della diagonale strettamente positivi}\}$; cioè $M = (a_{ij}) \in T_+(3)$ se e solo se: $a_{ij} = 0$ se $i > j$ e $a_{ii} > 0$, $1 \leq i \leq 3$. Mostrare che se $M \in T_+(3)$ allora M è invertibile e $M^{-1} \in T_+(3)$. Dedurne che $(T_+(3), \cdot)$ è un sottogruppo di $(GL(3), \cdot)$ (suggerimento per mostrare che $M^{-1} \in T_+(3)$: considerare un endomorfismo associato ad M).

8.9) Si considera \mathbb{R}^3 con il prodotto scalare usuale: se $v = (x, x', x'')$, $w = (y, y', y'')$ allora $(v|w) := xy + x'y' + x''y''$. Per la definizione di $T_+(3)$ vedere 8.8.

Sia $A \in GL(3)$ e siano c_1, c_2, c_3 i vettori colonna di A . Si pone:

$$f_1 := c_1, \quad f_2 := c_2 - (c_2|f_1)/(f_1|f_1).f_1, \quad f_3 := c_3 - (c_3|f_1)/(f_1|f_1).f_1 - (c_3|f_2)/(f_2|f_2).f_2.$$

(i) Dimostrare che $(f_i|f_j) = 0$, $1 \leq i, j \leq 3$, $i \neq j$.

(ii) Sia $F \in M_3(\mathbb{R})$ la matrice di vettori colonna f_1, f_2, f_3 . Mostrare che $F = A.T_1$ dove $T_1 \in T_+(3)$, T_1 avente solo degli 1 sulla diagonale.

(iii) Si pone $b_i := f_i / \|f_i\|$, $1 \leq i \leq 3$. Sia B la matrice di vettori colonna b_1, b_2, b_3 . Mostrare che $B = F.T_2$ dove T_2 è la matrice diagonale avente $1/\|f_i\|$, $1 \leq i \leq 3$, come termini non nulli. Dedurne che ogni matrice invertibile, A , si scrive $A = B.C$ con $B \in O(3)$, $C \in T_+(3)$.

(iv) Osservando che una matrice diagonale è uguale alla sua trasposta, dimostrare che $O(3) \cap T_+(3) = \{\text{Id}\}$. Concludere che ogni matrice invertibile, $A \in M_3(\mathbb{R})$, si scrive in modo unico come prodotto di una matrice ortogonale e di una matrice di $T_+(3)$.

8.10) Si generalizza l'esercizio precedente: dedurre dall'enunciato della Prop.7 (e usando Prop.14) che ogni matrice invertibile $M \in M_n(\mathbb{R})$ si scrive in modo unico nella forma $M = A.T$ dove $T \in T_+(n)$ e dove A è una matrice ortogonale.

9) Cenni sulla classificazione delle isometrie vettoriali.

Si classificano le isometrie vettoriali di uno spazio euclideo di dimensione al più tre e si dà un teorema di struttura nel caso generale (cfr Teorema 16).

ISOMETRIE VETTORIALI DELLA RETTA REALE.

Sia E un \mathbb{R} -spazio vettoriale di dimensione uno, $f : E \times E \rightarrow \mathbb{R}$ una forma bilineare simmetrica, definita positiva e $v : E \rightarrow E$ un'isometria vettoriale (i.e. $f(x,y) = f(v(x),v(y))$ per ogni x, y in E). Sia $e \in E$, $e \neq 0$ (quindi e è una base di E). Se $v(e) = \lambda e$ da §8, 15 segue che $\lambda = \pm 1$. In conclusione $O(E) = \{\pm \text{Id}_E\}$.

ISOMETRIE VETTORIALI DEL PIANO.

Sia E uno spazio vettoriale euclideo di dimensione due e $B = (e_1, e_2)$ una base ortonormale di E . Sia v un'isometria vettoriale di E e $M = \text{mat}(v; B, B)$ con $M = \begin{pmatrix} ac \\ bd \end{pmatrix}$.

1: Lemma: Con le notazioni precedenti:

$$M \in SO_2(\mathbb{R}) = O_2^+(\mathbb{R}) \Leftrightarrow M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ con } a^2 + b^2 = 1.$$

$$M \in O_2^-(\mathbb{R}) \Leftrightarrow M = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ con } a^2 + b^2 = 1.$$

Dim: Le implicazioni (\Leftarrow) risultano da una verifica immediata (${}^t M M = I_2$). Dimostriamo le altre. Sappiamo che $(v(e_1), v(e_2))$ è una base ortonormale di E (§8, Prop.14) quindi $(v(e_i) | v(e_j)) = \delta_{ij}$. In termini di coordinate otteniamo:

$$ac + bd = 0 \quad (\text{i}),$$

$$a^2 + b^2 = 1 \quad (\text{ii}),$$

$$c^2 + d^2 = 1 \quad (\text{iii}).$$

Equivalentemente, queste relazioni si possono ottenere scrivendo che M è ortogonale (${}^t M M = I_2$).

Se $a = 0$ da (ii) viene $b^2 = 1$ quindi $b = \epsilon$ ($\epsilon \in \{\pm 1\}$). Da (i) segue che $d = 0$ e da (iii) $c = \epsilon'$ ($\epsilon' \in \{\pm 1\}$). In conclusione se $a = 0$, $M = \begin{pmatrix} 0 & \epsilon \\ \epsilon' & 0 \end{pmatrix}$ e $\det(M) = -\epsilon \cdot \epsilon'$.

Supponiamo $a \neq 0$. Da (i): $c = -bd/a$. Da (iii) $d^2(a^2 + b^2)/a^2 = 1$; usando (i) viene $d^2 = a^2$ e quindi $d = \epsilon a$ ($\epsilon \in \{\pm 1\}$). Avevamo $c = -bd/a$ da cui $c = -\epsilon b$. In

conclusione $M = \begin{pmatrix} a & -\varepsilon b \\ b & \varepsilon a \end{pmatrix}$ (osservare che ponendo $a = 0$ e tenendo conto che $b^2 = 1$, si ritrova il caso precedente). Abbiamo $\det(M) = \varepsilon(a^2 + b^2) = \varepsilon$. Si conclude osservando che $M \in O_2^+(R)$ (risp. $M \in O_2^-(R)$) implica $\det(M) = 1$ (risp. $\det(M) = -1$). ♦

2: Corollario: $O_2^+(R)$ è un gruppo abeliano.

Dim: Si verifica facilmente, usando il lemma 1, che se M, N sono elementi di $O_2^+(R)$ allora $M.N = N.M$. ♦

Sia E uno spazio vettoriale euclideo di dimensione n e H un iperpiano vettoriale di E . Sappiamo che $E = H \oplus H^\perp$. Se $x \in E$ allora $x = y + z$ con y in H , z in H^\perp . Poniamo $r_H(x) = y - z$. Si verifica facilmente che $r_H : E \rightarrow E$ è un'applicazione lineare. Osserviamo che $\|x\| = (y+z|y+z|)^{1/2} = (\|y\|^2 + \|z\|^2)^{1/2} = \|r_H(x)\|$ quindi (§8,14), r_H è un'isometria vettoriale.

3: Definizione: Con le notazioni precedenti r_H si chiama la simmetria ortogonale rispetto ad H (si dice anche riflessione ortogonale rispetto ad H).

Tornando alle isometrie del piano:

4: Proposizione: Sia E un piano euclideo, se $v \in O^+(E)$ allora v è una simmetria ortogonale rispetto ad una retta vettoriale.

Dim: Sia B una base ortonormale e $A = \text{mat}(v; B, B)$. Dal lemma 1, $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ con $a^2 + b^2 = 1$ per opportuni a, b reali. In particolare A è simmetrica: ${}^t A = A$. Dalla relazione di ortogonalità ${}^t A \cdot A = I_2$ deduciamo $A^2 = I_2$. Sappiamo che in queste condizioni, A è diagonalizzabile e che se λ, μ sono i suoi autovalori allora $\lambda, \mu \in \{\pm 1\}$ (cfr Es.II.16.5). Siccome $\pm \text{Id}_E \in O^+(E)$, l'unica possibilità è $\lambda = 1$ e $\mu = -1$. Sia $B' = (e'_1, e'_2)$ una base di E fatta da autovettori di v : $v(e'_1) = e'_1$, $v(e'_2) = -e'_2$. Se $x = \alpha e'_1 + \beta e'_2$, allora $v(x) = \alpha e'_1 - \beta e'_2$; quindi v è una simmetria rispetto alla retta vettoriale $\langle e'_1 \rangle$. Questa simmetria è ortogonale ossia i vettori e'_1 e e'_2 sono ortogonali. Infatti da una parte: $(v(e'_1) \mid v(e'_2)) = (e'_1 \mid -e'_2) = -(e'_1 \mid e'_2)$ (e'_i sono autovettori di v); mentre dall'altra parte: $(v(e'_1) \mid v(e'_2)) = (e'_1 \mid e'_2)$ perché v è un'isometria. Confrontando: $(e'_1 \mid e'_2) = 0$. ♦

5: Definizione: Una rotazione è un'isometria appartenente a $O^+(E)$.

6: Proposizione: Ogni rotazione è il prodotto di due simmetrie rispetto a due rette di cui una può essere scelta arbitrariamente.

Dim: Sia $r \in O^+(E)$. Se $s \in O^-(E)$ è una simmetria qualsiasi allora rs e sr sono delle simmetrie (considerare i determinanti): $rs = \sigma$, $sr = \tau$. Visto che $s = s^{-1}$ abbiamo $r = \sigma s$, $r = s\tau$ ♦

Per approntare lo studio delle rotazioni del piano bisogna introdurre la nozione di angolo (cfr §10).

ISOMETRIE VETTORIALI DELLO SPAZIO EUCLIDEO.

Un accorgimento fondamentale per la classificazione delle isometrie vettoriali è la nozione di sottospazio invariante (o fisso), globalmente invariante:

7: Definizione: Sia E uno spazio euclideo e $v: E \rightarrow E$ un'isometria. Un sottospazio vettoriale $F \subseteq E$ è invariante ("punto per punto"), o fisso, sotto v se $v|_F = \text{Id}_F$; il sottospazio F è globalmente invariante (o stabile) sotto v se: $v(F) = F$.

7.1: Osservazione : Se $\dim(F) = 1$ e se F è globalmente invariante allora F è un autospazio e $v|_F = \pm \text{Id}_F$.

Abbiamo per esempio:

8: Lemma: Sia $v \neq \pm \text{Id}_E$ un'isometria vettoriale del piano euclideo E .

- (i) L'isometria v è positiva se e solo se non esiste nessun sottospazio globalmente invariante sotto v .
- (ii) L'isometria v è negativa se e solo se v ammette un sottospazio invariante.

Dim: Segue immediatamente da 1, 4 ♦

9: Lemma: Siano E uno spazio euclideo, $v: E \rightarrow E$ un'isometria, e $F \subseteq E$ un sottospazio stabile sotto v . Allora F^\perp è stabile sotto v (i.e. $v|_{F^\perp}: F^\perp \rightarrow F^\perp$ è un'isometria).

Dim: Segue da $(v(x) \mid v(y)) = (x \mid y)$ (§7, 8) ♦

10: Lemma: Sia $n \geq 1$ un numero intero dispari. Sia $A \in M_n(\mathbb{R})$ una matrice ortogonale.

- (i) Se $\det(A) = 1$ allora $\det(A - I_n) = 0$

(ii) Se $\det(A) = -1$ allora $\det(A+I_n) = 0$.

Dim: (i) Abbiamo ${}^t A \cdot A = I_n$ perché A è ortogonale. Pertanto $\det(A-I_n) = \det(A-{}^t A \cdot A) = \det((I_n-{}^t A) \cdot A) = \det(I_n-{}^t A) \cdot \det(A) = \det(I_n-{}^t A) = \det({}^t I_n-{}^t A) = \det({}^t(I_n-A)) = \det(I_n-A) = \det(-(A-I_n)) = (-1)^n \cdot \det(A-I_n) = -\det(A-I_n)$.

(ii) La dimostrazione è analoga♦

11: Lemma: Sia v un'isometria vettoriale, positiva dello spazio euclideo. Allora $v = \text{Id}$ o v è una rotazione di asse una retta vettoriale.

Dim: Un vettore è fisso se è un autovettore relativo all'autovalore $\lambda = 1$. Dal lemma 10, $\det(A-I_3) = 0$ quindi l'autospazio $E_v(1) =: F$ ha dimensione ≥ 1 . Se $\dim(F) = 3$, v è l'identità. Supponiamo, d'ora in poi, $v \neq \text{Id}$. Dal lemma 9, $v|_{F^\perp}$ è un'isometria di F^\perp . Se $\dim(F) = 2$ allora $\dim(F^\perp) = 1$ e $v|_{F^\perp} = \pm \text{Id}$ (cfr "isometrie della retta"). Non può essere $v|_{F^\perp} = \text{Id}$ perché $F = E_v(1)$ e $F \cap F^\perp = \{0\}$. Quindi $v|_{F^\perp} = -\text{Id}$. Prendendo una base di autovettori viene $\det(v) = -1$, contro l'ipotesi che v è positiva. Quindi, se $v \neq \text{Id}$, l'unica possibilità è $\dim(F) = 1$, $\dim(F^\perp) = 2$. Prendendo una base ortonormale di F^\perp , vediamo che $\det(v|_{F^\perp}) = 1$, quindi $v|_{F^\perp}$ è una rotazione. Pertanto v è una rotazione di asse la retta F ♦

11.1: Osservazione : Con le notazioni della dimostrazione precedente, se $B = (e_1, e_2, e_3)$ è una base ortonormale con $F = \langle e_1 \rangle$ allora $\text{mat}(v; B, B)$ è della forma $\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & -b & a \end{pmatrix}$ con $a^2 + b^2 = 1$.

12: Proposizione: Sia v un'isometria negativa dello spazio euclideo. Esiste una retta vettoriale, F , tale che $v|_F = -\text{Id}_F$ e tale che $v|_{F^\perp}$ sia una rotazione.

Dim: Es.1♦

12.1: Osservazione : Si usa distinguere tre casi; la rotazione $v|_{F^\perp}$ è: (i) l'identità, (ii) meno l'identità, (iii) $\neq \pm \text{Id}$. Nel caso (i) v è una simmetria rispetto al piano F^\perp , nel caso (ii) v è la simmetria rispetto all'origine.

Concluderemo lo studio delle isometrie vettoriali di uno spazio euclideo di dimensione n con un teorema di struttura.

13: Lemma: Sia E uno spazio euclideo di dimensione n . Ogni riflessione ortogonale appartiene a $O^-(E)$.

Dim: Sia r_H la riflessione ortogonale rispetto all'iperpiano H . Esiste una base ortonormale di E , $B = (e_1, \dots, e_{n-1}, e_n)$, con e_1, \dots, e_{n-1} in H e e_n in H^\perp (considerare la restrizione del prodotto scalare ad H : è ancora un prodotto scalare). Dal teorema di Sylvester, esiste una base ortonormale di H : (e_1, \dots, e_{n-1}) ; si conclude prendendo un vettore unitario, e_n , in H^\perp . Sia $M = \text{mat}(r_H; B, B)$

B) allora $M = \begin{pmatrix} I_{n-1} & 0 \\ & \vdots \\ & 0 \\ 0 \dots 0 & -1 \end{pmatrix}$. Quindi $\det(M) = \det(I_{n-1}) \cdot \det(-1) = -1$ e pertanto r_H appartiene ad $O^+(E) \blacklozenge$

14: Lemma: Siano E uno spazio vettoriale euclideo, v, w due vettori non nulli di E con $\|v\| = \|w\|$. Sia $H = \{x \in E / d(x, v) = d(w, x)\}$. Allora $H = \{x \in E / (x \mid v-w) = 0\}$. Pertanto H è un iperpiano vettoriale.

Dim: Sia $H' = \{x \in E / (x \mid v-w) = 0\}$. Se $x \in H'$ allora $(x \mid v) = (x \mid w)$. Abbiamo $d(x, v) = \|x-v\| = [\|x\|^2 + \|v\|^2 - 2(x \mid v)]^{1/2}$ e $d(x, w) = \|x-w\| = [\|x\|^2 + \|w\|^2 - 2(x \mid w)]^{1/2}$. Siccome $\|v\| = \|w\|$ per ipotesi, e $(x \mid v) = (x \mid w)$ perché x è in H' , vediamo che x appartiene ad H . Quindi $H' \subseteq H$. Viceversa se $x \in H$ allora $\|v\|^2 - 2(x \mid v) = \|w\|^2 - 2(x \mid w)$. Siccome $\|v\| = \|w\|$ segue che $(x \mid v) = (x \mid w)$ ossia $(x \mid v-w) = 0$. Pertanto $H \subseteq H'$. Finalmente H' è un iperpiano vettoriale perché è il ker della forma lineare $E \rightarrow \mathbb{R}: y \mapsto (y \mid v-w) \blacklozenge$

15: Definizione: Sia E uno spazio vettoriale euclideo. Se v, w sono due vettori non nulli di E , con $\|v\| = \|w\|$, l'iperpiano $H = \{x \in E / d(x, v) = d(x, w)\}$, si chiama il bisettore ortogonale di v e w .

15.1: Osservazione : Se H è il bisettore ortogonale di v e w , $v-w$ appartiene a H^\perp (cfr lemma 14) quindi $r_H(v-w) = w-v$. D'altra parte $(v+w \mid v-w) = (v \mid v) - (w \mid w) = \|v\|^2 - \|w\|^2 = 0$, quindi $v+w$ appartiene ad H e perciò $r_H(v+w) = v+w$. Combinando col risultato precedente: $r_H(v) = w$ e $r_H(w) = v$.

Adesso possiamo formulare e dimostrare il teorema di struttura:

16: Teorema: Sia E uno spazio vettoriale euclideo di dimensione $n \geq 2$. Ogni isometria vettoriale di E è il prodotto di al più n riflessioni ortogonali.

Sappiamo già che il teorema è vero se $n = 2$ (cfr Prop.6). Il teorema sarà conseguenza di un enunciato più generale:

17: Teorema: Siano E un spazio vettoriale euclideo di dimensione $n \geq 2$ e $v : E \rightarrow E$ un'isometria vettoriale. Se esiste un sottospazio vettoriale V di E con $\dim(V) = n-r$ e tale che la restrizione di v a V sia l'identità, allora v è un prodotto di al più r riflessioni ortogonali rispetto ad iperpiani contenenti V .

17.1: Osservazione : Il teorema 16 segue da 17 prendendo $V = \{0\}$.

Dimostrazione del teorema 17: Per induzione su r . Se $r = 0$ allora $v = \text{Id}_E$ e per ogni riflessione r_H , $\text{Id}_E = r_H^0$. Sia dunque $r > 0$ e assumiamo il teorema per $r-1$. Sia v un'isometria tale che $v|_V$ sia l'identità con $\dim(V) = n-r$. Abbiamo $E = V \oplus V^\perp$. Mostriamo che $v(V^\perp) \subseteq V^\perp$. Sia x in V^\perp , per ogni y in V abbiamo: $(v(x)|v(y)) = (x|y) = 0$ (la prima uguaglianza risulta dal fatto che v è un'isometria cfr §8, lemma 10; la seconda dal fatto che $x \in V^\perp$). Siccome $v|_V$ è l'identità abbiamo $v(y) = y$ quindi $(v(x)|y) = 0$ e perciò $v(x) \in V^\perp$. Questo dimostra $v(V^\perp) \subseteq V^\perp$ (in realtà $v(V^\perp) = V^\perp$ perché v è biettiva). Siccome la restrizione del prodotto scalare a V (risp. V^\perp) è ancora un prodotto scalare, prendendo delle basi ortonormali in V e V^\perp , e considerandone la riunione otteniamo una base ortonormale di E , $B = (e_1, \dots, e_n)$ tale che e_1, \dots, e_{n-r} sia una base di V e e_{n-r+1}, \dots, e_n sia una base di V^\perp . Sia H il bissettore ortogonale di $f := e_{n-r+1} e$, $v(e_{n-r+1}) = v(f)$ (osservare che $\|f\| = \|v(f)\|$ perché v è un'isometria). Sia v' l'isometria definita da: $v' = r_H \circ v$ e mostriamo che v' ristretto a $V' = V \oplus \langle e_{n-r+1} \rangle$ è l'identità. Per mostrare che $v'|_{V'}$ è l'identità basta mostrare che V' è contenuto in H (perché $v|_V$ è l'identità e r_H ristretto ad H è l'identità). Per definizione $H = \{x \in E / (x|e_{n-r+1}-v(e_{n-r+1})) = 0\}$. Se $x \in V'$ allora $(x|e_{n-r+1}-v(e_{n-r+1})) = 0$ perché $e_{n-r+1} \in V^\perp$ e $v(e_{n-r+1}) \in V^\perp$ (usare $v(V^\perp) \subseteq V^\perp$). Adesso $r_H \circ v(e_{n-r+1}) = r_H(v(e_{n-r+1})) = e_{n-r+1}$ (cfr 15.1). Quindi $v'|_{V'}$ è l'identità. Osserviamo che V' ha dimensione $n-(r-1)$ quindi, per ipotesi di induzione, $r_H \circ v$ è un prodotto di al più $r-1$ riflessioni ortogonali rispetto ad iperpiani contenenti V' (quindi V): $r_H \circ v = r_{H_1} \circ r_{H_2} \circ \dots \circ r_{H_k}$, $k \leq r-1$. Siccome $r_H = r_H^{-1}$, otteniamo $v = r_H \circ r_{H_1} \circ r_{H_2} \circ \dots \circ r_{H_k}$ dove l'iperpiano di ogni riflessione contiene V ♦

Esercizi:

9.1) Dimostrare la Proposizione 12: sia v un'isometria negativa dello spazio euclideo. Esiste una retta vettoriale, F , tale che $v|_F = -\text{Id}_F$ e tale che $v|_{F^\perp}$ sia una rotazione.

9.2) Dire se $O_3^+(\mathbb{R})$ è un gruppo abeliano (cfr Corollario 2). Cosa si può dire di $O_n^+(\mathbb{R})$, $n \geq 3$?

9.3) Classificare le isometrie del piano e dello spazio usando i teoremi **16, 17**.

9.4) Sia $u: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'endomorfismo dello spazio euclideo \mathbf{R}^3 (con il riferimento

standard) tale che $\text{mat}(u; C, C) = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$ dove C indica la base canonica.

(i) Dire se u è un'isometria.

(ii) Determinare $x \in \mathbf{R}^3$ tale $u(x) = {}^t(1, 2, -1)$.

(iii) Descrivere l'endomorfismo u .

9.5) Un'isometria v è involutiva se $v^2 = \text{Id}$.

(i) Determinare le isometrie involutive del piano.

(ii) Dimostrare che una rotazione, $r \neq \text{Id}$, dello spazio euclideo \mathbf{R}^3 è involutiva se e solo se è una simmetria rispetto ad una retta.

9.6) Sia il piano euclideo \mathbf{R}^2 con il riferimento standard. Siano u, v due vettori di \mathbf{R}^2 .

Dimostrare che l'area del parallelogramma costruito su u, v è uguale a $|\det(u, v)|$. In modo analogo mostrare che il volume del parallelepipedo costruito su tre vettori u, v, w di \mathbf{R}^3 è uguale a $|\det(u, v, w)|$. Sia f un'isometria dello spazio euclideo \mathbf{R}^n . Siano v_1, \dots, v_n vettori di \mathbf{R}^n . Dimostrare che $\det_C(v_1, \dots, v_n) = \pm \det_C(f(v_1), \dots, f(v_n))$, dove C indica la base canonica (si preciserà il segno secondo f). Conclusione?

10) Angoli.

In questo paragrafo si introduce la nozione di angolo (non orientato, orientato), e si discute il problema della misura degli angoli.

Angoli (non orientati).

Sia E uno spazio vettoriale euclideo di dimensione n e v, w due vettori non nulli di E . La diseguaglianza di Schwarz: $|v \cdot w| \leq \|v\| \cdot \|w\|$ è equivalente a: $-1 \leq \frac{v \cdot w}{\|v\| \cdot \|w\|} \leq 1$. Da quest'ultima espressione risulta che: $-1 \leq \cos(\theta)/\|v\| \cdot \|w\| \leq 1$. Dalle proprietà della funzione coseno discende che esiste un unico numero reale θ , tale che: $0 \leq \theta \leq \pi$ e $\cos(\theta) = (v \cdot w)/\|v\| \cdot \|w\|$.

1: Definizione: Diremo che $\theta := \arccos[(v \cdot w)/\|v\| \cdot \|w\|]$ è l'angolo, non orientato, formato da v e w .

1.1: Osservazione : (i) La definizione ha senso perché $\cos : [0, \pi] \rightarrow [-1, 1]$ è una biiezione. Ovviamente questa definizione dell'angolo non orientato tra due vettori presuppone la conoscenza della funzione coseno.

(ii) Si ha $(v \cdot w) = \cos(\theta) \cdot \|v\| \cdot \|w\|$, in particolare se $\|v\| = \|w\| = 1$, $(v \cdot w) = \cos(\theta)$. Si vede così che v e w sono ortogonali se e solo se formano un angolo non orientato uguale a $\pi/2$ (angolo retto).

ANGOLI NEL PIANO E ROTAZIONI.

2: Lemma: Siano E un piano vettoriale euclideo, v e w due vettori non nulli di E con $\|v\| = \|w\|$. Esiste un'unica rotazione, r , tale che $r(v) = w$.

Dim: Possiamo supporre $\|v\| = \|w\| = 1$ (altrimenti considerare $v/\|v\|$, $w/\|w\|$). Consideriamo la retta v^\perp . Esistono due vettori $v', v'' \in v^\perp$ di norma uguale a 1 (prendere un vettore, $x \neq 0$, in v^\perp e porre $v' = x/\|x\|$, $v'' = -v'$). In particolare $B = (v, v')$ è una base ortonormale di E . Abbiamo $w = av + bv'$ per opportuni scalari a, b . Inoltre $\|w\| = 1 = a^2 + b^2$.

La rotazione r tale che $\text{mat}(r; B, B) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ soddisfa $r(v) = w$. Se r' è una rotazione tale che $r'(v) = w$ allora $\text{mat}(r'; B, B)$ ha il primo vettore colonna uguale a quello di $\text{mat}(r; B, B)$. Da §9, lemma 1 segue che $\text{mat}(r; B, B) = \text{mat}(r'; B, B)$ quindi $r = r'$ (§8, Cor.13)♦

2.1: Osservazione : (i) Con le notazioni precedenti: $(v \mid w) = (v \mid av+bv') = a.(v \mid v) = a \cdot \|v\|^2 = a \cdot \|v\| \cdot \|w\|$ (perché $\|w\| = \|v'\|$). Quindi $(v \mid w) / \|v\| \cdot \|w\| = a = \cos(\theta)$ dove θ è l'angolo non orientato tra v e w .

(ii) Nella dimostrazione precedente, se invece di B prendiamo come base $B' = (v, v'')$ allora $\text{mat}(r; B', B') = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. Osserviamo che se $M = \text{mat}(\text{Id}_E; B, B')$ allora M è una matrice ortogonale con $\det(M) = -1$.

(iii) Sia r' l'unica rotazione tale che $r'(w) = v$ allora esiste una base ortonormale, $B = (w, f)$ tale che $\text{mat}(r'; B, B)$ sia $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Risulta da quanto precede che $\cos \theta$ (θ l'angolo non orientato tra v e w) si può definire tramite l'unica rotazione, r (risp. r'), che manda v in w (risp. w in v); l'ordine nel quale vengono presi i vettori non importa. Non importa neanche quale vettore unitario si scelga in v^\perp (v' o v'').

Detto ciò sembrerebbe naturale dire che r (risp. r') è una rotazione di angolo non orientato θ . Osserviamo, sempre con le notazioni della dimostrazione precedente, che $a = \cos(\theta)$ è univocamente determinato da v, w ($a = (v \mid w) / \|v\| \cdot \|w\|$) quindi da θ . Altrettanto non si può dire di b (dipende dalla scelta tra v' e v''). Dalla relazione $a^2 + b^2 = 1$ vorremo avere $b = \sin(\theta)$; purtroppo questo non è possibile, c'è un'ambiguità sul segno ($\sin(-x) = -\sin(x)$ mentre $\cos(-x) = \cos(x)!$)

Questa ambiguità si riflette nella:

3: Proposizione: Siano $v \in O^+(E)$, B una base ortonormale di E e $M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \text{mat}(v; B, B)$. Allora a e $|b|$ dipendono solo da v (e non da B).

Dim: Infatti $\text{tr}(M) = 2a$ e $a^2 + b^2 = 1$ ♦

Questa proposizione permette la:

3.1: Definizione: Noteremo ρ l'applicazione $O^+(E) \rightarrow [-1, 1] : r \rightarrow \rho(r) = a_{11}$ dove (a_{ij}) è la matrice di r in una base ortonormale qualsiasi.

L'applicazione ρ , come vedremo, è l'applicazione Cos; per definire l'applicazione Sen bisogna risolvere l'ambiguità sul segno e per questo bisogna orientare gli angoli e il piano euclideo.

ORIENTAZIONE DEL PIANO EUCLIDEO.

Siano B, B' due basi ortonormali di E . Se $A = \text{mat}(\text{Id}_E; B, B')$ allora A è una matrice ortogonale (cfr §8, 11 (iv)). Diremo che B e B' definiscono la stessa orientazione se $\det(A) = 1$, ossia se A è la matrice di una rotazione. Otteniamo in questo modo una relazione d'equivalenza sull'insieme delle basi ortonormali di E ; l'insieme quoziente di questa relazione d'equivalenza è costituito da due elementi (Es.II.14.6). Un'orientazione di E è un elemento di questo insieme quoziente. Sceglijamone uno; ogni base ortonormale appartenente a questa classe d'equivalenza sarà detta diretta (e le altre indirette).

Sia v un vettore con $\|v\| = 1$ e v', v'' i due vettori di norma uno in v^\perp . Una delle due basi (v, v') , (v, v'') è diretta, l'altra indiretta.

4: Proposizione: *Sia E un piano euclideo orientato e B una base diretta. Definiamo $\alpha_B : O^+(E) \rightarrow [-1,1] : r \rightarrow a_{21}$ dove $\text{mat}(r; B, B) = (a_{ij})$. Se B' è un'altra base diretta le applicazioni $\alpha_B, \alpha_{B'}$ sono uguali.*

Dim: Sia $P = \text{mat}(r; B, B)$ e $A = \text{mat}(\text{Id}_E; B, B')$. Per ipotesi A è la matrice di una rotazione. Usando il lemma 1 del §9, si verifica che se $A^{-1} \cdot M \cdot A = (a'_{ij})$, allora $a_{21} = a'_{21}$ ♦

4.1: Osservazione : In realtà se r è una rotazione e se B, B' sono due basi ortonormali dirette allora $\text{mat}(r; B, B) = \text{mat}(r; B', B')$ (cfr 8).

5: Notazione: Noteremo $\alpha : O^+(E) \rightarrow [-1,1]$ l'applicazione α_B dove B è una base diretta.

Siano v e w due vettori unitari (i.e. $\|v\| = \|w\| = 1$). Sia r l'unica rotazione tale che $r(v) = w$ (cfr lemma 2), vorremmo avere $\alpha(r) = \text{sen}(\theta)$ dove θ è l'angolo non orientato tra v e w . C'è però un piccolo problema: l'angolo non orientato θ soddisfa $0 \leq \theta \leq \pi$ e quindi $0 \leq \text{sen}(\theta) \leq 1$. Perciò se $B = (e_1, e_2)$ è una base diretta e se $v = e_1, w = (e_1 - e_2) / \|e_1 - e_2\| = (e_1 - e_2) / \sqrt{2}$; allora $\alpha(r) = -1/\sqrt{2}$ che è senz'altro diverso da $\text{sen}(\theta)$. Per risolvere questa difficoltà bisogna abbandonare la nozione di angolo non orientato ed orientare anche gli angoli.

ANGOLI ORIENTATI.

Sia $S^1(E)$ (o più semplicemente S^1) la circonferenza unità: $S^1(E) = \{x \in E / \|x\| = 1\}$. Se $r \in O^+(E)$ l'applicazione $S^1 \times S^1 \rightarrow S^1 \times S^1 : (x, y) \rightarrow (r(x), r(y))$ è una biiezione. Si definisce così un'operazione del gruppo $O^+(E)$ su $S^1 \times S^1$ (Es.III.6.1). L'angolo orientato $\langle x, y \rangle$ è l'orbita di (x, y) sotto $O^+(E)$: $\langle x, y \rangle = \{(r(x), r(y)) / r \in O^+(E)\}$. Questo

definisce l'angolo orientato tra due vettori di norma uno. Se v, w sono due vettori non nulli qualsiasi, l'angolo orientato $\langle v, w \rangle$ è per definizione uguale a $\langle v', w' \rangle$ dove $v' = v / \|v\|$, $w' = w / \|w\|$.

6: Notazione: Noteremo A l'insieme degli angoli orientati (o più semplicemente angoli).

7: Proposizione: Siano $\alpha \in A$; (v, w) , (x, y) due rappresentanti di α e r l'unica rotazione tale che $r(v) = w$ (risp. r' l'unica rotazione tale che $r'(x) = y$). Allora $r = r'$.

Dim: Esercizio ♦

8: Corollario: Esiste un'applicazione $\Phi: A \rightarrow O^+(E)$: $\alpha = \langle x, y \rangle \mapsto r$, dove r è l'unica rotazione tale che $r(x) = y$. L'applicazione Φ è una biiezione.

Dim: Esercizio ♦

8.1: Osservazione : Possiamo trasportare la struttura di gruppo di $O^+(E)$ su A tramite la biiezione Φ ; questo definisce la somma degli angoli.

9: Proposizione: Sia E un piano vettoriale euclideo orientato e $\varphi: O^+(E) \rightarrow \mathbb{R}^2$: $r \mapsto (\rho(r), \alpha(r))$. L'applicazione φ induce una biiezione tra $O^+(E)$ e S^1 .

Dim: Esercizio ♦

9.1: Osservazione : Per poter definire la biiezione φ bisogna aver orientato E (per definire α).

10: Definizione: Combinando 8 e 9 otteniamo una biiezione: $T := \varphi \circ \Phi: A \rightarrow S^1$. Se $\zeta \in A$ allora si pone $T(\zeta) := (\cos(\zeta), \sin(\zeta))$. La rotazione $\Phi(\zeta) = r$ si chiama la rotazione di angolo ζ .

La matrice di r in una base ortonormale diretta qualsiasi è: $\begin{pmatrix} \cos(\zeta) & -\sin(\zeta) \\ \sin(\zeta) & \cos(\zeta) \end{pmatrix}$; quest'ultimo fatto caratterizza le funzioni trigonometriche dell'angolo ζ e permette di ritrovare tutte le formule di trigonometria. Il contesto più naturale per esprimere tutto questo è quello dei numeri complessi. Indichiamo rapidamente come riallacciare quanto fatto finora con i numeri complessi.

Sia $j: \mathbb{R}^2 \rightarrow \mathbb{C}: (a, b) \mapsto a+ib$; j è una biiezione.

Poi abbiamo $i: C \rightarrow M_2(\mathbb{R})$: $a+ib \rightarrow \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$; i è un'iniezione; osservare che $i(z.z') = i(z).i(z')$ e $i(z+z') = i(z) + i(z')$; $i(C)$ è un sottocorpo di $M_2(\mathbb{R})$. Notiamo S l'insieme dei numeri complessi di modulo 1 (quindi S è l'immagine tramite j della circonferenza unità di \mathbb{R}^2 , e possiamo identificare S e S^1). Abbiamo $i(S^1) = O_2^+(\mathbb{R})$.

Come già visto, una volta scelta un'orientazione del piano vettoriale euclideo possiamo identificare A con, diciamo, S^1 l'insieme dei numeri complessi di modulo uno. Osserviamo che S^1 è un gruppo per la moltiplicazione.

Il problema della misura degli angoli consiste nel trovare un'applicazione $f: \mathbb{R} \rightarrow S^1$ che sia un isomorfismo tra il gruppo additivo $(\mathbb{R}, +)$ e il gruppo moltiplicativo (S^1, \cdot) e che sia bicontinuo; questo permetterebbe di caratterizzare un angolo tramite un unico numero reale ("la sua misura"), in modo che la misura della somma di due angoli sia la somma delle loro misure. Una tale applicazione f non esiste! (S^1 è compatto, \mathbb{R} non lo è). Bisogna quindi ripiegare e chiedere soltanto che f sia suriettiva. Si può dimostrare che in queste condizioni il $\text{Ker } f$ dev'essere necessariamente della forma $a.\mathbb{Z}$. Per definire un'applicazione con le caratteristiche di f , si definiscono le funzioni $\exp(z)$, $\cos(z)$, $\sin(z)$ per ogni numero complesso z , tramite serie di potenze:

$$\exp(z) := \sum_{n=0}^{\infty} (z^n / n!) =: e^z,$$

$$\cos(z) := \sum_{n=0}^{\infty} (-1)^n z^{2n} / (2n)! ; \sin(z) := \sum_{n=0}^{\infty} (-1)^n z^{2n+1} / (2n+1)!$$

Si verifica che se t è un numero reale allora $e^{it} = \cos(t) + i \sin(t)$.

Osservare la formula "magica" che collega i quattro numeri e , i , π e 1:

$$e^{i\pi} = -1$$

(π è la determinazione principale (vedere qui sotto) dell'angolo $\langle e_1, -e_1 \rangle$ dove $e_1 = (1,0)$).

La presentazione tramite numeri complessi permette di ritrovare tutte le formule di trigonometria. Per esempio $e^{i(t+t')} = \cos(t+t') + i \sin(t+t')$; ma d'altra parte: $e^{i(t+t')} = e^{it}.e^{it'} (\text{morfismo di gruppi}) = (\cos(t)+i \sin(t)).(\cos(t')+i \sin(t')) = \cos(t).\cos(t') - \sin(t).\sin(t') + i[\cos(t).\sin(t') + \sin(t).\cos(t')]$. Confrontando le parti reali e immaginarie otteniamo:

$$\cos(t+t') = \cos(t).\cos(t') - \sin(t).\sin(t')$$

$$e: \quad \sin(t+t') = \cos(t).\sin(t') + \sin(t).\cos(t).$$

Visto che $i(S^1) = O_2^+(\mathbb{R})$ (cfr qui sopra) questa presentazione è equivalente (ma più leggera) a quella che consiste nel calcolare con le matrici ortogonali speciali.

Osserviamo tra l'altro che se $z = a+ib$ è un numero complesso qualsiasi allora $z/|z|$ è un numero complesso di modulo uno ($|z| = (a^2+b^2)^{1/2}$). Siccome $z/|z|$ appartiene ad S^1 , possiamo scriverlo nella forma $\cos(\zeta) + i \sin(\zeta)$, e quindi $z =$

$r(\cos(\zeta) + i\sin(\zeta))$ dove $r = |z|$; questa è la rappresentazione polare del numero complesso z .

Prendiamo come base diretta di \mathbb{R}^2 la base canonica $e_1 = (1,0)$, $e_2 = (0,1)$ allora ζ è l'angolo (orientato) $\langle e_1, v \rangle$ dove v ha per coordinate $(a/|z|, b/|z|)$ nella base canonica; ζ è l'angolo dell'unica rotazione che manda e_1 su v .

Detto ciò possiamo definire un'applicazione f con le proprietà richieste: $f : \mathbb{R} \rightarrow S^1$: $t \rightarrow e^{it}$. Abbiamo $\text{Ker}(f) = 2\pi\mathbb{Z}$, (quindi otteniamo una bijezione tra $\mathbb{R}/2\pi\mathbb{Z}$ e S^1 o ancora tra A e $\mathbb{R}/2\pi\mathbb{Z}$). In conclusione abbiamo un'applicazione con le dovute proprietà: $f : \mathbb{R} \rightarrow A$; se ζ è un angolo, $f^{-1}(\zeta)$ è una classe d'equivalenza per la relazione d'equivalenza su \mathbb{R} : $x \sim y \Leftrightarrow$ esiste $k \in \mathbb{Z}$ tale che $x-y = 2k\pi$. Esiste un unico x in $f^{-1}(\zeta)$ tale che $0 \leq x < 2\pi$, x è la determinazione principale dell'angolo ζ , si dice anche che x è la misura (in radianti) dell'angolo ζ . Nella pratica si identifica spesso ζ con $f^{-1}(\zeta)$ o ancora con la sua determinazione principale.

L'applicazione $t \rightarrow e^{it}$ non è l'unico epimorfismo continuo tra $(\mathbb{R}, +)$ e (S^1, \cdot) ma si può dimostrare che ogni tale applicazione è della forma $t \rightarrow e^{i\lambda t}$, per qualche λ non nullo in \mathbb{R} . Per $\lambda = \pi/180$ (risp. $\lambda = \pi/200$) si ottiene la misura degli angoli in gradi (risp. gradianti).

Esercizi:

10.1) Dimostrare la proposizione 7, il corollario 8 e la proposizione 9.

11) Spazio affine euclideo.

Sia E uno spazio vettoriale euclideo. Possiamo considerare E anche con la sua struttura di spazio affine. Nello spazio affine E , grazie al prodotto scalare, sappiamo misurare le distanze e gli angoli, ottenendo così una struttura di spazio affine euclideo su E (più brevemente spazio euclideo). Ogni spazio euclideo di dimensione n è isomorfo a \mathbb{R}^n con il prodotto scalare usuale (e il riferimento standard).

1: Definizione: Il riferimento $(Q; e_1, \dots, e_n)$ dello spazio euclideo E è detto ortonormale (o cartesiano) se (e_1, \dots, e_n) è una base ortonormale.

1.1: Esempio : Il riferimento standard $(O; c_1, \dots, c_n)$ di \mathbb{R}^n dove $O = (0, \dots, 0)$ e dove (c_1, \dots, c_n) è la base canonica è un riferimento ortonormale (si dice anche sistema di coordinate cartesiane).

2: Definizione: Sia E uno spazio euclideo. Due sottospazi affini F, F' di E sono ortogonali se $\text{dir}(F)$ e $\text{dir}(F')$ sono due sottospazi vettoriali ortogonali (cfr §3, Def.4).

Due sottospazi affini F, F' sono perpendicolari se gli spazi vettoriali $\text{dir}(F)^\perp$ e $\text{dir}(F')^\perp$ sono ortogonali.

2.1: Osservazione : Con le notazioni precedenti F e F' sono ortogonali se $\text{dir}(F) \subseteq \text{dir}(F')^\perp$ (o $\text{dir}(F') \subseteq \text{dir}(F)^\perp$) mentre F e F' sono perpendicolari se $\text{dir}(F)^\perp \subseteq \text{dir}(F')$ (o $\text{dir}(F')^\perp \subseteq \text{dir}(F)$); quindi in generale le nozioni di ortogonalità e perpendicolarità sono diverse. Le due nozioni coincidono se F e F' hanno dimensioni complementari: $\dim F + \dim F' = \dim E$; questo è il caso in particolare se F è una retta e F' è un iperpiano.

3: Proposizione: Sia $H = a + F$ un'iperpiano affine dello spazio euclideo E . Per ogni punto p di E esiste una ed un'unica retta, R_p , passante per p e perpendicolare ad H . Sia $\pi: E \rightarrow H$ l'applicazione definita da $\pi(p) = H \cap R_p$.

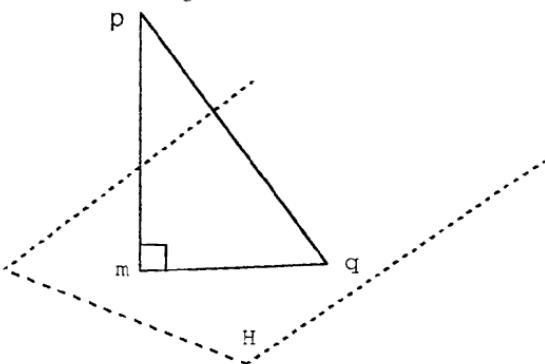
(i) π è un'applicazione affine ("proiezione ortogonale su H ")

(ii) la distanza, $d(p, H)$, dal punto p all'iperpiano H ($d(p, H) := \inf \{d(p, x) / x \in H\}$, cfr Es. IV.6.1) è uguale a $\|p - \pi(p)\|$.

Dim: La retta R_p è la retta $p + F^\perp$.

(i) Cfr Es.3.

(ii) Risulta dal teorema di Pitagora:



4: Definizione: Sia H un iperpiano dello spazio euclideo E . Un versore normale ad H è un vettore unitario (cioè di norma uno) appartenente ad $(\text{dir}H)^\perp$.

4.1: Osservazione : Siccome $\dim(\text{dir}H)^\perp = 1$ ci sono due versori normali ad H : se $(\text{dir}H)^\perp = \langle v \rangle$ allora $n := v / \|v\|$ e $-n$ sono i due versori normali ad H .

5: Lemma: Sia E uno spazio euclideo con un sistema di coordinate cartesiane e sia H l'iperpiano di E di equazione $\sum a_i X_i + d = 0$. I versori normali ad H sono i vettori $\pm a / \|a\|$ dove $a = (a_1, \dots, a_n)$.

Dim: La direzione di H è l'insieme delle soluzioni dell'equazione omogenea $\sum a_i X_i = 0$. Pertanto è chiaro che $(\text{dir}H)^\perp = \langle a \rangle$ ♦

6: Proposizione: Sia E uno spazio euclideo con un sistema di coordinate cartesiane e sia H l'iperpiano di E di equazione $\sum a_i X_i + d = 0$. La distanza dal punto $p = (p_1, \dots, p_n)$ all'iperpiano H è uguale a: $\|\sum a_i p_i + d\| / (\sum a_i^2)^{1/2}$.

Dim: La distanza da p ad H è $\delta = \|p - m\|$ dove m è la proiezione ortogonale di p su H (cfr Prop.3). Se n è un versore normale ad H : $\|(n \mid p-m)\| = |\cos\theta| \cdot \|n\| \cdot \|p-m\| = \|p-m\| = \delta$. D'altra parte, se q è un punto qualsiasi di H : $(n \mid p-q) = (n \mid (p-m)+(m-q)) = (n \mid p-m) + (n \mid m-q) = (n \mid p-m)$ (perché n è ortogonale a $m-q \in \text{dir}(H)$). In conclusione: $\delta = |(n \mid p-q)|$. Sia $n = a / \|a\|$ (cfr

Lema 5; abbiamo: $(a \mid p-q) = \sum a_i(p_i-q_i) = \sum a_i p_i - \sum a_i q_i = \sum a_i p_i + d$, e quindi $\delta = |(n \mid p-q)| = |\sum a_i p_i + d| / \|a\|$ ♦

6.1: Osservazione : Le proposizioni precedenti permettono di risolvere i problemi di distanza tra sottospazi affini nel piano: la distanza da un punto ad una retta è un caso particolare della Prop. 6, e per quanto riguarda la distanza tra due rette R, R' ci sono due casi: (i) $R \cap R' \neq \emptyset$, allora $d(R, R') = 0$, (ii) $R // R'$ allora $d(R, R') = d(p, R')$ dove p è un punto qualsiasi di R (cfr Es.6). Per quanto riguarda il caso dello spazio tridimensionale considerazioni analoghe e la proposizione seguente permettono di ottenere un'analisi completa.

7: Proposizione: Siano $R = a + \langle v \rangle$, $R' = a' + \langle v' \rangle$ due rette non parallele dello spazio tridimensionale euclideo E. Esiste una ed un'unica coppia di punti $(p, p') \in R \times R'$ tale che la retta $D = [p, p']$ sia ortogonale a R e a R' . La distanza da R ad R' è uguale a $\|p-p'\|$.

Dim: Osserviamo che la condizione D ortogonale a R e a R' è equivalente a $p-p' \in v^\perp \cap v'^\perp$. Siccome R e R' non sono parallele $v^\perp \cap v'^\perp$ ha dimensione uno (cfr Es.7) cioè $v^\perp \cap v'^\perp = \langle w \rangle$. Quindi stiamo cercando $p = a + \lambda v$, $p' = a' + \mu v'$ tali che $p-p' = a-a' + \lambda v - \mu v' = \alpha w$, per qualche α . Siccome v, v' e w sono linearmente indipendenti (cfr Es.7) esiste una ed un'unica terna (λ, μ, α) di scalari tale che $a-a = \lambda v + \mu v' - \alpha w$, e questo determina univocamente la coppia (p, p') . Dal teorema di Pitagora segue che $d(R, R') = \|p-p'\|$ ♦

Isometrie

Sia E uno spazio euclideo. Un'isometria è un'applicazione $g: E \rightarrow E$ che conserva le distanze (cfr §8, Def.9): $\|g(v) - g(w)\| = \|v - w\|$, per ogni v, w in E.

8: Lemma: Un'isometria dello spazio euclideo E è un'affinità. Più precisamente se $g: E \rightarrow E$ è un'isometria, allora $g = t_a \circ v$ dove v è un'isometria vettoriale.

Dim: Sia $a = g(0)$ e sia $v = t_{-a} \circ g$, dove t_{-a} è la traslazione di vettore $-a$. L'applicazione v è un'isometria perché $\|v(x) - v(y)\| = \|g(x) - g(y)\| = \|x -$

y 11. Siccome $v(0) = 0$, v è un'isometria vettoriale (§8, 12). Pertanto $g = t_a \circ v$ è un'affinità (III. §5, Lemma 4)♦

9: Definizione: Con le notazioni del lemma 8 l'isometria g si dice diretta se $\det(v) = 1$, e inversa se $\det(v) = -1$.

10: Lemma: Sia $\text{Isom}(E)$ l'insieme delle isometrie di E , allora $\text{Isom}(E)$ è un gruppo per la composizione delle applicazioni. L'insieme $\text{Isom}_+(E)$ delle isometrie dirette è un sottogruppo di $\text{Isom}(E)$.

Dim: Es.8♦

Isometrie del piano.

Sia E un piano euclideo con un riferimento ortonormale $(O; e_1, e_2)$. Ricordiamo (cfr III. §5, §6) che la scelta di questo riferimento conferisce al piano affine E una struttura di \mathbb{R} -spazio vettoriale e che per questa struttura l'origine O si identifica col vettore nullo, 0 . Equivalentemente possiamo pensare ad E come ad uno spazio vettoriale euclideo, munito di una base ortonormale (e_1, e_2) e considerare sottospazi affini di questo spazio vettoriale (cfr III); pertanto non faremo distinzioni tra il punto O e il vettore 0 , passando alternativamente da un linguaggio all'altro.

Sia adesso $g: E \rightarrow F$ un'isometria diretta che fissa un punto p : $g(p) = p$. Sia $f = t_{-p} \circ g \circ t_p$; f è un'isometria diretta (Es.8) e $f(0) = 0$ quindi f è una rotazione vettoriale, $f = r_\theta$ dove θ è l'angolo della rotazione (cfr §9, §10). Pertanto $g = t_p \circ r_\theta \circ t_{-p}$, e vediamo che g è una rotazione di centro p , angolo θ : $g = r_{p,\theta} = t_p \circ r_\theta \circ t_{-p}$. Viceversa se $g = t_a \circ v$ è un'isometria diretta, consideriamo $f = \text{Id} - v$.

Se $v \neq \text{Id}$ (cioè se g non è una traslazione) allora f è invertibile (1 non è autovalore di un'isometria vettoriale positiva $v \neq \text{Id}$, cfr §9, Lemma 8). Quindi esiste $x \in E$ tale che: $x - v(x) = a$, e g fissa il punto x . In conclusione abbiamo dimostrato:

11: Proposizione: Sia $g \neq \text{Id}$ un'isometria diretta del piano euclideo E . Se g fissa un punto, allora g è una rotazione di centro quel punto; altrimenti g non ha punti fissi ed è una traslazione di vettore non nullo.

Introduciamo una definizione prima di passare al caso delle isometrie inverse:

12: Definizione: Sia E uno spazio euclideo e sia H un iperpiano di E . Se $p \in E$, $r_H(p)$ è definito dalla relazione: $r_H(p) - n = -(p + n)$, dove n è la proiezione ortogonale di p su H . L'applicazione $r_H: E \rightarrow E$ si chiama la riflessione ortogonale di asse H .

12.1: Osservazione : Se H è un iperpiano passante per l'origine r_H è stata definita in §9, Def.3. Si verifica che una riflessione ortogonale è un'isometria (Es.9).

Sia adesso $g: E \rightarrow E$ un'isometria inversa, $g = t_a \circ v$. L'isometria vettoriale v è negativa ed è quindi una riflessione ortogonale (cfr §9) di asse E_+ dove $E_+ = E_v(1)$ è l'autospazio associato all'autovalore 1. Inoltre $E = E_+ \oplus E_-$, dove $E_- = E_v(-1)$. Ogni vettore q si scrive in modo unico $q = q_+ + q_-$, $v(q_+) = q_+$, $v(q_-) = -q_-$. Osserviamo che g ha un punto fisso se e solo se $a \in E_-$: infatti $g(x) = x \Leftrightarrow a+v(x) = x \Leftrightarrow a \in \text{Im}(Id-v)$ e $\text{Im}(Id-v) = E_-$ (perché $v(y-v(y)) = v(y) - v^2(y) = v(y) - y$). Se $a \in E_-$ la retta $D = a' + E_+$, con $a' = a/2$, è fissa. Infatti sia $x = a' + e \in D$, allora $g(x) = v(a'+e)+a = v(a/2) + e + a = -a/2 + e + a = x$. In queste condizioni g è la riflessione ortogonale di asse la retta D : se p è un punto qualsiasi, la proiezione ortogonale di p su D è $\pi_D(p) = a/2 + p_+$. Quindi l'immagine di p tramite la riflessione di asse D è $\pi_D(p) - p_- + a/2 = p_+ - p_- + a = g(p)$. Abbiamo dimostrato:

12: Proposizione: Se un'isometria inversa ha un punto fisso allora ha tutta una retta fissa ed è una riflessione ortogonale di asse la retta fissa.

Se l'isometria inversa g non ha punti fissi allora $a \notin E_-$, cioè $a_+ \neq 0$. Questa volta la retta $D = a/2 + E_+$ è globalmente invariante (ma non invariante punto per punto). Se $x = a/2 + e \in D$, $g(x) = v(a/2 + e) + a = a_+/2 - a_-/2 + e + a_+ + a_- = x + a_+$. Quindi $g|_D = t_{a_+}$. Pertanto l'isometria inversa $t_{a_+} \circ g$ fissa la retta D . Dalla proposizione 12 segue che $t_{a_+} \circ g = r_D$, la riflessione ortogonale di asse la retta D . Quindi $g = t_{a_+} \circ r_D$. Osserviamo che il vettore a_+ è parallelo a D ($\text{dir}(D) = E_+$).

13: Definizione: Una glissoriflessione è un'isometria del piano euclideo E ottenuta come composizione, $t_w \circ r_L$, di una riflessione ortogonale, r_L , di asse una retta L e di una traslazione, t_w , di vettore non nullo w , parallelo a L .

Riassumendo abbiamo il seguente teorema di classificazione delle isometrie del piano euclideo:

14: Teorema: (*Chasles, 1831*) *Un'isometria del piano euclideo che ammette almeno un punto fisso è una rotazione o una riflessione ortogonale, a seconda che sia diretta o inversa. Un'isometria del piano euclideo senza punti fissi è una traslazione o una glissoriflessione, a seconda che sia diretta o inversa.*

Nel caso dello spazio tridimensionale si ha un risultato analogo (dovuto a Eulero, 1776) che ci limiteremo ad enunciare. Una glissoriflessione è la composizione di una riflessione con una traslazione di vettore parallelo al piano della riflessione. Una glissorotazione è la composizione di una rotazione con una traslazione in una direzione parallela all'asse della rotazione. Una riflessione rotatoria è la composizione di una rotazione con una riflessione rispetto ad un piano perpendicolare all'asse della rotazione.

15: Teorema: *Ogni isometria dello spazio tridimensionale euclideo è di uno dei sei tipi seguenti: traslazione, rotazione, riflessione, glissoriflessione, glissorotazione, riflessione rotatoria.*

Esercizi:

11.1) Dimostrare che ogni spazio euclideo, E , di dimensione n è isomorfo a \mathbf{R}^n : più precisamente esiste un isomorfismo affine $f: E \rightarrow \mathbf{R}^n$ tale $d(x, y) = d(f(x), f(y))$ per ogni x, y in E (cioè f è un'isometria).

11.2) Sia E un piano euclideo con riferimento cartesiano $(O; e_1, e_2)$. Siano $P = (p, p')$, $Q = (q, q')$, $R = (r, r')$ tre punti affinamente indipendenti. Dimostrare che l'area del parallelogramma individuato da P, Q, R è uguale a $|\det(M)|$ dove $M = \begin{pmatrix} p-r & p'-r' \\ q-r & q'-r' \end{pmatrix}$.

Esplicitare nel caso $E = \mathbf{R}^2$ con il riferimento standard. Generalizzare nel caso $\dim(E) > 2$.

11.3) Dimostrare che la proiezione ortogonale su un iperpiano è un'applicazione affine.

11.4) In \mathbf{R}^2 calcolare la distanza dal punto $p = (1, 1)$ alla retta di equazione $x+y+2=0$.

11.5) In \mathbf{R}^3 determinare la proiezione ortogonale del punto $p = (1, 0, -1)$ sul piano H di equazione $x+y+z+1=0$. Calcolare $d(p, H)$.

11.6) Siano A, A' due sottospazi affini dello spazio euclideo E . Se $A // A'$ dimostrare che $d(A, A') = d(p, A')$ dove p è un punto qualsiasi di A .

11.7) In uno spazio euclideo tridimensionale riferito ad un sistema di coordinate cartesiane, siano $R = a + \langle v \rangle$, $R' = a' + \langle v' \rangle$, due rette non parallele.

(i) Dimostrare che $v^\perp \cap v'^\perp$ ha dimensione uno.

(ii) Con le notazioni precedenti, se $\langle w \rangle = v^\perp \cap v'^\perp$ dimostrare che v, v', w formano una base dello spazio vettoriale E .

(iii) Sia n un versore dell'unica retta ortogonale a R e R' , e intersecante R e R' . Dimostrare che $d(R, R') = |(b - q)|$ dove q e q' sono due punti qualsiasi di R e R' .

(iv) Sia $V = \begin{pmatrix} v_1 & v_2 & v_3 \\ v'_1 & v'_2 & v'_3 \end{pmatrix}$, dove $v = (v_1, v_2, v_3)$, $v' = (v'_1, v'_2, v'_3)$. Si nota δ_i il minore 2×2 ottenuto da V sopprimendo la i -esima colonna. Dimostrare che $N = (n_1, n_2, n_3)$ dove $n_i = (-1)^{i+1} \delta_i$ è ortogonale a v e a v' .

$$\begin{pmatrix} a_1 - a'_1 & a_2 - a'_2 & a_3 - a'_3 \\ v_1 & v_2 & v_3 \\ v'_1 & v'_2 & v'_3 \end{pmatrix}$$

(v) Concludere che $d(R, R') = |D|/t$, dove $D = \det \begin{pmatrix} a_1 - a'_1 & a_2 - a'_2 & a_3 - a'_3 \\ v_1 & v_2 & v_3 \\ v'_1 & v'_2 & v'_3 \end{pmatrix}$ (a_i e a'_i

sono le coordinate dei punti a, a'), e dove $t = (\delta_1^2 + \delta_2^2 + \delta_3^2)^{1/2}$.

11.8) Sia E uno spazio euclideo. Dimostrare che $\text{Isom}(E)$ è un gruppo per la composizione delle applicazioni, e che $\text{Isom}_+(E)$ è un sottogruppo di $\text{Isom}(E)$. Dimostrare che T_E , l'insieme delle traslazioni di E , è un sottogruppo di $\text{Isom}_+(E)$.

11.9) (i) Dimostrare che una riflessione ortogonale è un'isometria.

(ii) In \mathbb{R}^2 sia D la retta di equazione $x+y-1=0$. Dare le equazioni della riflessione ortogonale r_D .

(iii) In E sia H il piano di equazione $x+y-z+1=0$ e sia $p=(1, 1, 1)$. Determinare $r_H(p)$.

11.10) Sia E un piano euclideo. Un sottogruppo G di $\text{Isom}(E)$ si dice discontinuo se per ogni $p \in E$ esiste $r > 0$ tale che nessuno dei punti $g(p)$, $g \in G$ sia contenuto nel disco $D(p; r) = \{q / d(p, q) < r\}$.

(i) Mostrare che ogni sottogruppo finito di $\text{Isom}(E)$ è discontinuo.

(ii) Sia $v \in E$ fissato. Mostrare che $T_E(v) = \{t_{\alpha v} / \alpha \in \mathbb{Z}\}$ è un sottogruppo discontinuo di $\text{Isom}(E)$.

(iii) Un sottogruppo finito di $\text{Isom}(E)$ non contiene traslazioni diverse dall'identità, e quindi non contiene glissoriflessioni.

12) Operatori aggiunti e teorema spettrale.

In questo paragrafo si dimostra che ogni matrice simmetrica reale è congruente, tramite un cambiamento di base ortonormale, ad una matrice diagonale ("teorema spettrale"). Come conseguenza si ottiene che ogni matrice simmetrica reale è diagonalizzabile. Il teorema spettrale ha conseguenze geometriche importanti (cfr classificazione euclidea delle coniche, § 13). Si introduce inoltre la nozione di operatore aggiunto.

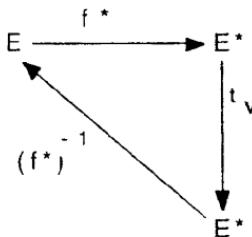
OPERATORE AGGIUNTO RISPETTO AD UNA FORMA BILINEARE NON DEGENERE.

Sia $f : E \times E \rightarrow k$ una forma bilineare simmetrica, non degenere. Abbiamo visto (cfr IV.§3) che si può stabilire, tramite f , un isomorfismo tra E e il suo duale E^* :

$$f^* : F \rightarrow E^*; x \rightarrow f^*(x) = f_x$$

dove f_x è la forma lineare: $E \rightarrow k$: $y \rightarrow f_x(y) = f(x, y)$.

1: Definizione: Sia $v : E \rightarrow E$ un endomorfismo. L'aggiunto (o il trasposto) di v rispetto a f è l'endomorfismo $v^* := (f^*)^{-1} \circ (t_v) \circ f^*$.



2: Proposizione: Con le notazioni precedenti:

- (i) $\forall (x, y) \in E^2: f(v(x), y) = f(x, v^*(y))$
- (ii) se $w : E \rightarrow E$ è un endomorfismo tale che per ogni x, y in E : $f(w(x), y) = f(x, w(y))$, allora $w = v^*$.

Dim: (i) Per ogni y in E , $(t_v \circ f^*)(y) = t_v(f_y) = f_y \circ v$. Quindi $(t_v \circ f^*)(y)$ è la forma lineare: $E \rightarrow k$: $x \rightarrow f(y, v(x))$. Visto che f^* è un isomorfismo esiste uno ed un'unico vettore z tale che $f^*(z) = (t_v \circ f^*)(y)$ (\dagger). Componendo con $(f^*)^{-1}$: z

$= [(f^*)^{-1} \circ (t_v) \circ (f^*)](y) = v^*(y)$. Da (\dagger) viene $f(v^*(y), x) = f(y, v(x))$ per ogni x, y in E .

(ii) Se $f(v(x), y) = f(x, w(y))$ allora $[(t_v \circ f^*)(y)](x) = [f^*(w(y))](x)$ per ogni x ; ossia le forme lineari $(t_v \circ f^*)(y)$ e $f^*(w(y))$ sono uguali. Componendo con $(f^*)^{-1}$: $v^*(y) = w(y)$ per ogni y , quindi $w = v^*$ ♦

3: **Proposizione:** Con le notazioni precedenti (in particolare f non degenere) sia B una base di E e $A = \text{mat}_B(f)$, $M = \text{mat}(v; B, B)$. Allora $\text{mat}(v^*; B, B) = A^{-1} t M A$.

Dim: Risulta dalla definizione e da §3), 5, 6♦

4: **Definizione:** Un endomorfismo $v: E \rightarrow E$ si dice autoaggiunto (o simmetrico) rispetto a f se $v = v^*$.

5: **Corollario:** Con le notazioni precedenti v è autoaggiunto rispetto a f se e solo se: $M = A^{-1} t M A$.

IL CASO EUCLIDEO.

Sia E uno spazio vettoriale euclideo. Ricordiamo che notiamo $(x|y)$ invece di $f(x,y)$; inoltre se B è una base ortonormale allora $\text{mat}_B(f) = I_n$. Dalla Prop.3 abbiamo:

6: **Corollario:** Sia $v: E \rightarrow E$ un endomorfismo, B una base ortonormale e $\text{mat}(v; B, B) = M$. Allora $\text{mat}(v^*; B, B) = t M$.

7: **Proposizione:** Con le notazioni precedenti v è autoaggiunto (o simmetrico) rispetto al prodotto scalare se e solo se $\text{mat}(v; B, B)$ è simmetrica per ogni base ortonormale di E .

Notiamo inoltre:

8: **Proposizione:** Sia E uno spazio vettoriale euclideo e v un endomorfismo di E ; v è un'isometria se e solo se $v^* = v^{-1}$.

Dim: Sia B una base ortonormale e $M = \text{mat}(v; B, B)$, $M^* = \text{mat}(v^*; B, B)$. Abbiamo $M^* = t M$ (cfr 6) d'altra parte v è un'isometria se e solo se $M^{-1} = t M$ ♦

DIAGONALIZZAZIONE DELLE MATRICI SIMMETRICHE REALI.

Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica reale. Il polinomio caratteristico di A , $P_A(X)$, è un polinomio a coefficienti reali; lo si può anche considerare come polinomio a coefficienti complessi e come tale ha tutte le sue radici in \mathbb{C} (perché \mathbb{C} è algebricamente chiuso). Si ricorda che un numero complesso z è reale se e solo se è uguale al suo coniugato: $z = z^*$.

9: Lemma: *Ogni radice di $P_A(X)$ è reale.*

Dim: Sia λ una radice di $P_A(X)$. Possiamo associare ad A un endomorfismo: $\mathbb{C}^n \rightarrow \mathbb{C}^n: Z \mapsto A.Z$, dove ${}^t Z = (z_1, \dots, z_n)$, $z_i \in \mathbb{C}$; ovviamente λ è un autovalore di questo endomorfismo. Sia Y un autovettore associato a λ : $A.Y = \lambda.Y$. Prendendo i coniugati coordinate per coordinate: $A^*.Y^* = \lambda^*.Y^*$. Ma A è a coefficienti reali quindi $A^* = A$ e perciò $A.Y^* = \lambda^*.Y^*$. Adesso calcoliamo ${}^t Y^*.A.Y$ in due modi diversi:

$${}^t Y^*.A.Y = ({}^t Y^*.A).Y = ({}^t Y^*.{}^t A).Y = {}^t (A.Y^*).Y = {}^t (\lambda^*.Y^*).Y = \lambda^*.({}^t Y^*.Y)$$

$${}^t Y^*.A.Y = {}^t Y^*(A.Y) = {}^t Y^*(\lambda.Y) = \lambda.({}^t Y^*.Y).$$

Osserviamo che se $Y = (y_1, \dots, y_n)$ allora $({}^t Y^*.Y) = |y_1|^2 + \dots + |y_n|^2 > 0$ perché Y è un vettore non nullo. Deduciamo che $\lambda^* = \lambda$ ossia λ è reale ♦

10: Teorema ("spettrale"): *Sia E uno spazio vettoriale euclideo e $v: E \rightarrow E$ un operatore simmetrico. Esiste una base ortonormale di E , B , tale che $\text{mat}(v; B, B)$ sia diagonale.*

Dim: Si procede per induzione su $n = \dim(E)$. Se $n = 1$ non c'è niente da dimostrare. Assumiamo il teorema per $n-1$ e dimostriamolo per n . Sia B' una base ortonormale di E . Abbiamo che $\text{mat}(v; B', B')$ è simmetrica (cfr lemma 7). Dal lemma precedente, esiste un autovalore λ e quindi un autovettore, x , associato. Sia $H = x^\perp$. Mostriamo che $v(H) \subseteq H$: se $y \in H$ allora $(v(y)|x) = (y|v^*(x))$ ma $v^* = v$ perché v è simmetrico, perciò: $(v(y)|x) = (y|v(x)) = \lambda(y|x)$ (perché x è un autovettore) e $(y|x) = 0$ perché $y \in H = x^\perp$. Quindi $v(H) \subseteq H$. La restrizione di v ad H è dunque un operatore simmetrico di H . Per ipotesi di induzione (H con la restrizione del prodotto scalare è uno spazio vettoriale euclideo) esiste una base ortonormale di H , B_H , tale che la matrice di $v|_H$ sia diagonale. Completiamo B_H ad una base,

B, di E prendendo un vettore di norma uno in $\langle x \rangle$. E' chiaro che $\text{mat}(v; B, B)$ è diagonale♦

Una formulazione equivalente:

11: Teorema: Per ogni matrice simmetrica reale, $A \in M_n(\mathbb{R})$, esiste una matrice ortogonale M tale che $M^{-1}AM$ sia diagonale.

O ancora:

12: Teorema: Per ogni forma quadratica $q: E \rightarrow \mathbb{R}$ su uno spazio vettoriale euclideo di dimensione finita, esiste una base ortonormale diagonalizzante.

12.1: Osservazione : Nella dimostrazione del teorema spettrale abbiamo usato il concetto di operatore aggiunto solo per dire che $(v(y)|x) = (y|v(x))$ dove $\text{mat}(v; B, B) = A$ è simmetrica (e B è ortonormale). Questo risulta molto semplicemente da: $(y|v(x)) = {}^tY.(A.X) = {}^tY.({}^tA.X) = ({}^tY.{}^tA).X = {}^t(A.Y).X = (v(y)|x)$.

Esercizi:

12.1) Siano a, b, c, f quattro numeri reali. Dimostrare che il polinomio $X^3 - (a+b+c)X^2 + X(bc - f^2 + ab + ac) - abc + af^2$, ha tre radici reali.

12.2) Determinare una matrice ortogonale, M , tale che tMAM sia diagonale dove $A =$

$$(a) \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, (b) \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 2 & 0 & -1 \end{pmatrix}, (c) \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

13) Applicazioni alla classificazione delle coniche.

Si mostra come i risultati precedenti permettano di stabilire la classificazione affine ed euclidea delle coniche piane. Va osservato, peraltro, che la classificazione delle coniche (e più generalmente delle quadriche) assume un aspetto più naturale nell'ambito della geometria proiettiva.

1: Definizione: Una conica, C , nel piano k^2 è l'insieme dei punti che soddisfano un'equazione di secondo grado a coefficienti in $k : C = \{(x, y) \in k^2 / P(x, y) = 0\}$ dove $P(X, Y) = a_{11}X^2 + a_{22}Y^2 + 2a_{12}XY + 2a_{01}X + 2a_{02}Y + a_{00}; a_{ij} \in k$.

1.1: Osservazione : La relazione $P(X, Y) = 0$ si può scrivere in forma

matriciale nel modo seguente: $(1, X, Y) \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = 0$, con $A = (a_{ij})$ matrice simmetrica (cioè $a_{ij} = a_{ji}$).

2: Classificazione affine delle coniche: Tramite cambiamenti di variabili affini (i.e. affinità del piano) si cercherà di scrivere l'equazione di una conica nella forma più semplice possibile (forma ridotta, canonica). Si otterrà così una lista (finita) di forme ridotte non equivalenti. Lo studio delle proprietà affini delle coniche si riduce quindi allo studio delle proprietà affini delle coniche in forma ridotta.

3: Cambiamenti di variabili affini: Si tratta di cambiamenti di variabili dati da un'affinità:

$$X = m_{11}X' + m_{12}Y' + c_1$$

$$Y = m_{21}X' + m_{22}Y' + c_2$$

Sotto forma matriciale: $\begin{pmatrix} X \\ Y \end{pmatrix} = M_0 \begin{pmatrix} X' \\ Y' \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, dove $M_0 = (m_{ij})$ e $\det(M_0) \neq 0$.

Osserviamo che questa relazione si può scrivere anche: $\begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = M \begin{pmatrix} 1 \\ X' \\ Y' \end{pmatrix}$ (*),

dove $M = \begin{pmatrix} 1 & 0 & 0 \\ c_1 & m_{11} & m_{12} \\ c_2 & m_{21} & m_{22} \end{pmatrix}$. Effettuando la sostituzione (*), $(1, X, Y)A^t(1, X, Y)$

diventa: $(1, X', Y').A'.{}^t(1, X', Y')$ con $A' = {}^tM.A.M$. Siccome M è invertibile (perché $\det(M_0) \neq 0$), A e A' sono congruenti. In particolare $\text{rango}(A) = \text{rango}(A')$ (IV. §1, lemma 9).

4: Definizione: Con le notazioni precedenti, il rango della conica C è il rango della matrice A . Il rango è un invariante affine. La conica C è non degenere se $\text{rgo}(C) = 3$, degenere se $\text{rgo}(C) \leq 2$.

Notiamo $A_0 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, la matrice della forma quadratica presente nell'equazione di C (parte omogenea di grado due): $(X, Y).A_0.{}^t(X, Y) = a_{11}X^2 + a_{22}Y^2 + 2a_{12}XY$ ($a_{21} = a_{12}$). Se consideriamo il prodotto ${}^tM.A.M = A'$ vediamo che $A'_0 = {}^tM_0.A_0.M_0$, dove A'_0 è la matrice della parte omogenea di grado due dell'equazione trasformata. Siccome M_0 è invertibile, A_0 e A'_0 sono congruenti e hanno lo stesso rango.

4.1: Osservazione : Risulta dalla discussione precedente che il rango di A_0 è un invariante affine della conica C . Inoltre se $k = \mathbb{R}$, il segno di $\det(A_0)$ è un invariante affine di C (perché $(\det M_0)^2 > 0$).

5: Definizione: Con le notazioni precedenti il rango di A_0 è un invariante affine.

- (i) Se $\det(A_0) = 0$, C è una parabola.
- (ii) Se $\det(A_0) \neq 0$, C è una conica a centro.
- (iii) Inoltre se $k = \mathbb{R}$ e se $\det(A_0) > 0$, C è un'ellisse; se $\det(A_0) < 0$, C è un'iperbole.

Cerchiamo adesso di giustificare queste definizioni, in particolare (ii).

6: Definizione: Un punto (x_0, y_0) è centro di simmetria della conica C se: $(x, y) \in C$ implica $(2x_0-x, 2y_0-y) \in C$.

6.1: Osservazione : Il punto $(2x_0-x, 2y_0-y)$ è il simmetrico di (x, y) rispetto a (x_0, y_0) .

7: Proposizione: Una conica non degenere ha un centro di simmetria se e solo se $\det A_0 \neq 0$. Se $\det A_0 \neq 0$, il centro di simmetria è unico e le sue coordinate sono le soluzioni del sistema di Cramer:

$$A_0 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_{01} \\ -a_{02} \end{pmatrix}.$$

Dim: Osserviamo che $P(x, y) = x(a_{11}x + a_{12}y + 2a_{01}) + y(a_{22}y + a_{12}x + 2a_{02}) + a_{00}$. Quindi $P(2x_0 - x, 2y_0 - y) = (2x_0 - x)[2(a_{11}x_0 + a_{12}y_0 + a_{01}) - a_{11}x - a_{12}y] + (2y_0 - y)[2(a_{22}y_0 + a_{12}x_0 + a_{02}) - a_{22}y - a_{12}x] + a_{00}$. Questo si può riscrivere: $P(2x_0 - x, 2y_0 - y) = 4x_0(x_0a_{11} + y_0a_{12} + a_{01}) + 4y_0(a_{22}y_0 + a_{12}x_0 + a_{02}) - 2x(x_0a_{11} + y_0a_{12} + a_{01}) - 2y(a_{22}y_0 + a_{12}x_0 + a_{02}) + a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + 2x(-a_{11}x_0 - y_0a_{12}) + 2y(-x_0a_{12} - y_0a_{22}) + a_{00}$. Uguagliando a $P(x, y)$ viene il sistema:

$$(\alpha) 4x_0(x_0a_{11} + y_0a_{12} + a_{01}) + 4y_0(a_{22}y_0 + a_{12}x_0 + a_{02}) = 0$$

$$(\beta) x_0a_{11} + y_0a_{12} + a_{01} = 0$$

$$(\gamma) x_0a_{12} + y_0a_{22} + a_{02} = 0$$

Le ultime due equazioni implicano la prima. Le equazioni $(\beta), (\gamma)$ si scrivono sotto forma matriciale $A_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} -a_{01} \\ -a_{02} \end{pmatrix}$. Se $\det A_0 = 0$ e se il sistema è compatibile allora $a_{11} = ta_{12}, a_{12} = ta_{22}, a_{01} = ta_{02}$; ossia $a_{11} = t^2a_{22}, a_{01} = ta_{02}, a_{12} = ta_{22}$. In queste condizioni le ultime due righe della matrice A sono proporzionali, quindi $\det A = 0$ e la conica C è degenere. Pertanto se C è non degenere e se il sistema è compatibile allora $\det A_0 \neq 0$, il sistema è di Cramer e il centro di simmetria è unico♦

7.1: Osservazione : Il sistema lineare $A_0 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_{01} \\ -a_{02} \end{pmatrix}$ si scrive anche $\partial P(x,y)/\partial x = 0, \partial P(x,y)/\partial y = 0$

La classificazione affine delle coniche è data da:

8: Teorema: Ogni conica non degenere di k^2 è affinamente equivalente ad una delle seguenti:

I) k algebricamente chiuso:

$$X^2 + Y^2 - 1 = 0 \text{ (a centro)}$$

$$Y^2 = X \text{ (parabola)}$$

II) $k = \mathbb{R}$

$$X^2 + Y^2 - 1 = 0 \text{ (ellisse)}$$

$$X^2 - Y^2 - 1 = 0 \text{ (iperbole)}$$

$$X = Y^2 \text{ (parabola)}$$

$$X^2 + Y^2 + 1 = 0 \text{ (ellisse senza punti reali).}$$

Le coniche di ognuno dei gruppi precedenti sono due a due non affinamente equivalenti.

Dim: 1º) La prima tappa consiste nell'eliminare il termine misto $2\alpha_{12}XY$. Grazie al teorema 4 di IV. §4, possiamo diagonalizzare la forma quadratica presente nell'equazione di C: esiste M_0 invertibile tale che: ${}^t M_0 \cdot A_0 \cdot M_0 = A'_0$ con A'_0 diagonale, $A'_0 = (\alpha_{ij})$. Dopo questa trasformazione possiamo supporre l'equazione della forma: $\alpha_{11}X^2 + \alpha_{22}Y^2 + 2\alpha_{01}X + 2\alpha_{02}Y + \alpha_{00} = 0$. Osserviamo che C è non degenere se e solo se: $-\alpha_{11}\alpha_{02}^2 + \alpha_{22}\alpha_{00}\alpha_{11} - \alpha_{22}\alpha_{01}^2 \neq 0$ (*).

2º) Adesso bisogna eliminare i termini lineari e il termine costante.

(a) Supponiamo $\alpha_{11} \neq 0$ e $\alpha_{22} \neq 0$ (quindi C è a centro):

Si procede per completamento del quadrato: $\alpha_{11}X^2 + 2\alpha_{01}X = \alpha_{11}(X + \alpha_{01}/\alpha_{11})^2 - (\alpha_{01}/\alpha_{11})^2$. Nello stesso modo: $\alpha_{22}Y^2 + 2\alpha_{02}Y = \alpha_{22}(Y + \alpha_{02}/\alpha_{22})^2 - (\alpha_{02}/\alpha_{22})^2$. Dopo il cambiamento di coordinate: $X' = X + \alpha_{01}/\alpha_{11}$, $Y' = Y + \alpha_{02}/\alpha_{22}$, l'equazione è della forma: $\alpha_{11}X'^2 + \alpha_{22}Y'^2 + c = 0$. Abbiamo $c \neq 0$ perché C è non degenere. Dividendo per $-c$, ci riconduciamo ad un'equazione del tipo: $aX^2 + bY^2 - 1 = 0$, $a \neq 0$, $b \neq 0$ (1).

(b) Supponiamo $\alpha_{11}\alpha_{22} = 0$. Non si può avere $\alpha_{11} = \alpha_{22} = 0$ perché C è non degenere (cfr (*)). Possiamo quindi assumere $\alpha_{11} = 0$, $\alpha_{22} \neq 0$. Effettuando la traslazione: $X' = X$, $Y' = Y + \alpha_{02}/\alpha_{22}$, l'equazione diventa: $\alpha_{22}Y'^2 + 2\alpha_{01}X' + d = 0$. Abbiamo $\alpha_{01} \neq 0$ perché C è non degenere. Ponendo $X'' = X' + d/2\alpha_{01}$, l'equazione diventa del tipo: $aY'^2 + 2bX' = 0$ con $ab \neq 0$ (2).

3º) Adesso si normalizzano i coefficienti, è qui che bisogna distinguere il caso k algebricamente chiuso dal caso $k = \mathbb{R}$.

(1) Sia C la conica di equazione $aX^2 + bY^2 - 1 = 0$ con $ab \neq 0$. Se k è algebricamente chiuso esistono α, β tali che $\alpha^2 = a$, $\beta^2 = b$. Ponendo $X' = \alpha X$, $Y' = \beta Y$, l'equazione diventa: $X'^2 + Y'^2 - 1 = 0$.

Se $k = \mathbb{R}$ e se $a < 0$ allora esiste α tale che $\alpha^2 = -a$. Stesso ragionamento se $b < 0$. Vediamo quindi che, secondo i segni di a, b , l'equazione si può mettere nella forma: $\pm X^2 \pm Y^2 - 1 = 0$. Quindi, scambiando semmai X' con Y' abbiamo i tre tipi seguenti: $-X^2 + Y^2 - 1 = 0$, $X^2 + Y^2 - 1 = 0$, $-X^2 - Y^2 - 1 = 0$.

(2) Sia C la conica di equazione $aY^2 + 2bX = 0$ con $ab \neq 0$.

Se k è algebricamente chiuso sia α tale che $\alpha^2 = a$. Ponendo $\alpha Y = Y'$, $-2bX = X'$, l'equazione diventa: $X' = Y'^2$. Se $k = \mathbb{R}$ possiamo sempre supporre $a > 0$ (moltiplicare eventualmente l'equazione per -1). Come prima ci riduciamo a: $X' = Y'^2$.

Rimane da mostrare che le coniche considerate sono due a due non equivalenti. Per questo consideriamo le matrici A_0 corrispondenti. Se k è algebricamente chiuso: la conica di equazione $X^2 - Y^2 - 1 = 0$ non è affinamente equivalente alla conica $Y^2 - X = 0$ perché nel primo caso $\text{rango}(A_0) = 2$ mentre nel secondo caso $\text{rango}(A_0) = 1$ (cfr 4.1).

Se $k = \mathbb{R}$ i quattro tipi: (A) $X^2 + Y^2 - 1 = 0$ (ellisse), (B) $X^2 - Y^2 - 1 = 0$ (iperbole), (C) $X = Y^2$ (parabola), (D) $X^2 + Y^2 + 1 = 0$ (ellisse senza punti reali), sono due a due non affinamente equivalenti. Infatti (C) è l'unico caso in cui $\text{rango}(A_0) = 1$, quindi non è equivalente a nessuno degli altri casi. Il caso (B) non è equivalente agli altri perché è l'unico con $\text{rango}(A_0) = 2$, $\det(A_0) < 0$ (cfr 4.1). Finalmente (A) e (D) non sono equivalenti perché hanno supporti diversi: il luogo dei punti di \mathbb{R}^2 soddisfacenti (A) è la circonferenza di raggio uno, mentre il supporto di (D) è vuoto ♦

8.1: Osservazione : Si può completare la classificazione con l'analisi dei casi degeneri: se k è algebricamente chiuso abbiamo tre tipi: $X^2 + Y^2 = 0$ (*conica a centro degenero*), $Y^2 = 1$ (*parabola degenera*), $Y^2 = 0$ (*conica doppiamente degenera*).

Se $k = \mathbb{R}$ si ottiene: $X^2 + Y^2 = 0$ (*ellisse degenera*), $-X^2 + Y^2 = 0$ (*iperbole degenera*), $Y^2 \pm 1 = 0$ (*parabole degeneri*), $Y^2 = 0$ (*conica doppiamente degenera*).

Classificazione euclidea.

Si procede come prima ma questa volta usando solo trasformazioni euclidee (isometrie). Si ottiene pertanto una classificazione diversa. Per esempio dal punto di vista affine la circonferenza di centro l'origine e di raggio a : $X^2 + Y^2 = a^2$ è equivalente alla circonferenza unità: ponendo $X' = X/a$, $Y' = Y/a$ l'equazione diventa $X'^2 + Y'^2 = 1$, e l'applicazione $f(X, Y) = (X/a, Y/a)$ è un'affinità. Però l'applicazione f non è un'isometria ($d(P, O) \neq d(f(P), f(O))$). In effetti le due circonferenze considerate non sono equivalenti dal punto di vista euclideo (Es. 2).

9: Teorema: *Ogni conica non degenera del piano euclideo \mathbb{R}^2 è congruente a una delle seguenti:*

$$X^2/a^2 + Y^2/b^2 = 1 \quad (a \geq b > 0) \quad (\text{ellisse})$$

$$X^2/a^2 - Y^2/b^2 = 1 \quad (a \geq b > 0) \quad (\text{iperbole})$$

$$Y^2 - 2pX = 0 \quad (p > 0) \quad (\text{parabola})$$

$$X^2/a^2 + Y^2/b^2 = -1 \quad (\text{ellisse a punti non reali}).$$

Le coniche precedenti sono due a due non congruenti.

Dim: Si riprende la dimostrazione del teorema 8. La prima tappa si può ripetere grazie al teorema spettrale (§12): esiste una matrice ortogonale, M_0 , tale che ${}^t M_0 A_0 M_0$ sia diagonale. La seconda tappa usa traslazioni che sono isometrie e quindi si può ripetere. A questo punto abbiamo le due forme: $aX^2 + bY^2 - 1 = 0$, $a \neq 0$, $b \neq 0$ (1), $aY^2 + 2bX = 0$ con $ab \neq 0$ (2). Nel caso (1) se $a > 0$ allora possiamo scrivere $a = \alpha^2$, se $a < 0$ allora $-a = \alpha^2$; idem per b . Secondo i segni otteniamo un'iperbole o un'ellisse (eventualmente senza punti reali). Nel caso (2): possiamo assumere $a > 0$ (altrimenti moltiplicare l'equazione per -1). Dividendo per a : $Y^2 + 2bX/a = 0$. Se $b < 0$ la relazione si scrive $Y^2 - 2pX = 0$ con $p = |b|/a$. Se $b > 0$ eseguiamo la trasformazione $X' = -X$, $Y' = Y$ (è un'isometria), e l'equazione diventa $Y^2 - 2pX' = 0$ con $p = b/a > 0$. Finalmente come nel caso affine si vede che le coniche considerate sono due a due non congruenti (Es.3)♦

9.1: Osservazione : Considerando i casi degeneri si hanno anche i seguenti tipi: $X^2/a^2 + Y^2/b^2 = 0$ (ellisse degenera), $X^2/a^2 - Y^2/b^2 = 0$ (iperbole degenera), $Y^2 \pm a^2 = 0$ (parabole degeneri), $Y^2 = 0$ (retta doppia, conica doppiamente degenera).

Esercizi:

13.1) Dimostrare le osservazioni **8.1**, **9.1**, e completare la dimostrazione del teorema 9.

13.2) Dimostrare direttamente che le circonferenze di centro l'origine, di raggio rispettivamente a , l sono congruenti se e solo se $a = l$.

13.3) Sia $C \subseteq \mathbf{R}^2$ una conica a centro di equazione $P(x, y) = 0$. Dimostrare che il centro c ha coordinate $(0, 0)$ se e solo se P non contiene termini di primo grado. Dedurne che se $c = (x_0, y_0)$ la traslazione $X = X' - x_0$, $Y = Y' - y_0$, trasforma C in una conica la cui equazione è priva di termini di primo grado.

13.4) Per ciascuna delle seguenti coniche $C \subseteq \mathbf{R}^2$ determinare se C è degenera, a centro oppure no, e nel caso C sia a centro, determinare le coordinate del centro di C :

$$(i) X^2 + Y^2 + XY + X + Y + 1 = 0$$

$$(ii) 3X^2 - 8XY - 3Y^2 + 8 = 0$$

$$(iii) 2X^2 + 4XY + 5Y^2 - 12 = 0.$$

13.5) Si considerino le coniche del piano euclideo \mathbf{R}^2 le cui equazioni sono quelle di **Esempio 13.4**. Per ognuna di esse determinare un'isometria che la trasformi in forma canonica e la forma canonica ottenuta.

Bibliografia

Berger, M.: "Geometry", Springer, Berlin 1980.

Godement, R.: "Cours d'algèbre", Hermann, Paris 1963.

Hilbert, D.: "Fondamenti della geometria", Feltrinelli, Milano 1970.

Lang, S.: "Algebra lineare", Bollati Boringhieri, Torino 1963.

Sernesi, E.: "Geometria I", Bollati Boringhieri, Torino 1989.

