Corso
di
Combinatorica Algebrica

Michela Montrasio

16 gennaio 2016

Capitolo 1

CENNI INTRODUTTIVI

1.1 Concetti introduttivi

- Ci occuperemo di *error correcting and detecting codes*, ovvero di codici che individuano o correggono errori verificatisi durante la trasmissione di informazioni in presenza di rumori o disturbi.
- Si parte da un insieme di oggetti detto *alfabeto*. Per *messaggio* si intende una sequenza concatenata di simboli presi dall'alfabeto.
- Si ha l'obbiettivo di usare strumenti di tipo matematico che consentono di individuare il messaggio originale trasmesso, detti *decodificatori*.
- Un esempio di decodificatore è quello di *trasmettere due volte il messaggio inviato*. Con l'alfabeto $F_2 = \{0, 1\}$, supponiamo di voler trasmettere 101, e in presenza di rumore alcune posizioni potrebbero essere alterate. Si usa la strategia di trasmettere il messaggio due volte, ovvero 101101, perché si suppone che le posizioni in cui ci sia rumore non siano sempre le stesse. Supponiamo di ottenere il messaggio 101111. Questo decodificatore non consente di ricostruire il messaggio trasmesso, ma permette di capire se c'è stato errore nella trasmissione. Se invece si ripete tre volte il messaggio inviato, supponiamo di inviare 101101101 e di ricevere 101111101. Con buona probabilità, il messaggio originario è 101 perché ha il maggior numero di occorrenze, e quindi quest'ultimo decodificatore consente anche di correggere l'errore.
- Fissato $n \in \mathbb{Z}$ e un alfabeto A, chiamo $A^n = A \times \cdots \times A$ (per n volte) (ad esempio se $A = \{a, b, c\}$, $cabb \in A^4$). $A^* = (A, \cdot)$ si chiama monoide delle parole.
- In informatica si definisce codice C un sottoinsieme del monoide delle parole. In teoria dei codici, un *codice* C su A è un sottoinsieme di A^n con n fissato. n viene detto *lunghezza* del codice. I messaggi che considereremo saranno le parole del codice.
- Il decodificatore non può ignorare la struttura fisica del canale. Al canale è associata una probabilità di commettere errori. Diremo che un canale è *simmetrico* se dato un simbolo x appartenente all'alfabeto, la probabilità di inviare x e ricevere $y \neq x$ è un numero p indipendente da x e da y. In simboli, dato x, $P(y|x) = p$ se $x \neq y$. Si utilizzano alfabeti con un numero finito di elementi. Un canale che soddisfa le ipotesi precedenti quando l'alfabeto è finito si dice *canale n-ario*.
- Dalla nascita della teoria dei codici l'esempio di alfabeto utilizzato è F_2. Noi utilizzeremo come alfabeti F_n con n primo.
- Definendo p come prima, nel caso di codici binari simmetrici, la probabilità di inviare un bit e ricevere lo stesso bit è $1-p = q$. Possono verificarsi le possibilità $q < p, q = p, q > p$. Ovviamente un canale binario simmetrico è buono se $q > p$.
- Esempio: Supponiamo di avere il messaggio 101, e di volerlo trasmettere attraverso un canale simmetrico binario. *Calcoliamo la probabilità di inviare il messaggio e ricevere* 111. La probabilità da calcolare è il prodotto tra le tre probabilità relative ad ognuno dei bit e quindi è q^2p (questo avviene per l'ipotesi di mancanza di memoria sul canale, per la quale gli eventi relativi a bit distinti sono indipendenti). Analogamente, la probabilità di ricevere 101 è q^3. Osservo che $q^3 > q^2p$ se e solo se $q > p$.
- Il cosidetto *algoritmo di decofifica di massima probabilità* funziona nel modo seguente:

ricevuto il messaggio, cerca tra tutte le parole di A^n quella che con massima probabilità è stata originariamente trasmessa. In generale, nell'ipotesi $q > p$, data la parola $c \in A^n$, la parola r che rende massima la probabilità di inviare c e ricevere r è esattamente c.

- Se l'alfabeto non è binario, per ogni x in A, la probabilità q di inviare x e ricevere x è $1-(m-1)p$, con m numero dei simboli dell'alfabeto, infatti

$$q = 1 - \sum_{y \in A, y \neq x} P(y|x) = 1 - (m-1) * p$$

(come prima gli eventi $x \mapsto y$ e $x \mapsto z$ con $y \neq z$ sono indipendenti per l'ipotesi di assenza di memoria sul canale). La condizione $q > p$ è soddisfatta se

$$1 - (m-1)p > p \longrightarrow 1 - mp > 0 \longrightarrow p < 1/m$$

Concludo che un canale è buono se il numero m di elementi nell'alfabeto è tale che $1/m > p$.

1.1.1 Passi della trasmissione

- Quando si invia un messaggio $i \in C$, se ne effettua una codifica che permetta di risalire più facilmente al messaggio inviato in caso di errori di trasmissione. Per *codifica* si intende una funzione f definita sul codice. Pongo $C' = f(C)$.
  Ad esempio, sia $A = F_2$ e $k = 3$ con k lunghezza delle parole, un esempio di algoritmo di codifica è la funzione $f(i) = i^3$ (ripeto tre volte il messaggio), quindi $f : C \subset A^3 \to A^9$ (si ottengono sequenze di 9 bit a tre a tre uguali).
- Trasmetto il messaggio codificato e chiamo c' ciò che ricevo.
- A questo punto si applica l'algoritmo di decodifica a c': la decodifica di c' è $c'' \in C$ dove i' è tale che con probabilità massima inviando c' ottengo c''.
  Ad esempio, se $i = 101$, la sua codifica è $w = 101101101$, e supponiamo di ricevere $c' = 101111101$. Ricevendo c' si capisce che c'è un errore di trasmissione, per decodificare la parola originariamente trasmessa devo cercare la parola di $C' \subseteq A^9$ che differisca il meno possibile da c', e risalire poi alla parola di partenza. In questo caso esiste ed è unica la parola $c'' \in C$ più vicina a c', che è 101101101, e decodificandola si ottiene la parola di partenza.
- Per fare in modo che la parola di massima probabilità esista e sia unica, si fanno le seguenti richieste:

  (i) Le parole di C devono essere sufficientemente numerose in modo che ogni parola $w \in A^n$ sia "abbastanza vicina" a qualche parola $c \in C$

  (ii) le parole del codice devono essere "abbastanza distanti" l'una dall'altra

  Da questo punto di vista l'algoritmo $f(i) = i$ non è efficace.
  Sia $A^2 = \{00, 01, 10, 11\}$ e $C = \{00, 11\}$. Ci sono situazioni di stallo: se ricevo 01 o 10, queste parole hanno la stessa distanza dalle parole del codice, quindi le richieste non sono soddisfatte. Un esempio di codice correttore di al più un errore è

$$C = \{000, 111\}$$

in F_2^3.

Il codice fiscale o il codice ISBN sono esempi di codice che evidenziano almeno un errore (informazioni sul codice fiscale nel file codicefiscale.txt).

8-10-2015

1.2 Esempi di codici

Si considerano messaggi su un alfabeto A. Si definisce una funzione di codifica $f : C \subseteq A^k \to A^n$ con $n \geq k$ (se $n < k$ si perde informazione). Definiamo *codice (a blocchi)* di lunghezza n il sottoinsieme $f(A^k) \subseteq A^n$, che ha la caratteristicha che tutte le parole hanno la stessa lunghezza. Riportiamo alcuni esempi di codici a blocchi in cui analizziamo il rapporto tra k e n.

1.2.1 Codici di ripetizione

Nel caso $k = 1$, fisso n qualsiasi e uso la notazione $A \times \cdots \times A = A^n$. Sia $C = \{a^n \, a \in A\}$: C è detto *codice di ripetizione di lunghezza n* sull'alfabeto A ($rep_n A$). Sia $\mathcal{C}$ un canale, esso è buono se $p < 1/m$, con m numero di elementi dell'alfabeto; indico il canale con ${}_{m,s}\mathcal{C}(p)$.

Proposizione 1.1. *Se il numero e di errori commessi in fase di trasmissione di a^n è $\leq \lfloor (n-1)/2 \rfloor$ allora a è il simbolo che ricorre con maggiore frequenza nella parola ricevuta r.*

Nel caso in cui n sia dispari, ovvero $n = 2t + 1$, il codice di ripetizione di lunghezza n corregge fino a t errori. Se invece n è pari, ovvero $n = 2t$, si ha lo svantaggio che quando gli errori sono esattamente t, non si riesce a risalire al messaggio originario.

Osservazione 1.2. E' importante capire quanto costa calcolare la funzione di codifica dal punto di vista del tempo e della memoria utilizzata. Nel caso di n pari si sfrutta memoria che in realtà non è necessaria, infatti, codici di lunghezza $2t+1$ o $2t+2$ correggono entrambi t errori ma il calcolo della funzione di codifica nel caso di codici di lunghezza pari è più dispendioso.

n va scelto accuratamente, infatti se n è eccessivamente grande le procedure di decodifica sono dispendiose, ma se n è troppo piccolo non si ha la possibilità di correggere molti errori.

Esempio 1.3. Il codice fiscale evidenzia un errore ma non lo corregge.

Codice fiscale = testa (15 caratteri alfanumerici) + coda (1 carattere di controllo)

Quanto frequentemente può accadere che codici fiscali aventi teste differenti hanno lo stesso carattere di controllo?

1.2.2 Codici con controllo di parità

Sull'alfabeto $A = F_2$, fisso un intero arbitrario n e costruisco il seguente sottoinsieme di A^n:

$$C = \{(a_1, \ldots, a_n), \, t.c. \, a_i \in F_2, \, \sum_i a_i = 0 \in F_2\}.$$

Osserviamo che $|A_n| = 2^n$ e $|C| = 1/2 * 2^n = 2^{n-1}$. Per ora non esplicitiamo la funzione di codifica tale che $A^k \mapsto C$. Per fissare condizioni su n, si chiede che $|A|^k \leq |C|$ ovvero $2^k \leq 2^{n-1}$ e quindi $n \geq k + 1$.

Questo codice individua la presenza di un errore, sia data infatti una stringa su F_2 di lunghezza n, della forma $(a_1, \ldots, a_n)$, e supponiamo di inviarla attraverso un canale, sia $(b_1, \ldots, b_n)$ la stringa ricevuta. Supponiamo che esiste al più un indice i tale che $a_i \neq b_i$, allora $b_i = 1 + a_i$, e sommando le componenti della parola ricevuta, segue che

$$\sum_j b_j = \sum_j a_j + 1 = 1$$

e quindi si va contro la condizione $\sum_i a_i = 0$. In realtà ci si rende conto se viene commesso un numero dispari di errori, ma se viene commesso un numero pari di errori non ce ne si può accorgere.

1.2.3 Codice di Hamming

Sia $n = 7$ e $A = F_2$, e considero il codice

$$H = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$$

Le sette righe della matrice sono costituite dallo sviluppo binario degli interi fra 1 e 7. Chiamo $C = \ker H$, ovvero l'insieme dei vettori riga $(a_1, \ldots, a_7)$ tali che $(a_1, \ldots, a_7) * H = (0, 0, 0)$. Il rango della matrice è 3 (scelgo come righe linearmente indipendenti i vettori $(0, 0, 1)$, $(0, 1, 0)$ e $(1, 0, 0)$), quindi il nucleo ha dimensione $7 - 3 = 4$ e ha ordine 2^4.

Esercizio 1.4. Quali sono le possibili scelte di 4 parametri che fissano univocamente i restanti tre?

Dimostrazione. Dato lo spazio vettoriale V su un campo K, chiamo n la dimensione di V su K. Dato un sottospazio vettoriale $W \leq V$, $\mathrm{Dim}W = k \leq \mathrm{Dim}V$ implica che esiste una base di k elementi per W. Un sottospazio W di V può essere considerato come il nucleo di una particolare applicazione lineare o di una matrice H $n \times r$ a coefficienti sul campo. La dimensione del ker è data da

$$k = \mathrm{Dim}W = n - \mathrm{Rg}H$$

con $\mathrm{Rg}H$ numero massimo di colonne linearmente indipendenti estraibili da H. Determino $\ker H$ riducendo a scala la matrice:

```
H=[0 0 1; 0 1 0; 0 1 1; 1 0 0; 1 0 1; 1 1 0; 1 1 1]
```

I parametri non liberi (che vengono determinati dai restanti quattro) si hanno in corrispondenza delle sottomatrici non singolari di ordine massimo (ovvero di ordine 3). Una sottomatrice non singolare di H è costituita dalle righe 1,2,4. Considero la relazione:

$$(v_1, \ldots, v_7) \times H = 0$$

Scambio la terza e la quarta riga di H, ovvero pongo `H=[0 0 1; 0 1 0; 1 0 0; 0 1 1; 1 0 1; 1 1 0; 1 1 1]` in modo da ottenere

$$(v_1, v_2, v_4, v_3, v_5, v_6, v_7) \times H' = 0$$

Chiamo A il blocco 3×3 non singolare di H', ovvero `A=[0 0 1; 0 1 0; 1 0 0]` chiamo B il blocco restante, e pongo

$$\bar{v} = (v_1, v_2, v_4),\ \bar{w} = (v_3, v_5, v_6, v_7)$$

Vale quindi la relazione

$$\bar{v} * A + \bar{w} * B = (0, 0, 0)$$
$$\bar{v}A = -\bar{w}B$$

e poiché A è non singolare posso moltiplicare per il suo inverso:

$$\bar{v} = -\bar{w}BA^{-1}$$

Concludo che fissati i parametri $3, 5, 6, 7$, essi determinano univocamente la prima, la seconda e la quarta componente del vettore. Si può ripetere questo ragionamento a tutte le sottomatrici non singolari, ovvero, determino tutte le terne scelte dall'insieme $\{1, \ldots, 7\}$ che realizzano una sottomatrice 3×3 non singolare. Il complementare di ognuna di queste terne fornisce i parametri liberi. Ci sono a priori 35 terne (7*6*5/6) □

H è detta *matrice di Hamming*, considero una parola c nel codice, e supponiamo di ricevere r, allora se ho commesso al più un errore, sono in grado di correggerlo e riottenere c da r. In caso di un errore, esiste un indice i tale che $r = c + e_i$. Per individuare quale sia l'indice, si moltiplica la parola ricevuta a destra per H; si ha che $r * H = c * H + e_i * H$, ma per ipotesi $c * H = 0$. La posizione in cui è stato commesso l'errore è data dal numero della riga della matrice che si ottiene dal prodotto $e_i * H$.

La capacità correttiva di questo codice è 1.

1.2.4 Codice di Hamming esteso

Su F_2, considero la matrice $\bar{H}$ che si ottiene aggiungendo alla matrice di Hamming una colonna di 1 e la riga 0001 in fondo, ovvero

$$\bar{H} = \begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{matrix}$$

Chiamo $\bar{C} = \ker \bar{H}$ (contiene vettori di 4 componenti), $\bar{C}$ è detto *codice di Hamming esteso*. La lunghezza è $\bar{n} = 8$, mentre la dimensione di $\ker \bar{H}$ è $\bar{k} = 4$.

Per un ragionamento analogo a quello fatto per il codice di Hamming, il codice di Hamming esteso corregge un errore, ma a differenza del codice di Hamming ne individua 2 (vedi esempio 1.11).

1.2.5 Distanza di Hamming

Introduciamo la seguente definizione:

Definizione 1.5 (distanza di Hamming)**.** Sia $A = F_n$ e considero due elementi $x, y \in A^n$. Definisco una funzione $d_H : A^n \times A^n \to \mathbb{N}$ tale che

$$d_H(x, y) = |\{i \in \{1, \dots, n\} \, t.c. \, x_i \neq y_i\}|$$

d_H viene detta *distanza di Hamming* su A^n e rispetto a questa distanza A^n è uno spazio metrico.

d_H **è una distanza** infatti, $\forall x, y, z \in A^n$:

(i) $d(x, y) = 0$ se e solo se $x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ (disuguaglianza triangolare). Per mostrare quest'ultima affermazione, sia $a = d(x, z)$ e $b = d(z, y)$, allora cambiando al più $a + b$ coordinate di x ottengo y, per questo $d(x, y) \leq a + b$.

Date due parole del codice c e d, e dato $* \in A^n$, ci si chiede quanto questo elemento sia vicino alle parole del codice. Se le parole del codice sono abbastanza distanziate e non ci sono troppi errori, le sfere di influenza degli elementi del codice sono disgiunte.

Definizione 1.6. Dato un codice a blocchi C di lunghezza n, la *distanza minima* di C è definita come

$$\min_{c,d \in C,\, c \neq d} d_H(c, d)$$

(l'ipotesi che $c \neq d$ serve per escludere 0 che altrimenti sarebbe sempre il minimo)

Calcoliamo la distanza minima negli esempi precedenti:

(i) nel caso del codice C di ripetizione di lunghezza n su un alfabeto A, vogliamo determinare la distanza minima d_C. Dati due oggetti distinti $c, d \in C$, della forma $c = a^n, d = b^n$ con $a \neq b$, $d_C = n$ (tutti gli elementi sono distinti).

(ii) Nel codice di controllo di parità sull'alfabeto F_2, $d_C \leq n$ e $d_C > 0$. Dati due elementi del codice, la loro distanza non può essere 1 perché altrimenti la somma delle compolo faremo in seguitonenti di uno dei due elementi non sarebbe 0 (osservo che i vettori della base standard non rientrano in questo codice). Si ha $d_C \geq 2$ e vale l'uguaglianza realizzata dalla coppia $c = (1, 1, 0, 0, \dots, 0)$ e d uguale al vettore nullo.

(iii) Si può dimostrare che la distanza minima del codice di Hamming è uguale a 3 (1.18).

12-10-2015

1.2.6 Osservazioni sulla dimensione del codice di ripetizione

Dato un alfabeto A, considero su A^n il codice di ripetizione di lunghezza n. Allora la distanza minima di questo codice è n. Se $A = K$ è un campo, C non è solo un sottoinsieme di K^n, ma è un sottospazio vettoriale di K^n, e scriveremo $C \leq K^n$. Dati due elementi di C

λ^n, μ^n, con $\lambda, \mu \in K$, la loro somma è definita come $(\lambda + \mu)^n = (\lambda + \mu, \ldots, \lambda + \mu)$. Inoltre dato $\alpha \in K$, $\alpha\lambda^n := (\alpha\lambda)^n$.

La dimensione di C è 1. Dal punto di vista dell'algebra lineare, i sottospazi V di K^n che possono considerarsi "equivalenti" a C sono quelli per cui esiste un isomorfismo $\phi : V \to C$. Si può dimostrare che questa definizione induce una relazione di equivalenza. Le classi di equivalenza di un sottospazio W di K^n sono costituite da tutti i sottospazi V di K^n con la stessa dimensione di W, infatti tra due spazi U, W della stessa dimensione si può definire un isomorfismo nel seguente modo: date una base $\{u_1, \ldots, u_k\}$ per U e $\{w_1, \ldots, w_k\}$ per W, allora esiste uno e un solo automorfismo $\phi t.c. \phi(u_i) = w_i$, e le due basi si possono completare a basi di K^n.

Allora, in base a quanto detto, $\text{Rep}_{n,k}$ **è equivalente a** $\text{Span}\{e_1\} = \{(1, 0, \ldots, 0, 0) * \lambda\}$. Considero un automorfismo ϕ di K^n, con la base standard sia in partenza che in arrivo, tale che $\phi(C) = \text{Span}\{e_1\}$, allora $\phi(1, 1, \ldots, 1) = e_1$. Chiamo T la matrice associata a ϕ. Questa matrice ha come righe le coordinate rispetto ai vettori della base in arrivo delle immagini dei vettori della base in partenza. Poiché per $i \geq 2$, $\phi(e_i) = e_i$, tutte le righe della matrice eccetto la prima sono uguali a $e_2, \ldots, e_k$ rispettivamente. Determiniamo la prima riga: sappiamo che $(1, 1, \ldots, 1)T = e_1$, e poiché $(1, 1, \ldots, 1) = \sum_i e_i$ si ha

$$(1, 1, \ldots, 1)T = (e_1 + \sum_{i=2}^{n} e_i)T = e_1T + \sum_{i=2}^{n} e_i = e_1$$

quindi $e_1T = e_1 - \sum_{i=2}^{n} e_i$ e la prima riga di T è data da $(1, -1, \ldots, -1)$.

Osservazione 1.7. Nonostante $\text{Span}\{e_1\}$ e $C = \text{Rep}_{n,k}$ siano equivalenti come spazi vettoriali, non lo sono come codici. Infatti, la distanza minima del codice di ripetizione è n, mentre la distanza minima sullo spazio isomorfo $\text{Span}\{e_1\}$ è 1. In particolare, T corregge 0 errori, mentre il codice di ripetizione ne corregge molti di più.

1.3 Codici e-perfetti e limiti sulla cardinalità

1.3.1 Strumenti geometrici per descrivere la bontà di un codice

Definizione 1.8. Dato uno spazio metrico, si chiama *geodetica* che congiunge due punti A e B una "linea di minima lunghezza" tra A e B.

Sia $A = F_3$ e $n = 2$, allora $A \times A$ con la distanza di Hamming è uno spazio metrico. Dato $P_0 = (0, 0)$ e $P_1 = (1, 1)$, la linea di minima lunghezza che li congiunge è formata da due segmenti, il primo conciunge P_0 a $(0, 1)$ e l'altra unisce $(0, 1)$ a P_1, in particolare tale linea ha lunghezza 2 (l'unica retta passante per i due punti, che sarebbe la curva di minima lunghezza in $\mathbb{R}^2$, non rientra in $A \times A$).

Ora forniremo una descrizione della bontà di un codice dal punto di vista geometrico.

Introduco un algoritmo di decodifica, data una parola $r \in A^n$ fissato $C \subseteq A^n$, decodifico r con c, dove c, se esiste, è l'unica parola del codice che è più vicina ad r rispetto alla distanza di Hamming. Se la parola di minima distanza non esiste unica, non si ha la possibilità di decodificare la parola ricevuta.

Definizione 1.9. Dato un elemento $w \in A^n$ e un numero reale $\rho > 0$, chiamo *bolla* di raggio ρ centrata in w e indicata con $B_\rho(w)$ l'insieme

$$\{u \in A^n \, t.c. \, d(u, w) \leq \rho\}.$$

Un codice è buono se le bolle relative alle varie parole sono opportunamente distanziate.

Il principio di decodicica precedente viene chiamato *principio di collassamento delle bolle* (o sphere shrinking method, indicato con s.s.e.), ovvero si identifica r con d se e solo se d è l'unico elemento in C con $r \in B_e(d)$ con $e = \min_{t \in C} d(r, t)$. Se questo avviene, si restringe la bolla facendola collassare sul suo centro, ovvero identificando r con c.

Proposizione 1.10. *Le seguenti affermazioni sono equivalenti:*

(i) *nell'algoritmo s.s.e. decodifico $r \in A^n$ con l'unica parola del codice che dista al più e da r;*

(ii) *per ogni coppia di elementi $x, y \in C$ le bolle di raggio e centrate nei due elementi risultano essere disgiunte.*

(iii) *$d \geq 2e + 1$, con d distanza minima.*

Commento: detta d la distanza minima del codice, la sua capacità correttiva e_C (ovvero il massimo numero di errori che può correggere) si ottiene cercando il massimo e tale che $2e + 1 \leq d$, ovvero $e_C = \lfloor (d-1)/2 \rfloor$.

Dimostrazione. 1 $\iff$ 2: l'algoritmo di decodifica funziona se esiste unica la parola c nel codice la cui distanza di Hamming dalla parola ricevuta risulta essere $\leq e$, e quindi le bolle devono essere disgiunte, altrimenti un elemento potrebbe avere distanza minima $\leq e$ da più di una parola del codice.

2 $\longrightarrow$ 3: supponiamo che le bolle siano tutte disgiunte, e supponiamo per assurdo che esistano $x, y \in C$ con $d = d(x, y) \leq 2e$. Supponiamo senza perdita di generalità che x e y differiscano nelle prime d componenti. Chiamo z l'elemento che si ottiene a partire da x modificandone le prime e componenti (non sta nel codice), allora modificando le successive $d - e$ componenti di z si può ottenere y.

Sia $x = (x_e, x_{d-e}, x_{n-d})$ (dove x_i indica una porzione di vettore di lunghezza i) e $y = (y_e, y_{d-e}, x_{n-d})$. Per costruzione $z = (y_e, x_{d-e}, x_{n-d})$. Allora z sta nell'intersezione tra le bolle $B_e(x)$ e $B_e(y)$ (infatti per l'ipotesi $B_{d-e}(y) \subseteq B_e(y)$), contro l'ipotesi.

3 $\longrightarrow$ 2: viceversa, supponiamo che valga 3 ma non valga 2, allora esistono x, y distinti nel codice con z contenuto nell'intersezione tra le due bolle. Questo implica $d(x, y) \leq d(x, z) + d(z, y) \leq 2e$, ovvero $d_C \leq 2e$, contro l'ipotesi. □

Se la distanza tra due elementi del codice è dispari, le bolle centrate nei due elementi o si intersecano o sono disgiunte. Se invece la distanza è pari potrebbe verificarsi la situazione in cui due bolle di raggio e sono tangenti in un punto.

Nei codici con distanza minima d pari sono in grado di evidenziare $d/2 - 1$ errori, e ne correggo $\lfloor (d-1)/2 \rfloor$.

Esempio 1.11. Ricordando il codice di Hamming esteso, considero la sottomatrice 4×4 data dai vettori $(0, 0, 1, 1), (0, 1, 0, 1), (1, 0, 0, 1), (1, 0, 1, 1)$, che sono indipendenti. Segue che $\mathrm{Rg}\bar{H} = 4$, allora $\mathrm{Dim}\bar{H} = 8 - 4 = 4$.

Supponiamo che nella trasmissione di c siano stati commessi due errori, e che si riceva $\hat{r} = c+e_i+e_j$. Considero la sindrome $\bar{r}$ di questo vettore, ovvero $\hat{r}*\bar{H}$ (ha quattro componenti). Poiché $c\bar{H} = 0$ si ha

$$(c + e_i + e_j)\bar{H} = (e_i + e_j)\hat{H} = \bar{H}_i + \bar{H}_j$$

In particolare, poiché ogni riga di $\bar{H}$ termina con 1, $\bar{H}_i + \bar{H}_j$ ha l'ultima componente uguale a 0. Concludo che se l'ultima componente di $\hat{r}\bar{H}$ è nulla, allora è stato commesso un numero pari di errori. In particolare se $\hat{r} * \bar{H} = 0$ posso aver commesso 0 o 4 errori invece se $v \neq 0$ ho commesso almeno due errori. Quindi **il codice di Hamming esteso è in grado di stabilire se sono stati commessi almeno due errori**.

Tuttavia **tale codice non corregge due errori**, infatti esistono coppie $(k, l) \neq (i, j)$ tale che $\bar{H}_k + \bar{H}_l = \bar{H}_i + \bar{H}_j$ e quindi esistono due sindromi uguali. Basta considerare ad esempio:

$$\bar{H}_k = (0, 0, 1, 1), \bar{H}_l = (0, 1, 0, 1), \bar{H}_i = (1, 0, 1, 1), \bar{H}_j = (1, 1, 0, 1)$$

Questo è un esempio di codice con distanza minima 4.

1.3.2 Cardinalità delle bolle

Contiamo i punti all'interno di una bolla di un dato raggio e. Considero un alfabeto con m elementi, e sia w una parola di lunghezza n, allora la bolla $B_e(w)$ ha

$$|B_e(c)| = \sum_{i=0}^{e} \binom{n}{i} (m-1)^i \tag{1.1}$$

elementi.

Significato della formula: considero una parola qualsiasi $u \in B_e(w)$, e sia $h = d(u, w)$, considero $w = (\omega_1, \ldots, \omega_n)$. Devo scegliere i posizioni (e questo può essere fatto in $\binom{n}{i}$ modi) in cui le due parole differiscono, inoltre la scelta del simbolo può essere fatta in $m - 1$ modi per ognuna delle posizioni in cui si ha l'errore.

In particolare, chiamo $g_i(w) = \{v \in A^n \, t.c. d(v, w) = i\}$ (guscio), allora

$$B_e(w) = \bigcup_{i \leq e} g_i(w)$$

1.3.3 Sphere packing bound

Notazione: indico con $A_m(n, d)$ la cardinalità massima di un codice m-ario di parametri n e d)

Teorema 1.12 (sphere packing bound)**.** *$A_m(n, d)$ soddisfa il cosidetto* sphere packing bound*:*

$$A_m(n, d) \leq \frac{m^n}{\sum_{i \in [0, \lfloor (d-1)/2 \rfloor]} \binom{n}{i} (m-1)^i}.$$

Dimostrazione. Sia C come sopra, allora la sua capacità correttiva di errore è $e_C = \lfloor (d - 1)/2 \rfloor$, e le bolle di raggio e_C centrate in elementi del codice risultano essere disgiunte: allora

il prodotto tra il numero di bolle e la cardinalità di ogni bolla risulta essere minore della cardinalità di A^n, quindi, sfruttando la formula (1.1)

$$|C| * \sum_{i=1}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i} (m-1)^i \leq |A^n| = m^n$$

e dividendo per la somma si ottiene l'asserto. □

1.3.4 Codici e-perfetti

Si vuole stabilire l'esistenza o meno di codici che realizzino tale limite. Introduciamo la seguente definizione:

Definizione 1.13. Un codice si dice *e-perfetto* se soddisfa le seguenti condizioni:

(i) per ogni coppia (x, y), $B_e(x) \cap B_e(y) = \emptyset$ (ovvero C corregge fino a e errori)

(ii) ogni elemento di A^n sta in $B_e(x)$ per qualche $x \in C$.

Le due condizioni implicano che $A^n = \bigcup_{x \in C} B_e(x)$ dove l'unione è disgiunta. Questo significa che per un codice e-perfetto

$$|A^n| = m^n = |B_e(x)| * |C|$$

ovvero

$$m^n = \sum_{i=1}^{e} \binom{n}{i} (m-1)^i |C|.$$

Definizione 1.14. Dato un codice $C \subseteq A^n$, si chiama *raggio di ricoprimento* (covering radius)

$$cr_C = \min\{r \, t.c. \, A^n = \bigcup_{x \in C} B_r(x)\}.$$

Nel caso di un codice e-perfetto, la definizione è ben posta, perché l'insieme di r che soddisfano la condizione non è vuoto poiché contiene almeno e.

Osservazione 1.15. Una nozione equivalente di codici e_C-perfetti è data dal richiedere che la sua capacità correttiva e_C risulti essere uguale a cr_C (in questo modo segue che le bolle sono disgiunte e coprono tutto F^n).

Teorema 1.16. *Un codice e_C-perfetto deve avere distanza minima dispari, ovvero $d = 2l + 1$.*

Dimostrazione. Dato un codice con distanza minima pari, della forma $d = 2l$, prendo due elementi x, y tali che $d(x, y) = 2l$ con $x, y \in C$. Allora preso z tale che $d(z, x) = d(z, y) = l$, si ha che z sta nell'intersezione delle due sfere di raggio l centrare nei due punti, quindi $e_C = l - 1$. In questo modo però $z \notin \bigcup_{t \in C} B_{l-1}(t)$ ovvero $rc(C) \neq e_C$, contro la definizione. □

1.3.5 Esempi di codici perfetti

Esempio 1.17. Considero il codice di ripetizione, affinché il codice abbia più di un elemento richiedo che $|A| \geq 2$. Affinché la distanza minima sia dispari, prendo $n = 2e + 1$, e quindi

$e_C = \lfloor (d-1)/2 \rfloor = e$. Imponendo che il codice sia e-perfetto, lo sphere packing bound implica che

$$m^n = \sum_{i=0}^{e} \binom{n}{i} (m-1)^i |C|$$

da cui, poiché $|C| = m$:

$$m^{n-1} = \sum_{i=0}^{e} \binom{n}{i} (m-1)^i,$$

e scegliendo $m = 2$:

$$2^{n-1} = \sum_{i=0}^{e} \binom{n}{i} 1^i = \sum_{i=0}^{e} \binom{n}{i}$$

Allora poiché il codice ha n elementi, posso porre

$$s = \sum_{i=0}^{e} \binom{n}{i} = \sum_{i=0}^{e} \binom{n}{n-i}$$

e ponendo $k = n - i$:

$$s = \sum_{k=n-e}^{n} \binom{n}{k} = \sum_{i=e+1}^{n} \binom{n}{i}$$

perché $n - e = e + 1$. Allora

$$2s = \sum_{i=0}^{n} \binom{n}{i} = 2^n$$

ovvero $s = 2^{n-1}$. Nel caso di $n = 3$ (e quindi $e = 1$) e $A = F_2$ la disuguaglianza è verificata.

Problema: stabilire se esistono valori di $m > 2$ tali che sia verificata la condizione

$$m^{n-1} = \sum_{n=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{i} (m-1)^i$$

ovvero tale che il limite sia raggiunto.

Esempio 1.18. Considero il codice di Hamming di lunghezza 7 su F_2. Allora $m = 2$, $n = 7$, e vogliamo stabilire la distanza minima del codice. Poiché questo codice corregge un errore, o $d = 3$ o $d = 4$ (se $d = 5$ il codice correggerebbe $\lfloor (d-1)/2 \rfloor = 2$ errori e questo non avviene). Se considero il vettore $(1, 1, 1, 0, 0, 0, 0)$ esso sta in $\ker H$ e ha distanza uguale a 3 dal vettore nullo, quindi $d = 3$.

Lo sphere packing bound è dato da

$$2^7 = \sum_{i=0}^{1} \binom{7}{i} * 2^4$$

$$\sum_{i=0}^{1} \binom{7}{i} = 1 + 7 =^? 2^3$$

e questo è vero, quindi si raggiunge il limite di Hamming, e **il codice di Hamming di lunghezza 7 su F_2 è 1-perfetto**.

Vedremo che esiste una famiglia infinita di codici dipendenti da due parametri q, r, che sono tutti 1-perfetti.

Esistono altri codici 1-perfetti che non sono di Hamming ma non esiste una loro classificazione. I *codici di Golay* sono codici definiti su F_3 o su F_2 e correggono rispettivamente 2 o 3 errori. Questi sono gli unici codici e-perfetti per $e_C > 1$.

1.3.6 Esistenza di codici e-perfetti

Il problema dell'esistenza di codici e-perfetti si riconduce al seguente problema: **per quali valori** m, n, e **vale che la cardinalità della bolla** $B_e(x)$, $x \in A^n$ **(con** $|A| = m$**) divide** m^n? Per $m = 2, 3, 4, 5$, si riescono sempre a trovare m, i che soddisfano le condizioni mentre per $m = 6$ questo non avviene.

Osservazione 1.19 (curiosità)**.** Problema dei 36 ufficiali: ho 36 ufficiali provenienti da 6 reggimenti diversi con 6 gradi diversi. Li vogliamo disporre su una matrice quadrata 6x6 in modo da avere su qualunque riga e su qualunque colonna tutti i reggimenti e tutti i gradi. Si dimostra che per tale problema non ci sono soluzioni.

La teoria dei codici è legata alla teoria dei disegni che ha applicazioni in statistica e in esperimenti.

Se $|A| = p$, la traduzione aritmetica del problema dell'esistenza di codici perfetti è la seguente:

$$|B_e(x)| = \sum_{i=0}^{e} \binom{n}{i} (p-1)^i \mid p^m$$

in particolare la quantità a sinistra dev'essere una potenza di p della forma p^t, con $t < m$, ovvero

$$1 + \sum_{i=1}^{e} \binom{n}{i} (p-1)^i = p^t$$

$$n * (p-1) + \binom{n}{2} * (p-1)^2 + \cdots = p^t - 1$$

e dividendo per $p - 1$:

$$n + P = \frac{p^t - 1}{p-1} = p^{t-1} + \cdots + p + 1$$

con P somma di multipli di $p - 1$, ovvero

$$P = \sum_{i=2}^{e} \binom{n}{i} (p-1)^{i-2}.$$

allora $n \equiv t \mod p - 1$.????

1.3.7 Limite inferiore

Si vuole stabilire se la funzione $A_m(n, d)$ ammetta anche un limite inferiore.

Teorema 1.20 (Gilbert-Varshamor bound)**.** *Dato un codice m-ario di lunghezza n e distanza minima d,*

$$A_m(d, n) \geq \frac{m^n}{\sum_{i=0}^{d-1} \binom{n}{i} (m-1)^i}.$$

esempio di algoritmo ingordo. Costruisco un codice costituito da un unico elemento all'interno di A^n. Considero la bolla $B_{d-1}(c_1)$; se esiste $c_2 \notin B_{d-1}(c_2)$, costruisco un nuovo codice aggiungendo c_2 al codice di partenza. Se l'unione $B_{d-1}(c_2) \cup B_{d-1}(c_1)$ non esaurisce tutto lo spazio ambiente, considero c_3 al loro esterno e lo aggiungo al codice, ripeto il ragionamento fino ad arrivare al passo $i-1$. Costruisco in questo modo un codice con distanza minima $\geq d$. Posso procedere finché è verificata la condizione $\bigcup_k B_{d-1}(c_k) \neq A^n$. Questa condizione in particolare è verificata nell'ipotesi più forte $\sum_{c\in C} |B_{d-1}(c)| < m^n$ (l'ipotesi è più forte perché non è detto che le bolle siano disgiunte). Sicuramente la disuguaglianza è verificata se $|B_{d-1}(c)| * (i-1) < m^n$, ovvero

$$(i-1)\sum_{i=0}^{d-1} \binom{n}{i}(m-1)^i < m^n$$

quindi, affinché non si possano più aggiungere elementi al codice dev'essere

$$|C| \geq \frac{m^n}{\sum_{i=0}^{d-1} \binom{n}{i}(m-1)^i}$$

Notiamo che per costruzione la distanza minima di tale codice è $\geq d$. □

Ci si chiede quale sia la distanza tra il limite di Hamming e il limite di Gilbert-Varshamor.

Esempio 1.21. Nel caso $n = 2, m = 90$ ed $e = 2$, la sfera di raggio 2 ha cardinalità esattamente una potenza di 2. (tuttavia non esiste un codice e-perfetto con questi parametri)

$$\frac{2^{90}}{B_2(c)}$$

è un limite superiore.

$$e = 2 \longrightarrow d \geq 5$$

Allora $|B_4(x)| = 2676767$, allora il limite inferiore è $4,62 * 2^{20}$.

Introdurremo concetti geometrici per studiare in maniera efficiente dei codici.

1.3.8 Teoria dell'informazione

L'entropia, importante in termodinamica, misura il disordine di un sistema fisico. Il problema di misurare la complessità di un insieme di dati si riconduce allo studio dell'entropia.

Definizione 1.22. Dato un codice C m-ario (definito su un alfabeto con m elementi), finito, definisco *dimensione* del codice, la quantità $k_C := \log_m |C|$. (a volte poiché k_C non è intero si considera come dimensione il minimo $k \in \mathbb{Z}$ tale che $k_C \leq k$.

In generale, di un codice m-ario si dichiarano la lunghezza n_C, la dimensione k_C (eventualmente approssimata) e la distanza minima d_C. Si parla di n, M, k-codici con m numero di elementi nel codice ($M = m^k$).

Definizione 1.23. Definisco *tasso di informazione* (information rate) o dimensione normalizzata $\chi = k/n$ e chiamo $\delta = d/n$ la *distanza normalizzata.*

Questi due valori sono sempre compresi tra 0 e 1, in particolare $0 \leq \delta \leq \chi \leq 1$.

Con argomenti di tipo probabilistico e stime di tipo analitico mostreremo che esiste una funzione che suddivide i codici "buoni" da quelli "cattivi", anche se la divisione non è netta.

19-10-2015

1.3.9 Considerazioni probabilistiche

(i) Dato un codice C e un elemento $x \in C$, considero l'algoritmo di decodifica d e chiamo $p_{x,\mathrm{d}}$ la **probabilità di commettere errore inviando x decodificandolo con l'algoritmo** d. Si vuole limitare $p_{x,\mathrm{d}}$.

(ii) Se C è finito, definisco la media

$$p_{C,\mathrm{d}} := 1/|C| \sum_{x \in C} p_{x,\mathrm{d}}$$

$p_{C,\mathrm{d}}$ rappresenta la **probabilità media di commettere errori in fase di decodifica inviando un elemento del codice**.

(iii) Pongo poi $p_C = \min_{\mathrm{d}} p_{\mathrm{d},C}$, dove il minimo viene fatto sull'insieme dei possibili algoritmi di decodifica.

Esempio 1.24. Consideriamo l'algoritmo di decodifica s.s.e., per codici correttori di e errori (e quindi con distanza minima che $e = \lfloor (d-1)/2 \rfloor$) $p_{\mathcal{C},s.s.e.}$ dipende dalla struttura del canale di trasmissione. Considereremo canali simmetrici, e indico con p la probabilità di ricevere $b \neq a$ inviando a nell'alfabeto (un altro esempio di canali sono i canali gaussiani, dove la distribuzione dell'errore è data da una gaussiana, ma non li considereremo). Allora

(i) la **probabilità di commettere errore nell'invio di un elemento a nell'alfabeto** è data da $s = (m-1)p$ (con m numero di elementi dell'alfabeto).

(ii) Indico con $q = 1 - s = 1 - (m-1)p$ la probabilità dell'evento complementare, ovvero la **probabilità di non commettere errore nell.invio di $a \in A$.**

(iii) La **probabilità di commettere t errori nell'invio di $x \in A^n$** è

$$\binom{n}{i} s^i q^{n-i}$$

(per la scelta delle posizioni in cui avviene l'errore ho $\binom{n}{i}$ possibilità, s^i è la probabilità che i componenti siano corrette e q^{n-i} è la probabilità che le rastanti siano sbagliate).

(iv) La **probabilità di commettere al più e errori in fase di trasmissioni usando un canale simetrico con probabilità p** è

$$\sum_{i=0}^{e} \binom{n}{i} ((1-m)p)^i q^{n-i}$$

Siccome il canale è simmetrico, se $\mathrm{d} = s.s.e.$, $\mathcal{P}_{x,\mathrm{d}} = p_{C,\mathrm{d}}$ (il valore $p_{x,\mathrm{d}}$ è uguale per ogni x).

(v) La **probabilità di commettere più di e errori** è data da:

$$1 - \sum_{i=0}^{e} \binom{n}{i} ((1-m)p)^i q^{n-i}$$

e, usando l'algoritmo s.s.e., coincide con $P_C(\mathrm{d}) = 1/|C| \sum_{x \in \mathrm{d}} p_C(x)$, perché con tale algoritmo una parola viene decodificata in modo errato solo se sono stati commessi più di e errori.

1.3.10 Famiglia di Shannon

Il problema è *determinare una famiglia (infinita) F di codici C tali che fissato $\varepsilon > 0$ esista sempre un codice in F con $p_C \leq \varepsilon$.*

Questo conduce alla seguente definizione:

Definizione 1.25. Una famiglia F di codici si dice *famiglia di Shannon* se soddisfa la richiesta precedente ($p_C \leq \varepsilon$). Inoltre tale famiglia deve comprendere codici con $\chi = k/n$ non tendente a 0.

Esempio 1.26. Su $A = F_2$, considero la famiglia dei codici di ripetizione al variare di $n \in \mathbb{N}$. Questa famiglia comprende codici con capacità correttiva tendente a infinito. Per qualunque ε fissato si trova un codice per cui la probabilità di commettere errori sia sufficientemente piccola.

Teorema 1.27 (teorema di Shannon)**.** *Chiamo F_χ l'insieme dei codici C tali che $\chi_C \geq \chi$ con $\chi \in [0,1]$ fissato. Allora esiste una funzione $C_m(p)$ tale che, se $\chi \leq C_m(p)$, F_χ è di Shannon, altrimenti non lo è. (in particolare $C_m(p)$ è definita per $p \in [0, 1/m]$ perché dev'essere soddisfatta la condizione di bontà del canale $p < 1/m$).*

In altre parole, in base a questo teorema la condizione $\chi \leq C_m(p)$ limita la famiglia dei codici buoni.

Definizione 1.28. Per $n = 2$, chiamo *capacità del canale* la quantità:

$$1 + p * \log_2 p + q \log_2 q.$$

La funzione $H_2(p) = -p \log_2 p - q \log_2 q$ è anche detta *funzione di entropia* associata al canale (in questo caso al canale binario simmetrico di probabilità p). Tra la capacità c_C e l'entropia vale la relazione $c_C = 1 - H_m$

Definizione 1.29 (definizione generale di entropia)**.** In meccanica statistica, dato un sistema fisico con un numero finito di stati $s_1, \ldots, s_n$, con probabilità di realizzarsi $p_1, \ldots, p_n$, allora l'entropia del sistema è definita come

$$\S := -\sum_i p_i \log_2 p_i, \text{ definizione di Boltzmann.}$$

Quando $|A| = m$, chiamo

$$H_m(x) = -x \log_m(\frac{x}{m-1}) - (1-x) \log_m(1-x),$$

e per $m = 2$ si ottiene la definizione

$$H_2(x) = -x \log_2 x - (1-x) \log_2(1-x).$$

Estendendo per continuità questa funzione, $H_m(0) = 0$, infatti

$$\lim x \to 0 - x \log_m \frac{x}{m-1} - (1-x)\log_m(1-x) = 0.$$

Inoltre $H_m(\frac{m-1}{m}) = 1$, infatti

$$\begin{aligned} H_m(\frac{m-1}{m}) &= -\frac{m-1}{m}\log_m \frac{m-1}{m*(m-1)} - (1-\frac{m-1}{m})\log_m(1-\frac{m-1}{m}) \\ &= \frac{m-1}{m}\log_m m - \frac{1}{m}\log_m \frac{1}{m} \\ &= \frac{m-1}{m} + \frac{1}{m} = 1. \end{aligned}$$

Per proseguire è necessario il seguente risultato di analisi:

Teorema 1.30. *Sia $|A| = m$, e $\sigma \in [0, (m-1)/m)$. Allora*

$$\frac{1}{n}\log_m |B_{\sigma n}(x)| \to H_m(\sigma). \tag{1.2}$$

cenni di dimostrazione. Sappiamo che

$$|B_{\sigma_n}(x)| = \sum_{i=0}^{\lfloor \sigma n \rfloor} \binom{n}{i} (m-1)^i.$$

Si approssima questa cardinalità dando una stima del coefficiente binomiale e approssimando i fattoriali con la formula di Stirling.

$$n! \sim (n/e)^n \sqrt{2\pi n}$$

e si ottiene che

$$\sum_{i=0}^{\lfloor \sigma n \rfloor} \binom{n}{i} (m-1)^i \sim \binom{n}{n\sigma} (m-1)^{n\sigma}.$$

La dimostrazione coinvolge i numeri di Bernoulli che permettono di dare una stima della serie

$$\sum_{n=1}^{m} n^k = 1^k +^k + \cdots + m^k$$

□

dimostrazione del teorema di Shannon. Dato un codice correttore di e errori, determiniamo **quanti errori commetto mediamente trasmettendo una parola di lunghezza n con un canale simmetrico**. Poiché il canale è privo di memoria, ogni componente della parola ha la stessa probabilità s di venire modificata, quindi alla fine del processo di trasmissione la probabilità media che una componente sia stata modificata è sn. Considero codici la cui capacità correttiva è $e \sim sn$. Applico a questi codici il limite di Hamming ovvero

$$|C| * |B_{sn}(x)| \leq m^n$$

e applicando i logaritmi

$$\log_m(|C| * |B_{sn}(x)|) \leq n$$

$$n^{-1}\log_m(|C| * |B_{sn}(x)|) \leq 1$$
$$n^{-1}(\log_m |C| + \log_m |B_{sn}(x)|) \leq 1$$

e per $n \to \infty$ e per il risultato (1.2):

$$n^{-1}\log_m |B_{sn}(x)| \to H_m(s)$$

si ha

$$n^{-1}\log_m |C| + H_m(s) \leq 1$$

e ricordando che $\chi = 1/n \log_m |C|$ si ha

$$\chi \leq 1 - H_m(s) = 1 - H_m((m-1)p)$$

e pongo $C_m(p) = 1 - H_m((m-1)p)$. □

Il teorema di Shannon assicura l'esistenza di codici buoni, e ci si pone il problema di come costruirli. La dimostrazione del teorema di Shannon però non è costruttiva.

Il teorema vale anche per codici C definiti su un alfabeto A che sia un campo, questi codici vengono detti *lineari*.

1.3.11 Regione di codici m-ari

Ricordiamo che dato un codice, si associano ad esso le quantità $\delta_C = d/n$ (distanza normalizzata) e $\chi_C = k/n$ (tasso di informazione). $(\delta, \chi) \in [0,1] \times [0,1]$.

Non tutti i codici con parametri nel quadrato $[0,1] \times [0,1]$ sono utili dal punto di vista pratico (se la lunghezza non è sufficientemente elevata il codice non può correggere molti errori). Per questo motivo si dà la seguente definizione.

Definizione 1.31. Fissato un alfabeto con m elementi, si chiama *regione dei codici m-ari* l'insieme

$$\{(\chi, \delta)\, t.c.\, \exists C_n \in A^n\, t.c.\, \lim_{n\to\infty} \chi_{C_n} = \chi,\ \lim_{n\to\infty} \delta_{C_n} = \delta\}.$$

ovvero è l'insieme dei codici di parametri (χ, δ) tali per cui esistono codici di lunghezza crescente con parametri convergenti a (χ, δ).

A parte un numero finito di codici, la funzione $C_m(p)$ definita prima limita superiormente la famiglia di codici che si vogliono caratterizzare.

Vale inoltre il seguente

Teorema 1.32 (teorema di Manin)**.** *Esiste una funzione $\alpha_m(\delta)$, non crescente e continua, tale che*

$$\alpha_m(\delta) = \limsup_{n\to\infty} \frac{\log_m A_m(n, \delta_n)}{n}$$

in modo che la regione dei codici m-ari si parametrizza come:

$$\{(\delta, \chi)\, t.c.\, \delta \leq 1,\ 0 \leq \chi \leq \alpha_m(\delta)\}$$

(ricordo infatti che χ è il logaritmo della cardinalità del codice diviso per n). Quindi per $n \to \infty$, $\chi \to \alpha_m(\delta)$.

Esempio 1.33. Considero il segmento verticale di lunghezza 1. Se $C_n = \mathrm{Rep}_n A$,

$$\lim_{n\to\infty} \delta_n = \lim_{n\to\infty} n/n = 1.$$

$$\lim_{n\to\infty} \chi_n = \lim_{n\to\infty} 1/n = 0.$$

allora $(1,0)$ è un punto che sta nella regione R, ovvero $\alpha_m(1) = 0$.

Anche $(0,1)$ è di accumulazione. Infatti, se A è un gruppo abeliano, posso definire $C_n = \{a \in A_n \, t.c. \, \sum_i a_i = 0\}$. Allora la dimensione di questo codice è $(n-1)$ (gli elementi sono determinati da $(1,1,\ldots,1,-\sum_i a_i)$). Allora $\chi = (n-1)/n = 1 - 1/n \to 1$. (un altro modo per determinare k è usare direttamente la formula

$$k = \log_m |C| = \log_m m^{n-1} = n-1.$$

) Invece $\delta_n = 2/n \to 0$. Anche $(0,1)$ sta nella regione m-aria, ovvero $\alpha_m(0) = 1$.

Esistono stime per $\alpha_m(\delta)$?

Teorema 1.34 (teorema di Singleton)**.** *Sia C un codice di lunghezza n su un alfabeto A con distanza minima d, allora la sua dimensione k è $\leq n - d + 1$.*

Dimostrazione. Siano $x, y \in C$ due elementi aventi le prime $n - d + 1$ componenti uguali, ovvero $x = (x_1, \ldots, x_{n-d+1}, x_{n-d+2}, \ldots, x_n)$ e $y = (x_1, \ldots, x_{n-d+1}, y_{n-d+2}, \ldots, y_n)$, la coda diversa nei due elementi ha lunghezza $n - (n - d + 1) = d - 1$, e quindi $d(x, y) \leq d - 1$, ma poiché la distanza minima del codice è uguale a d, segue che $x = y$, ovvero due elementi con la stessa testa sono uguali. Allora dette h_x e t_x la testa e la coda di x, segue che la funzione $x \mapsto h_x$, è iniettiva, allora $|C| \leq |A|^{n-d+1}$, allora

$$k = \log_m |C| = \log_m A_m(n,d) \leq n - d + 1$$

□

In altre parole, per un generico codice, $k + d \leq n + 1$, e la dimensione e la distanza minima sono dinamicamente contrastanti.

Un altro modo per enunciare il teorema di Singleton è il seguente:

$$A_m(n,d) \leq m^{n-d+1}.$$

La limitazione su $\alpha_m(\delta)$ si traduce nel modo seguente. Dato un codice (n,k,d), se pongo $\delta = d/n$, $d = \delta * n$, allora

$$\alpha_m(\delta) = \limsup_{n\to\infty} n^{-1} \log_m A_m(n, \delta n)$$

e per il teorema di Singleton:

$$\leq \lim_{n\to\infty} (n - d + 1)/n = 1 - \delta.$$

La versione asintotica del limite di Singleton afferma che $\alpha_m(\delta) \leq 1 - \delta$.

In generale, ogni limite inferiore o superiore su $A_m(n, \delta n)$ si può tradurre in un limite su $\alpha_m(\delta)$.

Inoltre:

(i) Il limite di Hamming afferma che

$$A_m(n, \delta n) \leq \frac{m^n}{\sum_{i=0}^{\delta n/2} \binom{n}{i}(m-1)^i}$$

si dimostra che una versione asintotica di tale limite è:

$$\alpha_m(\delta) \leq 1 - H_m(\delta/2).$$

Dimostrazione.

$$\begin{aligned}
\alpha_m(\delta) &= n^{-1} \log_m A(m, \delta n) \\
&\leq n^{-1} \log_m \frac{m^n}{\sum_{i=0}^{\delta n/2} \binom{n}{i}(m-1)^i} \\
&\leq n^{-1}(\log_m m^n - \log_m \sum_{i=0}^{\delta n/2} \binom{n}{i}(m-1)^i) \\
&\leq 1 - n^{-1} \log_m |B_{n\delta/2}(x)| \leq 1 - H_m(\delta/2)
\end{aligned}$$

dove nell'ultimo passaggio si applica la formula (1.2). □

(ii) Il limite inferiore, ovvero il limite di Gilbert-Varshamor, afferma che

$$\frac{m^n}{\sum_{i=0}^{\delta n-1} \binom{n}{i}(m-1)^i} \leq A_m(n, \delta n)$$

e la versione asintotica di questa stima è:

$$\alpha_m(\delta) \geq 1 - H_m(\delta)$$

Negli anni ottanta si pensava che $1 - H_m(\delta) = \alpha_m(\delta)$, perché questa stima era vera per tutti i codici costruiti fino a quel momento. Si è poi dimostrato che questa stima è falsa. Si costruiscono codici usando famiglie di divisori su una varietà algebrica. Definita una funzione meromorfa sulla compattificazione del campo dei complessi (sfera di Riemann), il numero di zeri e di poli contati con le relative molteplicità sono uguali. Se $p_i = z_i$, in un suo intorno la funzione si scrive come $f(z) = f_i * (z - z_i)^{n_i}$. Una varietà algebrica può essere considerata come un insieme di punti in uno spazio che soddisfano una certa condizione algebrica. Un divisore è l'insieme dei punti che annullano un funzionale meromorfo su S^1. Usando questo linguaggio si riuscì a costruire codici sofisticati che non soddisfano l'uguaglianza nel limite di Gilbert-Varshamor. In particolare, per $n = 49$ esiste una famiglia di codici per i quali $\alpha_m(\delta) > 1 - H_m(\delta)$.

Si possono enunciare anche alcuni limiti superiori oltre a quello di Hamming, basati sulla programmazione lineare intera: fissati alcuni parametri a_{ij} positivi e date variabili $x_1, \ldots, x_n$ e parametri $p_1, \ldots, p_n$, si vuole risolvere un sistema lineare di disequazioni, non cercando esplicitamente le soluzioni ma cercando x_i che soddisfano $x_i \geq 0$. Imponendo $a_{ij} \in \mathbb{Z}, x_i \in \mathbb{Z}$ il problema si complica.

Capitolo 2

CODICI LINEARI

2.1 Parametri di un codice lineare

Ci poniamo il problema di cercare codici per cui i processi di codifica e decodifica siano semplici.

2.1.1 Cardinalità

Consideriamo una classe di codici in cui l'alfabeto A sia un campo K.

Considero la mappa $\phi : \mathbb{Z} \to K$ con K campo, tale che $z \mapsto z * 1_A$, segue che $\ker \phi = (m)$ con m caratteristica del campo. Questa mappa è un omomorfismo di anelli, $\frac{\mathbb{Z}}{\ker \phi}$ dev'essere privo di divisori dello zero, quindi m dev'essere primo (che chiamiamo p), e $\frac{\mathbb{Z}}{(m)} \cong F_m$ è contenuto in K. Posso dare una struttura di F_m-spazio vettoriale a K. Data $n = \mathrm{Dim}_{F_m} K$, segue che $o(K) = p^n$. (p^n viene indicata con q).

Definizione 2.1. Un sottoinsieme C di K^n si dice un *codice lineare* se C è un sottospazio vettoriale di K^n.

2.1.2 Distanza

Siccome C è uno spazio vettoriale contiene sempre il vettore nullo.

Definizione 2.2. Chiamo *peso di Hamming* di un vettore $c = (c_1, \ldots, c_n)$ la quantità

$$d(c, 0) := \#\{i \, t.c. \, c_i \neq 0\}$$

Proposizione 2.3. *Sia C un codice lineare in F_{p^n}, allora se definiamo*

$$w_C := \min_{c \in C \setminus 0} w(c)$$

segue che $d_C = w_C$.

Dimostrazione. $d \leq w_C$ perché w_C in effetti è la distanza tra due elementi del codice. D'altra parte, date $c, t \in C$,

$$d(c, t) = d(c - t, 0) = w(c - t), \; c - t \in C.$$

quindi $w_C \leq d$, □

Concludo che **in un codice lineare, distanza minima e peso minimo coincidono**.

Il vantaggio di calcolare il peso minimo invece della distanza minima di un codice è la minor complessità computazionale: il numero di elementi di cui devo calcolare il peso è inferiore al numero delle coppie di elementi tra cui devo calcolare la distanza.

2.1.3 Descrizioni di codici lineari

Esistono almeno due descrizioni di sottospazi di F_{p^n}:

(i) descrizione estensionale: per evitare di descrivere tutti gli elementi se ne fornisce una base.

(ii) approccio duale: si considera C come nucleo di un'applicazione lineare, in particolare $C = \ker H$. Se C è un (n, k, d)-codice, allora esiste una matrice H con $r = n - k$ colonne, che rappresenta la ridondanza di componenti che non danno informazione ma servono per rimediare a errori di trasmissione. r è anche il rango della matrice H. In questo modo

$$C = \{v \in F_{p^n} \, t.c. \, cH = 0\}$$

Sia $C^\perp$ l'insieme dei vettori

$$\{vt.c.v^t c = 0 \forall c \in C\}$$

$C^\perp$ è anch'esso un sottospazio vettoriale di F_{p^n}. Inoltre $n = \mathrm{Dim}V + \mathrm{Dim}V^\perp$. La dimensione di $C^\perp$ è $n - k$, allora esiste una base $\{h_1, \ldots, h_r\}$ per $C^\perp$.

H è la matrice le cui colonne sono i trasposti dei vettori della base di $C^\perp$. H viene detta una *matrice di controllo* per il codice C.

Esempio 2.4 (esempi di codici lineari)**.** (i) un codice di ripetizione C è un codice lineare: C è lo spazio dei multipli scalari del vettore con componenti uguali a 1. H deve avere n righe e $n - 1$ colonne, e le colonne sono i vettori $e_1 - e_i$, $i \in 2, \ldots, n$.

(ii) il duale del codice di ripetizione è ancora un codice lineare (è il codice del controllo di parità, è l'insieme dei vettori per cui la somma delle componenti è uguale a 0).

(iii) il codice di Hamming è un codice lineare con H che ha 7 righe e 3 colonne.

29-10-2015

2.1.4 Riepilogo

Dall'altra volta, abbiamo visto che dato $A = F$ campo (finito, in generale $F = \mathbb{F}_q$), allora $C \leq F^n$ è detto codice lineare, con $n = n(C)$ lunghezza del codice. Inoltre ci sono 2 modi per descrivere un sottospazio vettoriale:

(i) Estensionale, tramite una base;

(ii) Duale, ovvero come nucleo di una matrice H, $C = \ker H$. In particolare se k è la dimensione del codice, allora $H \in \mathrm{Mat}_{n \times r}(F)$ e $r = n - k$ può essere visto o come il rango di H o come la ridondanza del codice.

Capiremo se partendo da H si possono trovare delle informazioni su C.

Teorema 2.5. *Sia $C = \ker H$ un codice lineare, si indichi con h_i la i-esima riga di H. Allora*

$$d_C = \min\{|S| : S \subseteq \{1, \ldots, n\} \, t.c. \, \{h_s : s \in S\} \; sono \quad linearmente \quad dip.\},$$

in altre parole la distanza minima è uguale al numero minimo di righe linearmente dipendenti estraibili da H.

Dimostrazione. Si indichino conn $d := d_C$ e m il secondo membro dell'uguaglianza che si vuole dimostrare.

Sappiamo già che

$$d = w_C = \min\{w_H(c) : c \in C \setminus \{0\}\}.$$

Sia $c \in C$ t.c $w_H(c) = d$. Senza perdita di generalità siano $c_1, \ldots, c_d \neq 0$ e $c_{d+1} = \ldots = c_n = 0$, dove $c = (c_1, \ldots, c_n)$. Allora poiché $c \in C$ ne segue che $cH = 0$, ma anche

$$(c_1, \ldots, c_n)H = (c_1, \ldots, c_n) \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} = c_1 h_1 + \ldots + c_d h_d.$$

Poiché $(c_1, \ldots, c_n) \neq 0$, $(c_1, \ldots, c_d) \neq (0, \ldots, 0)$ quindi le d righe $h_1, \ldots, h_d$ di H sono linearmente dipendenti, ovvero $m \leq d$.

Viceversa siano senza perdita di generalità $h_1, \ldots, h_m$ righe linearmenti dipendenti di H, allora $\exists \lambda_1, \ldots, \lambda_m \in F$ scalari non tutti nulli t.c $\sum_{i=1}^{m} \lambda_i h_i = 0_r$. Costruiamo allora un vettore c il cui peso sia $\leq m$, ovvero

$$c := (\lambda_1, \ldots, \lambda_m, 0, \ldots, 0).$$

Allora $cH = \sum_{i=1}^{m} \lambda_i h_i = 0_r \Rightarrow c \in C$ e $c \neq 0$ quindi $w_H(c) \leq m$ (perché qualche λ_i potrebbe essere $= 0$, ma non possono esserlo tutti), e dato che $d \leq w_H(c)$ poichè d è il minimo peso allora $d \leq m$. □

Questo è il duale del concetto di rango della matrice, poichè ricordo che

Definizione 2.6 (rango)**.** Il *rango* di una matrice di n righe è

$$\max\{|S| : S \subseteq \{1, \ldots, n\} \text{t.c.} \{h_s : s \in S\} \text{sono linearm indip}\}.$$

Vediamo ora qualche esempio sul perché sia utile questo teorema.

Esempio 2.7 (Codice di ripetizione)**.** Sia $C := \text{Rep}_n(F)$. Nella versione estensionale si ha che $C = \text{Span}(1, \ldots, 1)$. Cerco $H \in \text{Mat}_(n \times r)$ dove $r = n - 1$ ($\text{Dim}C = 1$), tale che $C = \ker H$.

Un esempio di matrice di controllo è

$$H = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -1 & -1 & -1 & \ldots & -1 \end{bmatrix}$$

Poniamo $\tilde{C} = \ker H$, sicuramente Rango$H = n - 1$, perchè le prime $n - 1$ righe costituiscono la matrice identità che ha rango massimo.

Inoltre $\text{Dim}(\ker H) = n - r = n - (n - 1) = 1 \Rightarrow \text{Dim}\tilde{C} = 1$, che è proprio quello che si vuole ottenere (vogliamo far vedere che il codice $\tilde{C}$ generato da questa matrice è proprio C).

In particolare osservo che se $c = (c_1, \ldots, c_n) \in \tilde{C}$, l'i-esima componente di cH è il prodotto di c con l'i-esima colonna di H:

$$(cH)_i = c_i - c_n = 0$$

allora per ogni i, $c_i = c_n$. Concludo che $c_1 = c_2 = \ldots = c_{n-1} = c_n \Rightarrow c = (c_1, \ldots, c_n) = c_n(1, \ldots, 1) \in C$. Finalmente possiamo concludere che H è un matrice di controllo per $C = \text{Rep}_n(F)$.

Usiamo quindi H per trovare d_C, che sappiamo già essere n. Devo provare che esistono n righe lineamente dipendenti, ma che ogni scelta di $n-1$ righe è un insieme indipendente.

E' ovvio che scegliendo n righe, queste sono linearmente dipendenti, perchè $\{h_1, \dots, h_r\} \subseteq F^{n-1}$, quindi sono in numero maggiore rispetto alla dimensione del campo.

Per scegliere $n-1$ righe ho le seguenti possibilità:

(i) scegliere le prime $n-1$, ovvero la matriche identica I_{n-1}, che è non singolare perchè $\det I_{n-1} \neq 0$, e quindi le righe sono lin indip;

(ii) scartare una delle prime $n-1$ righe e rimpiazzarla con l'ultima. Per farlo, parto da I_{n-1} e permuto riga per riga fino a portare h_n al posto i-esimo: questi scambi cambiano il segno al determinante della matrice di partenza, in particolare il determinante della matrice finale con h_n al posto i-esimo sarà $(-1)^{i+1}$. In ogni caso il determinante sarà diverso da $0 \Rightarrow$ le righe sono lin indip.

Possiamo finalmente concludere che $d_C = n$.

Osservazione 2.8. Per concludere che $d_C = n$, abbiamo dovuto calcolare $\binom{n}{n-1} = n$ determinanti, ovvero uno per ogni possibile scelta delle $n-1$ righe tra le n presenti. Quindi dal punto di vista

(i) estensionale ho che $\mathrm{Dim} C = 1$,

(ii) duale ho che $\mathrm{Dim} C^\perp = n-1$

$\Rightarrow$ se ho un codice C di dimensioni piccole, lavoro su C, se ho un codice di dimensione grande, lavoro su $C^\perp$.

Esempio 2.9 (parity check code)**.** Sia

$$D = (\mathrm{Rep}_n(F))^\perp = \{v \in F^n : \sum v_i = 0, v = (v_1, \dots, v_n)\},$$

ovvero il parity-check code. La matrice di controllo di questo codice è $H \in \mathrm{Mat}_{n \times r}(F)$ con $r = n - (n-1) = 1$ (perchè $\mathrm{Dim} C = n-1$), quindi H è un vettore colonna. Se considero

$$H = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

segue che se $v \in D$

$$v \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 0$$

quindi H è una matrice di controllo per D.

Calcoliamo ora d_D che coincide con il minimo numero delle righe linearmente dipendenti di H.

Attenzione: se prendo una sola riga, ho un "vettore" formato da un'unica componente, ed è linearmente dipendente solo se è 0. Ma nel nostro caso ci sono solo righe composte da 1, quindi sono righe linearmente indipendenti.

Invece due righe sono sempre linearmente dipendenti, infatti, dati $i, j \in \{1, \dots, n\}$:

$$h_i - h_j = (1) - (1) = 0.$$

pertanto $d_D = 2$.

Osservazione 2.10. In questi due esempi non esiste una scelta privilegiata per le posizioni che forniscono righe linearmente dipendenti.

Vediamo ora un esempio più interessante.

Esempio 2.11 (Codice di Hamming)**.** Per ora sappiamo che il codice di Hamming è un $(7, 4, ?)$-codice su $\mathbb{F}_2$, ovvero non ne abbiamo ancora determinato la distanza minima. Lo facciamo ora con questo nuovo strumento.

Ricordiamoci la matrice H:

$$H = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$$

Osserviamo che

- $d_C = 1$ avviene nella condizione che una riga da sola sia linearmente dipendente, ma questo avviene se la riga è nulla, quindi $d_C \geq 2$.
- $d_C = 2$ si verifica se esistono due righe distinte h_i, h_j che siano linearmente dipendenti, ovvero se $\exists a, b \in F_2$ non entrambi nulli t.c $ah_i + bh_j = 0$. Ma per il caso precedente devo avere $(a, b) \neq (a, 0), (a, b) \neq (0, b) \Rightarrow a, b \neq 0$, e quindi

  $$h_i = -\frac{b}{a} h_j.$$

  Ma in F_2 $-\frac{b}{a} = 1$, cioè $h_i = h_j$. Allora non esistono due righe di H distinte linearmente dipendenti perciò $d_C \geq 3$.
- A questo punto osservo che $h_1 + h_2 + h_3 = 0$, ovvero ho trovato 3 vettori linearmente dipendenti quindi $d_C = 3$.

Esercizio 2.12. Trovare altre scelte per terne linearmente dipententi.

Dimostrazione. Precedentemente (vedi esercizio 1.4) abbiamo determinato le terne per cui la corrispondente sottomatrice 3×3 di H è non singolare, l'insieme complementare dà le terne linearmente dipendenti. □

Attenzione: Sto cercando il minimo numero, non il massimo, quindi mi basta trovare solo una terna linearmente dipendente e ho finito!!!

Esempio 2.13 (Codice di Hamming esteso)**.** Questo codice è ottenuto con l'estensione della matrice H, che indichiamo con $\hat{H}$. Detto C il codice di Hamming, l'insieme di generatori per il codice di Hamming esteso $\hat{C}$ è dato da

$$\hat{C} = \{\hat{c} \in F^8 : \hat{c} = (c, -\sum_{i=1}^{7} c_i), \text{ dove } c = (c_1, \dots, c_7) \in C\} = \ker \hat{H}.$$

Avevamo visto che $n = 8, k = 4$. Mostriamo che $d_C = 4$.

- $d_C \geq 3$, infatti consideriamo la proiezione $\pi : F^4 \to F^3$ t.c. $(x_1, x_2, x_3, x_4) \mapsto (x_1, x_2, x_3)$; prima abbiamo mostrato che $ah_i + bh_j = 0_3$ è impossibile, ora bisogna considerare però anche il caso $a\hat{h}_8 + b\hat{h}_j = 0_4$, ma se proietto questa relazione ottengo $bh_j = 0$, che è ancora impossibile.
- Supponiamo $d_C = 3$, ovvero che $\exists \hat{h}_i, \hat{h}_j, \hat{h}_k$ t.c. $\hat{h}_i + \hat{h}_j + \hat{h}_k = 0_4$. Consideriamo la proiezione $\sigma : F^4 \to F$ t.c. $(x, x_4) \mapsto x_4$, osservo che $\sigma(\hat{h}_i + \hat{h}_j + \hat{h}_k) = 1 \neq 0 = \sigma(0_4)$ (questo avviene perché ho la somma di tre vettori la cui ultima componente è 1). Allora $d_C \neq 3 \Rightarrow d_C \geq 4$. (Per il teorema di Singleton d può essere massimo 5.)
- Osserviamo che $\hat{h}_4 + \hat{h}_2 + \hat{h}_1 + \hat{h}_6 = 0_4 \Rightarrow d_C = 4$.

Possiamo concludere che $\hat{C}$ è un $(8, 4, 4)$-codice. Ne segue anche che poichè e_C è la parte intera di $(d-1)/2$, si ha $e_C = \lfloor \frac{4-1}{2} \rfloor = 1$. Avevamo poi anche visto che questo codice può individuare 2 errori (esempio 1.11), ma non ne corregge due, perché per alcune parole ricevute r esistono due parole del codice c e d equidistanti da r. Questo succede quando la distanza minima è pari, proprio come in questo caso.

2.1.5 Codici autoortogonali

Definizione 2.14. Dato un codice (lineare) $C \leq F^n$, si dice che C è *autoortogonale* (self-orthogonal) se

$$C \leq C^\perp = \{v \in F^n | c \cdot v = 0 \forall c \in C\}.$$

Si può definire $C^\perp$ anche quando C non è sottospazio di F^n. Si può dimostrare che in ogni caso $C^\perp$ è un sottospazio.

Definizione 2.15. Dato un codice (lineare) $C \leq F^n$, si dice che C è *autoduale* (self-dual) se $C = C^\perp$.

Un'ultima precisazione: valgono le definizioni

Definizione 2.16. Sia V spazio vettoriale a prodotto interno su cui è definito un prodotto scalare, dato $W \leq V$ si definisce il suo *complemento ortogonale*

$$W^\perp := \{v \in V : w \cdot v = 0 \forall v \in W\}.$$

e anche

Definizione 2.17. Dato $W \leq V$, $U \leq V$ è il *complemento* di W se $W + U = V$ e $W \cap V = \{0\}$.

Attenzione: il duale $C^\perp$ visto finora NON è il complemento di C.
2-11-2015

2.1.6 Esistenza di codici autoortogonali su campi finiti

Osservazione 2.18. Su $\mathbb{R}$ non esistono codici autoduali o autoortogonali, infatti se per assurdo fosse $C \leq C^{\perp}$, preso $c = (c_1, \ldots, c_n) \in C$ si avrebbe

$$c \cdot c = \sum_{i=1}^{n} c_i^2 = 0$$

ma poiché tutti i quadrati sono positivi questo avviene se e solo se tutte le componenti di c sono nulle.

Consideriamo il caso di campi finiti di cardinalità q:

(i) Se $n = 1$, sia γ l'unica componente di un codice candidato ad essere autoortogonale, allora segue che $\gamma^2 = 0$ ovvero $\gamma = 0$, e si riottiene il codice ridotto al solo 0.

(ii) Se $n = 2$, considero la parola $c = (\alpha, \beta)$ contenuta in un codice autoortogonale. Allora $\alpha^2 + \beta^2 = 0$. Supponiamo $\alpha, \beta \neq 0$, allora la relazione che deve valere si traduce in $(\alpha/\beta)^2 = -1$, affinché questo avvenga deve esistere un elemento i nel campo che gioca il ruolo dell'unità immaginaria sui complessi. $\beta \neq 0$, quindi sta nel gruppo moltiplicativo G del campo, tale elemento deve avere ordine 4, allora $4 \mid q - 1$. Vale anche viceversa: infatti per un risultato dovuto a Gauss, il gruppo moltiplicativo di un campo finito è ciclico (e quindi, se $4 \mid q-1$, esiste sempre un elemento di ordine 4 ed è dato da $g^{(q-1)/4}$. Supponiamo che esista i tale che $i^2 = -1$. Allora posso considerare la retta generata da un elemento $v = (\alpha, \beta)$ tale che $v^2 = 0$, allora ogni elemento λv della retta ha ancora la proprietà che $(\lambda v)^2 = 0$ e quindi tale retta è un codice autoduale.

(iii) Per $n \geq 3$, esistono vettori isotropi (ovvero tali che $v^2 = 0$) $v \in F_{q^n}$ per ogni campo finito, infatti vale il seguente lemma.

Lemma 2.19. *In ogni campo finito -1 è sempre somma di due quadrati, ovvero esistono $\alpha, \beta \in F_q$ tali che $\alpha^2 + \beta = 2 = -1$.*

Conseguentemente, moltiplicando la relazione per un qualsiasi γ^2, si ottiene

$$\gamma^2\alpha^2 + \gamma^2\beta^2 = -\gamma^2$$

e si ha una terna $\delta = \gamma\alpha$, $\varepsilon = \gamma\beta$, γ la cui somma di quadrati sia nulla, e quindi un vettore con $(\delta, \gamma, \varepsilon, 0, \ldots, 0)$ soddisfa la condizione.

dimostrazione del lemma. Sia

$$S = \{\alpha^2,\ \alpha \in F_q\}$$

In un campo di cardinalità q, ogni elemento annulla il polinomio $x^q - x$. Distinguiamo due casi:

(i) Se q è pari, per ogni elemento del campo F_q si può scrivere $\alpha = (\alpha^{q/2})^2$, quindi $S = F_q$, allora la relazione da dimostrare vale poiché $((-1)^{q/2})^2 + 0^2 = -1$.

(ii) Affermo che **se q è dispari, i quadrati sono** $(q + 1)/2$. Infatti, i quadrati sono le immagini della mappa $\sigma : G \to G$ tale che $g \mapsto g^2$. Questa mappa è un omomorfismo di gruppi. Mostro che il prodotto viene preservato:

$$\sigma(ab) = (ab)^2 = abab \stackrel{?}{=} \sigma(a)\sigma(b)$$

e l'uguaglianza è vera perché essendo in un campo il prodotto è commutativo.

$$\sigma(G) = S \setminus 0$$

Osservo che l'immagine è un gruppo, isomorfo a $\frac{G}{\ker \S}$. $\ker \S$ è dato dagli elementi di ordine 2. $\S(g) = 1$ ovvero $g^2 = 1$ ha come soluzioni $g = \pm 1$ (nel caso di q dispari queste due soluzioni sono distinte). Poiché il gruppo è ciclico, non ci sono altre soluzioni (infatti per ogni divisore dell'ordine del gruppo esiste un unico sottogruppo di tale ordine). Allora $o(\ker \S) = 2$. Quindi

$$o(S \setminus 0) = o(\sigma(G)) = o(\frac{G}{\ker \S}) = (q-1)/2$$

quindi

$$o(S) = (q-1)/2 + 1 = (q+1)/2,$$

come abbiamo asserito.
Definisco poi il traslato di S:

$$T = \{-1 - \beta^2,\, \beta \in F_q\}$$

esso ha ancora $(q+1)/2$ elementi. Mostriamo che $S \cap T \neq \emptyset$. Se $S \cap T = \emptyset$,

$$q = o(S \cup T) = o(S) + o(T) = (q+1)/2 + (q+1)/2 = q + 1,$$

assurdo. Allora dev'esistere un elemento γ in $S \cap T$: $\gamma \in S$ implica $\gamma = \alpha^2$ per un certo α, inoltre $\gamma \in T$ implica $\gamma = -1 - \beta^2$ per un certo β, allora

$$\alpha^2 = -1 - \beta^2 \longrightarrow \alpha^2 + \beta^2 = -1.$$

□

Dato un codice che sia uno spazio vettoriale che ammette una forma bilineare alternante, si dice che V è uno spazio simplettico. Esistono sottospazi W detti isotropi che coincidono con il proprio ortogonale. Questi spazi giocano lo stesso ruolo di codici autoduali nella teoria dei codici.

2.1.7 Esempio di applicazione del teorema sui codici

Esempio 2.20. Sia $K = F_3$, costruisco una matrice H dotata di 4 righe e 2 colonne.

$$H = \begin{matrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{matrix}$$

(le righe sono date dalla rappresentazione ternaria dei numeri 1,2,4,5) Allora $C = \ker H$ che è un sottospazio vettoriale di F_4 corregge un errore. La seconda e la prima riga di questa matrice costituiscono la matrice identica, quindi il rango è uguale a 2 e la dimensione è $4 - 2 = 2$. La distanza è il minimo numero di righe linearmente dipendenti estraibili da H. In particolare

(i) La distanza non può essere 1 perché non ci sono righe nulle.

(ii) Se fosse $d = 2$, esistono h_i, h_j righe linearmente dipendenti, ovvero esiste $\lambda \neq 0$ tale per cui $h_i = \lambda h_j$. Esclusa la prima riga che è sicuramente indipendente dalle altre, si ha una situazione del tipo $\lambda 1 = 1$ per la prima componente e $\lambda 1 \neq 1$ per la seconda, il che è assurdo. Allora $d \geq 3$.

(iii) $d = 3$, infatti $(0,1), (1,0), (1,1)$ sono linearmente dipendenti.

Questo è un (4,2,3)-codice chiamato *codice di Hamming ternario.*

2.1.8 Dimostrazione algebrica del teorema di Singleton

Ricordiamo il limite di Singleton: un qualsiasi codice soddisfa la condizione $k_C + d_C \leq n_C + 1$. Ne diamo una dimostrazione utilizzando l'algebra lineare:

Dimostrazione. Sia H una matrice di controllo per C (con C lineare), allora H ha n righe e $r = n - k$ colonne. La relazione da dimostrare, $d \leq n - k + 1$ si traduce in $d \leq r + 1$. Per il teorema sui codici lineari $d = \min |S|$ con S insieme di righe linearmente dipendenti. Ogni scelta di $r + 1$ righe di H fornisce un insieme linearmente dipendente (si considerano $r + 1$ elementi in uno spazio standard di vettori di lunghezza r), e quindi la relazione da mostrare è vera. □

Definizione 2.21. Un codice C di parametri n, k, d si dice m.d.s. (maximum distance separator code), se $d + k = n + 1$.

Un esempio di codice m.d.s. è il codice ternario di Hamming $(2 + 3 = 4 + 1)$.

Non esistono codici troppo lunghi che siano m.d.s., ovvero devono valere le condizioni $n \leq q + 1$ se q è dispari e $n \leq q + 2$ se q è pari.

2.1.9 Analisi delle matrici di controllo

Come sono fatte tutte e sole le matrici di controllo di un fissato codice?

Dato un codice $C = \ker H$ con H di rango massimale, quali sono altre possibili candidate ad essere matrici di controllo? Ovvero se $C = \ker H = \ker H'$, che relazione c'è tra H e H'?

Implicazione 1: se $H' = HM$ con M $r \times r$ invertibile, allora $\ker H = \ker HM$. Infatti, $\ker H \subseteq \ker HM$ perché dato $c \in C$,

$$0 = cH = cHM.$$

Inoltre $\ker HM \subseteq \ker H$, infatti se per assurdo esiste $v \notin C, v \in HM$, segue che $w = cH \neq 0 \in \ker M$, ma questo è assurdo perché M è invertibile.

Implicazione 2: Viceversa, se $\ker H = \ker H'$, allora $H' = HM$ per qualche M matrice invertibile $r \times r$ a coefficienti nel campo.

Le colonne di H, $t^1, \ldots, t^r$ sono r vettori riga trasposti. Se considero il codice duale, esso ammette come base i vettori $t_1, \ldots, t_r$ (è l'insieme dei vettori ortogonali a tutti i vettori del codice e in particolare gli elementi del codice sono ortogonali alle colonne di H). Poiché la matrice ha rango massimo le colonne sono linearmente indipendenti e quindi costituiscono una base. Se chiamo $\bar{H}$ la seconda matrice di controllo e $\bar{t}_1, \ldots, \bar{t}_r$ le sue colonne, anch'esse sono una

base per il duale. Se voglio esprimere $\bar{t}_i$ in funzione dei t_j, esisteranno scalari $m_{ij}, j = 1, \dots, r$ tali che $\bar{t}_i = \sum_j m_{ij} t_j$ e in particolare esiste una matrice M invertibile che manda una base nell'altra, e tale matrice dev'essere invertibile.

2.2 Descrizione estensionale dei codici

Considero uno spazio vettoriale $C = F^n$, e fisso una base di vettori per C data da vettori $c_1, \dots, c_k$. Essi possono essere codificati con una matrice G costituita dalla concatenazione dei vettori, detta *matrice generatrice*. G è associata a $f : F^k \to F^n$, e rappresenta un processo di codifica. L'informazione viene codificata per permettere la correzione di errori, e gli elementi di F^k vengono mandati in elementi di F^n. Date due matrici generatrici $G, \bar{G}$, segue che $\bar{G} = NG$, dove N è una matrice invertibile $k \times k$ (per passare da una matrice generatrice all'altra basta fare un cambio di base. Anche qui come nel caso precedente il rango delle due matrici dev'essere massimale.

Supponiamo che G non sia quadrata, e consideriamo le matrici della forma PGQ con P $k \times k$ e Q $n \times n$, entrambe invertibili.

Osserviamo che per una particolare scelta di P e Q si può ottenere la matrice uguale all'identità nel blocco $r \times r$ in alto a sinistra, e nulla altrove (r è la dimensione dello spazio generato dalla matrice di partenza).

Siano $w_1, \dots, w_r \in F^k$ tali che $w_i G = v_i$. Il nucleo ha dimensione $k - r$. Allora esistono $w'_{r+1}, \dots, w'_k \in \ker G$. Si nota che $\{w_1, \dots, w_r\} \cap \{w'_{r+1}, \dots, w'_k\} = \emptyset$, allora l'unione di questi vettori è una base per F^k.

Scegliamo vettori in modo che

$$\{v_1, \dots, v_r, v'_{r+1}, \dots, v'_n\}$$

sia una base per F^n. Determiniamo la matrice associata a L_G rispetto a queste due basi.

Chiamo M la matrice che esprime le immagini dei vettori della base iniziale come combinazione lineare di quelli della base di arrivo. M ha il blocco $r \times r$ a sinistra uguale all'identità (per il fatto che $w_i G = v_i$) e 0 altrove (perché w'_i stanno nel ker).

Se chiamo P la matrice che esprime i w_i in funzione della base canonica, e Q la matrice che esprime i vettori della base canonica in funzione dei v_i, allora $PGQ = M$, e M ha la forma descritta precedentemente.

Se fosse possibile modificare ogni matrice generatrice G di un codice C moltiplicando a destra e a sinistra per opportune matrici invertibili, la matrice avrebbe la forma descritta sopra e quindi lo studio di tale codice non sarebbe interessante.

2.2.1 Decodifica di un messaggio

Dato un vettore w e considerato l'elemento del codice wG, supponiamo che dopo la trasmisione si riceva v. Vogliamo decodificare v e ottenere come risultato un elemento di F^k.

Definizione 2.22. Si dice che G è *in forma standard* se G ha la matrice identica nel blocco $k \times k$ a sinistra (questa matrice ha k righe e n colonne).

Dato un messaggio $m \in F^k$, se G è in forma standard, la sua codifica è $mG = (mI, mA)$. Se la trasmissione è stata effettuata senza rumore, guardando le prime k componenti del vettore ricevuto si risale subito all'informazione che è stata codificata.

Definizione 2.23. G si dice *sistematica* se esistono k colonne in G coincidenti con quelle della matrice identità.

Esempio 2.24. Esempio di matrice sistematica:

$$\begin{matrix} 0 & 1 & -1 & 1 \\ 1 & 3 & 1 & 0 \end{matrix}$$

Se chiamo $i_1, \ldots, i_k$ le posizioni delle colonne uguali a quelle dell'identità prese nell'ordine (ovvero i_i corrisponde all'i-esima colonna dell'identità), segue che nel vettore che si ottiene le componenti corrispondenti alle posizioni $i_1, \ldots, i_k$ prese nell'ordine danno il vettore originario trasmesso.

Data una matrice sistematica, devo riordinare le colonne di G usando qualche trasformazione che coinvolga matrici. Allora la moltiplico a destra per una matrice di permutazione, che ha un solo elemento uguale a 1 per ogni riga e per ogni colonna, e i restanti elementi sono uguali a 0.

Permutando le righe della matrice si ottiene un nuovo codice (associato a GQ con Q matrice di permutazione), che non coincide necessariamente con C (associato alla matrice iniziale G). Da un punto di vista dei parametri (distanza minima, capacità correttiva, dimensione) essi rimangono uguali nei due codici.

Esercizio: Nell'esempio precedente $d = 3$ se la caratteristica è ≥ 2 e altrimenti è 2.

Esiste una trasformazione lineare che ha la proprietà di mantenere i pesi dei vettori, ovvero $w(c) = w(cG)$. Dato un vettore di componenti $(\gamma_1, \ldots, \gamma_n)$, se le moltiplico per scalari $\lambda_1, \ldots, \lambda_n$ non nulli, il peso del vettore viene preservato. Combinando la moltiplicazione per uno scalare alla permutazione delle componenti, si ha che

$$(\gamma_1, \ldots, \gamma_n)$$

e

$$(\lambda_1\gamma(\pi(1)), \ldots, \lambda_n\gamma(\pi(n)))$$

hanno lo stesso peso.

Le posizioni in cui mettere 1 in una matrice di permutazione sono determinate dalla scelta della permutazione π.

Il gruppo delle matrici invertibili $n \times n$ su un campo F è detto *gruppo monomiale*. Si può considerare come prodotto intrecciato del gruppo moltiplicativo del campo con il gruppo simmetrico.

Teorema 2.25 (teorema di Francy-Macwilliams)**.** *Sia f una mappa lineare che preserva i pesi, allora f è monomiale.*

(se $f : C \to \bar{C}$ con $C, \bar{C}$ sottospazi di F^n con la stessa dimensione, f è lineare e $w_C = w_{f(C)}$, allora esiste $\bar{f} : F^n \to F^n$ lineare e monomiale tale per cui $\bar{f}(C) = f$)

5-11-2015

2.3 Alcuni esercizi

2.3.1 Codici autoduali

Esercizio 2.26. Dimostrare che il codice di Hamming $(8, 4, 4)$ è autoduale.

Il codice è autoduale se Dim$C$ è pari alla metà della lunghezza del codice (qui questa condizione è soddisfatta). La seconda condizione per l'esistenza di codici autoduali è l'esistenza di sottospazi costituiti da vettori isotropi, ovvero vettori $(v_1, \ldots, v_n)$ tali che la somma delle componenti al quadrato sia nulla. Nel caso in cui Car$F = 2$,

$$(a + b)^2 = a^2 + b^2 + 2ab = a^2 + b^2$$

ovvero in un campo di caratteristica 2, $a^2+b^2 = 0$ equivale ad $(a+b)^2 = 0$. Quindi, se il campo ha q elementi e $C = F^n$ i vettori isotropi sono q^{n-1} (è l'iperpiano ottenuto considerando il duale, perché è dato dagli elementi del codice che elevati al quadrato danno 0).

Una volta fatta l'ipotesi sulla dimensione, per verificare se un codice è autoduale basta verificare che $\omega_i \cdot \omega_j = 0$ per ogni ω_i, ω_j vettori della base. Questo implica che dati vettori qualsiasi $v, w \in C$, essi sono combinazioni lineari degli elementi della base, in particolare

$$v = \sum_i c_i\omega_i,\ w = \sum_j c_j\omega_j$$

quindi

$$v \cdot w = (\sum_i c_i\omega_i) \cdot (\sum_j c_j\omega_j) = \sum_{i,j} c_ic_i\omega_i \cdot \omega_j$$

e se tutti i prodotti scalari sono nulli anche $v \cdot w = 0$. Qui Dim$C$ = Dim$C^\perp = 4$, quindi mostrando l'ortogonalità tra i vettori della base segue subito che i due spazi sono uguali.

Osservo che le colonne di $\hat{H}$, dette $h_1^t, \ldots, h_4^t$ costituiscono una base per il codice duale, esse infatti sono linearmente indipendenti perché il rango di $\hat{H}$ è 4, inoltre stanno nel duale perché $c_ihi^t = 0 \forall i$. Basta mostrare che il duale di $\hat{C}$ è autoduale, ovvero $(C^\perp)^\perp = C$.

$$h_1 = (1, 1, 1, 1, 0, 0, 0, 0)$$

$$h_2 = (0, 1, 1, 0, 0, 1, 1, 0)$$

$$h_3 = (1, 0, 1, 0, 1, 0, 1, 0)$$

$$h_4 = (1, 1, 1, 1, 1, 1, 1, 1)$$

Il peso dei vettori di questa base è congruo a 0 modulo 4. Allora $h_i^2 = 0 \forall i$. Inoltre per ogni i, $h_i \cdot h_4 \equiv 0 \mod 2$ (infatti è dato dalla somma delle componenti di h_i poiché sto moltiplicando per un vettore che ha tutte le entrate uguali a 1) Negli altri casi si svolgono i calcoli:

$$h_1 \cdot h_2 = 2 \equiv 0 \mod 2$$

$$h_1 \cdot h_3 = 2 \equiv 0 \mod 2$$

$$h_2 \cdot h_3 = 2 \equiv 0$$

e quindi il codice è autoduale.

Osservazione 2.27. Considero $C = \mathrm{Rep}_2(F_2)$, allora $\mathrm{Dim}C = 1 = n/2$. La base di questo codice è costituita solo da $(1,1)$, che ha prodotto scalare nullo con sé stesso, quindi il codice è autoduale.

Su $F = F_2$, si distinguono due famiglie di codici autoduali:

(i) si richiede che $w_H(c) \equiv 0 \mod 4 \forall c \in C$.

(ii) si richiede che $w_H(c) \cong 1 \mod 4 \forall c \in C$

Esercizio 2.28. Dimostrare che se C è un codice autoduale definito su F_2, allora ogni elemento del codice ha peso di Hamming pari.

Dimostrazione. Dato $c = (\gamma_1, \ldots, \gamma_n)$, segue che

$$c \cdot c = \sum_i \gamma_i^2 = |\{i : \gamma_i = 1\}| = w_H(c) \equiv 0 \mod 2$$

per la dualità, quindi $w_H(c)$ dev'essere pari. □

Un codice appartenente alla prima famiglia è il cosidetto codice di Golay (ha lunghezza 23 e dimensione 12). Il conteggio del numero di parole di un dato peso nei codici autoduali della prima famiglia è univocamente determinato dall'informazione sul codice di Golay e sul codice di Hamming esteso.

Esercizio 2.29. Dimostrare che il codice di Hamming ternario (definito su F_3) è autoduale.

Dimostrazione.

$$H = \begin{matrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{matrix}$$

Mostriamo che $C = \ker H$ è autoduale. Le colonne della matrice sono date da:

$$h_1 = (0,1,1,1)$$

$$h_2 = (1,0,1,2)$$

$$h_1 \cdot h_1 = 3 \equiv 0 \in F_3$$

$$h_1 \cdot h_2 = 0$$

$$h_2 \cdot h_2 = 1 + 1 + 4 = 6 \equiv 0 \mod 3$$

Questo è un m.d.s. codice, ovvero $k + d = n + 1$, infatti $2 + 3 = 4 + 1$. La lunghezza di m.d.s. codici dev'essere uguale a $q+1$ se q è dispari e a $q+2$ se q è pari (qui $4 = 3+1$ quindi questo codice m.d.s. è massimale). □

Esercizio 2.30. Mostrare che ogni matrice generatrice per un codice C di parametri $(4,2,3)$ su F_3 ha la proprietà che ogni sua sottomatrice 2×2 è non singolare.

Dimostrazione. Poiché il codice è autoduale, la matrice H considerata prima coincide con la sua matrice generatrice (trasposta), quindi

$$G = \begin{matrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{matrix}$$

Distinguo due casi: ci sono sottomatrici della forma:

$$\begin{pmatrix} 0 & 1 \\ 1 & \nu \end{pmatrix}$$

che hanno determinante -1, quindi non sono singolari. Altre sottomatrici sono della forma

$$\begin{pmatrix} 1 & \nu \\ 1 & \mu \end{pmatrix}$$

con determinante $\mu - \nu \neq 0$.

Proviamo ora che **l'assenza di sottomatrici singolari non dipende dalla particolare scelta della matrice, ma è verificata per una qualsiasi matrice generatrice** $\bar{G}$. Se $\bar{G}$ è un'altra matrice generatrice, essa si ottiene a partire da G attraverso la moltiplicazione per una matrice invertibile, quindi esiste una matrice P invertibile tale che $\bar{G} = PG$. Se chiamo $\bar{G}^i$ l'i-esima colonna di $\bar{G}$, si ha $\bar{G}^i = PG^i$. In particolare le sottomatrici 2×2 di $\bar{G}$ sono della forma:

$$\begin{matrix} \bar{G}^i & \bar{G}^j \end{matrix} = P \begin{matrix} G^i & G^j \end{matrix}$$

ma poiché la matrice di colonne (G^i, G^j) è non singolare per i conti precedenti, e P è non singolare, anche il determinante della matrice al primo membro è diverso da 0 per il teorema di Binet. □

2.3.2 Information set

Definizione 2.31. Sia C un codice lineare di lunghezza n su un campo F e di dimensione k, e sia G una matrice generatrice. Un sottoinsieme S di $\{1, \ldots, n\}$ viene detto *information set* se ha esattamente k elementi e se il determinante della sottomatrice data dalle colonne di G corrispondenti agli indici di S è non singolare.

E' possibile ottenere m in F^k a partire dalla scelta delle componenti del vettore codificato se e solo se S è un information set.

Osservazione 2.32. L'ordine di S non può essere inferiore a k, altrimenti non sarebbe possibile avere una mappa suriettiva da $F^{|S|} \to F^k$. Allora la cardinalità minima è k.

Considero G che abbia due blocchi, il primo di dimensione $k \times k$ che chiamo A e il secondo di dimensione $k \times (n-k)$. La corrispondenza $m \mapsto mA$ è invertibile se e solo se A è non singolare. Se chiamo c_1, c_2 le porzioni della parola codificata corrispondenti rispettivamente alle prime k e alle ultime $n-k$ componenti di c, riesco a risalire a m moltiplicando c_1 per A^{-1}.

Dall'esercizio precedente segue che per il codice (4,2,3) di Hamming su F_3 ogni sottoinsieme di $\{1, \ldots, 4\}$ di cardinalità 2 è un information set.

Dato un codice (4,2,3), si possono ottenere altri codici (4,2,3) permutando le colonne della sua matrice.

Si può mostrare anche che gli information sets di un dato codice non dipendono dalla scelta di una sua matrice generatrice. Infatti abbiamo mostrato che dato un information set relativo a G, la sottomatrice estratta da $\bar{G}$ è non singolare.

2.3.3 Struttura del gruppo monomiale

Esercizio 2.33. Dimostrare che il gruppo $GL(2, F_2)$ è isomorfo al gruppo simmetrico su 3 elementi, e determinare i sottogruppo monomiale e di permutazione.

Dimostrazione. Il sottogruppo monomiale è l'insieme di tutte le matrici che si ottengono moltiplicando una matrice diagonale invertibile per una matrice di permutazione, mentre il sottogruppo di permutazione è l'insieme delle matrici di permutazione.

$$GL(2, F_2) = \{ \begin{matrix} a & b \\ c & d \end{matrix} \ t.c.\, ad - bc \neq 0,\ a, b, c, d \in F_2 \}.$$

Allora $GL(2, F_2)$ è in corrispondenza biunivoca con l'insieme delle radici in F_2^4 dell'equazione polinomiale $xy - zw = 1$, che è a coefficienti in F_2 (varietà algebrica affine).

(i) Caso 1: $a = 0$ implica $bc = 1$, ovvero $b = c = 1$, e si ottengono le due matrici

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

(ii) Caso 2: se $a = 1$, $d = 1 + bc$, e ci sono quattro matrici possibili (corrispondono ai possibili modi in cui si può scegliere la coppia (b, c)):

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Allora $GL(2, F_2)$ ha esattamente 6 elementi, come S_3. In un gruppo di ordine 6, gli elementi possono avere ordine 1,2,3. Tra le matrici elencate precedentemente, le matrici triangolari superiori vengono dette blocchi di Jordan. Ad esempio un blocco di Jordan è la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = (id + e_{12})$$

Per calcolare A^2 applico la formula del binomio di Newton:

$$(id + E_{12})^2 = id^2 + E_{12}^2 + 2E_{12}id = id + 0 + 2E_{12} = id$$

quindi le matrici di Jordan hanno ordine 2.

Le matrici

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

sono dette matrici compagne del polinomio x^2+x+1, si dimostra che data la matrice compagna di un polinomio, il suo polinomio caratteristico coincide con tale polinomio. x^2+x+1 è fattore di $x^3 - 1$, quindi $m^2 + m + 1 = 0$ implica $m^3 = 1$, e quindi tali matrici hanno ordine 3.

Un isomorfismo tra due gruppi mantiene gli ordini degli elementi. Per mostrare che c'è un isomorfismo si applica il teorema di Cayley. Si mostra che G non è normale, e si considera l'insieme delle classi laterali H, HM, HM^2 con

$$M = \begin{matrix} 0 & 1 \\ 1 & 1 \end{matrix}$$

Si costruisce una corrispondenza tra gli elementi del gruppo e le permutazioni su Ω. □

9-11-2015

2.4 Criterio di equivalenza per codici lineari

2.4.1 Sottogruppo diagonale

Definizione 2.34. Siano C, D codici nello spazio F^n. Si dice che due codici sono *equivalenti* se esiste una trasformazione lineare $\phi : C \to D$ invertibile tale che per ogni $c \in C$, $w_H(c) = w_H(\phi(c))$.

Una nozione più debole di equivalenza consiste nel richiedere che i codici abbiano la stessa distanza minima e la stessa dimensione.

Definizione 2.35. Sia G il gruppo $GL(n, F)$. Si definisce *sottogruppo diagonale* $D_{n,F}$ l'insieme di tutte le matrici diagonali $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ con $\lambda_i \in F \setminus 0 \forall i$.

Mostro che **il sottogruppo diagonale è davvero un sottogruppo**, ovvero che dati $g_1, g_2 \in D$, $g_1 g_2^{-1} \in G$. Considero due matrici: $A = \text{Diag}(\lambda_1, \ldots, \lambda_n)$ e $B = \text{Diag}(\mu_1, \ldots, \mu_n)$. Allora

$$AB^{-1} = \text{Diag}(\lambda_1, \ldots, \lambda_n) * \text{Diag}(\mu_1^{-1}, \ldots, \mu_n^{-1}) = \text{Diag}(\lambda_1\mu_1^{-1}, \ldots, \lambda_n\mu_n^{-1})$$

In particolare si può costruire una corrispondenza biunivoca tra il sottogruppo delle matrici diagonali e $(F^*)^n$. Il prodotto di matrici diagonali si traduce nel prodotto componente per componente in $F^* \times \cdots \times F^*$.

2.4.2 Matrici di permutazioni

Fissiamo una base per F^n, ad esempio la base standard $\{e_1, \ldots, e_n\}$. Prendo una permutazione su n lettere (gli indici dei vettori della base) e definisco la mappa P tale che

$$P(e_i) = e_{i\pi}$$

dove $i\pi$ significa $\pi(i)$. P definisce una sola trasformazione lineare di F^n in sé.

P è invertibile infatti se a π corrisponde P, a π^{-1} corrisponde P^{-1}. Allora anche la matrice in $\mathbf{Mat}(n, F)$ associata a P è invertibile.

Applicando in sequenza due permutazioni π, σ:

$$P_\sigma P_\pi(e_i) = P_\sigma(e_{i\pi}) = e_{i\pi\sigma}.$$

Allora a $\pi \circ \sigma$ corrisponde il prodotto delle matrici associate a tali permutazioni, ovvero PS.

Se chiamo $\psi : S_n \to GL(n, F)$ l'applicazione che a una permutazione π associa la corrispondente matrice P definita come sopra, segue che ψ **è un morfismo iniettivo**, infatti

$$\ker \psi = \{\pi \, t.c. P_\pi = id, \longrightarrow \pi(i) = i \forall i = 1, \ldots, n\}$$

Il gruppo delle matrici appena descritte viene detto *gruppo delle matrici di permutazione.*

La matrice P associata alla permutazione π è una matrice che ha su ogni riga e su ogni colonna un unico elemento uguale a 1, in particolare se $\pi(i) = j$, la posizione (i, j) è uguale a 1 (se assumo che ci siano due elementi uguali a 1 sulla stessa colonna ci sarebbero due elementi con la stessa immagine e questo non può avvenire per l'iniettività della permutazione).

Il gruppo simmetrico ha $n!$ elementi, mentre il gruppo diagonale ne ha $(q-1)^n$.

2.4.3 Relazione tra gruppo diagonale e gruppo di permutazioni

Sia $D_{n,F}$ sia S_n contengono elementi che preservano i pesi dei vettori.

Dati sottogruppi A, B, il sottogruppo che li contiene entrambi è descritto nel seguente modo:

$$< A, B >= \{a_1 b_1 + \cdots + a_n b_n, \, a_i \in A, \, b_i \in B, \, n \in N\}$$

Data una matrice diagonale $\text{Diag}(\lambda_1, \ldots, \lambda_n)$, essa manda e_i in $\lambda_i e_i$. Osservo che

$$e_i \mapsto^P e_{i\pi} \mapsto^D \lambda_{i\pi} e_{i\pi} \mapsto^{P^{-1}} \lambda_{i\pi} e_i$$

quindi

$$P^{-1}DP = \mathrm{Diag}(\lambda_{1\pi}, \ldots, \lambda_{n\pi}).$$

Quindi $P_1^{-1}D_1P_1 = D_1' \longrightarrow D_1P_1 = P_1D_1'$. Vale la scrittura

$$P_1D_1P_2D_2 = D_1'(P_1P_2)D_2 = D_1'D_2'P_1P_2$$

Allora il sottogruppo generato da S_n e da $D_{F,n}$ si può indicare con

$$\{DP, D \text{ diagonale}, P \text{ di permutazione}\}$$

(in pratica si considera $n = 1$ nella scrittura precedente)

Si introduce una nozione di prodotto nell'insieme delle coppie (D, P) ponendo:

$$(D, P)(\bar{D}, \bar{P}) = (DP\bar{D}P^{-1}, P\bar{D})$$

Definizione 2.36. Si chiama *sottogruppo monomiale* di $GL(n, F)$ il gruppo $MN(F)$ generato dalle matrici diagonali e dalle matrici di permutazione. In particolare, questo gruppo è l'insieme dei prodotti di una matrice diagonale con una matrice di permutazione.

(i) Se $n = 2, q = 2$, il gruppo monimiale è isomorfo a S_2. La matrice identica e la matrice che corrisponde allo scambio $(1, 2)$ costituiscono il gruppo di permutazione. Il gruppo diagonale ha solo un elemento, l'identità. Il prodotto tra i due gruppi coincide con il gruppo costituito dall'identità e da $(1, 2)$.

(ii) Se $q = 3, n = 2$, il gruppo monomiale è isomorfo a S_3: il gruppo simmetrico rimane uguale a prima, il gruppo diagonale è costituito da 4 elementi,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

ed è isomorfo a $C_2 \times C_2$ (si associa ad ogni elemento z $\log_{-1} z$).

2.4.4 Enunciato del teorema di Macwilliams

Teorema 2.37. *Siano C e D due codici lineari contenute in F^n, con F campo finito; sia ϕ un isomorfismo lineare tra C e D che preserva i pesi. Allora ϕ si può estendere ad una trasformazione lineare invertibile di tutto F^n* che sia monomiale.

In realtà, una mappa lineare che sia invertibile su un sottospazio di F^n si può sempre estendere a una mappa invertibile su tutto F^n, la novità di questo teorema è che la mappa sia monomiale.

Si dice che ε è un'estensione di ϕ se $\varepsilon : F^n \to F^n$ è tale che $\varepsilon(c) = \phi(c) \forall c \in C$. Non è detto che un'estensione sia unica, come mostra il seguente esempio.

Esempio 2.38. Sia $C = D$, e cerco una trasformazione ϕ che sia l'identità su C. Prendo un complemento K di C, su tale complemento considero una mappa invertibile ψ che manda K in sé. Allora la mappa che manda $c + k$ in $c + \psi(k)$ è un'estensione della mappa di partenza.

2.4.5 Risultati ausiliari per la dimostrazione

L'idea della dimostrazione è quella di fissare $\mathrm{Dim}C = \mathrm{Dim}D = k$, e di considerare F^k. Si vuole determinare nello spazio vettoriale l'insieme di tutte le rette (o i punti proiettivi di F^k).

Lemma 2.39. *Il numero di tali rette è* $\mu_k = \frac{q^k-1}{q-1}$.

Dimostrazione. Per individuare una retta scelgo $v \neq 0 \in F^k$, e ho $q^k - 1$ scelte. Ogni classe di equivalenza (ovvero ogni classe di vettori che generano la stessa retta), contiene $q-1$ elementi (una classe contiene gli elementi $k * v$, $k \in F_q$, $k \neq 0$). □

Fisso un ordinamento sulle rette $l_1, \ldots, l_{\mu_k}$. Supponiamo che la retta i-esima sia generata da un rappresentante rispetto alla relazione di equivalenza indicata prima. Si costruisce una matrice a coefficienti interi, della forma $\mu_k \times \mu_k$, simmetrica, e dette A_{ij} le sue entrate

$$A_{ij} = \begin{cases} 0 & \iff l_i \perp l_j \\ 1 & \text{altrimenti} \end{cases}$$

dove $l_i \perp l_j$ se il prodotto scalare tra i generatori è nullo (l'ortogonalità non dipende dalla scelta del rappresentante).

Questa viene chiamata matrice di adiacenza del grafo dei sottospazi di F_n rispetto all'ortogonalità.

Esempio 2.40. Sia $l_1 =< (1,0) >$, $l_2 =< (0,1) >$, $l_3 =< (1,1) >$. Allora la matrice A ha la forma

$$A = \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{matrix}$$

Il vettore che ha come componenti le somme delle righe di A è $(2,2,2)$. Fissato j come indice di colonna, considero la sottomatrice formata solo dalle righe con 0 in posizione j-esima. Poi considero il vettore che ha come componenti la somma delle colonne della matrice: qui tale vettore ha somma $(0,1,1)$ per $j = 1$, $(1,0,1)$ per $j = 2$, $(1,1,0)$ per $j = 3$. In particolare, per ogni j, c'è una sola riga con 0 nella j-esima posizione. Lo spazio vettoriale generato dalle righe di A è F_3. Infatti $(1,1,1)$ e $(0,1,1)$ stanno in S, allora anche la loro differenza sta in S, ovvero $(0,0,1) \in S$, con S spazio generato dalle righe della matrice.

Teorema 2.41 (generalizzazione dell'esempio precedente)**.** *Considero il piano di Fano, che è un grafo che unisce due rette se non sono ortogonali, e la matrice A ad essa associata definita sopra: allora*

(i) *il vettore che ha come componenti la somma delle colonne di A, che chiamo K è il vettore di componenti* $\mu_k - \mu_{k-1}$.

(ii) *Il vettore che ha come componenti la somma delle colonne della sottomatrice di A formata dalle righe con 0 in posizione j-esima, indicato con y^j, ha la forma $(a, \dots, 0, a, \dots, a)$ con $a = \mu_{k-1} - \mu_{k-2}$ e 0 in posizione j-esima.*

(iii) *la matrice è invertibile.*

Dimostrazione. (i) La somma dell'i-esima colonna di A è data da

$$|\{j \, t.c. l_i \not\perp l_j\}|$$

ed è la differenza tra il numero totale di rette (μ_k) e il numero delle rette l_j ortogonali alla retta data. Il duale della retta fissata è uno spazio vettoriale di dimensione $k-1$, e quindi le rette nel duale sono μ_{k-1}. Allora la somma cercata è $\mu_k - \mu_{k-1}$

(ii) Fisso j e considero la sottomatrice formata da tutte le righe con 0 in posizione j-esima, ovvero è formata dalle righe di indice m con $l_m \perp l_j$. La somma dell'i-esima colonna di tale matrice, con $i \neq j$, è data da

$$|\{kt.c. l_k \not\perp l_i,\, l_k \perp l_j\}|$$

quindi come prima, questa cardinalità è la differenza tra il numero totale di rette e il numero di rette perpendicolari a l_i e l_j contemporaneamente, ovvero

$$= \mu_{k-1} - |\{kt.c.\, l_k \perp l_j, l_k \perp l_i\}| = \mu_{k-1} - \mu_{k-2}$$

poiché il sottospazio generato da l_i e l_j ha dimensione $k-2$.

(iii) Se chiamo L il Q-spazio vettoriale generato dalle righe di A, ottengo che il vettore $(1, \dots, 1)$ sta in L perché è uguale a $\frac{1}{\mu_k - \mu_{k-1}} K$. Inoltre per il punto 2, $(1, 1, \dots, 0, \dots, 1) = (1, 1, \dots, 1) - e_j$ sta in L essendo uguale a

$$\frac{1}{\mu_{k-1} - \mu_{k-2}} y^j$$

allora anche la differenza fra questi due vettori, che è e_j, sta in L. Allora poiché tutti i vettori della base canonica stanno nello spazio generato dalle righe, questa matrice è invertibile.

□

Lemma 2.42. *Sia G una matrice generatrice di un codice C, di k righe e n colonne. Introduco un vettore w con μ_k componenti, che dipende dal codice, dove*

$$w_i = \#\{G^i \, t.c. \, (G^i)^t \in l_i\}$$

con G^i colonne della matrice generatrice. Allora

$$(Aw^t)_i = w_H(v_i G),$$

dove v_i è il generatore della retta l_i.

Dimostrazione. Per definizione di prodotto

$$(Aw^t)_i = \sum_{j=1}^{k} A_{ij} w_j = \sum_{j \in S_i} w_j$$

dove $S_i = \{j\, t.c.\, l_j \not\perp l_i\}$.

$$= \sum_{j \in S_i} |\{G^j\, t.c. (G^j)^t \in l_j\}|$$
$$= |\{G^j\, t.c. (G^j)^t \not\perp l_i\}|$$

D'altra parte, se considero la decodifica di v_i, v_iG è il vettore che ha come componenti $v_i(G^j)^t$ al variare di G^j nell'insieme delle colonne della matrice. Quindi $W(v_iG)$ è pari al numero di colonne non ortogonali a l_i, quindi, eguagliando le due relazioni

$$(Aw^t)_i = w_H(v_iG)$$

□

2.4.6 Dimostrazione del teorema di Macwilliams

Dimostrazione. Considero la mappa $\phi : C \to D$, e sia f la mappa di codifica su C, allora

$$F^k \mapsto^f C \mapsto^\phi D$$

Se definisco $g := \phi \circ f$, g è una mappa di codifica sul codice D e la matrice G' con colonne $g(e_1), \ldots, g(e_k)$ è una matrice generatrice per il codice D. La matrice generatrice G per C è la matrice che ha colonne $f(e_1), \ldots, f(e_k)$.

Definisco i vettori w, w' con le seguenti componenti:

$$w_i = |\{G^i\, t.c.\, (G^i)^t \in l_i\}|$$

$$w'_i = |\{G'^i\, t.c.\, (G'^i)^t \in l_i\}|$$

Allora, per il lemma 2.42

$$(Aw^t)_i = w_H(v_if) = w_H(f(v_i))$$

e poiché ϕ preserva i pesi:

$$= w_H(\phi(f(v_i)) = w_H(g(v_i)) = (Aw'^t)_i \forall i = 1, \ldots, \mu_k$$

Allora $Aw^t = Aw'^t$. Siccome A è invertibile,

$$Aw^t = Aw'^t \longrightarrow w^t = A^{-1}Aw'^t \longrightarrow w = w'$$

e quindi i due vettori coincidono componente per componente:

$$w_i = w'_i \longrightarrow |\{G^i t.c. (G^i)^t \in l_i\}| = |\{G'^i t.c. (G'^i)^t \in l_i\}|$$

Fissata unna retta nello spazio F^k, il numero di colonne delle due matrici generatrici che appartengono a tale retta è uguale. Applicando una permutazione posso riordinare le colonne di G o di G' in modo che le colonne con la proprietà di appartenere a l_i siano nelle stesse posizioni. Di conseguenza, dette $c_1, \ldots, c_n, c'_1, \ldots, c'_n$ le colonne delle due matrici, dopo aver applicato una matrice di permutazione, si avrà che, fissato un indice j, le colonne c_j e c'_j hanno la proprietà di appartenere ad una stessa retta l_i, ovvero esistono scalari λ_j, λ'_j tali che

$$\lambda_j c_j = \lambda'_j c'_j,\ \forall j = 1, \ldots, \mu_k.$$

Quindi

$$c'_j = (\lambda'_j)^{-1}\lambda_j c_j$$

ovvero le colonne della prima matrice si ottengono dalle colonne della seconda per moltiplicazione per uno scalare, e questo equivale a moltiplicare G per una matrice diagonale. Allora esiste una trasformazione monomiale (la composizione della matrice di permutazione e di quella diagonale) che trasforma le colonne di G in quelle di G'. □

Lemma 2.43 (lemma di Wit)**.** *Dato uno spazio vettoriale con una forma bilineare non degenere, e un'isometria (trasformazione lineare invertibile tra due sottospazi di V che conserva le distanze), allora f si estende a un'isometria di tutto lo spazio.*

12-11-2015

2.5 Decodifica con codici lineari

Consideriamo la seguente situazione:

$$m \in F^k \to mG \in F^n \rightsquigarrow r \in F^n \to c \in F^n \dashrightarrow m' \in F^k$$

Con $F^kG = C \leq F^n$. Si cerca un protocollo per decodificare $r \in F^n$ mediante $c \in C$, ovvero scelgo c tale che

$$d_H(r, c) = \min\{d_H(r, t) : t \in C\}.$$

Potrebbe però accadere che durante il processo di trasmissione siano stati commessi troppi errori e quindi $Dec(r)$ è sbagliato, oppure esistono più parole che hanno la stessa distanza (minima) da r. Il processo di decodifica descritto sopra funziona se il numero di errori commessi è minore o uguale a e_C. Bisogna determinare quindi le parole $r \in F^n$ tali che: $r \in \bigcup_{c\in C} B_e(c)$ (unione disgiunta), poiché per tali parole l'algoritmo di decodifica descritto sopra funziona.

Preso $C \leq F^n$: durante il processo di trasmissione $c \rightsquigarrow r = c + \varepsilon$, dove ε viene chiamato *errore*.

Sezionando F^n rispetto alla relazione di equivalenza E, con $vEw \Leftrightarrow v - w \in C$, le classi di equivalenza sono le classi laterali

$$[v]_E = \{w : wEv\} = \{w \in F^n : w = v + C\},$$

dove $v + C$ è il traslato di C mediante v, ossia un sottospazio affine (In F^2 gli spazi affini rispetto ad uno spazio vettoriale sono tutte le rette parallele).

Se $r = c + \varepsilon$, $r + C = \varepsilon + c + C = \varepsilon + C$. All'interno di una classe di equivalenza bisogna cercare elementi di peso "piccolo", in particolare sono interessanti solo le classi dove "piccolo" significa minore o uguale della capacità correttiva del codice.

Osserviamo che possono esistere classi $\varepsilon + C$ dove $w_H(\varepsilon + C) > e\ \forall c \in C$ (basta prendere $\varepsilon \notin \bigcup_{c\in C} B_e(c)$, e un tale elemento esiste se il codice non è $e-$perfetto).

Proposizione 2.44. *Se $r + C$ ammette un elemento ε tale che $w_H(\varepsilon) \leq e_C$, allora esso è unico e $Dec(r) = r - \varepsilon$.*

Dimostrazione. Supponiamo che esistano $\varepsilon_1, \varepsilon_2 \in r + C$, tali che $w_H(\varepsilon_i) \leq e_C$. Allora risulta $\varepsilon_1 = r + c_1$, $\varepsilon_2 = r + c_2$ ed $\varepsilon_1 - \varepsilon_2 = c_1 - c_2 \in C$. Si ha poi

$$w_H(\varepsilon_1 - \varepsilon_2) \leq w_H(\varepsilon_1) + w_H(\varepsilon_2) \leq 2e_C \leq d \longrightarrow \varepsilon_1 = \varepsilon_2$$

infatti non può esistere un elemento in C con distanza minore della distanza minima del codice. □

Definizione 2.45. Data una classe laterale $r + C$ si definisce *coset leader*, se esiste, l'unico elemento di peso minimo $\varepsilon \in r + C$.

L'idea è di far variare $\varepsilon \in F^n$ con $w_H(\varepsilon) \leq e_C$ e determinare tutte le classe $\varepsilon + C$. Memorizzate queste classi, ricevuto $r \in \bigcup_{w_H(\varepsilon) \leq e_C} \varepsilon + C$, segue che $Dec(r) = r - \varepsilon$ se $r \in \varepsilon + C$. Se r non sta nell'unione delle classi $\varepsilon + C$ al variare di ε con peso $\leq e_C$, non è possibile decodificare r.

Sia H una matrice di controllo per C ($cH = \underline{0} \forall c \in C$. Allora il rango di H è $n - k$ con $k = \mathrm{Dim}C$. Se $r = \varepsilon + c$, risulta

$$rH = (\varepsilon + c)H = \varepsilon H,$$

$s = rH$ è detto la sindrome di r; cerco (se esistono) vettori ε di peso piccolo aventi la stessa sindrome s di r e pongo $Dec(r) = r - \varepsilon$.

Con questo approccio vi sono due vantaggi:

(i) $s \in F^{n-k}$ con $n - k < n$.

(ii) Basta guardare gli s che provengono da ε con $w_H(\varepsilon) \leq e_C$, ovvero basta memorizzare le sindromi dei vettori dell'insieme

$$\xi = \{\varepsilon \in F^n : \, w_H(\varepsilon) \leq e_C\},$$

ovvero l'insieme delle sindromi $\mathcal{S} = \xi H$.

Ricordando che il raggio di ricoprimento di un codice C (covering radius, $cr(C)$) è il minimo intero t tale che $F^n = \bigcup_{c \in C} B_t(c)$ vale la seguente proposizione:

Proposizione 2.46. *$t = rc(C)$ è il minimo numero di righe di una matrice di controllo H per C necessarie per ottenere ogni elemento di F^r come loro combinazione lineare (H ha n righe e $n - k$ colonne).*

Dimostrazione. Preso $s \in F^r$, $v \in F^n$, con $v \in \bigcup B_t(c)$, allora $v = c + \varepsilon$ con $w_H(\varepsilon) \leq t$, $\varepsilon = (0, ..., 0, \varepsilon_{i_1}, 0, ..., \varepsilon_{i_t}, 0, ...0)$. Risulta

$$s = vH = \varepsilon H = \sum_{j=1}^{t} \varepsilon_{i_j} h_{i_j}$$

con h_k = riga $k-$esima di H. Poiché il rango di H è r, esistono r righe di H linearmente indipendenti, ad esempio posso assumere che siano le prime r, $h_1, ..., h_r$ e $\oplus_{i=1}^{r} F_{h_i} = F^r$ e quindi $cr(C) \leq r$. □

Se si costruiscono famiglie di codici dotati di elevata simmetria: $C \leq F^n$ $\varphi \in \mathrm{End}_F(F^n)$, $\varphi \in GL_n(F)$, sappiamo che per il teorema di MacWilliams si può considerare $\varphi \in Mn(F)$.

Definizione 2.47. $\varphi \in Mn(F)$ viene detto un *automorfismo* del codice C se $\varphi(C) = C$. $Aut(C)$ è un sottogruppo di $Mn(F)$.

Capitolo 3

CODICI DI HAMMING

3.1 Definizione generale dei codici di Hamming

3.1.1 Spazi e punti proiettivi

Sia F campo, $|F| = q$, $r \in \mathbb{N}$.
Preso $V = F^r$, e posto $V^{\#} = \frac{F^r}{\{0\}}$, introduco la seguente relazione

$$vEw \iff \exists \lambda \in F^* \, t.c. \, v = \lambda w$$

(è ovviamente una relazione d'equivalenza).

Definizione 3.1. L'insieme

$$[v]_E = \{\lambda v : \lambda \in F^*\}$$

è detto il *punto proiettivo* rappresentato da v, invece $\{[v]_E : \ v \in V^{\#}\}$ è detto *spazio proiettivo* $(r-1)-$dimensionale su F e si denota con $\mathcal{P}^{r-1}(F)$.

Esempio 3.2. $\mathcal{P}^1(\mathbb{R})$ è una retta privata dell'origine, una semicirconferenza con estremi identificati codifica tutti e soli i rappresentanti di punti proiettivi.

Uso $\mathcal{P}^{r-1}(F)$ per costruire $H \in \mathbf{Mat}_{n\times r}(F)$, scelgo un sistema di rappresentanti per tutti gli n punti proiettivi, risulta

$$n = \frac{q^r - 1}{q - 1} = q^{r-1} + ... + 1.$$

Se $v = (v_1, ..., v_r)$ identifico $[v]_E = (v_1 : ... : v_r)$. Se prendo un elemento $w = (\mu_1, ..., \mu_r)$ nella classe, si ha $\mu_i = \lambda v_i \forall i \Rightarrow \frac{\mu_i}{\mu_j} = \frac{v_i}{v_j}$. Seguendo tale ragionamento, se $v = (v_1, ...v_r)$ con $v_1 \neq 0$ posso prendere come rappresentante di $[v]_E$ il vettore $(1 : \frac{v_2}{v_1} : ... : \frac{v_r}{v_1})$, (osserviamo che un rappresentante di tale forma è unico, se ce ne fosse un altro lo scalare di passaggio sarebbe uguale ad 1).

- I punti aventi rappresentanti $(v_1, ..., v_r)$, con $v_1 \neq 0$ sono q^{r-1}, e lo spazio contenente tali vettori, indicato con $\mathbb{A}^{r-1}(F)$, è uno spazio affine di dimensione $r-1$ su F. Infatti si può scrivere

  $$(1, v_2, ..., v_r) = (1, 0, ..., 0) + (0, v_2, ..., v_r).$$

  L'insieme

  $$(1, 0, ..., 0) + \{v \in F^r : v_1 = 0\}$$

  è un iperpiano di dimensione $r-1$.

- Iterando il procedimento su $(0, v_2, ..., v_r)$ con $v_2 \neq 0$ si ha

  $$(0 : 1 : v_3 ... : v_r) = (0, 1, \ldots, 0) + (0, 0, v_3, ..., v_r)$$

  e lo spazio

  $$(0, 1, 0, ..., 0) + \{v \in F^r : v_1 = v_2 = 0\}$$

  ha q^{r-2} elementi e si indica con $\mathbb{A}^{r-2}(F)$ ed è un iperpiano di dimensione $r-2$.

- Si arriva a concludere che

  $$\mathcal{P}^{r-1}(F) = \oplus_{i=0}^{r-1} \mathbb{A}^i(F).$$

3.1.2 Costruzione del codice di Hamming

Preso T sistema di rappresentanti per $\mathcal{P}^{r-1}(F)$, pongo $n := |T| = \frac{q^r-1}{q-1}$. Sottointendendo un ordinamento su T costruiamo la matrice H che ha come righe $t \in T$.

Definizione 3.3. Un codice $C = \ker H$ viene detto un *codice di Hamming* di parametri q, r con $n_C = \frac{q^r-1}{q-1}$, $k_C = n_C - r$, ed inoltre $r = n_C - k_C$ viene detta *ridondanza* di C.

16-11-2015 Se chiamo e_i l'i-esimo vettore della base standard, un rappresentante della sua classe di equivalenza sarà della forma $t_i = (0, 0, \dots, \lambda_i, 0, \dots, 0)$ con λ_i scalare non nullo in posizione i-esima, di lunghezza r. Nella matrice H ordino i vettori in modo che all'i-esima riga corrisponda il rappresentante dell'i-esimo vettore della base standard. Detta G la sottomatrice $r \times r$ in alto, si ha che $\det G = \prod_i \lambda_i$, quindi tale sottomatrice è non singolare, e r è il rango di H e $k = \frac{q^n-1}{q-1} - r$.

Determiniamo la distanza minima dei codici di Hamming. Devo mostrare che nessuna coppia di righe di H è linearmente dipendente. Se in H ci fossero due righe linearmente dipendenti, dette h_i, h_j, allora $ah_i + bh_j = 0$ con $a, b \neq 0$. Se uno dei due scalari fosse nullo seguirebbe automaticamente che anche l'altro lo è: supponiamo allora $a \neq 0$, da cui

$$h_i = -b/ah_j$$

ma questo significa che i punti proiettivi rappresentati da h_i e h_j coincidono, e quindi questo implica $i = j$. Se considero i vettori riga

$$\lambda e_1,\ \mu e_2,\ \nu(e_1 + e_2)$$

essi sono tre righe di H per qualche scelta di scalari λ, μ, ν non nulli. Tali vettori ovviamente sono linearmente dipendenti, e questo assicura che la minima distanza di codici di Hamming è $d = 3$.

Questi codici prendono il nome di *single error correcting codes*.

Esempio 3.4. Se $q = 3, r = 2$, ottengo un codice di lunghezza

$$\frac{3^2 - 1}{3 - 1} = 4$$
$$k = 4 - 2 = 2,\ d = 3$$

Si può creare un omomorfismo tra i punti proiettivi e il campo degli scalari. Associo a punti della forma $x : y$ lo scalare y/x. Questa mappa è definita solo per $x \neq 0$, e in tal caso è ben posta, perché un vettore equivalente a $[x : y]$ sarà della forma $[\lambda x : \lambda y]$ e quindi la quantità associata ad esso è sempre y/x. L'unico punto con la prima coordinata nulla viene chiamato punto infinito della retta proiettiva e indicato con $-\infty$.

Esempio 3.5. Consideriamo la matrice di controllo per il codice di Hamming ternario

$$H = \begin{matrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{matrix}$$

Questo codice di Hamming ha parametri $q = 3, r = 2,\ d = 3, e = 1$, e le righe della matrice sono punti proiettivi.

Proposizione 3.6. *Il raggio di ricoprimento dei codici di Hamming è uguale a 1.*

Dimostrazione. Ricordiamo che

$$RC = \min\{r t.c. \bigcup_{c\in C} B_r(c) = F^4\}$$

$$= \min\{s\, t.c. \exists\, i_1, \ldots, i_s,\, t.c. v = \sum_{j=1}^{s} \lambda_{i_j} h_{i_j}\}$$

Poiché le righe di H sono tutti i punti proiettivi, dato $v \in F_{3,2}$, se $v \neq 0$ esiste i tale che $[v] = [h_i]$, ovvero $v = \lambda h_i$ per qualche λ, e quindi il raggio di ricoprimento è uguale a 1. □

Come indicato nell'osservazione 1.15, se C è un codice allora C è perfetto se e solo se

$$e_C := \max\{tt.c. B_t(c) \cap B_t(c') = \emptyset\}$$

coincide con il suo raggio di ricoprimento

$$RC := \min\{t\, t.c.\, F^n = \bigcup_{c\in C} B_t(c)\}.$$

Teorema 3.7. *Il raggio di ricoprimento dei codici di Hamming è uguale a 1 (quindi tali codici sono e perfetti).*

Dimostrazione. Si può scegliere di dimostrare che i codici di Hamming di parametri q ed r sono 1-perfetti replicando la dimostrazione precedente fornita per il codice relativo ai parametri $3, 2$.

Alternativamente si ha la seguente dimostrazione aritmetica: F_{q^n} ha q^n punti, vogliamo contare gli elementi del codice. In una bolla di raggio 1, ci sono $1 + n * (q-1)$ (il centro, a cui sommo tutti gli elementi che distano da esso per una coordinata). Sfruttando la definizione di n:

$$1 + n(q-1) = 1 + \frac{q^r - 1}{q-1} * (q-1) = 1 + q^r - 1 = q^r.$$

Se il codice è e-perfetto, tutte le bolle sono disgiunte. Poiché i punti del codice sono q^k, moltiplicando tale numero per la cardinalità di una bolla si ha

$$q^k * q^r = q^{k+r} = q^n$$

e questo avviene se e solo se $k = n - r$, e questo è verificato. Quindi l'unione delle bolle copre tutti gli elementi del codice, quindi il raggio di ricoprimento è 1. □

Ricordiamo che i codici e-perfetti hanno sempre distanza minima dispari, e quindi coincidente con $2e + 1$.

Nel 1971 fu dimostrato il seguente fatto:

Teorema 3.8. *Sia C un codice non necessariamente lineare, e-perfetto su un alfabeto con q elementi, allora C ha gli stessi parametri di un codice di Hamming (lunghezza $\frac{q^r-1}{q-1}$, cardinalità $q^{\frac{q^r-1}{q-1}-r}$). Non è dimostrato che q debba essere potenza di un primo, ma non esistono esempi con q che non lo sia. Se aggiungo l'ipotesi che C sia lineare allora C dev'essere un codice di Hamming.*

Esistono codici non lineari e non traslati di Hamming 1-perfetti.

Se $e = 2$, esiste un unico codice perfetto (codice di Golay su F_3) con lunghezza 11, dimensione 6, distanza minima 5. Analogamente, se $e = 3$, esiste un codice di Golay con lunghezza 23, dimensione 12 e distanza minima 7 che è 3-perfetto.

3.1.3 Codici di Hamming di parametri q, r monomialmente equivalenti

Considero una matrice H^T con T insieme ordinato di rappresentanti per P^{r-1}. Dato $S \neq T$ e detta H^S la matrice corrispondente, H^T si ottiene moltiplicando H^S per una matrice monomiale. Infatti se $T = \{T_1, \dots, T_n\}$ e $S = \{S_1, \dots, S_n\}$, esiste una permutazione π tale che S_i sia equivalente a $T_{i\pi}$, ovvero

$$S_i = \lambda T_{i\pi}$$

e quindi considero $M = DP$ con

$$D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

segue che $h^S = MH^T$.

Inoltre, se considero $c_S \in C^S$, esso viene ucciso da H^S, quindi

$$0 = c_S H^S = c_S M H^T$$

quindi $c_S M \in C^T$, ovvero C^S, C^T **sono monomialmente equivalenti**.

Viceversa, **dato un codice di Hamming e** M **una matrice monomiale, mostro che** $C^S := C^t M$ **è ancora un codice di Hamming**. Infatti se $M = DP$, $M^{-1} = P^{-1}D^{-1} = P^{-1}D^{-1}PP^{-1}$, dove se P è associata alla permutazione σ, P^{-1} è associata a σ^{-1}. Mostro che $P^{-1}D^{-1}P$ è ancora una matrice diagonale, infatti

$$e_i \mapsto^{P^{-1}} e_{i\sigma^{-1}} \mapsto^{D^{-1}} (\lambda_{i\sigma^{-1}})^{-1} e_{i\sigma^{-1}} \mapsto^{P^{-1}} (\lambda^{-1})_{i\sigma^{-1}} e_i$$

allora M^{-1} è monomiale e C^S è ancora un codice di Hamming.

3.1.4 Gruppo degli automorfismi di un codice

Vogliamo ricavare informazioni sul codice di Hamming studiando le loro simmetrie, e determinare la cardinalità di $H_{q,r}$, insieme di tutti i codici di Hamming di parametri q, r. Nel linguaggio della teoria dei gruppi, $H_{q,r}$ è l'orbita di un qualsiasi suo elemento rispetto all'azione del gruppo monomiale. Gli elementi di un'orbita sono dati da $\frac{o(G)}{S}$ dove S è lo stabilizzante, ovvero il gruppo degli automorfismi del codice stesso (è l'insieme delle matrici monomiali che mandano il codice in sé).

Cercare il gruppo di automorfismi di un codice equivale a cercare matrici monomiali che fissano questo insieme.

Matrici generatrici: Sia $C = F^k * G$ con G matrice generatrice. Per ottenere un codice equivalente a C, esso deve avere la stessa dimensione e la trasformazione deve preservare i pesi, e quindi si deve avere $C' = CM$ (vedi osservazione precedente). Da un punto di vista delle matrici generatrici, $G' = GM$. Se M manda C in sé, e quindi se $C = C'$, GM è ancora una matrice generatrice per C, allora per quanto osservato nella sezione 2.2 deve esistere $Nk \times k$ invertibile tale che $GM = NG$.

Data una matrice generatrice G di C, le matrici della forma NG con N invertibili sono ancora matrici generatrici per C. GM è ancora una matrice generatrice per C se e solo se $GM = NG$ con N matrice invertibile $k \times k$. Quindi, **fissata una matrice generatrice G per C, il gruppo dei suoi automorfismi è dato dal gruppo di tutte le matrici monomiali M per cui esiste una matrice N invertibile tale che $GM = NG$.**

MATRICI DI CONTROLLO: In maniera analoga, dato un codice C con parametri n, k, d, ci si chiede come è fatta una matrice di controllo per un codice $C' = CM^{-1}$: se $c \in C$,

$$0 = cH = cM^{-1}MH = c'MH$$

I ranghi di H e MH coincidono perché ho moltiplicato per una matrice invertibile. Allora MH è una matrice di controllo per C'. Se $C = C'$, MH è una matrice di controllo per C, allora deve esistere N invertibile tale che $MH = HN$. Allora **il gruppo degli automorfismi è l'insieme delle matrici monomiali tali per cui esiste una matrice non singolare N tale che $MH = HN$.**

Teorema 3.9 (teorema di Wilson)**.** *Sia C un (n,k,d)-codice lineare con distanza minima d e $r = n - k$, allora*

(i) *esiste un omomorfismo $\theta : Aut(C) \to GL(q, r)$.*

(ii) *θ è iniettivo se $d \geq 3$.*

Dimostrazione. (i) Data una matrice monomiale, M essa sta in $Aut(C)$ se e solo se esiste N invertibile tale che $MH = HN$ per le osservazioni precedenti. Inoltre $H' = MH$ ha rango r, siano $h'_1, \ldots, h'_r$ le sue colonne. Inoltre $h'_1, \ldots, h'_r$ sono linearmente indipendenti. Sia $H'' = HN$, allora le sue colonne sono combinazioni lineari delle colonne di H. Esiste un'unica trasformazione lineare da S (sottospazio di dimensione r di F^n formato dalle colonne di H) che manda la i-esima colonna di H nella i-esima colonna di H'. Allora esiste ed è unica la trasformazione lineare associata ad N, quindi la mappa che manda M in N è ben posta.
Mostriamo che θ **che manda M in N è un omomorfismo**: siano $M_1, M_2 \in Aut(C)$ matrici monomiali nel gruppo degli automorfismi, legati dalla relazione

$$M_i H = HN_i.$$

Allora

$$M_2 M_1 H = M_2 H N_1 = H N_2 N_1$$

quindi $\theta(M_2M_1) = \theta(M_2)\theta(M_1)$, ovvero $\theta : Aut(C) \to GL(n, \mathbb{R})$ è un omomorfismo.
Osservazione: In generale, dato un gruppo A e un omomorfismo $\alpha : A \to B$ con B gruppo noto, segue che $\alpha(A)$ è un sottogruppo di B e B è isomorfo a $\frac{A}{\ker \alpha}$, e si ha la massima informazione su A quando α è iniettivo.

(ii) Determinare il nucleo di θ equivale a cercare matrici M che vengano mandate in id, ovvero tali che

$$MH = HN, \longrightarrow HM = H.$$

Siano $m_1, \ldots, m_n$ le righe della matrice MH, allora esse sono della forma $\lambda_1 r(\pi(1)), \ldots, \lambda_n r(\pi(n))$ dove π è una permutazione degli indici di riga. Allora se M sta nel nucleo

$$\lambda_1 r(\pi(i)) = r(i), \forall i$$

Se esistesse i tale che $\pi(i) \neq i$, $r(i)$ e $r(\pi(i))$ sono righe dipendenti di H, allora $d \leq 2$, contro l'ipotesi. Segue quindi che $\pi = id$. Allora $\lambda_i r_i = r_i$, ma poiché $r_i \neq 0$ segue che $\lambda_i = 1$, e quindi M è la matrice identica $n \times n$.

□

Il vantaggio è che attraverso θ si passa da matrici monomiali in $GL(n, q)$ a matrici in $GL(r, q)$. Se il codice è un codice di Hamming, si ha una riduzione sensibile della dimensione dello spazio di arrivo.

Come sono fatti i sottogruppi di $GL(r, q)$? La risposta dipende sia da r che da q. Dato un gruppo γ finito di ordine r, esso può essere considerato un sottogruppo del gruppo simmetrico. Siccome gli elementi del gruppo simmetrico possono essere rappresentati come matrici di permutazione, si ha un isomorfismo con un sottogruppo di $GL(n, q)$.

19-11-2015

3.2 Esercizi

Esercizio 3.10. Ideare un algoritmo di recupero di errori per codici di Hamming di parametri q e r (ricordiamo che, poiché i codici di Hamming hanno distanza minima 3, hanno capacità correttiva uguale ad 1).

Dimostrazione. Il codice di Hamming è definito come nucleo di una matrice H $n \times r$, con $n = \frac{q^r-1}{q-1}$, e le righe di H rappresentano tutti e soli i punti proiettivi nello spazio di dimensione $r-1$ sul campo. Sia V un generico spazio vettoriale di dimensione n sopra q, e sia $v \in V$. Vale l'identità $v = c + \varepsilon$ con ε vettore di errore. Si definisce *sindrome* il valore $v * H$, che è un vettore con r componenti. Si ha

$$vH = (c + \varepsilon)H = cH + \varepsilon H = \varepsilon H$$

quindi, a partire da v, si può ricavare ε con la condizione che $w_H(\varepsilon) \leq 1$ (infatti il codice corregge un errore). Se $w_H(\varepsilon) \leq 1$, o ε è il vettore nullo, oppure $\varepsilon H = \lambda H_t$ con H_t t-esima riga di H (infatti sto moltiplicando H per un vettore con una sola componente, la t-esima, diversa da 0). In particolare λ, t sono unici, infatti se esistesse un'altra coppia (μ, s) che soddisfa la relazione, allora $\mu H_s = \lambda H_t$, ovvero $[H_t] = [H_s]$ e questo implica $t = s$, ovvero le due coppie coincidono. La sindrome, vettore di lunghezza r, sarà della forma μH_i perché le righe di H coprono tutto l'insieme dei punti proiettivi, allora

$$S = \mu H_i = \varepsilon H = \lambda H_t$$

segue che $i = t$, quindi a partire dalla sindrome, si può risalire alla posizione in cui è avvenuto l'errore, inoltre $\mu = \lambda$, e quindi si conosce anche l'entità dell'errore.

Riassumendo, l'algoritmo che si applica è il seguente:

(i) calcolo la sindrome S relativa al vettore v;

(ii) determino l'unico i per cui S è proporzionale all'i-esima riga di H;

(iii) determino λ che realizza $S = \lambda H_i$

(iv) la decodifica di v è data da $v - \lambda e_i$ con e_i vettore della base standard di F_{q^n}.

□

Esercizio 3.11. Il gruppo di automorfismi di un qualsiasi codice di Hamming di parametri q ed r è un sottogruppo di $GL(r,q)$, infatti si può definire la mappa θ tale che $M \mapsto N$ con $M \in Aut(C) \subseteq MN(q)$ e $N \in GL(r,q)$ con $MH = HN$, se $d = 3$ la mappa risulta essere iniettiva. Dimostrare che nel caso dei codici di Hamming θ è suriettiva (in modo che $GL(r,q) \cong Aut(C)$), ovvero per qualunque scelta di N in $GL(r,q)$ esiste M monomiale tale che $MH = HN$.

Dimostrazione. Considero la matrice di controllo H con righe $r_1, \ldots, r_n$ dove gli r_i sono vettori a coefficienti in F_q. Le righe di HN sono date da H_iN (la j-esima componente è data dal prodotto dell'i-esima riga per la j-esima colonna).

Preso l'insieme delle classi di equivalenza delle n righe di H, esso coincide con tutto lo spazio proiettivo. Affermo che vale lo stesso per le righe di HN, basta mostrare che N **induce una permutazione sugli elementi dello spazio proiettivo**. Supponiamo che due delle classi di equivalenza delle righe di HN coincidono, ovvero $[R_iN] = [R_jN]$, e quindi

$$R_iN = \lambda R_jN$$

e poiché N è invertibile, segue che $\lambda R_i = R_j$ ovvero $[R_i] = [R_j]$ e $i = j$. Siccome i due insiemi hanno lo stesso numero di oggetti, N induce una permutazione. Fissato un indice i, deve esistere j tale che sia verificata $[R_iN] = [R_j]$, e si può pensare che $j = \pi(i)$. Allora esiste λ_i tale che

$$R_iN = \lambda_i R_{\pi(i)}$$

Considero allora una matrice M la cui i-esima riga ha un unico elemento $\lambda_i \neq 0$, che sta esattamente in posizione $\pi(i)$. Moltiplicando M per H, si ottiene

$$(MH)_i = \lambda_i R_{\pi(i)} = R_iN = (HN)_i$$

e quindi θ è suriettiva (M definita come sopra è la preimmagine di N), pertanto il gruppo degli automorfismi di codici di Hamming di parametri q, r è isomorfo a $GL(r,q)$. □

Esercizio 3.12. Sia $q = 2, r = 3$. Determinare l'unica matrice monomiale M tale che $\theta(M) = N$ con

$$N = \begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{matrix}$$

Dimostrazione. N è un blocco di Jordan, dato da $id + A = id + (e_{12} + e_{23})$, e A è idempotente. N **è invertibile**, infatti

$$N = id + A$$

con

$$e_1A = e_2;\ e_2A = e_3;\ e_3A = 0$$

$$e_1A^2 = e_3;\ e_2A^2 = 0;\ e_3A^2 = 0$$

Sia x una variabile, si vuole calcolare $(1+x)^{-1}$. Devo cercare una serie di potenze tale che

$$(1+x) \circ \sum a_i x^i = 1$$

Considero la serie geometrica di ragione $-x$, allora

$$(1+x) * \sum_i (-x)^i = (x+1)^{-1} \frac{1}{1+x} = 1$$

quindi l'inversa di $1+x$ è la serie geometrica di ragione $-x$.

Tornando al caso delle matrici, sostituendo x con A, poiché A^3 è 0, si svolge il prodotto

$$(1+A)(1-A+A^2) = 1$$

e l'inversa di $(1+A)^{-1}$ è $1-A+A^2$.

Nota: non è sempre lecito fare questa sostituzione e passare dal caso di variabili al caso di matrici. Si consideri ad esempio il teorema di Cayley-Hamilton: Se sostituisco la matrice A alla variabile x nel polinomio caratteristico $p_A(x) = \det(A - xid)$ si ottiene 0. Sostituendo A a x l'asserto seguirebbe immediatamente, in realtà quest'operazione qui non è consentita perché le matrici non commutano.

Nel caso considerato però la sostituzione funziona, perché l'identità commuta con le altre matrici. In particolare allora N è invertibile, e N^{-1} è calcolabile.

Devo costruire H tale che $C = \ker H$, devo quindi scegliere i rappresentanti dei punti proiettivi e l'ordinamento. Qui la classe di equivalenza di ogni punto proiettivo contiene solo un elemento, quindi si ha solo da compiere la scelta sull'ordinamento. Considero come ordinamento lo sviluppo binario dei numeri compresi tra 1 e 7:

$$H = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$$

$$HN = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{matrix}$$

ovvero, le righe di HN corrispondono alle righe di H prese nel seguente ordine: $(1, 3, 2, 6, 7, 5, 4)$. Allora la permutazione cercata è

$$\mu = (2, 3)(4, 6, 5, 7)$$

quindi

$$M = \begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{matrix}$$

La permutazione μ corrisponde in maniera biunivoca ad M e N. Si ha quindi un isomorfismo tra le matrici monomiali che fissano il codice e le permutazioni che fissano le righe della matrice. L'ordine di un elemento è il minimo intero per cui $A^k = 1$. Poiché A corrisponde alla permutazione μ, l'ordine di A coincide con l'ordine di μ, che è 4 (minimo comune multiplo tra le lunghezze dei cicli).

Calcolando esplicitamente l'ordine di A

$$(A + id)^4 = id^4 + A^4 = 0$$

poiché $A^3 = 0$. □

Esercizio 3.13. Dato il codice di Hamming di parametri $2, 3$, dimostrare che se D è il duale di questo codice, ogni suo elemento ha peso 0, oppure 4, ovvero questo codice ha distribuzione costante dei pesi.

Dimostrazione. Considero come generatrice del codice duale la trasposta di H, ovvero

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$$

C è un (7,4,3)-codice mentre D è un (7,3,4)-codice, e ha cardinalità 2^3. Siano v_1, v_2, v_3 gli elementi della matrice generatrice, che hanno peso 4. Considero tutte le combinazioni lineari dei tre elementi e ne calcolo il peso. Una generica combinazione lineare è data da

$$\alpha v_1 + \beta v_2 + \gamma v_3$$

$$= (\gamma, \beta, \gamma + \beta, \alpha, \alpha + \gamma, \alpha + \beta, \alpha + \gamma + \beta)$$

Considero tutte le possibili scelte di $\alpha, \beta, \gamma \in F_2$:

- le combinazioni in cui un'unica componente della terna (α, β, γ) è diversa da 0 sono i vettori della matrice generatrice che hanno peso 4
- il multiinsieme delle componentidi una generica combinazione lineare è

  $$\{\alpha, \beta, \gamma, \alpha + \beta, \alpha + \gamma, \beta + \gamma, \alpha + \beta + \gamma\}$$

  e la somma di tali elementi determina il peso della combinazione lineare considerata. Il multiinsieme è invariante rispetto all'azione del gruppo che scambia tra loro α, β, γ. Allora il multiinsieme delle componenti relative alle terne $(1, 0, 1), (0, 1, 1), (1, 1, 0)$ rimane

uguale, e quindi anche il peso delle relative combinazioni lineari. Calcoliamo ad esempio quello della terna $(1,1,0)$:

$$\{(1,1,0),(0,0,1)(1)\}$$

e la somma è 4.

- Infine controllo la combinazione relativa alla scelta dei parametri $(1,1,1)$, e ottengo il vettore

$$(1,1,1,0,0,0,1)$$

che ha peso 4.

□

Esercizio 3.14. Dimostrare che D, duale del codice di Hamming con $q = 2, r = 3$ è autoortogonale.

Dimostrazione. Considero i prodotti tra i possibili elementi di D. Fare il prodotti di un vettore con sé stesso equivale a calcolarne il peso, che per l'esercizio precedente è $4 \equiv 0 \mod 2$. Mostro poi l'ortogonalità tra v_1, v_2, v_3. In generale, dati due vettori $v, u \in F_{2^n}$ con $n = 7$, si ha che $v \cdot u \equiv w_H(v \cap u) \mod 2$ dove

$$(v \cap u)_i = \begin{cases} 1 & \iff v_i = u_i = 1 \\ 0 & \text{altrimenti} \end{cases}$$

e questo segue dal fatto che

$$v \cdot u = \sum_j u_j v_j$$

e se u_j o v_j sono nulli, il fattore $v_j u_j$ è nullo. Qui

$$v_1 \cap v_2 = (0,0,0,0,0,1,1)$$

$$v_1 \cap v_3 = (0,0,0,0,1,0,1)$$

$$v_2 \cap v_3 = (0,1,1,0,0,0,0))$$

(questi non sono vettori del codice) Tali vettori hanno peso pari quindi tutti i prodotti scalari sono nulli. □

Proposizione 3.15. *Se D definito su F_2 è autoortogonale e ammette una base di vettori v_i i cui pesi sono congrui a 0 modulo 4, allora per ogni $v \in C$, $w_H(v) \equiv 0 \mod 4$.*

Considerato il duale del codice di Hamming, esso ha base che soddisfa questa condizione. Allora per questo risultato ogni vettore ha un peso multiplo di 4. Siccome la lunghezza del codice è 7, i pesi possono essere 0 o 4.

Esiste un teorema che descrive tutti e soli i codici di peso costante. Essi sono ripetizioni del duale di un codice di Hamming, ovvero dato un elemento che sta nel duale di un codice di Hamming, una ripetizione di questo codice è un codice la cui lunghezza è un multiplo t della lunghezza del codice di partenza.

23-11-2015

3.3 Risultati conclusivi sui codici di Hamming

Proposizione 3.16. *Sia C un codice definito su F_2 autoortogonale, e generato da vettori di peso 4, allora ogni elemento $v \in C$ è tale che il suo peso di Hamming sia un multiplo di 4.*

(abbiamo già dimostrato l'asserto per il codice di Hamming di parametri 2 e 3)
Per la dimostrazione serve il seguente

Lemma 3.17. *Sia D un codice su F_2 autoortogonale, e considero $x, y \in D$, allora*

$$w_H(x+y) = w_H(x) + w_H(y) - 2w_H(x \cap y)$$

dove

$$(x \cap y)_i = \begin{cases} 1 & \iff x_i = y_i = 1 \\ 0 & \textit{altrimenti} \end{cases}$$

Dato un campo F e uno spazio vettoriale di dimensione n su F, si dice *prodotto di Hadamard* il prodotto $*$ tale che

$$(a_1, a_2, \ldots, a_n) * (b_1, b_2, \ldots, b_n) = (a_1 b_1, a_2 b_2, \ldots, a_n b_n).$$

(è un prodotto tra liste)
Su F_2,

$$(x \cap y)_i = x * y$$

dimostrazione del lemma. Siano $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, allora

$$w_H(x+y) = \sum_{x_i + y_i = 1} (x_i + y_i)$$

$$= \sum_{\{x_i=1, y_i=0\} \cap \{x_i=0, y_i=1\}} (x_i + y_i)$$

Siano

$$X = \{i \, t.c. \, x_i = 1\}$$

$$Y = \{i \, t.c. \, y_i = 1\}$$

$$w_H(x) = \#X, \; w_H(y) = \#Y$$

$$w_H(X+Y) = \#X + \#Y - 2\#(X \cap Y)$$

(gli elementi di $X \cap Y$ vengono contati due volte) □

Lemma 3.18. *Sia D un codice su F_2, allora*

$$w_H(x \cap y) \equiv x \cdot y \mod 2.$$

Dimostrazione.

$$x \cdot y = \sum x_i y_i = \sum_{i \in X \cap Y} x_i y_i = \#(X \cap Y) = w_H(X \cap Y)$$

□

dimostrazione del risultato. La dimostrazione è per induzione sul numero di elementi della base di D che compaiono nella scrittura di v. Siano $\{v_1, \ldots, v_k\}$ i vettori di una base per D con la proprietà che $w_H(v_i)$ è un multiplo di 4. Sia $v \in D$, e sia s il numero di vettori nella base che sono coinvolti nella scrittura lineare di v come combinazione di vettori della base. L'asserto vale per $s = 0$, in cui si ottiene il vettore nullo, e per $s = 1$, in cui si ottiene un multiplo di un generatore che ha peso 4 per ipotesi. Sia $s = 2$, e $v = x + y$, allora per il primo lemma

$$w_H(x + y) = w_H(x) + w_H(y) - 2w_H(x \cap y) \equiv -2w_H(x \cap y) \mod 4$$

per l'ipotesi. Per il secondo lemma, $w_H(x \cap y) \equiv x \cdot y \mod 2$, allora

$$w_H(x + y) \equiv -2x \cdot y \equiv 0 \mod 4$$

dove si sfrutta l'autoortogonalità del codice.

Passo induttivo: supponiamo che l'asserto valga per $s - 1$ e lo dimostro per s. Sia $v = x_{i_1} + \cdots + x_{i_s}$. Poniamo $x = v_{i_1} + \cdots + v_{i_{s-1}}$, e $y = v_s$, allora per induzione $w_H(x) \equiv 0 \mod 4$, e per ipotesi $w_H(y) \equiv 0 \mod 4$, e poi si applica il ragionamento fatto per $s = 2$. □

3.3.1 Codici δ-divisibili

Definizione 3.19. Dato $\delta > 0$, sia C un codice definito su un campo finito F_q, C viene detto un *codice δ-divisibile* se per ogni $c \in C$, $\delta \mid w_H(c)$.

Il codice di Hamming di parametri $q = 2, r = 3$ è un codice 4-divisibile.

Definizione 3.20. Sia C un codice di Hamming qualsiasi di parametri q,r, allora il suo duale viene detto *codice simplesso* (symplex code).

Il (4,2,3)-codice di Hamming è autoduale, quindi è anche un symplex code.

Proposizione 3.21. *Considero un codice di Hamming e il suo duale D, allora ogni elemento $v \in D$ non nullo ha peso q^{r-1}.*

Se $q = 2, r = 3$, tutti i pesi sono congrui a $4 = 2^{3-1}$.

Dimostrazione. Considero una matrice generatrice G per il symplex code. Essa è una trasposta della matrice di controllo H per C, le cui righe sono date da tutti e soli i punti proiettivi in $P^{r-1}(q)$. Gli elementi di D sono della forma $x * G$, $x \in F^k$ con k dimensione del codice simplesso. Affermoche **il peso di Hamming di un generico elemento del codice è dato da $w_H(x) = n - s$ dove s è il numero di componenti nulle in xG.** Sia $x = 0$, allora $n = s$ e l'asserto vale. Sia $x \neq 0$, allora le colonne y^t di G che forniscono componente nulla sono colonne ortogonali a x. Notiamo che se $[y] = [z]$, allora $x \cdot y = 0$ se e solo se $x \cdot z = 0$ (i due prodotti differiscono per uno scalare). La dimensione dell'ortogonale di x è $r - 1$. Cerco punti proiettivi che provengono da vettori nello spazio ortogonale alla retta generata da x. Tale numero è $s = \frac{q^{r-1}-1}{q-1}$. Inoltre

$$n = \frac{q^r - 1}{q - 1}$$

Allora

$$n - s = \frac{q^r - 1}{q - 1} - \frac{q^{r-1} - 1}{q - 1} = \frac{q^r - q^{r-1}}{q - 1}$$

$$= q^{r-1}\frac{q-1}{q-1} = q^{r-1}.$$

□

Esistono forti restrizioni su codici δ-divisibili di dimensione pari.

Teorema 3.22. *Sia C un (2m,m)-codice su F_q δ-divisibile, allora $q, \delta = 2, 3, 4$.*

3.3.2 m-ripetizione

Definizione 3.23. Dato un codice C e un intero m, si chiama *m-ripetizione* di C il codice $\mathbf{Rep}_m(C)$ definito come l'insieme delle liste le cui componenti sono vettori del codice ripetuti m volte. In particolare, gli elementi di $\mathbf{Rep}_m(C)$ hanno lunghezza mn.

Il "vecchio" codice di ripetizione si ottiene quando $n = 1$.

Osservazione 3.24. Se indico con c^m la lista delle componenti di c ripetute m volte, segue che il peso di Hamming si comporta come una funzione esponenziale, ovvero $w_H(c^m) = m * w_H(c)$. Quindi, se C è un codice a peso costante, lo stesso avviene per $\mathbf{Rep}_m(C)$.

Considero $c \in C$ di peso w, vogliamo costruire una nuova parola $(\hat{c}_1, \ldots, \hat{c}_m)$ con $c_i \in F^n$, con $w_H(\hat{c}) = mw$. Se considero m endomorfismi $\phi_1, \ldots, \phi_m$ che preservano i pesi e chiamo $\hat{c} = (\phi_1(c), \ldots, \phi_m(c))$, segue che $w_H(c) = w_H(\hat{c})$. Il teorema di McWilliams afferma che le mappe ϕ_i sono (restrizioni di) mappe monomiali. In particolare, se C ha peso costante,

$$\hat{C} = \{(\phi_1(c), \ldots, \phi_m(c)),\ c \in C\}$$

ha peso costante. $\hat{C}$ viene comunque detto una m-ripetizione di C.

Teorema 3.25. *Sia C un codice lineare di peso costante su un campo fissato F_q, allora C è una m-ripetizione del duale di un codice di Hamming.*

Dimostrazione. Sia G una matrice generatrice per il codice, e siano $g_1^t, \ldots, g_n^t \in F^k$ le sue colonne (con k dimensione del codice). Per ogni classe di equivalenza, conto il numero di elementi che compaiono come colonne della matrice. Sia g^t una colonna di G, che compare con i suoi equivalenti il massimo numero di volte M. Sia $x \neq 0$ una colonna, allora esiste $B \in GL(k, F)$ tale che $Bg^t = x^t$. Allora, $G' = BG$ è ancora una matrice generatrice del codice di partenza e contiamo quante volte contiene $x^t = Bg^t$ o i suoi equivalenti. Se esiste $y^t = \lambda x^t$, si ha $y^t = \lambda Bg^t$, quindi $y^t = Bh^t$ con $h^t = \lambda g^t$. Allora h^t è una colonna di $G = B^{-1}BG$, quindi x^t e i suoi equivalenti compaiono esattamente M volte.

Per il teorema di Mcwilliams, dato un codice lineare invertibile che preserva i pesi, allora $\phi : C \to C$ si estende a una mappa $\hat{\phi}$ monomiale da F^n in sé. La mappa è lineare, è invertibile, preserva i pesi (siccome i pesi degli elementi di C non nulli sono costanti, la trasformazione $\phi : xG \mapsto xBG$ li preserva). Allora per il teorema di Mcwilliams ϕ è una trasformazione monomiale, ovvero esiste M monomiale di ordine n tale che $BG = GM$. Le righe della matrice GM sono $r_1M, \ldots, r_kM$, con $r_1, \ldots, r_k$ righe di G. La matrice GM ha colonne che sono una permutazione delle colonne di G, $(...g_i^t...)M = (...\lambda_i g_{\pi(i)}^t$. G e G' hanno la stessa distribuzione di occorrenze di un dato punto proiettivo tra le colonne, e vale la relazione

$$n_v(G) := \#\{it.c.[g^i] = [v]\} = n_v(GM)$$

(inoltre dato un vettore esso compare tante volte quanto il suo equivalente) Per le osservazioni iniziali, $N_v(G) = N_{Bv}(G)$, allora tutti i punti proiettivi di $P^{k-1}(q)$ compaiono esattamente M volte come colonne di G. Se chiamo $N = \frac{q^k-1}{q-1}$, posso riordinare le colonne di G nel modo seguente

$$[\lambda_1 g_1^t..g_N^t], [\lambda_2 g_1^t, ...g_N^t], ...[\lambda_m g_1^t, ...g_N^t]$$

Questa è la matrice generatrice del duale di un codice di Hamming di parametri q,k. A meno di scambiare le posizioni di colonna, e di moltiplicare per opportuni scalari, la matrice generatrice G è la ripetizione m volte di una matrice generatrice G_0, dove G_0^t è la matrice di controllo per un codice di Hamming di parametri q,k. □

Conclusione: dato un codice di peso costante, esso è monomialmente equivalente alla m-ripetizione classica del duale di un codice lineare.

Esercizio 3.26. Analizzare il codice di Hamming di parametri 3,2 e mostrare che tutte le parole del duale hanno peso uguale a 3.

Dimostrazione. Una matrice generatrice del duale è data da:

$$\begin{matrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{matrix}$$

I generatori $v_1 = (0, 1, 1, 1)$ e $v_2 = (1, 0, 1, 2)$ hanno peso 3. Tutti gli elementi del codice sono della forma:

$$\alpha v_1 + \beta v_2 = (\beta, \alpha, \alpha + \beta, \alpha + 2\beta$$

Considero le varie possibilità:

(i) in corrispondenza delle coppie (0,1) e (1,0) si ottengono i generatori del codice, che hanno peso 3 per quanto detto prima.

(ii) se $\alpha = 0, \beta = 2$

$$w = (\beta, \alpha, \alpha + \beta, \alpha + 2\beta = (1, 0, 2, 1)$$

che ha peso 3.

(iii) Se $\alpha = 2, \beta = 0$,

$$v = (0, 2, 2, 2)$$

che ha peso 3

(iv) se $\alpha = \beta = 1$

$$v = (1, 1, 2, 0)$$

che ha peso 3.

(v) Se $\alpha = 2, \beta = 1$

$$v = (1, 2, 0, 1)$$

che ha peso 3

(vi) Se $\alpha = 1, \beta = 2$

$$(2, 1, 0, 2)$$

che ha peso 3.

(vii) Infine, se $\alpha = \beta = 2$

$$v = (2, 2, 1, 0)$$

che ha peso 3.

□

Capitolo 4

CAMPI

Finora abbiamo introdotto codici con capacità correttiva uguale a 1, cerchiamo invece codici con una capacità correttiva più elevata.

4.1 Ripasso sulla fattorizzazione di polinomi a coefficienti su un campo finito

Siamo interessati a fattorizzare sul campo F_p il polinomio $f(x) = x^n - 1$, con n lunghezza del codice, nell'ipotesi che $p \nmid n$.

Valgono i seguenti fatti:

(i) Se F è finito, allora $o(F) = p^d$ con p primo e d intero.

(ii) il gruppo moltiplicativo di un campo è ciclico.

Dimostrazione. La dimostrazione è basata sul teorema di struttura dei gruppi abeliani: ogni gruppo abeliano finitamente generato è prodotto di gruppi ciclici. Per questo motivo ci si può ridurre al caso in cui G, gruppo moltiplicativo del campo, abbia ordine una potenza di un primo, p^s.
Dato un polinomio $x^d - 1$, esso ammette al più d radici, e il numero di elementi del campo in cui il polinomio ha coefficienti è $\leq d$ (applicazione ripetuta del teorema di Ruffini). Sia

$$\psi_d = \#\{g \in G \, t.c. \, o(g) = d\}$$

Sicuramente $\psi_d = 0$ se $d \nmid n$ (Infatti l'ordine di un elemento del gruppo deve dividere l'ordine del gruppo), ma a priori si potrebbe avere $\psi_d = 0$ anche in qualche altro caso. Quando invece $\psi_d \neq 0$, segue che $\psi_d = \varphi(d)$: infatti, in tal caso esiste g elemento di ordine d, allora $< g >$ ha d elementi $\{1, g, g^2, \ldots, g^{d-1}\}$, e ognuno di questi elementi è tale che $(g^i)^d = 1$ ovvero è una radice del polinomio $x^d - 1$. Poiché $< g >$ ha d elementi e $x^d - 1$ ha al più d radici, l'insieme delle radici di $x^d - 1$ coincide esattamente con $< g >$. Un elemento della forma g^j ha ordine $\frac{d}{M.C.D.(d,j)}$, allora gli elementi di $< g >$ di ordine d sono $\varphi(d)$ (sono gli elementi per cui $M.C.D.(d, j) = 1$).
Da quanto appena mostrato segue che i termini $\psi_d - \text{Var}\phi(d)$ sono ≤ 0.
Poiché dato un elemento di ordine n le sue potenze hanno ordine d con d divisore di n, segue che

$$\sum_{dt.c.d|n} \psi_d = n.$$

Inoltre per la funzione di Eulero vale la formula:

$$\sum_{dt.c.d|n} \varphi(d) = n.$$

Sottraggo termine a termine le due espressioni, allora

$$\sum_{d|n} (\psi_d - \varphi(d)) = 0$$

L'unica possibilità è che i termini, che a priori sono ≤ 0, siano tutti nulli, ovvero $\psi_d = \varphi(d) \forall d$ (ovvero l'ipotesi $\psi_d \neq 0$ vale per ogni divisore d di n). In particolare $\psi_n =$

$\varphi(n)$ (infatti $n \mid n$), allora ψ_n è positivo, ovvero esiste almeno un elemento di ordine n che genera G. Gli elementi che soddisfano la condizione $o(g) = n$ sono detti *elementi primitivi.* □

(iii) Il gruppo degli automorfismi di un campo F di ordine p^d è generato da un unico automorfismo di ordine d, detto *mappa di Frobenius* e definito ponendo $\phi_p : f \mapsto f^p$.

Dimostrazione. Mostro innanzitutto che ϕ_p **è un automorfismo**: preserva il prodotto e la somma, infatti, dati $\alpha, \beta \in F$

$$(\alpha\beta)^p = \alpha^p\beta^p$$

$$(\alpha+\beta)^p = \sum_{i=1}^{p} \binom{p}{i} \alpha^i \beta^{p-i} = \alpha^p + \beta^p$$

perché i coefficienti binomiali intermedi sono congrui a 1 modulo p. Inoltre **la mappa è biunivoca** perché $\alpha^p = 0$ implica $\alpha = 0$ per la legge di annullamento del prodotto in un campo. Affermo che $o(\phi) = d$, infatti, dato $\alpha \in F$ e $q = p^d$, se $\alpha = 0$, $\alpha^q = \alpha$. Se $\alpha \neq 0$, $\alpha^{q-1} = 1$ per il teorema di Lagrange, allora $\alpha^q = \alpha \forall \alpha$, quindi $\phi_p^d(\alpha) = \alpha$ ($\phi_p^d(\alpha)$, che è la d-esima potenza di ϕ_p, è dato da $\phi_p^d(\alpha) : \alpha \mapsto \alpha^{p^d}$). Abbiamo allora provato che $o(\phi_p) \leq d$. Sia $m < d$, e supponiamo che $\phi_p^m = id_F$, allora $\alpha^{p^m} = \alpha \forall \alpha \in F$, segue allora che ogni α è una radice del polinomio $f(x) = x^{p^m} - 1$. Questo però non può avvenire, perché tale polinomio ha al più p^m radici mentre gli elementi di F sono $p^d > p^m$.
Concludo che **il gruppo ciclico generato da ϕ_p coincide con gli automorfismi che fissano** F_p: infatti, dato un campo F con un sottocampo F_p, gli automorfismi di F che fissano tutti gli elementi di F_p sono al più $\mathrm{Dim}_{F_p} F = d$. Siccome il gruppo ciclico $< \phi_p >$ ha esattamente d elementi, segue che i due gruppi coincidono. □

(iv) F_q è il campo con il minor numero di elementi in cui $x^q - x$ si fattorizza completamente, ovvero in cui $x^q - x$ si può scrivere come il prodotto di fattori di grado 1 della forma $x - \alpha$. Si dice che F_q è un *campo di spezzamento* per $x^q - x$. Tale campo è unico a meno di isomorfismi

(v) Descriviamo il reticolo dei sottocampi di $F = F_q$. Ogni sottocampo L di F_q è costruito come uno spazio vettoriale su F_p, pertanto ha ordine $p^t < q$, ovvero $t \leq d$. In realtà mostriamo che t dev'essere un divisore di d.

Dimostrazione. Una prima dimostrazione segue immediatamente dal teorema del grado, ovvero $\mathrm{Dim}_{F_p} F = \mathrm{Dim}_L F * \mathrm{Dim}_{F_p} L$. Alternativamente, divido d per t e si può scrivere $d = st + r$. Allora

$$\frac{p^d - 1}{p^t - 1} = \frac{p^{t*s+r} - 1}{p^t - 1}$$

$$= \frac{p^{t*s+r} - p^r + p^r - 1}{p^t - 1}$$

$$= p^r * \frac{p^{ts} - 1}{p^t - 1} + \frac{p^r - 1}{p^t - 1}$$

Se L è un sottocampo di F, questa quantità dev'essere un intero. Poiché $0 \leq r < t$, l'unica possibilità per cui il secondo addendo sia un intero è che $r = 0$, ovvero $t \mid d$.

Viceversa, dato un divisore della dimensione di F, esiste unico un sottocampo L di F avente ordine p^t. □

26-11-2015

4.1.1 Elementi primitivi

Facendo riferimento al fatto che il gruppo moltiplicativo di un campo è ciclico, un qualsiasi generatore di G viene detto un elemento primitivo. Inoltre se $K \geq F$ $\gamma \in K$ si dice primitivo se $K = F[\gamma]$.

Proviamo che **un elemento primitivo nel primo senso è primitivo nel secondo senso**, ovvero che

$$F \geq F_p,\ G = F^* =< \gamma > \longrightarrow\ F = F_p[\gamma].$$

Dimostrazione. Per definizione

$$F = \{0\} \cup G = \{0\} \cup \{\gamma^i : i \in \mathbb{Z}\}.$$

Basta notare che si può sempre scrivere $0 = 0(\gamma)$, $\gamma^i = x^i(\gamma)$ con $x^i \in F_p[x]$ e questo prova l'asserto. □

Cerchiamo ora elementi primitivi in maniera probabilistica. È noto che se $|F| = q$ esistono $\varphi(q-1)$ elementi primitivi in F (si esclude lo zero).

Esempio 4.1. Se $q = 13$, $q - 1 = 12 = 2^2 \cdot 3$. Ricordando che φ è debolmente moltiplicativa ($\varphi(ab) = \varphi(a)\varphi(b)$ se $M.C.D.(a,b) = 1$ e $\varphi(p^a) = p^{a-1}(p-1)$) risulta $\varphi(12) = \varphi(4)\varphi(3) = 4$. Di conseguenza prendendo un elemento in F^* vi è un terzo della probabilità che esso sia primitivo (gli elementi primitivi sono 4 su 12). La probabilità del complementare è data da:

$$\mathrm{P}(\alpha \in G | \alpha \text{ non primitivo}) = \frac{2}{3}.$$

Se $\alpha = 2$, $o_{13}(2) \in \{1,2,3,4,6,12\}$ (dev'essere un divisore di 12). Facendo esplicitamente i calcoli:

$$2^2 = 4,\ 2^4 \equiv_{13} 3 \not\equiv_{13} 1$$

$$2^6 = 2^4 * 2^2 \equiv_{13} 3 \cdot 4 \equiv_{13} 12 \not\equiv_{13} 1$$

Questo basta a provare che 2 è primitivo, questo fatto è giustificato dal teorema seguente.

Teorema 4.2. *$\gamma \in F_q^*$ è un elemento primitivo per F_q se e solo se, detto $q - 1 = p_1^{a_1} \cdots p_r^{a_r}$, con p_i primi distinti, segue che $\gamma^{\frac{q-1}{p_i}} \not\equiv 1$ in F_q, per ogni i.*

(facendo riferimento all'esempio precedente, $12 = 3 * 2^2$, perchò per mostrare che 2 è primitivo è bastato mostrare che $2^{12/2} = 2^6$ e $2^{12/3} = 4$ non sono congrui a 13)

Congettura di Artin: Sia $\gamma \in \mathbb{Z}$ diverso da una potenza propria, allora esistono infiniti primi p tali che$F_p^* =< \gamma >$. *Congettura di Riemann*: Dato p primo esiste γ primitivo con $\gamma \leq C(\log_2 p)^2$.

Osservazione 4.3. Sia A un gruppo abeliano e $a, b \in A$ con $m = o(a), n = o(b)$ tali che $M.C.D.(m,n) = 1$. Allora $o(ab) = nm$.

Dimostrazione. Sicuramente

$$(ab)^{mn} = (a^m)^n (b^n)^m = 1.$$

Sia ora $t = o(ab)$, per il calcolo precedente $t \mid mn$. Supponiamo per assurdo che $t \neq mn$, ovvero t è un divisore proprio di mn, allora per un certo p divisore di mn deve valere $t = \frac{mn}{p}$. Supponiamo che $p \mid m$ (quindi $p \nmid n$ essendo $M.C.D.(m,n) = 1$). Allora

$$(ab)^{\frac{mn}{p}} = (a^{\frac{m}{p}})^n \cdot (b^n)^{\frac{m}{p}} = (a^{\frac{m}{p}})^n.$$

Detto $c = a^{\frac{m}{p}}$, $o(c) = p$. Poiché $M.C.D.(n,p) = 1$, $\exists u, v \in \mathbb{Z}$ tali che $nu + vp = 1$ (identità di Bezout). Allora

$$c = c^1 = c^{nu+vp} = (c^n)^u (c^p)^v = (c^n)^u$$

Dev'essere allora $c = 1$, infatti se $c^n \neq 1$ si ottiene l'assurdo (infatti $c^n = (a^{m/p})^n = 1$ per ipotesi assurda). □

Esercizio 4.4. Determinare l'elemento primitivo di F_7.

Dimostrazione. Un elemento γ in F_7 dev'essere tale che $o_7(\gamma) \in \{1, 2, 3, 6\}$. Sia $\gamma = 5$:

$$5 \equiv_7 12 \equiv_7 2 \cdot 6.$$

6 ha ordine 2 ($6 \equiv_7 -1$) e $o_7(2) = 3$, allora applicando il teorema precedente con $a = 2, b = 6, m = 3, n = 2$ segue che $o_7(5) = 2 * 3 = 6$, e quindi la classe del 5 è un elemento primitivo per F_7 (infatti $6 = 2 * 3$ e avviene che $5^{6/2} \neq 1$ e $5^{6/3} \neq 1$). □

4.1.2 Elementi algebrici

Definizione 4.5. Dati due campi $K \geq F$ $\alpha \in K$ si dice *algebrico* su F se $\exists g(x) \in F[x]$, $g(x) \neq 0$, tale che $g(\alpha) = 0$.

Esercizio 4.6. Dimostrare che se K è un campo finito con $|K| = q = p^d$, ogni elemento $\alpha \in K$ è algebrico su $F = F_p$.

Dimostrazione. Osserviamo che K ha la struttura di $F-$spazio vettoriale, con $\mathrm{Dim}_F K = d$. Sia $\alpha \in K$, gli elementi $1, \alpha, ..., \alpha^d$ sono sicuramente dipendenti perché sono $d+1$ elementi in uno spazio di dimensione d. Esistono quindi $g_0, ..., g_d \in F$ non tutti nulli tali che

$$g_0 + g_1\alpha ... + g_d\alpha^d = 0_K.$$

Preso quindi $g(x) = \sum_{i=0}^{d} g_i x^i$, $g(\alpha) = 0$, ovvero α è radice di un polinomio a coefficienti in F_p. □

Esercizio 4.7. Sia K finito tale che $|K| = q = p^d$. Allora $Aut(K) =< \varphi >$, con $\varphi : \alpha \mapsto \alpha^p$.

Dimostrazione. Sicuramente $Aut(K) \geq < \varphi >$, $| < \varphi >= d|$. Sia $\tau \in Aut(K)$, osserviamo che dato γ elemento primitivo per K, $K^* =< \gamma >$. Ogni $\alpha \in K = F[\gamma]$ si può scrivere come $\alpha = a(\gamma),$, con $a(x) \in F[x]$ della forma

$$a(x) = \sum_{i=0}^{n} a_i x^i,\ a_i \in F = F_p.$$

Allora:

$$\tau(\alpha) = \tau(a(\gamma)) = \tau(\sum_i a_i\gamma^i) = \sum_i \tau(a_i\gamma^i) = \sum_i \tau(a_i)\tau(\gamma)^i.$$

Osserviamo che dato $b \in F = F_p$: $\tau(b) = b$ infatti, ponendo $b = b \cdot 1_F$:

$$\tau(b) = \tau(b \cdot 1_F) = b \cdot \tau(1_F) = b \cdot 1_F = b.$$

Segue quindi che la relazione sopra si riscrive come

$$\tau(\alpha) = \sum_i a_i\tau(\gamma)^i.$$

ovvero $\tau(\alpha) = a(\tau(\gamma))$.

Notiamo che γ è algebrico su F e esiste un polinomio $g(x)$ di grado minore o uguale a d che viene annullato da γ. Applicando il precedente ragionamento a $\alpha = 0$ (e sostituendo g con a) segue che

$$0 = \tau(0) = g(\tau(\gamma)),$$

quindi $\tau(\gamma)$ deve essere radice di $g(x)$, che ha al più d radici. Quindi $|Aut(K)| \leq d = | < \varphi > |$, e quindi la tesi. □

Esercizio 4.8. Determinare un campo di ordine 8 e descrivere i suoi elementi pritivi.

Dimostrazione. L'idea è di costruire K come quoziente della forma $\frac{F_2[x]}{g(x)F_2[x]}$ (dove $g(x)F_2[x]$ è l'ideale generato da $g(x)$). Sia $I = g(x)F_2[x]$ e $R = F_2[x]$, Mostriamo che $R/I = |K| = 2^d$, con $d = \mathrm{Gr}(g(x))$. Infatti

$$R/I = \{h + I : h \in R\},$$

con $h(x) = g(x)q(x) + r(x)$ con $\mathrm{Gr}r(x) < d$ ed $h + I = gq + r + I = r + I$.

Inoltre, $r(x)$ è l'unico polinomio di grado $< d$ che vive in $H + I$. Se infatti ci fossero due polinomi di grado $< d$, r_1 e r_2 in $H + I$, segue che $r_1 - r_2 \in I = (g(x))$, allora dev'essere $r_1 - r_2 = p(x) * g(x)$, ma questo è assurdo perché $\mathrm{Gr}g(x) > d$ mentre $\mathrm{Gr}(r_1 - r_2) < d$, allora l'unica possibilità è che sia $r_1 - r_2 = 0$.

In questo modo

$$F^d \cong \{r(x) = r_0 + r_1x + ... + r_{d-1}x^{d-1}\}$$

Per queste biezioni $|R/I| = |F^d|$.

$K = R/I$ è un campo se I è massimale, e questo avviene se $g(x)$ è irriducibile. Determiniamo un polinomio irriducibile in $\mathbb{Z}_3$. Un generico polinomio ha la forma $g(x) = x^3 + ax^2 + bx + c$. Se g è irriducibile sicuramente $c = 1$ altrimenti x sarebbe un divisore. Poichè un polinomio di terzo grado è irriducibile se e solo se non ha radici nel campo, dobbiamo fare in modo che $g(1) = a + b \neq 0$ e ciò avviene per $a = 1,\ b = 0$ o $a = 0,\ b = 1$. Preso quindi $g(x) = x^3 + x^2 + 1$, esso è irriducibile e quindi $K = \frac{F_2[x]}{g(x)F_2[x]}$. Ogni elemento di $K = R/I$ è della forma $r(x) + I$ con $\mathrm{Gr}(r) < 3$, e quindi ha la forma

$$r_0 + r_1x + r_2x^2 + I = r_0(1 + I) + r_1(x + I) + r_2(x + I)^2 = r_0 + r_1\alpha + r_2\alpha^2$$

(osservando che $x + I = \alpha$). Gli elementi sono del campo sono $0, 1, \alpha, \alpha + 1, \alpha^2, \alpha^2 + 1, \alpha^2 + \alpha, \alpha^2 + \alpha + 1$. Poichè $\varphi(7) = 6$ tutti gli elementi diversi da $0, 1$ sono primitivi. □

Esercizio 4.9. Determinare una fattorizzazione del polinomio di grado 3, $g(x) = x^3 + x^2 + 1$ di cui α è radice.

Dimostrazione. Per quanto dimostrato precedentemente, se α è radice di un polinomio g, dato $\tau \in Aut(K) =< \phi >$, segue che $\tau(\alpha)$è radice di g. Qui $\phi = \alpha \mapsto \alpha^2$ (infatti $p = 2$),

$$< \phi >= \{\alpha \mapsto \alpha^2; \alpha \mapsto \alpha^4; id\}$$

quindi $\tau(\alpha) \in \{\alpha, \alpha^2, \alpha^4\}$, quindi

$$x^3 + x^2 + 1 = (x - \alpha)(x - \alpha^2)(x - \alpha^4).$$

Controlliamo esplicitamente che $(x-\alpha)(x-\alpha^4)(x-\alpha^2) = g(x)$ e che $\alpha^4 \in r(\alpha)$, con $\mathrm{Grr}(x) < 3$.

Divisione 1:

$$\frac{x^3 + x^2 + 1}{x - \alpha} =$$

$$q_1 = x^2,\ r_1 = x^3 + x^2 + 1 - x^2(x - \alpha) = (1 + \alpha)x^2 + 1$$

$$q_2 = (1 + \alpha)x,\ r_2 = (1 + \alpha)x^2 + 1 - (1 + \alpha)x(x - \alpha) = (\alpha + \alpha^2)x + 1$$

$$q_3 = \alpha + \alpha^2,\ r_3 = (\alpha + \alpha^2)x + 1 - (\alpha + \alpha^2)(x - \alpha) = 1 + \alpha^2 + \alpha^3 = 0$$

Quindi

$$g(x) = (x - \alpha)(x^2 + (1 + \alpha)x + \alpha + \alpha^2)$$

Divisione 2:

$$\frac{x^2 + (1 + \alpha)x + \alpha^2 + \alpha}{x - \alpha^2} =$$

$$q_1 = x,\ r_1 = (1 + \alpha)x + \alpha^2 + \alpha + x\alpha^2 = (1 + \alpha + \alpha^2)x + \alpha^2 + \alpha$$

$$q_2 = 1 + \alpha^2 + \alpha,\ r_2 = \alpha^2 + \alpha + \alpha^2(1 + \alpha^2 + \alpha) = \alpha^2 + \alpha + \alpha^2 + \alpha^4 + \alpha^3 = \alpha(\alpha^3 + \alpha^2 + 1) = 0$$

quindi

$$g(x) = (x + 1 + \alpha^2 + \alpha)(x - \alpha)(x - \alpha^2)$$

Infine

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

$$q_1 = x,\ r_1 = x^4 - x(x^3 + x^2 + 1) = x^3 + x$$

$$q_2 = 1,\ r_2 = x^3 + x - x^3 - x^2 - 1 = x^2 + x + 1$$

quindi il polinomio che rappresenta x^4 è $x^2 + x + 1$, e la condizione

$$g(x) = (x - \alpha)(x - \alpha^2)(x - \alpha^4)$$

è verificata. □

30-11-2015

4.1.3 Sottogruppi di $Aut(E)$

Se $F = F_p$ e E è un campo di cardinalità p^n che lo contiene, allora $Aut(E)$ coincide con $< \phi >$, dove $\phi : e \mapsto e^p$. $A = Aut(E)$ ha esattamente n elementi.

Sia C_F il centralizzante (stabilizzatore) dell'azione di A su E, tale che $(\varepsilon, \alpha) \mapsto \alpha(\varepsilon)$, dove $\alpha \in A$, $\varepsilon \in F$. Dato $\alpha \in A$,

$$\alpha(f) = \alpha(f * 1_E) = f\alpha(1) = f$$

ovvero ogni elemento del sottocampo viene fissato dagli elementi di A. A è generato dall'automorfismo di Frobenius, e da un punto di vista astratto è isomorfo a $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Per descrivere il reticolo dei sottogruppi di A, fissato un divisore d, esiste unico un sottogruppo che ha come ordine quel divisore (questo avviene perché il gruppo di partenza è ciclico). Per trovare un gruppo di ordine d, devo quindi cercare una potenza di ϕ che abbia come ordine d e che lo generi.

Il reticolo dei sottogruppi di A è isomorfo al reticolo dei divisori di n, e in tale corrispondenza vengono preservate le inclusioni.

Abbiamo dimostrato che dato un campo intermedio L tra E e F, segue che $|L| = p^s$, con $s \mid n$. Mostreremo che L **è l'insieme dei punti fissati da un elemento di** A.

Osservazione 4.10. In analogia con quanto avviene per i gruppi ciclici, preso un divisore d di n, esiste un unico sottocampo che ha come cardinalità $p^{n/d}$. Abbiamo già dimostrato l'esistenza di tale campo (è il campo di spezzamento di $f(x) = x^{p^{n/d}} - x = x^{q_0} - x$, le radici di questo polinomio sono un sottocampo di E)

Viceversa, dato un sottocampo di ordine q_0 in E, esso deve contenere le radici del polinomio $x^{q_0} - x$, ma questo implica che L è l'unico sottocampo di E con cardinalità uguale a q_0.

Teorema 4.11. *Dato E di ordine p^n, esiste ed è unico un sottocampo L di E di ordine p^d con d divisore di n (il rapporto tra il gruppo totale e quello di ordine d e tra il campo totale e quello di ordine p^d è lo stesso a livello di "distanza"). In particolare, se L ha ordine p^t, L è ottenuto come insieme dei punti $\lambda \in E$ tali che $\lambda^{p^t} = \lambda$, ovvero è l'insieme degli elementi fissati da ϕ^t, potenza t-esima della mappa di Frobenius e quindi da tutto il sottogruppo generato da quest'ultima.*

Notiamo che la t-esima potenza di Frobenius ha ordine $\frac{n}{M.C.D.(n,t)}$.

L'insieme dei punti di E fissati da un gruppo di automorfismi è sempre un sottocampo di E per ogni divisore t di n.

4.2 Polinomi minimi

Sia ε algebrico su F. Osserviamo che $F[x]$ è un dominio a ideali principali, e l'insieme

$$\mathrm{Ann}_{F[x]}\varepsilon = \{g(x) \in F[x] \, t.c. \, g(\varepsilon) = 0\}$$

è un ideale. Allora esiste unico un polinomio monico $g(x)$, che indico con $m_\varepsilon(x)$, che genera tale ideale, esso è il *polinomio minimo* di ε. $m_\varepsilon(x)$ divide ogni polinomio a coefficienti in F che si annulla in ε.

Il polinomio minimo è irriducibile (altrimenti avrebbe un fattore, che si annulla ancora in ε), ed essendo un generatore ha grado minimo. Ci poniamo l'obiettivo di determinare la struttura del polinomio minimo.

Teorema 4.12. *Sia $m(x)$ il polinomio minimo, di ε (quindi $m(x)$ si annulla in ε); allora anche $m(\phi^i(\varepsilon)) = 0$.*

Dimostrazione. Applicando ϕ all'espressione

$$0 = m(\varepsilon)$$

si ha

$$\phi(0) = 0 = \phi(m(\varepsilon)).$$

Sia $m(x) = \sum_{i=0}^{t} m_i x^i$, allora, poiché ϕ è un automorfismo, commuta con la sommatoria, quindi

$$\phi(m(\varepsilon)) = \sum_{i=1}^{n} \phi(m_i \varepsilon^i)$$

$$= \sum_{i=1}^{t} \phi(m_i)(\phi(\varepsilon))^i$$

e poiché gli m_i vengono fissati:

$$= \sum_{i=1}^{t} m_i(\phi(\varepsilon))^i = m(\phi(\varepsilon))$$

allora

$$0 = m(\phi(\varepsilon))$$

allora iterando il procedimento segue che $\phi(\varepsilon), \phi^2(\varepsilon), \ldots, \phi^i(\varepsilon)$ sono radici del polinomio minimo. □

Definizione 4.13. Gli elementi $\phi^i(\varepsilon)$ vengono detti *coniugati* di ε.

4.2.1 Fattorizzazione del polinomio minimo

Mostriamo che **i coniugati sono tutte e sole le radici del polinomio minimo**.

Osservazione 4.14. Sia t il minimo intero tale che $\varepsilon = \phi^t(\varepsilon)$ (al massimo $t = n$). Allora $< \phi^t t >$ è contenuto nello stabilizzante in A di ε (tutti gli automorfismi di A che fissano ε), che indico con $C_A(\varepsilon)$. t dev'essere un divisore di n perché il numero di elementi nell'orbita di ε è uguale all'indice nel gruppo degli automorfismi del centralizzante di ε. Il centralizzante deve avere n/t elementi.

Teorema 4.15. *Il polinomio minimo su F di ε è prodotto di $x - \phi^i(\varepsilon)$ con $i = 0, \ldots, t-1$.*

Dimostrazione. Sappiamo già che $x - \phi^i(\varepsilon)$ sono fattori del polinomio minimo (per il teorema 4.12), Dobbiamo però mostrare che i coefficienti del polinomio

$$\prod_{i=0}^{t-1}(x - \phi^i(\varepsilon))$$

stanno in F. Estendo l'automorfismo $\phi : E \to E$ a $\bar{\phi} : E[x] \to E[x]$ ponendo

$$\bar{\phi}(\sum_i f_i x^i) = \sum_i \phi(f_i) x^i.$$

Tutti e soli i polinomi che vengono fissati da $\bar{\phi}$ sono i polinomi a coefficienti in F_p: ovviamente tali polinomi vengono fissati da $\bar{\phi}$, viceversa, sia $f(x)$ un polinomio a coefficienti in E fissato da ϕ, e supponiamo che sia fissato da $\bar{\phi}$, mostro allora che i suoi coefficienti f_i stanno in F. Deve valere

$$\phi(f) = f \longrightarrow \sum_i \phi(f_i) x^i = \sum_i f_i x^i$$

e quindi $\phi(f_i) = f_i \forall i$, ovvero, usando la definizione della mappa di Frobenius, $f_i^p = f_i$, ovvero sono le radici di $x^p - x$ (l'insieme di queste radici è l'unico sottocampo di ordine p), e quindi che stanno in F_p.

Sia $f(x)$ come nell'enunciato, ovvero

$$f(x) = \prod_{i=0}^{t-1} (x - \phi^i(\varepsilon))$$

e mostriamo che tale polinomio viene fissato da $\bar{\phi}$:

$$\bar{\phi}(f(x)) = \prod_{i=0}^{t-1} \bar{\phi}(x - \phi^i(\varepsilon))$$

$$= \prod_{i=0}^{t-1} (x - \phi^{i+1}(\varepsilon)) = f(x)$$

(le radici sono solo riordinate in un altro modo nei due polinomi)

Poiché $\phi(f(x)) = f(x)$, il polinomio sta in $F[x]$ per il punto precedente. Tale polinomio è il polinomio minimo perché è monico, si annulla in ε, ha coefficienti in $F[x]$ e ha grado minimo. □

In particolare, il grado del polinomio minimo è t, con t numero di elementi che stanno nell'orbita di ε rispetto agli automorfismi ϕ^i.

Riprendiamo il seguente esercizio:

Esercizio 4.16. Costruire una copia del campo con 8 elementi considerando il quoziente $\frac{F_2[x]}{(x^3+x+1)}$.

Dimostrazione. Posto $\varepsilon = x + I$ con $I = (x^3 + x + 1)$, si mostra che $\varepsilon^3 + \varepsilon + 1 = 0$. Il polinomio minimo su F di ε è proprio $f(x)$ (è stato scelto irriducibile). Calcoliamo il prodotto:

$$= \prod_{i=0}^{2} (x - \phi^i(\varepsilon)) = (x - \varepsilon)(x - \varepsilon^2)(x - \varepsilon^4)$$

$$= x^3 + \varepsilon(1 + \varepsilon + \varepsilon^3)x^2 + (\varepsilon^3 + \varepsilon^5 + \varepsilon^6)x + \varepsilon^7$$

Riduciamo al rappresentante canonico quest'espressione:

$$= x^3 + \varepsilon(1 + \varepsilon + \varepsilon^3)x^2 + \varepsilon^3(1 + \varepsilon^2 + \varepsilon^3)x + \varepsilon^7$$

Allora:

- il coefficiente di x è nullo poiché $1 + \varepsilon + \varepsilon^3 = 0$ perché ε è una radice di tale polinomio.
- poiché $\varepsilon^3 = \varepsilon + 1$, il coefficiente di x^2 si riscrive come:

$$\varepsilon^3(1 + \varepsilon^2 + \varepsilon^3) = (1 + \varepsilon)(1 + \varepsilon^2 + 1 + \varepsilon)$$

$$= (1 + \varepsilon)(\varepsilon + \varepsilon^2) = \varepsilon * (1 + \varepsilon^2) = \varepsilon + \varepsilon^3 = 1.$$

- Infine

$$\varepsilon^7 = (\varepsilon + 1)(\varepsilon + 1)\varepsilon = (\varepsilon^2 + 2\varepsilon + 1)\varepsilon = \varepsilon^3 + \varepsilon = 1.$$

Alternativamente, si può applicare l'algoritmo della divisione dividendo $f(x)$ per i vari fattori. □

4.2.2 Relazione tra polinomio minimo e morfismo di valutazione

Notiamo che data $F = F_p$ e E un'estensione di F_p di ordine p^n, se l'orbita di ε, che chiamo ε^A, ha t elementi, il polinomio minimo di ε è

$$m(x) = \prod_{\beta \in \varepsilon^A} (x - \beta)$$

Consideriamo la mappa di valutazione da $F[x]$ in E, definita come $f(x) \in F[x] \mapsto f(\varepsilon) \in E$. L'immagine di questo morfismo di anelli è isomorfa al quoziente $L = \frac{F[x]}{\ker \phi}$, dove $\ker \phi = \mathrm{Ann}_{F[x]}\varepsilon$, che è l'ideale generato dal polinomio minimo su F di ε, allora $L = \frac{F[x]}{(m(x))}$, e poiché il polinomio minimo è irriducibile, L è un campo.

L contiene una radice del polinomio minimo di ε, che chiamo ancora ε. Osservo che L **è un sottocampo intermedio tra** E **e** F, e $\mathrm{Dim}_F L$ è il grado del polinomio minimo, che è anche il numero di elementi coniugati a ε.

In particolare

$$L = \{g(\varepsilon)\, g(x) \in F[x]\}$$

ovvero è il sottocampo di E generato da ε su F ed è il minimo sottocampo di E contenente F ed ε.

E è anche uguale all'insieme delle radici di $f(x) = x^q - x$. In particolare, anche ε, che sta in E è radice di $f(x)$, e $f(x)$ ha coefficienti in F ed è monico, allora $m(x) \mid x^q - x$.

Più precisamente, in generale, sia t un divisore di n e $g(x)$ un polinomio a coefficienti in F irriducibile, di grado t. Se chiamo $L = \frac{F[x]}{(g(x))}$, esso è un campo di ordine p^t. E ha p^n elementi, allora per quanto detto in precedenza esiste un suo sottocampo L_0 di ordine p^t. I due campi L e L_0 hanno lo stesso ordine e sono isomorfi. Sia $\varepsilon_0 \in E$ l'immagine di ε mediante tale isomorfismo, e $m(x)$ è ancora il polinomio minimo di ε_0, ma allora $g(x) \mid x^q - x$.

Teorema 4.17. *Tutti e soli i fattori irriducibili monici di $x^q - x$ considerato come polinomio in $F[x]$, con $q = p^n$, sono tutti e soli i polinomi monici irriducibili di grado t con $t \mid n$ in $F[x]$.*

4.3 Prodotto di convoluzione

Date due funzioni f, g, definisco il *prodotto di convoluzione di Cauchy* ponendo

$$f * g(y) = \int f(x)g(y - x)\, dx$$

Dato un anello commutativo R, definisco il *prodotto di convoluzione di Dirichlet* ponendo

$$f * g(n) = \sum_{d|n} f(d)g(n/d)$$

Considero una funzone

$$f(n) = \sum_{d|n} g(d)$$

Se definiamo N_0 la funzione che vale 1 sugli interi e zero altrove, usando la convoluzione possiamo scrivere

$$f(n) = \sum_{d} g(d)N_0(n/d) = g * N_0$$

(infatti n/d è un intero solo se $d \mid n$, e N_0 è nulla negli altri casi)

Per poter invertire la formula, ed esprimere g in funzione di f, introduciamo la *funzione di Moebius* definita ponendo

$$\mu(n) = \begin{cases} 1 & \iff n = 1 \\ (-1)^s & \iff n = p_1 * p_2 * \cdots * p_s,\ p_i \neq p_j \text{ se } i \neq j \\ 0 & \text{altrimenti} \end{cases}$$

Proposizione 4.18 (formula di invertibilità di Moebius)**.** *Vale la formula:*

$$g = f * \mu.$$

Dimostrazione. Passo 1: In primo luogo mostriamo che

$$\mu * N_0 = \begin{cases} 1 & \iff n = 1 \\ 0 & \text{altrimenti} \end{cases} \tag{4.1}$$

$$\mu * N_0(n) = \sum_{d|n} \mu(d)N_0(n/d) = \sum_{d|n} \mu(d)$$

Ovviamente, se $n = 1$, $\mu * N_0(1) = 1$. Sia $n > 1$, allora se $n = p_1^{d_1} * p_2^{d_2} * \cdots * p_r^{d_r}$. Sui divisori della forma $p_1^{a_1} * \cdots * p_r^{a_r}$, $a_i \leq d_i$ con $a_i > 1$ per qualche i, la funzione μ è nulla perché nella fattorizzazione di d, non tutti i primi sono distinti. Allora si può limitare la sommatoria ai divisori della forma $p_1^{a_1} * \cdots * p_r^{a_r}$ con $a_i \in \{0, 1\} \forall i$. Supponiamo che in un certo divisore d compaiano k elementi della sommatoria con esponenti uguali a 1 (fissato $k \in \{1, \ldots, r\}$ ho $\binom{r}{k}$ possibilità di scelta per determinare le posizioni degli esponenti uguali a 1).Quindi La funzione di Moebius su un divisore di questo tipo vale $(-1)^k$. Allora

$$n > 1 \longrightarrow \mu_0 * N(n) = \sum_{k=1}^{r} \binom{r}{k} (-1)^k$$

$$= \sum_{k=1}^{r} \binom{r}{k} (-1)^k 1^{r-k} = (1-1)^k = 0$$

ovvero vale la formula (4.1).

Passo 2: osservo ora che data una funzione f qualsiasi, la convoluzione di f con la funzione indicatrice di 1 (che qui indico con I_1) restituisce la stessa f. Infatti

$$f * I_1(n) = \sum_{d|n} f(d) I_1(n/d) = f(n)$$

infatti l'unico termine che sopravvive è quello in cui $d = n$.

Passo 3: Per definizione

$$f = g * N_0$$

e facendo la convoluzione con μ a entrambi i membri:

$$f * \mu = g * N_0 * \mu$$

e poiché il prodotto di convoluzione è associativo

$$f * \mu = g * (N_0 * \mu) = g * I_1 = g$$

dove negli ultimi passaggi ho sfruttato i passi 1 e 2. □

4.3.1 Applicazione ai campi

Applichiamo questa formula ai conti precedenti. Sia $q = p^n$ e $N_F(t)$ la cardinalità dell'insieme dei polinomi il cui grado è t. Abbiamo mostrato che

$$x^q - x = \prod_{t|n} g_1(t) * \cdots * g_s(t)$$

con $g_i(t)$ polinomi monici irriducibili di grado t. Eguagliando i gradi

$$p^n = \sum_{t|n} t * N_F(t)$$

Dette $g(t) = t * N_F(t)$, e $f(n) = p^n$, la relazione sopra si scrive come

$$f(n) = \sum_{t|n} g(t)$$

allora per la formula di inversione di Moebius

$$g(n) = f * \mu(n) \longrightarrow n * N_F(n) = \sum_{d|n} p^d * \mu(n/d) \tag{4.2}$$

e quest'ultima formula fornisce il numero di polinomi irriducibili con grado d.

Corollario 4.19. *Per ogni $n \geq 1$, $N_F(n)$ è positivo, e quindi esistono polinomi irriducibili di grado n sul campo.*

Dimostrazione. Nel peggiore dei casi, ovvero quando n è pari, $n/2$ è un suo divisore e quindi

$$n * N_F(n) = p^n - p^{n/2} - \cdots \sim p^{n/2} * (p^{n/2} - 1 - \dots) > 0$$

□

Esempio 4.20. Dato un campo con $8 = 2^3$ elementi, consideriamo il polinomio $x^8 - x$. Applichiamo la formula (4.2), e mostriamo che $x^8 - x$ è prodotto dei polinomi irriducibili di grado d con $d \mid n = 3$. i polinomi irriducibili di grado 1 su F_2 sono 2: x e $x-1$, come conferma la formula:

$$1 * N_F(1) = \sum_{d|1} \mu(1/d) * 2^d = \mu(1) * 2 = 2$$

$$3 * N_F(3) = \sum_{d|3} \mu(3/d) * 2^d =$$

$$3 * N_F(3) = 1 * 2^3 - 1 * 2^1 = 6, \longrightarrow N_F(3) = 2$$

I due polinomi irriducibili di grado 3 sono $x^3 + x + 1$, $x^3 + x^2 + 1$. Svolgiamo il prodotto dei polinomi:

$$x * (x-1)(x^3 + x + 1)(x^3 + x^2 + 1) = x * (x-1) * (x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)$$

$$= x * (x^7 - 1) = x^8 - x.$$

Per costruire i polinomi irriducibili ci sono varie tecniche universali. Per fattorizzare il polinomio $x^q - x$ una delle strade possibili è l'algoritmo di Berlecamp.

4.3.2 Informazioni sui campi di ordine q^n

Sia F un campo di ordine $q = p^a$, e sia $E = F_{q^n}$. Estendiamo i risultati precedenti ai campi intermedi tra E e F in questo caso (la differenza è che la cardinalità di F non è un primo ma è la potenza di un primo).

(i) Un sottocampo L di F_{q^n} ha ordine q^d con d divisore di n, perché L è un F-spazio vettoriale su F_q. In generale, tutti e soli i campi intermedi L tra F e E hanno ordine q^d con $q = o(F)$ e $d \mid n$, e il gruppo di automorfismi di E che fissano F, che chiamo A, ha ordine p^{an} (questi automorfismi non fissano necessariamente ogni elemento di F, ma complessivamente fissano il campo ovvero sono tali che $\phi(F) = F$).

(ii) Dato $\varepsilon \in E$, e calcolato $m(\varepsilon)$ su F, osservo che F è l'insieme degli elementi di E fissati da ψ, con ψ opportuna potenza della mappa di Frobenius, inoltre

$$m(\psi(\varepsilon)) = 0$$

ψ infatti non sposta i coefficienti di m.

(iii) Vale la formula

$$\prod_{\beta \in A_F} (x - \beta) = m(x)$$

dove i coniugati di ε sono le q^i-esime potenze di ε. Determino il minimo t tale che $\varepsilon = \psi^t(\varepsilon)$, ovvero $\varepsilon = \varepsilon^{q^t}$. In particolare, dato γ primitivo nel gruppo moltiplicativo di E, ε sarà della forma γ^s, con s un intero minore di $q^n - 1$, e si ha

$$\varepsilon = \varepsilon^{q^t} \longrightarrow \gamma^s = \gamma^{s*q^t}, \gamma^{s(1-q^t)} = 1$$

quindi $q^{n-1} \mid s(q^t - 1)$.

3-12-2015

4.4 Fattorizzazione di polinomi

Problema: Fattorizzare $x^n - 1 \in F[x]$ con $|F| = q = p^a$, p primo. Determinare le radici di questo polinomio. Se $F = \mathbb{C}$, le radici hanno una rappresentazione geometrica: se disegno nel piano complesso la circonferenza $|z| = 1$, le radici sono i vertici di un n-agono regolare, $z_k = e^{\frac{2\pi i \cdot k}{n}}$, con $k \in \mathbb{Z}_n$. Se $F = F_q$ vorrei determinare i fattori irriducibili, di $x^n - 1$, ovvero vorrei determinare polinomi g_i tali che

$$x^n - 1 = \prod g_i(x)^{m_i}.$$

Teorema 4.21. *$x^n - 1$ non ammette fattori moltiplicativi se e solo se $M.C.D.(n, p) = 1$, ovvero se e solo se p non divide n.*

Dimostrazione. $1 \longrightarrow 2$: supponiamo che $x^n - 1$ non ammetta fattorizzazione. Se per assurdo p divide n, allora posso scrivere $n = pm$ e otterremmo

$$x^n - 1 = (x^m - 1)^p$$

infatti

$$(x^m - 1)^p = \sum_{j=0}^{p} \binom{p}{j} x^{mj}(-1)^{p-j} = x^{mp} - 1 = x^n - 1.$$

perché $\binom{p}{j} \equiv 0 \mod p \forall j = 1, \ldots, p-1$.

Supponiamo che m sia ancora divisibile per p, ovvero $m = m' * p$ allora posso iterare il ragionamento su $x^m - 1 = (x^{m'} - 1)^p$. Sia b il massimo esponente tale che $p^b || n$ ossia $p^b \mid n$ ma p^{b+1} non divide n, si ha $x^n - 1 = (x^l - 1)^{p^b}$, ovvero $x^n - 1$ ammette una fattorizzazione, contro l'ipotesi.

$2 \longrightarrow 1$: posticipiamo la dimostrazione di questa implicazione, per la quale servono alcuni risultati preliminari. □

4.4.1 Derivata di un polinomio

Dimostreremo che, dato un campo F non necessariamente finito, $f(x) \in F[x]$ è privo di quadrati se e solo se $M.C.D.(f(x), Df(x)) = 1$, dove se $f(x) = \sum_j f_j x^j$, $Df(x) = \sum_j f_j j x^{j-1}$.

D può essere visto come un operatore lineare $D : F[x] \to F[x]$ tale che $x^j \mapsto j x^{j-1}$ e $x^0 \mapsto 0$.

Proposizione 4.22. *Vale la formula*

$$D(fg) = Df \cdot g + f \cdot Dg.$$

Dimostrazione. Siano

$$f(x) = \sum f_j x^j; \qquad g(x) = \sum g_k x^k$$

allora

$$h(x) = f(x)g(x) = \sum h_l x^l$$

dove $h_l = \sum_{j+k=l} f_j g_k$. Allora $Dh(x) = \sum l h_l x^{l-1}$. Il coefficiente di x^{l-1} in $Df \cdot g + f \cdot Dg$ è:

$$\sum_{j+k=l} j f_j g_k + \sum_{k+j=l} f_j k g_k = l \sum_{j+k=l} f_j g_k = l h_l$$

□

Proposizione 4.23. *Vale la formula*

$$Dg^m = mg^{m-1}Dg.$$

Dimostrazione. Dimostro l'asserto per induzione.

Passo base: se $g = \sum_j g_j x^j$, $Dg = \sum_j g_j x^{j-1} = j * g^0 * Dg$ quindi l'asserto vale.

Passo induttivo: supponiamo che valga $Dg^{m-1} = (m-1)g^{m-1}Dg$ e mostriamo l'asserto per m.

$$Dg^m = D(g^{m-1} * g) = Dg^{m-1} \cdot g + g^{m-1} \cdot Dg$$

$$= (m-1)g^{m-2}Dg \cdot g + g^{m-1} \cdot Dg = (m-1)g^{m-1} \cdot Dg + g \cdot Dg = mg^{m-1}Dg.$$

□

Definizione 4.24. Un campo di caratteristica p si dice *perfetto* se $F = F^p$.

Proposizione 4.25. *Se F è un campo di caratteristica 0 o è un campo perfetto, $f(x) = g(x)^m h(x)$ con $g(x)$ irriducibile e $m > 1$ se e solo se $g \mid M.C.D.(f, Df)$.*

Dimostrazione. $1 \longrightarrow 2$: applicando la formula sulla derivata del prodotto e la proposizione precedente

$$Df = D(g^m \cdot h) = Dg^m \cdot h + g^m \cdot Dh = mg^{m-1}Dg \cdot h + g^m \cdot Dh$$

quindi $g(x) \mid Df$, ossia $g(x) \mid M.C.D.(f, Df)$.

$2 \longrightarrow 1$: sia $M.C.D.(f, Df) = d(x) \neq 1$, e supponiamo che $g(x) \mid d(x)$ con $g(x)$ irriducibile. Ovviamente si ha $f(x) = g(x)^m h(x)$ perché $g(x) \mid f(x)$. Se fosse $m = 1$, allora

$$Df = Dg \cdot h + g \cdot Dh$$

allora $g(x) \mid Dg \cdot h$, ovvero, poiché h non è divisibile per g, $g(x) \mid Dg(x)$ che è assurdo. $t = deg$, allora $Dg = 0$, ossia $\sum k g_k x^{k-1} = 0$.

(i) se $\mathrm{Car}F = 0$, $tg_t = 0$ se e solo se $g_t = 0 \forall t$, assurdo.

(ii) se $\mathrm{Car}F = p > 0$, allora $kg_k = 0$ se si verifica una delle seguenti possibilità: k è congruo a 0 modulo p, oppure $g_k = 0$. Allora $g(x) = \sum g_{pl} x^{pl}$. Se chiamo $m(x) = \sum g_{pl} x^l$, $g(x) = m(x^p)$. Nel caso in cui F sia perfetto, $g_{pl} = m_l^p \forall l$ con $m_l \in F$, allora $g(x) = \sum m_l^p x^{pl} = (\sum m_l x^l)^p$, assurdo perchè per ipotesi $g(x)$ è irriducibile.

□

Nel nostro caso $f(x) = x^n - 1$ e $Df = nx^{n-1}$, quindi $M.C.D.(f(x), Df(x)) \neq 1$ se e solo se $p \mid n$. Quindi se dimostro che F è perfetto, dalla proposizione precedente segue che $f(x)$ non ha fattori multipli e rimane così provata la seconda implicazione del teorema 4.21. Quando $|F| = q = p^n$, effettivamente $\phi : F \to F$ t.c. $\phi : \varepsilon \to \varepsilon^p$ è un automorfismo, infatti:

- ϕ è F_p-lineare: dato $\lambda \in F$,

$$\phi(\lambda\varepsilon) = (\lambda\varepsilon)^p = \lambda^p \varepsilon^p = \lambda^p \phi(\varepsilon) = \lambda\phi(\varepsilon).$$

- ϕ è iniettiva infatti
$$\ker(\phi) = \{\varepsilon : \phi(\varepsilon) = \varepsilon^p = 0\}$$
allora $\ker\phi = 0$, inoltre poiché F è un insieme finito segue che ϕ è anche suriettiva.

Osservazione 4.26 (Principio di Hankel)**.** Considero l'applicazione $\mathrm{Gr} : F[x] \setminus \{0\} \to \mathbb{N}$; posso estendere la definizione su tutto $F[x]$, e considerare $\mathrm{Gr} : F[x] \to \mathbb{N}$, che deve soddisfare le seguenti proprietà:
$$\mathrm{Gr}(ab) = \mathrm{Gr}(a) + \mathrm{Gr}(b)$$
$$\mathrm{Gr}(0) = \mathrm{Gr}(a) + \mathrm{Gr}(0)$$
quindi $\mathrm{Gr}(0) = -\infty$ è una convenzione sensata.

Osservazione: Se prendo $F = \mathbb{Z}_6$ che NON è un campo
$$(2x+1)(3x+1) = 5x+1$$
I polinomi $a, b \in \mathbb{Z}_6[x]$ con $a \mid b$ ma $\mathrm{Gr}(a) > \mathrm{Gr}(b)$ e $b \neq 0$.

4.4.2 Determinazione del campo di spezzamento

Problema: Penso $x^n - 1 \in F[x]$ con $|F| = q = p^a$ e p che non divide n. Determinare il più piccolo campo $E \geq F$ su cui $x^n - 1$ si spezza in fattori lineari. Basta aggiungere ad F un elemento γ tale che $o(\gamma) = n$, ossia $\gamma^n = 1$ (radici di $x^n - 1$); $|\{\gamma^i : i \in \mathbb{Z}\}|$ e $(\gamma^i)^n = 1$ e sono tutte radici di $x^n - 1$, quindi $x^n - 1 = \prod(x - \gamma^i)$. $|E| = q^t$. Esiste $\gamma \in E^*$ t.c. $o(\gamma) = n$ se e solo se n divide $|E^*| = q^t - 1$. Se $p|n$, allora p non divide $q^t - 1$, perciò non esiste γ tale che $o(\gamma) = n$. Se p non divide n, esiste t tale che $n|q^t - 1$ e quindi q^t è congruo ad 1 modulo n con $q \in \mathbb{Z}_n$. Poichè $MCD(n, q) = 1$, si ha $q \in (\mathbb{Z}_n)^*$. Siccome $|(\mathbb{Z}_n)^*| = \phi(n)$, ogni elemento $n \in (\mathbb{Z}_n)^*$ ammette inverso d tale che $nd = 1$, quindi $t = o_n(q)$.

Esempio 4.27. Siano $n = 3$, $q = 2$ (sono coprimi), fattorizzo $x^3 - 1 \in F_9[x]$. $t = o_3(2) = 2$, quindi
$$E = F_{q^t} = F_4 \cong \frac{F[x]}{< x^2 + x + 1 >}$$
$|E^*| = 4 - 1 = 3$, allora $\alpha := x + I$, segue che
$$x^3 - 1 = (x-1)(x-\alpha)(x-\alpha^2).$$

Esempio 4.28. Siano $n = 7$, $q = 2$ (sono coprimi). $t = ord_7(2) = 3$, quindi
$$E = F_{2^3} = F_8 = \frac{F[x]}{< x^3 + x + 1 >}.$$
(è il campo con 8 elementi, quindi è isomorfo al quoziente di $F[x]$ per l'ideale generato da un polinomio irriducibile di grado 3)

$|E^*| = 7$, $\alpha = x + I$, $\alpha^3 + \alpha + 1 = 0$. α è primitivo perchè lo sono tutti gli elementi di $\mathbb{Z}_8^*$, allora
$$x^7 - 1 = \prod_{i=0}^{6}(x - \alpha^i)$$

in $E[x]$, ma sarò interessato a trovare la fattorizzazione in $F[x]$. Se $g(x) \mid x^7 - 1$ in $F[x]$ e $g(\varepsilon) = 0$, $g(x) = \prod_{\beta \in \varepsilon^A}(x - \beta)$. Se prendo α con $i = 0$, $A =< e >$ e $\{\alpha, \alpha^2, \alpha^4\} = \alpha^A$, quindi $g(x) = (x - \alpha)(x - \alpha^2)(x - \alpha^4) \in F[x]$. Prendo α^3, $\{\alpha^3, \alpha^6, \alpha^5\} = \alpha^{3^A}$, quindi $(x - \alpha^3)(x - \alpha^5)(x - \alpha^6) \in F[x]$. Se prendo 1, $\{1\} = 1^A$, $x - 1$ è irriducibile.

Capitolo 5

CODICI CILCICI

10-12-2015

5.1 Definizione di codici ciclici

Introduciamo una nuova classe di codici, che hanno la proprietà di contenere nel loro gruppo di automorfismi sottogruppi ciclici di dimensioni significative.

Definizione 5.1. Sia F un campo e n un intero. Dico che un sottospazio vettoriale di F^n è un *codice ciclico* se data una qualsiasi parola $c = (\gamma_0, \gamma_1, \dots, \gamma_{n-1})$ nel codice, anche la parola che si ottiene spostando in avanti le componenti, ovvero $(\gamma_{n-1}, \gamma_0, \dots, \gamma_{n-2})$ appartiene al codice.

In altre parole, un codice ciclico è invariante per spostamento in avanti delle componenti dei suoi elementi.

Dato uno spazio vettoriale F^n, definisco lo shift in avanti $\sigma : (\gamma_0, \gamma_1, \dots, \gamma_{n-1}) \mapsto (\gamma_{n-1}, \gamma_0, \dots, \gamma_{n-2})$.
σ è un endomorfismo F-lineare, infatti

(i) dato $\lambda \in F$:

$$\sigma(\lambda(\gamma_0, \gamma_1, \dots, \gamma_{n-1})) = (\lambda\gamma_{n-1}, \lambda\gamma_0, \dots, \lambda\gamma_{n-2}) = \lambda\sigma((\gamma_0, \gamma_1, \dots, \gamma_{n-1})$$

(ii) dati $c, c' \in C$, $\sigma(c + c') = \sigma(c) + \sigma(c')$.

Rispetto alla base standard, con vettori $\{e_0, e_1, \dots, e_{n-1}\}$ (qui per convenzione e_i è il vettore con la $i+1$-esiam componente uguale a 1 e le altre nulle), si ha che

$$\sigma(e_i) = e_{i+1} \forall i = 0, \dots, n-2; \quad \sigma(e_{n-1}) = e_0$$

La matrice che rappresenta tale applicazione rispetto alla base standard ha come colonne $e_1, e_2, \dots, e_{n-1}, e_0$. In particolare σ è invertibile ed è un automorfismo di F^n in sé.

Possiamo riformulare la definizione di codice ciclico nel seguente modo:

Definizione 5.2. Un sottospazio vettoriale C di F^n è ciclico se e solo se $\sigma(C) = C$.

σ ha ordine n, inoltre si verifica che, nel caso di codici ciclici

$$C = \sigma^n(C) = \sigma^{n-1}(C) = \dots = \sigma(C)$$

Esempio 5.3. Il codice di ripetizione di lunghezza n è ciclico (ogni vettore ha tutte le componenti uguali tra loro). Di più, questo codice è l'insieme dei vettori di F^n che vengono fissati punto per punto dalla mappa σ, ovvero **il codice di ripetizione di lunghezza n il centralizzante in F^n di σ**. La prima implicazione è ovvia perché un elemento del codice di ripetizione viene fissato da σ. Viceversa, preso un punto del centralizzante $c = (\gamma_0, \gamma_1, \dots, \gamma_{n-1})$, mostriamo che sta nel codice di ripetizione.

$$c = \sigma(c) \longrightarrow (\gamma_0, \gamma_1, \dots, \gamma_{n-1}) = (\gamma_{n-1}, \gamma_0, \dots, \gamma_{n-2})$$

$$\longrightarrow \gamma_1 = \gamma_0,\ \gamma_2 = \gamma_1,\ \dots, \gamma_{i+1} = \gamma_i, \dots$$

ovvero $\gamma_i = \lambda \forall i$, e quindi $c = (\lambda, \lambda, \dots, \lambda) \in \mathrm{Rep}_n F$.

Esercizio 5.4. Verificare se il codice di Hamming è ciclico oppure no.

Dimostrazione. Una matrice generatrice del codice di Hamming è data da

$$G = \begin{matrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{matrix}$$

Per ottenerla applico il seguente procedimento: voglio portare H nella forma $(-A^t, I_3)$, e per farlo scambio la prima e la settima riga, la seconda e la sesta riga, la quarta e la quinta riga

$$H' = \begin{matrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix}$$

Allora ottengo:

$$G' = (I_4, A)$$

$$= \begin{matrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{matrix}$$

e rifacendo gli scambi appena effettuati:

$$= \begin{matrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{matrix}$$

Chiamo g_1, g_2, g_3, g_4 le righe della matrice, che sono i generatori del codice. Ogni elemento del codice è della forma

$$\alpha g_1 + \beta g_2 + \gamma g_3 + \delta g_4$$

$$\alpha(1,1,0,1,0,0,1) + \beta(0,1,0,1,0,1,0) + \gamma(1,1,1,0,0,0,0) + \delta(1,0,0,1,1,0,0)$$

$$(\alpha + \gamma + \delta, \alpha + \beta + \gamma, \gamma, \alpha + \beta + \delta, \delta, \beta, \alpha)$$

□

5.1.1 Studio dei gruppi ciclici

Osservo che σ preserva i pesi, ovvero $\omega(c) = \omega(\sigma(c)) \forall c$ (infatti la matrice associata a σ è di permutazione e in particolare è monomiale).

Il gruppo generato da σ è ciclico, e ha ordine n. Sappiamo già che $\sigma^n = id$, e mostriamo che per ogni $m < n$, $\sigma^m \neq id$: questo è vero perché se considero e_0, $\sigma^m(e_0) = e_m$ che è diverso da e_0 se $m < n$.

Negli anni 50 fu determinata la relazione tra codici ciclici e polinomi. Costruisco un isomorfismo di spazi vettoriali tra F^n e un sottospazio di $F[x]$ (il primo ha dimensione n e il secondo ha dimensione infinita). Considero

$$F_n[x] = \{f(x) \in F[x] \, t.c. \, \mathrm{Gr} f(x) < n\}$$

allora $F^n \cong F_n[x]$. Basta osservare che i due spazi hanno la stessa dimensione su F, allora segue che la base $\{1, x, \dots, x^{n-1}\}$ di $F_n[x]$ può essere messa in corrispondenza con $\{e_0, e_1, \dots, e_{n-1}\}$. Dato $c = (\gamma_0, \gamma_1, \dots, \gamma_{n-1})$ definisco ψ tale che

$$c = \sum_i \gamma_i e_i \mapsto^{\psi} c(x) := \sum_i \gamma_i x^i.$$

Supponiamo che in $F_n[x]$ l'applicazione di σ corrisponda alla moltiplicazione per x. Osservo però che moltiplicando per x il polinomio $c(x)$, si ottiene

$$xc(x) = \gamma_0 x + \gamma_1 x^2 + \cdots + \gamma_{n-2} x^{n-1} + \gamma_{n-1} x^n$$

che ha grado uguale a n, e quindi non rientra più in $F_n[x]$, quindi quest'interpretazione di σ non è corretta.

Si vorrebbe interpretare il coefficiente di x^n come il termine noto, allora si può definire una relazione di equivalenza tale che $x^n \mapsto 1$.

Per risolvere il problema, **l'ambiente giusto su cui lavorare per studiare la struttura dei codici ciclici è il quoziente** $R := \frac{F[x]}{(x^n-1)}$ (qui $x^n - 1 \equiv 0$ quindi $x^n \equiv 1$).

Proposizione 5.5. *R è un F-spazio di dimensione n.*

Dimostrazione. **Descrizione degli elementi di** R: $r \in R$ è della forma $f(x) + (x^n - 1)$ con $f(x) \in F_n[x]$ (sono classi di equivalenza). Dato $f(x)$ di grado maggiore o uguale a n, si può scrivere

$$f(x) = (x^n - 1) * q(x) + r(x), \longrightarrow [r(x)] = [f(x)]$$

e $\mathrm{Gr} r(x) < n$.

R **ha una struttura di** F**-spazio** con le seguenti operazioni:

somma : $[f(x)] + [g(x)] = [f(x) + g(x)]$, con elemento neutro $[0]$.

prodotto per scalari : $(\lambda, f(x) + I) \mapsto \lambda f(x) + I = [\lambda f(x)]$.

Verifichiamo che le definizioni di somma e prodotto siano ben poste, e che se $g(x) = f(x) + i(x)$ è un altro rappresentante di $[f(x)]$, il risultato non cambia, ovvero $[\lambda g(x)] = [\lambda f(x)]$. Questo è vero perché dato $i(x) \in I$, si ha ancora $\lambda i(x) \in I$ perché un ideale contiene tutti i multipli polinomiali dei suoi elementi.

Mostriamo che $\mathrm{Dim} R = n$. Siano $f(x), g(x) \in F_n[x]$ con la stessa classe di equivalenza, allora $f(x) - g(x) \in I$, e quindi $x^n - 1 \mid f(x) - g(x)$, ma siccome ci si trova in un dominio di integrità, questo può avvenire se e solo se $f(x) - g(x) = 0$. Allora esiste un unico polinomio $r(x)$ di grado $< n$ in ogni classe di equivalenza, che viene detto *rappresentante canonico*. Si ha che $r(x) = r_0 + r_1 x + \cdots + r_{n-1} x^{n-1}$, $r_i \in F$, allora

$$r(x) + I = \sum_i r_i x^i + I = \sum_i (r_i x^i + I) = \sum_i r_i (x^i + I)$$

in R si ha anche una struttura di anello, e si può definire un prodotto moltiplicando i rappresentanti fra di loro, ovvero

$$(f(x) + I)) * (g(x) + I) = f(x)g(x) + I$$

quindi $x^i+I = (x+I)^i$. Se $\alpha = [x]$, segue che ogni $r \in R$ si può scrivere come $\sum_i r_i\alpha^i$. Tornando ai conti precedenti, ogni elemento di R si scrive in modo unico come F-combinazione lineare di $\alpha^0, \alpha^1, \ldots, \alpha^{n-1}$, e questi oggetti sono una F-base per R. □

Corollario 5.6. *$F_n[x]$ e R sono isomorfi come spazi vettoriali.*

Per risolvere il problema evidenziato precedentemente ($x^n \notin F_n[x]$), definisco $\psi : F^n \to R$ e cerco un modo per codificare σ. In notazione funzionale, interpreto σ come $\psi \circ \sigma \circ \psi^{-1}$, dove ψ è il morfismo definito prima che agisce sui vettori della base: $e_i \mapsto x^i$. Mostro che $\psi\sigma\psi^{-1} = \mu_\alpha$ con $\mu(\alpha) : R \to R : r \mapsto r\alpha$. Sia $v = (\omega_0, \omega_1, \ldots, \omega_{n-1})$, ovvero

$$v = \sum_{i=0}^{n-1} \omega_i e_i$$

allora

$$\psi(v) = \sum_{i=0}^{n-1} \omega_i \alpha^i \in R$$

Inoltre

$$\sigma(v) = \omega_{n-1}e_0 + \omega_0 e_1 + \cdots + \omega_{n-2}e_{n-1}$$

$$\psi(\sigma(v)) = \omega_{n-1}1 + \omega_0\alpha + \cdots + \omega_{n-2}\alpha^{n-1}$$

poiché $\alpha^n = 1$:

$$= \omega_0\alpha + \omega_1\alpha^2 + \cdots + \omega_{n-2}\alpha^{n-1} + \omega_{n-1}\alpha^n = \alpha\psi(v)$$

ovvero μ_α corrisponde a σ in R.

Per inciso, R è una F-algebra, ovvero è contemporaneamente un anello e uno spazio vettoriale e vale la proprietà di compatibilità $\lambda * rs = (\lambda * r)s = r(\lambda * s)$), ovvero $f * id_R$ è un sottoinsieme del centro di R.

5.1.2 Teorema fondamentale sui codici ciclici

Definizione 5.7. Un ideale J di un anello A deve soddisfare le seguenti condizioni:

(i) J è un sottogruppo di $(A, +)$

(ii) per ogni $a \in A, j \in J$, $aj \in A$.

Teorema 5.8. *Un codice è ciclico se e solo se $\psi(C)$ risulta essere un ideale di R.*

Dimostrazione. $1 \longrightarrow 2$: sia C un codice ciclico, e sia $J = \psi(C)$. Poiché ψ è un isomorfismo, J è un sottospazio di R. Inoltre, poiché $\sigma(C) = C$ vale l'equazione

$$\psi(c) = \psi(\sigma(c)) = \mu_\alpha * \psi(c)$$

quindi $\alpha c \in J$ per ogni $c \in J$. Ripetendo il ragionamento, per induzione segue che $\alpha^i j \in J$ per ogni i.

Siano $r \in R, j \in J$, allora

$$r = \sum_i r_i \alpha^i$$

$$rj = \sum_i r_i \alpha^i j$$

e per quanto detto prima $\alpha^i j \in J$, e poiché J è un sottospazio, la somma dei suoi elementi sta in J, quindi $rJ \in J$. J soddisfa le condizioni della definizione, quindi è un ideale.

$2 \longrightarrow 1$: considero C tale che $\psi(C) = J$ sia un ideale di R. In particolare, dato $\alpha \in R$, $\alpha j \in J$, e $\alpha^n j = idj \in J$. Allora $\mu_\alpha \psi(C) = \psi(C)$ inoltre

$$\mu_\alpha \psi(c) = \psi(\sigma(c))$$

unendo le due condizioni $\psi(c) = \psi(\sigma(c))$ ovvero, semplificando per $\psi(c)$, $\sigma(C) = C$ e C è ciclico. □

In generale, se A è un anello commutativo e I è un suo ideale, detto $\bar{A} = A/I$, esiste una corrispondenza biunivoca tra gli ideali del quoziente e gli ideali dell'anello di partenza che contengono I. Infatti, detta π la mappa canonica $a \mapsto a + I$, costruisco una corrispondenza nel seguente modo:

- preso M ideale di A, $\pi(M)$ è un ideale del quoziente.
- Viceversa, sia M un ideale del quoziente, e considero la controimmagine $A = \pi(M)^{-1}$, A è un ideale, poiché $\pi(M)$ contiene lo zero, A contiene I. Ovviamente A contiene anche M allora

$$\pi(a - m) = 0, \longrightarrow a - m \in I, \longrightarrow A = M \oplus I$$

Se M, N contengono I, essi non possono essere mappati nello stesso ideale, quindi la mappa indotta da π^{-1} (insieme delle controimmagini) e la mappa π danno una corrispondenza biunivoca tra le famiglie di ideali.

Quali sono gli ideali di R che sono anche sottospazi?

Considero gli ideali dell'anello $\frac{F[x]}{I}$. Per il teorema appena richiamato, c'è una corrispondenza biunivoca tra gli ideali di R e gli ideali di $F[x]$ che contengono I. Ma $F[x]$ è a ideali principali, (ovvero J è l'insieme dei multipli polinomiali di un certo polinomio $g(x)$), allora $x^n - 1 \in I \subset J$ dev'essere un multiplo di $j(x)$, ovvero $j(x) \mid x^n - 1$. Allora gli ideali $\bar{J}$ di R sono della forma $j(\alpha)R$, con $j(x) \mid x^n - 1$.

L'analisi dei codici ciclici in F^n si trasforma nello studio dei fattori di $x^n - 1$ in un campo finito (perché equivale allo studio di ideali del quoziente).

Esempio 5.9. Tornando agli esempi precedenti, il codice di ripetizione è il sottospazio vettoriale generato da $(1, 1, \ldots, 1)$. Il polinomio che genera l'ideale corrispondente al codice è

$$\psi(1, 1, \ldots, 1) = \sum_i \alpha^i$$

quindi $j(x) = \sum_{i=0}^{n-1} x^i$. Tale polinomio è un divisore di $x^n - 1$.

Esempio 5.10. Il codice nullo proviene dal polinomio $x^n - 1$. Invece, il codice ciclico costituito da F^n è generato da 1.

Esercizio 5.11. Dimostrare che $\psi^{-1}(J)$ con $J = (1 + \alpha + \alpha^3)$ e $F = F_2$, $n = 7$ è equivalente al codice di Hamming $(7, 4, 3)$. Esistono altri polinomi fattori di $x^7 - 1$ in F_2 che realizzano codici di Hamming?

14-12-2015

5.1.3 Riepilogo

Sia dato un codice ciclico n-dimensionale su un campo finito F di dimensione q; esso può essere messo in corrispondenza con la sua immagine in $R = \frac{F[x]}{(x^n-1)}$ attraverso la corrispondenza che, data $c \in C$, della forma

$$c = \sum_{i=0}^{n-1} \gamma_i e_i$$

lo manda in

$$\psi(c) = \sum_{i=0}^{n-1} \gamma_i x^i \in R.$$

$\psi(C)$ è un ideale in R, generato da un certo polinomio $g(\alpha)$, con $g(x)$ divisore di $x^n - 1$.

La classificazione dei codici ciclici equivale alla determinazione degli ideali di R, e quindi alla determinazione degli ideali di $F[x]$ che contengono $I =< x^n - 1 >$.

5.1.4 Passo 1: riduzione alla fattorizzazione di $x^m - 1$

Supponiamo che sia $q = p^a$ con p primo, e $n = p^b * m$ dove p^b è la massima potenza di p che divide n, allora si può scrivere

$$x^n - 1 = (x^m - 1)^b$$

e quindi ci si riconduce a determinare i divisori di $x^m - 1$. Il derivato di $x^m - 1$ è $mx^{m-1} \neq 0$ poiché p è coprimo con m, e $M.C.D.(f, f') = 1$, quindi $x^m - 1$ non ha radici multiple in qualsiasi estensione K di F.

Cerchiamo algoritmi per fattorizzare polinomi della forma $x^m - 1$ con m coprimo con p. L'algoritmo di Berlecamp fornisce esplicitamente i fattori g_i di $x^m - 1$, mentre altri algoritmi danno informazioni sui fattori di questo polinomio e sul loro numero senza darne un'espressione esplicita.

5.1.5 Passo 2: ricerca di una radice α

Se cerco una radice α di $x^m - 1$, non è detto che essa sia in F^*, campo dei coefficienti di $x^m - 1$. Ci chiediamo quando esiste un'estensione di F che contiene α radice di $x^m - 1$, ovvero **cerchiamo K che contenga α con $o(\alpha) = m \mid o(K^*)$.**

Quindi, si cerca il minimo campo K tale che $m \mid o(K^*)$. Poiché K è un F-spazio vettoriale, $o(K^*) = q^t - 1$, dove $t = \mathrm{Dim}_F K$. Allora la condizione $m \mid o(K^*)$ significa $q^t \equiv 1 \mod m$, e poiché m, q sono coprimi, questo significa che $t \equiv o(q) \mod m$.

Ci si chiede se **esiste t universale tale che $a^t \equiv 1 \mod m$ per ogni a coprimo con m.** La risposta è affermativa, infatti

- per il teorema di Eulero-Fermat, per ogni a coprimo con m,

$$a^{\varphi(m)} \equiv 1 \mod m,$$

- Sia $\varphi(m) = p_1 * p_2 * \cdots * p_r$, verifico se esiste $t < \varphi(m)$ che soddisfa la proprietà, ovvero verifico se per qualche indice i, $a^{\varphi(m)/p_i} \equiv 1 \mod m$ per ogni a coprimo con m.

• Se tale indice esiste, ripeto il ragionamento su $\varphi(m)/p_i$. Altrimenti, $t = \varphi(m)$.

Se $m = r_1^{m_1} * \cdots * r_s^{m_s}$ dove gli r_i sono primi, per il teorema cinese dei resti $\mathbb{Z}_m \cong \oplus_i \mathbb{Z}/(r_i^{m_i})$, attraverso l'isomorfismo $z \mapsto (z \mod m_1^{r_1}, \ldots, z \mod r_n^{m_n})$.

Se $r = 2$, gli elementi invertibili modulo 2^m dipendono dalla grandezza di m: inatti

• Se $m = 0$, 1 è l'unico elemento invertibile in $\mathbb{Z}_1$.

• se $m = 2$, ci sono 2 elementi invertibili in $\mathbb{Z}_{2^m} = \mathbb{Z}_4$,

• se $m \geq 3$, gli elementi invertibili sono $\varphi(2^m) = 2^{m-1}$;

Introduciamo la *funzione di Carmichael* λ che, dato $m = m_1^{r_1} * \cdots * m_r^{r_r}$, è definita come

$$\lambda(m) = l.c.m.(\varphi(m_i^{r_i})),$$

Esempio 5.12. Se $m = 15 = 3 * 5$, $\varphi(15) = 2 * 4 = 8$, mentre

$$\lambda_m = l.c.m.(\varphi(3), \varphi(5)) = 4.$$

Osserviamo che tutti gli ordini degli elementi coprimi con 15 dividono $\lambda(m)$, infatti gli elementi coprimi con 15 sono contenuti nell'insieme

$$\{1, 2, 4, 7, 8, 11, 13, 14\}$$

$$o(1) = 1, o(2) = 4,\ o(4) = 2, o(7) = 4, \ldots,$$

e gli ordini di questi elementi sono divisori di 4.

In conclusione, per determinare $\alpha \in K^*$:

• determino t tra i divisori della funzione di Carmichael o di Eulero;

• considero un elemento primitivo $\gamma \in K^*$, tale che $o(\gamma) = q^t - 1$;

• per definizione di t, $q^t \equiv 1 \mod m$ quindi $q^t - 1 = m * l$, allora pongo $\alpha = \gamma^l$;

• α soddisfa la condizione richiesta, ovvero $o(\alpha) = m$, infatti

$$\alpha^m = \gamma^{ml} = \gamma^{q^t - 1} = 1$$

• m soddisfa la condizione richiesta, infatti $m \mid o(K^*) = q^t - 1$.

5.1.6 Passo 3: fattorizzazione del polinomio

Una volta determinato α, il polinomio $x^m - 1$ si scrive come il prodotto di fattori di grado 1 in $K[x]$: in particolare $x^m - 1 = \prod_{i \in \mathbb{Z}_m}(x - \alpha^i)$, perché $o(\alpha^i)$ è un divisore di m, e la cardinalità di questi elementi è m. **Cerchiamo di accorpare questi fattori $x - \alpha^i$ in modo tale che il loro prodotto sia un elemento di $F[x]$.**

Fisso $\beta = \alpha^s$, $s \in \mathbb{Z}_m$. Sia β un elemento algebrico su F, allora i polinomi che si annullano in β costituiscono un ideale di $F[x]$, generato dal polinomio minimo di β, $m_\beta(x)$ che è irriducibile. Inoltre, β^q, immagine di β attraverso la mappa di Frobenius ϕ_q, è ancora una radice del polinomio $m_\beta(x)$. Infatti, se $m_\beta(x) = \sum_i k_i x^i$,

$$\phi_q(m(\beta)) = \phi_q(0) = 0 = \sum_i k_i \phi_q(\beta)^i$$

ovvero $\phi_q(\beta)$ è una radice di $m_\beta(x)$.

Considero l'orbita di β rispetto all'azione di ϕ, ovvero:

$$O_\beta = \{\beta, \phi(\beta), \ldots, \phi^{u-1}(\beta)\}$$

allora **il polinomio**

$$b(x) = \prod_{\delta \in O} (x - \delta)$$

è a coefficienti in F, infatti mostriamo che viene fissato da ϕ:

$$\phi(b(x)) = \prod_{\delta \in O} \phi((x - \delta)) = \prod_{\delta \in O} (x - \phi(\delta))$$

ma O e $\{\phi(\delta) : \delta \in O\}$ coincidono, quindi $\phi(b) = b$.

Si ha anche

$$b(\beta) = 0 \longrightarrow m_\beta(x) \mid b(x)$$

inoltre entrambi i polinomi sono monici e hanno $x-\beta, x-\phi(\beta), \ldots, x-\phi^{n-1}(\beta)$ come fattori, e hanno lo stesso grado, allora coincidono.

Concludo che $\mathrm{Gr} m_\beta(x) = o(O_\beta)$, dove

$$o(O_\beta) = \min\{u \, t.c. \, \phi^u(\beta) = \beta\},$$

ovvero il minimo u tale che $\beta^{q^u} = \beta$, e quindi

$$(\alpha^s)^{q^u} = \alpha^s \longrightarrow (\alpha^s)^{q^u - 1} = 1,$$

e poiché $m = o(\alpha)$, si cerca u tale che $m \mid s(q^u - 1)$. Il grado del polinomio minimo su F di α^s è

$$\min\{k t.c. s(q^u - 1) \equiv 0 \mod m\}.$$

Concludo che $x^m - 1$ **è prodotto dei polinomi minimi di** α^s **dove** α^s **sono i rappresentanti distinti delle orbite mediante** ϕ_q **in** K**, il grado del polinomio minimo di** α^s **è dato da**

$$\min\{u \, t.c. s(q^u - 1) \equiv 0 \mod m\}.$$

Esempio 5.13. Nel caso $n = 7, q = 2$, anche $m = 7$, e determiniamo $t = o(2) \mod 7$. Poiché 7 è primo, la funzione di Eulero e di Carmichael coincidono, quindi

$$\lambda(7) = \varphi(7) = 6$$

I divisori di 6 sono 2,3. Qui $2^3 \equiv 1 \mod 7$, quindi $t = 3$. Allora K è il campo con 2^3 elementi. Vogliamo fattorizzare $x^7 - 1$, le sue radici sono della forma α^s, $s \in \mathbb{Z}_7$.

(i) Se $s = 0$, $\beta = \alpha^0 = 1$,

$$\min\{u \, t.c. \, 0 * (q^u - 1) \equiv 0 \mod 7\}$$

Basta prendere $u = 1$, quindi $m_1(x) = x - 1$.

(ii) Se $s = 1$, $\beta = \alpha$, allora

$$\mathrm{Gr}m_\alpha(x) = \min\{k\, t.c.\, 2^k \equiv 1 \mod 7\}$$

e $k = 3$. Allora $m_\alpha(x)$ ha grado 3. Le sue radici sono $\alpha, \phi(\alpha) = \alpha^2, \phi(\alpha^2) = \alpha^4$. Quindi $m_\alpha(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^4)$.

(iii) Se $s = 3$ (parto dal primo esponente non ancora coinvolto), $\beta = \alpha^3$, allora

$$\mathrm{Gr}m_\beta(x) = \min\{k\, t.c.\, 3(2^k - 1) \equiv 0 \mod 7\}$$

e ancora $k = 3$. Allora le radici del polinomio minimo di α^3 sono α^3, α^6, $\alpha^{12} = \alpha^5$.

I polinomi irriducibili in F_2 di grado 3 sono dati dalla formula:

$$3n_3 = \sum_{d|3} \mu(3/d) * 2^d = 8 - 2 = 6, \longrightarrow n_3 = 2.$$

Ci sono due polinomi monici irriducibili di grado 3 su F_2, e tali polinomi coincidono con i polinomi minimi elencati sopra. Ci sono due codici ciclici corrispondenti ai quozienti $\frac{F_2[x]}{(x^3+x+1)}$ e $\frac{F_2[x]}{(x^3+x^2+1)}$.

Supponiamo di realizzare K come $\frac{F_2[x]}{(x^3+x+1)}$, e di chiamare $J = (x^3+x+1)$, allora $\alpha = x+I$ e mostriamo che

$$(x-\alpha)(x-\alpha^2)(x-\alpha^4) = x^3 + x + 1$$

con l'unica informazione che $\alpha^3 + \alpha + 1 = 0$.

- Il termine di grado massimo è x^3;
- il termine noto è

$$\alpha^7 = \alpha^6 * \alpha = (\alpha^3)^2 * \alpha = (\alpha+1)^2\alpha = (\alpha^2+1)\alpha = \alpha^3 + \alpha = 1.$$

($x^3 + x + 1$ è un divisore di $x^7 - 1$)

- Il coefficiente del termine di grado 2 è

$$-\alpha - \alpha^2 - \alpha^4 = -\alpha(1 + \alpha + \alpha^3) = 0$$

- Infine, il coefficiente di x è

$$\alpha * \alpha^2 + \alpha * \alpha^4 + \alpha^2 * \alpha^4 = \alpha^3 + \alpha^5 + \alpha^6 = \alpha^3 * (1 + \alpha^2 + \alpha^3)$$

$$= \alpha^3 * (\alpha^2 + \alpha) = \alpha^4 * (\alpha + 1) = \alpha^4 * \alpha^3 = 1$$

per i conti precedenti.

quindi $x^3 + x + 1 = (x-\alpha)(x-\alpha^2)(x-\alpha^4)$.

Esercizio 5.14. Con questo processo dimostrare che

$$(x-\alpha^3)(x-\alpha^5)(x-\alpha^6) = x^3 + x^2 + 1$$

Dimostrazione. • il coefficiente di grado 3 è 1.

- il coefficiente del termine x^2 è la somma delle radici, ovvero

$$-(\alpha^3+\alpha^5+\alpha^6) = -\alpha^3 * (1+\alpha^2+\alpha^3)$$

e poiché $\alpha^3 = \alpha + 1$

$$= \alpha^3 * (1+\alpha^2+1+\alpha) = (1+\alpha)(\alpha^2+\alpha) = \alpha(\alpha+1)^2$$

$$= \alpha(\alpha^2+1) = \alpha^3+\alpha = 1$$

- il coefficiente del termine di grado 1 è

$$-(\alpha^6 * \alpha^5 + \alpha^6 * \alpha^3 + \alpha^5 * \alpha^3) = -(\alpha^{11}+\alpha^9+\alpha^8) = -(\alpha^4+\alpha^2+\alpha)$$

$$= -\alpha(\alpha^3+\alpha+1) = 1$$

- il termine noto è

$$\alpha^{14} = \alpha^7 * \alpha^7 = 1.$$

□

Se $n = m * p^b$, con p^b potenza massima di p che divide n,

$$f(x) = x^n - 1 = (x^m-1)^{p^b}$$

La distanza minima dei codici associati a x^m-1 e x^n-1 rimane la stessa, quindi non ha senso considerare codici di lunghezza n perché si aumenta la lunghezza ma la distanza minima non migliora.

5.1.7 q-classi ciclotomiche

Definizione 5.15. Sia $s \in \mathbb{Z}_n$ una classe di resto, allora si chiama *q-classe ciclotomica* l'insieme degli esponenti degli elementi dell'orbita di α^s rispetto alla mappa di Frobenius ϕ_q, ovvero $s, qs, q^2s, \ldots,$.

La lunghezza di una q-classe ciclotomica è al più t. $\mathbb{Z}_n^*$ è unione delle q-classi ciclotomiche.

Detto r il numero delle q-classi ciclotomiche in $\mathbb{Z}_n$ e $\{s_i\}_{i=1}^r$ sono i rappresentanti.

Se chiamo u_i l'ordine della classe i-esima, se $M.C.D.(n,q) = 1$, allora x^n-1 è prodotto di r fattori irriducibili di grado u_i.

Esempio 5.16 (primo codice di Golé)**.** Questo codice ha parametri $n = 11, k = 6, d = 5, q = 3$. Analizzo le tre classi ciclotomiche in Z_{11}. Determiniamo $o(3)_{11}$. Poiché 11 è primo, $\lambda(11) = \varphi(11) = 10$.

$$3^2 \not\equiv 1 \mod 11$$

$$3^5 = 243 \equiv 1 \mod 11$$

quindi $t = 5$. Analizzo l'azione del gruppo di trasformazioni di Z_{11}.

$$O(0) = 0$$

$$O(1) = \{s, qs, q^2s, q^3s, q^4s\} = \{1, 3, 9, 5, 4\}$$

$$O(2) = \{2, 6, 7, 10, 8\}$$

(i risultati ottenuti sono stati ridotti modulo 11) NB: le orbite non hanno ordini casuali, infatti l'ordine dell'orbita è l'indice del centralizzante e quindi divide l'ordine del gruppo.

Questo conduce a due fattori irriducibili monici di grado 5, del polinomio $x^{11} - 1$.

Esempio 5.17 (secondo codice di Golay)**.** Questo codice ha parametri $n = 23, k = 12, d = 7, q = 2$. Le classi ciclotomiche sono

$$O(1) = \{1, 2, 4, 8, 16, 9, 18, 13, 3, 6, 12, 1\}$$

$$O(5) = \{5, 10, 20, 17, 11, 22, 21, 19, 15, 7, 14\}$$

e le orbite hanno lunghezza 11. Allora il polinomio si fattorizza come

$$(x - 1)g_1(x)g_2(x)$$

dove g_1, g_2 hanno grado 11.

Siccome $g(x) \mid x^n - 1$, $g(x) \in F[x]$, se α è un elemento di ordine n, e $t = o(q) \mod n$, esiste $\alpha \in F_{q^t}$ tale che $o(\alpha) = n$. Supponiamo di aver trovato s tale che $g(\alpha^s) = 0$, e sia S_g l'insieme di esponenti per α che forniscono radici di g. Divido $g(x)$ per il polinomio minimo di α^s con $s = s_1$, allora o i due polinomi coincidono oppure esiste $s_2 \notin \gamma(s_1)$ (dove $\gamma(s_1)$ è la q-classe contenente s_1) tale che $g(\alpha^{s_2}) = 0$, e il polinomio minimo di questo elemento è diverso dal precedente. In conclusione

$$g(x) = \prod_{i=1}^{v} m_{\alpha^{s_i}}(x)$$

dove gli s_i sono rappresentanti di q-classi ciclotomiche distinte, con α^{s_i} radice di g.

Siano $r(g) = \{\beta \in F_{q^t} \, t.c. \, g(\beta) = 0\}$, e $t(g) = \{s \in \mathbb{Z}_n \, t.c. \, g(\alpha^s) = 0\}$. Allora $t(g)$ è unione delle classi ciclotomiche. La distanza minima del codice associato a g è $\geq \delta + 1$ dove δ è il numero di elementi consecutivi che compaiono in $t(g)$.

Proposizione 5.18. *Sia C un codice ciclico generato da un certo polinomio, $\psi(C) = g(\alpha) * R$, allora* $\mathrm{Dim}C = n - \mathrm{Gr}g(x)$*, ovvero* $\mathrm{Gr}g(x) = n - k$*. (significato intuitivo: più grande è il codice, più piccolo è il grado del polinomio minimo associato)*

Dimostrazione. Sia $n - k = \mathrm{Gr}g(x)$, mostro che gli elementi $\{g(\alpha), \alpha g(\alpha), \ldots, \alpha^{k-1} g(\alpha)\}$ sono una base per $\psi(C)$. Poiché

$$< x^n - 1 > \subset < g(x) > \subset < 1 > = F[x]$$

posso costruire la mappa π che manda

$$R = \frac{\frac{F[x]}{<x^n - 1>}}{< g(x) >};$$

nel quoziente $\frac{F[x]}{<g(x)>}$. Per il teorema di isomorfismo, questo è un omomorfismo suriettivo di F-spazi vettoriali. Si ha che $\ker \pi =< g(x) >$ e $\mathrm{Dim}R = n$, allora la dimensione di $\ker \pi$ è $\mathrm{Gr}g(x) = n - k$, allora per il teorema di nullità più rango

$$\mathrm{Dim}C = \mathrm{Dim}\pi(R) = n - (n - k) = k$$

ovvero il codice ha dimensione k.

Mostriamo che **gli elementi** $\{g(\alpha), \alpha g(\alpha), \dots, \alpha^{k-1} g(\alpha)\}$ **sono una base**. Supponiamo che

$$\sum_i g(\alpha)\lambda_i \alpha^i = 0 \in R$$

e se $l(\alpha) = \sum_i \lambda_i \alpha^i$ segue che $l(\alpha)g(\alpha) = 0$, e $\mathrm{Gr}g(x)l(x) \leq n$. Il prodotto vale 0 nel quoziente solo se $g(x) * l(x) = 0 \in F[x]$, allora $l(x) = 0$ e tutti i suoi coefficienti λ_i sono nulli, e quindi si ha una relazione di indipendenza lineare degli elementi. □

Dato $g(x) = \sum_i g_i x^i$, il vettore ad esso corrispondente è $g = (g_0, g_1, \dots, g_{n-k}, 0, \dots, 0)$ (aggiungo k zeri per avere una lista di lunghezza n). Poiché il codice è ciclico, $\sigma(g) \in C$, con $\sigma(g)$ uguale allo shift in avanti del vettore, applicando σ^k a g gli elementi non nulli vengono spostati in fondo. Affermo che i $\sigma^i(g)$ sono linearmente indipendenti: siccome $\mathrm{Dim}C = k$, ottengo una nuova base.

L'indipendenza segue dal fatto che la matrice di colonne $g, \sigma(g), \sigma^2(g), \dots, \sigma^{k-1}(g)$, ovvero

$$\begin{matrix} g_0 & 0 & 0 & 0 \\ \dots & g_0 & 0 & 0 \\ g_{n-k} & \dots & g_0 & 0 \\ 0 & g_{n-k} & \dots & g_0 \\ 0 & 0 & g_{n-k} & \dots \\ 0 & 0 & 0 & g_0 \end{matrix}$$

ammette una sottomatrice $k \times k$ non singolare (quella in alto). Infatti $g_0 \neq 0$ (è il termine noto del polinomio) altrimenti x divide $g(x)$, ma questo non è possibile perché $x \nmid x^n - 1$. Questa matrice ha g_0 sulla diagonale, è triangolare superiore, con determinante $g_0^k \neq 0$, allora la matrice è non singolare. Quindi $\mathrm{Rg}G = k$, allora G è una matrice generatrice per C.

Si potrebbe trasformare la matrice in forma standard: i codici con matrice generatrice standard ammettono una facile decodifica nel senso che dato un vettore si risale facilmente al messaggio da cui proviene in base alle componenti del vettore ricevuto.

Teorema 5.19. *Se C è ciclico, allora il duale è anch'esso ciclico.*

Dimostrazione. Sia $c = (\gamma_0, \gamma_1, \dots, \gamma_{n-1}) \in C$, e $h = (\eta_0, \eta_1, \dots, \eta_{n-1}) \in C^\perp$.

$$\sigma(c) \cdot \sigma(h) = c \cdot h = \sum_i \gamma_i \eta_i = 0, \, \forall c \in C$$

e poiché $C = \sigma(C)$, $\sigma(H) \perp C$ quindi $\sigma(H) \subseteq C^\perp$, ossia $\sigma(C^\perp) \subset C^\perp$. □

$\psi(C^\perp)$ è generata da un certo polinomio $h(\alpha)$.

Esercizio 5.20. Come si traduce il prodotto scalare tra vettori quando passo ai corrispondenti in R?

17-12-2015

5.1.8 Informazioni sulla struttura dell'anello quoziente

Considero $R = \frac{F[x]}{<x^n-1>}$ e un codice ciclico C associato all'ideale $< g(x) >$ con $g(x) \mid x^n - 1$ nel quoziente. **Dati due ideali nel quoziente,** $g(x)F[x]$ **e** $h(x)F[x]$ **ci si chiede quando sono uguali**. Osserviamo che in $F[x]$ $\lambda g(x)$ e $g(x)$ generano lo stesso ideale, poiché tutti gli ideali sono principali hanno un unico generatore a meno di moltiplicazioni per una costante. Nel quoziente, se considero $I = g(\alpha) * R = f(\alpha) * R$, ci si chiede quale sia il legame tra $g(x)$ e $f(x)$: deve valere la relazione $g(\alpha) \mid f(\alpha)$ e $f(\alpha) \mid g(\alpha)$, ovvero esiste $u(x)$ tale che $u(\alpha) \in R^*$, ovvero $u(\alpha)$ è unitario.

Sia $x^n - 1 = \prod_{i=1}^{r} p_i(x)$, allora $R \cong \oplus_{i=1}^{r} \frac{F[x]}{(p_i(x))}$. Poiché n, p sono coprimi, non ci sono fattori multipli; inoltre $\frac{F[x]}{<p_i(x)>}$ è un campo perché i p_i sono irriducibili, ed è isomorfo a $F_{p^{d_i}}$ dove $d_i = \mathrm{Gr}p_i$ e dove p è la caratteristica del campo di partenza. Detto R^* l'insieme degli elementi invertibili di R, segue che $R^* = \prod_{i=1}^{r} E_i^*$, con $E_i = \frac{F[x]}{<p_i>}$.

$R \cong \oplus_{i=1}^{r} \frac{F[x]}{<p_i>}$ attraverso l'omomorfismo τ che manda un polinomio $p(x) \in R$ nella n-upla che ha come i-esima componenti i resti della divisione di $p(x)$ per p_i. Supponiamo che per $i = 0, \ldots, s-1$ i p_i siano fattori di $g(x)$, allora il codice associato all'ideale $< g(x) >$ è l'insieme degli elementi che sono multipli di $g(x)$ nel quoziente, ovvero tali che $g(\alpha) \mid c(\alpha)$, quindi le immagini degli elementi del codice mediante questo omomorfismo hanno zero nelle prime s componenti, quindi sono della forma $v = (0, 0, \ldots, 0, \gamma_s, \ldots, \gamma_{r-1})$.

Osserviamo che $\varepsilon = (0, \ldots, 0, 1, \ldots, 1)$ (le posizioni nulle sono s) si comporta come l'identità su $\tau(\psi(C))$. Segue quindi che ogni $u(\alpha)$ tale che $\tau(u(\alpha)) = \varepsilon$ è l'unità su $\psi(C)$.

Concludo che **ogni generatore di** $\psi(C) \in R$ **è della forma** $g(\alpha)u(\alpha)$ **con** u **invertibile**.

5.1.9 Generatore del duale di un codice ciclico

Sia $g(x)$ il polinomio monico di grado minimo tale che $g(\alpha)$ generi $\psi(C)$. Se C è ciclico, allora $C^\perp$ è ancora ciclico per il teorema 5.19, e quindi esiste un polinomio $f(x) \in F[x]$ tale che l'immagine di $C^\perp$ attraverso ψ è generata da $f(\alpha)$ nell'anello quoziente.

Determiniamo il generatore del codice duale: facciamo le seguenti osservazioni

(i) il polinomio generatore del duale deve avere grado $k = n - \mathrm{Gr}g$, perché $\mathrm{Dim}C + \mathrm{Dim}C^\perp = n$.

(ii) Sia $h(x) t.c. x^n - 1 = g(x) * h(x)$. Sia $h(x) = \sum_{j=0}^{k} \eta_j x^j$, e $g(x) = \sum_{j=0}^{n-k} \gamma_j x^j$. Ovviamente $x^n - 1 = g(x) * h(x) = 0 \in R$, e i coefficienti del prodotto sono dati da

$$0 = \mu_l = \sum i + m = l \gamma_i \eta_m = \sum_{i+m=l} \gamma_i \eta_{l-i}$$

(iii) Un generatore H^* del duale deve soddisfare la relazione

$$\sum_j \gamma_j \eta_j^* = 0$$

Mostro che il polinomio

$$h^* = h(x^{-1}) * x^k = \sum_{j=0}^{k} \eta_{k-j} x^j = \sum_{j=0}^{k} \eta_j^* x^j$$

soddisfa tale relazione.

$$g(x)h^*(x) = \sum_l \nu_l x^l$$

con

$$\nu_l = \sum_{i+k-j=l} \gamma_i \eta_{k-j} \longrightarrow i+k = l+j,$$

$$\longrightarrow \sigma^k(g) * \sigma^l(h^*) = 0$$

e poiché il prodotto scalare è invariante rispetto a σ:

$$g \cdot \sigma^{l-k}(h^*) = 0$$

allora $\sigma^{l-k}(h^*) \in C^\perp = \sigma(C^\perp)$ quindi $h^* \in C^\perp$.

$h^*(\alpha)$ è l'immagine in R di un polinomio di grado k, quindi l'immagine del codice duale è generato da h^*, che viene detto *polinomio reciproco* di g.

Osservazione 5.21. Se chiamo $\Gamma = (\gamma_0, \ldots, \gamma_r)$ i coefficienti del polinomio generatore, una matrice generatrice per questo codice è la matrice che ha per colonne $\Gamma, \sigma(\Gamma), \sigma^2(\Gamma), \ldots, \sigma^{r-1}(\Gamma)$. Se $h(x) = \frac{x^n-1}{g(x)}$ e chiamo η_i i suoi coefficienti, segue che

$$h^* = \sum_{i=0}^{k} \eta_{k-i} h^i$$

genera il codice duale, e la matrice di colonne $\eta^*, \sigma(\eta^*), \ldots, \sigma^{k-1}(\eta^*$ è una matrice di controllo per C.

5.1.10 Ciclicità dei codici di Hamming

Problema: quali codici di Hamming di parametri q, r sono ciclici?

Un codice è ciclico se e solo se il suo duale è ciclico. Per questo motivo, poiché i codici di Hamming sono descritti attraverso la matrice di controllo, è più conveniente determinare quando il loro duale è ciclico.

Esempio 5.22. Nel caso $q = 2, r = 3$, la matrice generatrice del duale è

$$\begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$$

Detti v_1, v_2, v_3 i polinomi associati alle righe di H^t, considero $R = \frac{F[x]}{<x^7-1>}$. Sia $g(x)$ generatore del codice duale, allora $v_i(\alpha) \in \psi(C^\perp)$ se e solo se $g(x) \mid v_i(x)$ in $F[x]$ (dove posso scegliere $\mathrm{Gr} v_i(x) < n$), che equivale a $g(\alpha) * u(\alpha) = v_i(\alpha)$. Osservo che

- $v_1(x) = x^3 + x^4 + x^5 + x^6$ $v_1(x) \in C^\perp$ se $v_1(\alpha) \in \psi(C^\perp)$, e lavoro con il polinomio $1 + x + x^2 + x^3$, ottenuto moltiplicando v_1 per α^{-3} (se il primo polinomio sta nell'ideale deve starci anche il secondo che è dato dalla moltiplicazione a sinistra di un elemento dell'ideale per un elemento dell'anello). Allora $1 + x + x^2 + x^3 \mod x^n - 1$ dev'essere un multiplo di $g(x)$.
- il secondo generatore è $v_2(x) = x+x^2+x^5+x^6$ e dividendo per α, ottengo $1+\alpha+\alpha^4+\alpha^5$, anch'esso deve essere multiplo di $g(x)$ modulo $x^n - 1$.
- Cerchiamo $M.C.D.(1 + x + x^2 + x^3, g(x))$. Se l'M.C.D. è 1, allora g è il polinomio costantemente uguale a 1 e si ottiene $R = F[x]$.
- Per quanto già dimostrato precedentemente:

  $$x^7 - 1 = (x - 1)(x^3 + x^2 + 1)(x^3 + x + 1)$$

  (i fattori corrispondono alle 2-classi ciclotomiche)
- Confronto $w(x) = 1 + x + x^2 + x^3$ con i fattori di $x^n - 1$. Si ha che $x - 1 \mid w(x)$, mentre $w(x)$ è coprimo con gli altri fattori. e non ci sono altri fattori comuni perché dividendo per $x - 1$ si ha un polinomio di grado 2 mentre gli altri hanno grado 3. Segue che $M.C.D.(x^7 - 1, w_1(x)) = x - 1$.
- Se ci fosse un ulteriore generatore per $\psi(C)$, esso deve dividere $M.C.D.(x^7 - 1, w_1(x))$, ovvero $g(x) \mid x - 1$, e questo non avviene. Se chiamo D il codice associato a $x - 1$, $\mathrm{Dim}D = 6$, $\mathrm{Dim}D^\perp \geq 6$, e questo è assurdo perché $\mathrm{Dim}D^\perp = 3$.

Concludo che **il duale del codice di Hamming di parametri $q = 2, r = 3$ non è ciclico**.

La risposta al problema precedente dipende quindi dalla matrice di controllo.

Esercizio 5.23. Se F è un campo finito con q elementi, $M.C.D.(n, q) = 1$ e $D \leq F^n$ è il codice la cui immagine è generata da $x - 1$, determinare esplicitamente D.

Dimostrazione. Basta osservare che $\mathrm{Dim}D = n - 1$ (devo sottrarre il grado del generatore), e si deduce che tale codice è il codice di parità. Infatti, gli elementi di $\psi(D)$ sono i multipli in R di $\alpha - 1$. Se considero un polinomio $h(x) = \sum_i \eta_i x^i$,

$$h(x)(x - 1) = \sum_i (\eta_{i-1} - \eta_i)x^i$$

e sommando le componenti si ottiene 0. □

Esercizio 5.24. Sia C un codice ciclico di F^n, $g(x)$ un suo generatore, e chiamo $Z = \{i\, t.c. g(\alpha^i) = 0\}$, dove α è una radice n-esima di 1. Chiamo $t = o_n(q)$. Dimostrare che $c \in C$ se e solo se $c(\alpha^i) = 0 \forall i \in Z$.

Dimostrazione. $1 \longrightarrow 2$: $c \in C$ se e solo se $g(x) \mid c(x)$ in $F[x]$. Siccome $g(\alpha^i) = 0$, anche $c(\alpha^i) = 0 \forall i \in Z$.

$2 \longrightarrow 1$: se $c(x)$ è tale che $c(\alpha^i) = 0 \forall i \in Z$, $x - \alpha^i \mid c(x) \in E[x]$, dove $E \supset F$ e $o(E) = q^t$. Questo implica che $\prod(x - \alpha^i) \mid c(x)$, ma $\prod(x - \alpha^i)$ coincide con $g(x)$, quindi $g \mid c$ e $c \in C$. □

Osservazione 5.25. Tornando all'esercizio precedente, $c \in D$ se $c(\alpha)$ è un multiplo di $g(\alpha)$. Poiché $g(x) = x - 1$, si deve avere $c(1) = 0$. Ma $c(1) = \sum_i \gamma_i$ dove $c = \sum_i \gamma_i x^i$. Quindi $c(1) = 0$ implica che c sta nel codice a somma 0.

Esercizio 5.26. Determinare i codici ciclici di F^n con $M.C.D.(o(F), n) = 1$, di codimensione 1.

Dimostrazione. Si può applicare direttamente la definizione di ciclicità e la descrizione dei codici di codimensione 1 (l'unico vettore ortogonale a tutti gli elementi del codice ha tutte le componenti uguali. Determinare $g(x) \mid x^n - 1$ con grado 1.

□

$C \in F^n$ è ciclico se ogni vettore della base viene fissato da σ.

Esercizio 5.27. Dimostrare che $\sigma^k(g)$ è una combinazione lineare degli spostamenti ciclici del vettore corrispondente, ovvero

$$\sigma^k(g) = \sum_{i=0}^{k-1} \lambda_i \sigma^i(g)$$

e determinare i λ_i.

Dimostrazione. Osservo che

$$\sigma^k(g) = (\gamma_{n-k}, \gamma_{n-k+1}, \dots, \gamma_n, \gamma_0, \gamma_1, \dots, \gamma_{n-k-1})$$

$$i = 0,\ \gamma_0\lambda_0 + \gamma_n\lambda_1 + \gamma_{n-1}\lambda_2 + \cdots + \gamma_{n-k-1}\lambda_{k-1} = \gamma_{n-k}$$

$$= \Lambda \cdot (\gamma_0, \gamma_n, \gamma_{n-1}, \dots, \gamma_{n-k-1})$$

$$i = 1,\ \gamma_n\lambda_0 + \gamma_0\lambda_1 + \gamma_1\lambda_2 + \cdots + \gamma_{n-k-2}\lambda_{k-1} = \gamma_{n-k+1}$$

□

21-12-2015

5.2 Caratterizzazione del duale

Se C è ciclico, anche il suo duale è ciclico e generato da $c^* = x^k * c(x^{-1})$. Se chiamo $c(x)$ un polinomio che sta nel codice generato da $g(x)$, esso soddisfa la relazione $c(x) = 0 \forall x \in Z$, con

$$Z = \{i \in \mathbb{Z}_n \, t.c. \, g(\alpha^i) = 0\}$$

con α radice n-esima primitiva fissata.

Vogliamo caratterizzare $Z^\perp$, e a tale proposito enunciamo la seguente proposizione.

Proposizione 5.28. *Sia Z l'insieme degli esponenti tali che α^i sia una radice di g generatore di un codice ciclico, allora $-Z^c$ è l'insieme degli esponenti tali che α^i siano radici di h^*.*

Osservazione 5.29. Dati due codici ciclici C_1, C_2, con generatori g_1, g_2, denotiamo con Z_1, Z_2 gli insiemi degli esponenti che danno radici di g_1 e g_2 rispettivamente. $g_1R \subset g_2R$ se $g_2 \mid g_1 \in F[x]$, segue allora che $Z_2 \subset Z_1$. In particolare, C è autoortogonale se e solo se $Z \supset Z^\perp = -Z^c$, ed è autoduale se $Z = Z^\perp$, e in quest'ultimo caso $n = 2 * |Z|$.

5.2.1 Limite BCH

Teorema 5.30. *Sia C un codice ciclico, e Z definito come sopra, supponiamo che Z contenga esponenti consecutivi, ad esempio possiamo supporre che Z contenga gli interi dell'intervallo $[b, b+t-1]$. Allora la distanza minima di C è almeno $t+1$.*

Casi limite:

(i) Se $Z = N$, ci sono n termini consecutivi, in questo caso $g(x) = x^n - 1 = 0 \in R$, ovvero il codice corrispondente è ridotto al vettore nullo.

(ii) Se $Z = \emptyset$, $g(x)$ è il polinomio costantemente uguale a 1, e il codice è F^n, che ha distanza minima 1, in questo caso $t = 0$.

Per la dimostrazione di questo fatto, si usano le proprietà della matrice di Vandermonde.

Definizione 5.31. Si dice *matrice di Vandermonde* una matrice quadrata dipendente da w parametri $\beta_1, \ldots, \beta_w$, della forma

$$\begin{matrix} 1 & 1 & \ldots & 1 \\ \beta_1 & \beta_2 & \ldots & \beta_w \\ \ldots & \ldots & \ldots & \ldots \\ \beta_1^{w-1} & \beta_2^{w-1} & \ldots & \beta_w^{w-1} \end{matrix}$$

Per induzione

$$\det W = \prod_{i<j,\, 1\leq i,j\leq w} (\beta_i - \beta_j)$$

e questa matrice è non singolare se e solo se i β_i sono distinti.

dimostrazione del teorema 5.30. Sia C un codice ciclico, e supponiamo che Z contenga $[b, b+1, \ldots, b+t-1]$. Vogliamo mostrare che **se** $c = (\gamma_0, \ldots, \gamma_{n-1}) \in C$ **ha peso** $w \leq t$**, allora** $c = 0$. Supponiamo che $i_1, \ldots, i_w$ siano le posizioni in cui $\gamma_{i_j} \neq 0$. Per ipotesi

$$c(\alpha^i) = 0 \forall i \in Z$$

e poiché Z contiene l'insieme di termini consecutivi,

$$c(\alpha^i) = 0 \forall i \in [b, b+1, \ldots, b+t-1].$$

Inoltre

$$c(x) = \sum_{j=1}^{w} \gamma_{i_j} x^{i_j}$$

e scrivo esplicitamente la condizione $c(\alpha^i) = 0 \forall i = b, \ldots, b+t-1$.

$$c(\alpha^b) = 0 \longrightarrow \gamma_{i_1}\alpha^{i_1 b} + \gamma_{i_2}\alpha^{i_2 b} + \cdots + \gamma_{i_w}\alpha^{i_w b} = 0$$

$$c(\alpha^{b+1}) = 0 \longrightarrow \gamma_{i_1}\alpha^{i_1 (b+1)} + \cdots + \gamma_{i_w}\alpha^{i_w (b+1)} = 0$$

$$\cdots\cdots\cdots$$

$$c(\alpha^{b+t-1}) = 0 \longrightarrow \gamma_{i_1}\alpha^{i_1 (b+t-1)} + \cdots + \gamma_{i_w}\alpha^{i_w (b+t-1)} = 0$$

Costruisco la matrice che ha per colonne i coefficienti dei polinomi rispetto al vettore $\gamma_{i_1}, \ldots, \gamma_{i_w}$.

$$\begin{matrix} \alpha^{i_1 b} & \alpha^{i_2 b} & \ldots & \alpha^{i_w b} \\ \alpha^{i_1 (b+1)} & \alpha^{i_2 (b+1)} & \ldots & \alpha^{i_w (b+1)} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha^{i_1 (b+t-1)} & \alpha^{i_2 (b+t-1)} & \ldots & \alpha^{i_w (b+t-1)} \end{matrix}$$

e ponendo $w_j = \alpha^{i_j}$:

$$\begin{matrix} \omega_1^b & \omega_2^b & \ldots & \omega_w^b \\ \omega_1^{b+1} & \omega_2^{b+1} & \ldots & \omega_w^{b+1} \\ \ldots & \ldots & \ldots & \ldots \\ \omega_1^{b+t-1} & \omega_2^{b+t-1} & \ldots & \omega_w^{b+t-1} \end{matrix}$$

che si riscrive come

$$\begin{matrix} 1 & 1 & \ldots & 1 \\ \omega_1 & \omega_2 & \ldots & \omega_w \\ \ldots & \ldots & \ldots & \ldots \\ \omega_1^{t-1} & \omega_2^{t-1} & \ldots & \omega_w^{t-1} \end{matrix} \times \begin{matrix} \omega_1^b \\ \omega_2^b \\ \ldots \\ \omega_w^b \end{matrix}$$

e quindi compare la matrice di Vandermonde.

$$\det M = \prod_j \beta_j^b \det W_{\beta_1, \ldots, \beta_w}.$$

I β_j sono non nulli e distinti perché α è primitiva, allora $\det M \neq 0$, e la soluzione del sistema associato a questa matrice è unica, ed è il vettore nullo, ovvero $\gamma_j = 0 \forall j$. Segue che

$$c(\alpha^k) = 0 \forall k = b, \ldots b + w$$

Questo è assurdo, perché per ipotesi tutti gli elementi γ_j sono non nulli, mentre la non singolarità di M implica che siano tutti nulli. E' assurdo assumere che esista una parola di peso $w \neq 0$ che sia non nulla, allora $w(c) > t + 1$. □

Definizione 5.32. Un codice C ciclico si dice *di distanza assegnata* $t + 1$ se Z contiene un insieme con $t + 1$ termini consecutivi in $\mathbb{Z}_n$.

Queste sono alcune applicazioni del teorema appena mostrato:

Esempio 5.33. Se considero il codice di ripetizione $Rep_n(F)$, esso è ciclico e ha distanza n, il suo generatore è $\frac{x^n - 1}{x - 1}$. Allora $Z = \{1, 2, \ldots, n - 1\}$ (non devo considerare 0 che corrisponde alla radice 1). Segue che $d \geq CHB = n - 1$

Esempio 5.34. Considero il codice ciclico generato da $x^3 + x + 1$ su F_2, con lunghezza $n = 7$. $3 = o_7(2)$, quindi considero $E = F_8$ che contiene le radici primitive di 7. $k = 7 - \mathrm{Gr}g = 4$, gli zeri di questo polinomio sono ottenuti considerando l'unione di due classi ciclotomiche.

$$C_1 = \{1, 2, 4\},\ C_2 = \{3, 5, 6\},\ C_3 = \{0\}$$

Se prendo $\alpha = x + I$, con $I = (x^3 + x + 1)$ generato su $F_2[x]$, segue che $Z = \{1, 2, 4\}$. In questa classe ciclotomica ci sono due termini consecutivi, quindi $d \geq 3$. Per il teorema di Singleton, $d = 3, 4$, ma $d \neq 4$ perché la cardinalità delle bolle di raggio 1 è ≥ 8, e questo non è possibile.

C è un (7,4,3)-codice, e affermo che C **è il codice di Hamming**. Infatti, una sua matrice di controllo H ha 7 righe e 3 colonne, ma $7 = \frac{2^3-1}{2-1}$ è l'ordine del piano proiettivo su F_2. Allora queste righe non possono essere proporzionali perché $d = 3$. Segue allora che H è la matrice di un codice di Hamming.

Esercizio 5.35. Costruire una permutazione (o una trasformazione monomiale) che porta C nel codice di Hamming di parametri $3, 2$ con righe che codificano lo sviluppo binario dei numeri da 1 a 7 (tale codice non è ciclico per una dimostrazione precedente).

Problema: BCH è il meglio possibile?

Teorema 5.36 (limite di Hartmann e Tseng)**.** *Se Z contiene s sequenze di elementi consecutivi, allora $d \geq t + s$.*

La scelta di α gioca un ruolo fondamentale: se α è una radice n-esima primitiva, α^a è ancora n-esima primitiva se $M.C.D.(a, n) = 1$.

Supponiamo che $Z = \{a, 2a, 3a\}$, se $M.C.D.(a, n) = 1$, Z contiene $1, 2, 3$. Allora il limite BCH si rienuncia dicendo che la distanzaminima di un codice ciclico è $\geq t + 1$ dove t è il massimo numero di elementi consecutivi modulo n in $\mathbb{Z} * a$, $a \in \mathbb{Z}_n$.

5.3 Codice di Gole

Il primo codice di Gole è un (23,12,7)-codice e quindi ha capacità correttiva $e = 3$. Il gruppo degli automorfismi di G_i è una variante del gruppo di Matthew.

I gruppi di Matthew sono sottogruppi del gruppo simmetrico su i elementi. M_{24} è un gruppo 5-transitivo, ovvero per ogni lista $(a_1, \ldots, a_5), (b_1, \ldots, b_5)$, con $a_i \in \{1, \ldots, 24\}$ esiste un unico elemento in M_{24} che manda ogni a_i in b_i ordinatamente. M_{12} invece è un gruppo 4-transitivo.

Se G è un gruppo n-transitivo, esiste $g \in G$ tale che $o(g) = n$.

Esempio 5.37. Il gruppo simmetrico su m lettere con $m \geq n$, è n-transitivo, perché la permutazione $a_1 \mapsto b_1, \ldots, a_n \mapsto b_n$ ha ordine n.

Gli unici gruppi 4- o 5-transitivi oltre al gruppo simmetrico e al gruppo alterno sono M_{24}, M_{23}, M_{12}.

Mostriamo che è possibile costruire un codice ciclico su F_2 di lunghezza 23 che risulti essere 3-perfetto. Per un esercizio precedente, sappiamo che

$$x^{23} - 1 = (x - 1)g_1(x)g_5(x)$$

dove gli zeri di $g_1(x)$ sono gli elementi della 2-classe ciclotomica C_1 e gli zeri di g_2 sono gli elementi di C_5 in Z_{23}. In particolare

$$C_1 = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$$

Ci sono $t = 4$ termini consecutivi in C_1, quindi, detto C_1 il codice generato da g_1, $d_C \geq 5$.

$$C_5 = \{5, 7, 10, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22\}$$

Anche qui $t = 4$. Sappiamo che $\varphi(23) = 22$ che ha come divisori 2 e 11. Inoltre $2^{11} \equiv 1 \mod 23$, quindi $t = 11$ e considero $E = F_{2^{11}}$. $\alpha = \gamma^{\frac{2^{11}-1}{23}}$ con γ elemento primitivo. 2 ha ordine 11, se considero $-2 = 21$, il suo ordine è 22, quindi -2 è un generatore del gruppo.

Esercizio 5.38. Dimostrare che ogni elemento di C_5 è un generatore di Z_{23}^*.

La variante del limite BCH non è applicabile.

5.3.1 Insieme degli elementi di modulo pari

Dato un codice ciclico C, definisco

$$C_e = \{c \in C \, t.c. \, w(c) \equiv 0 \mod 2\}$$

C_e è ancora ciclico perché gli shiftati di un vettore hanno lo stesso peso del vettore di partenza, inoltre è un sottocodice di C, allora il suo generatore $g_e(x)$ è divisibile per $g(x)$.

Dato $c \in C_e$,

$$c = \sum_{i=0}^{n-1} \gamma_i x^i$$

allora $w(C)$ è pari solo se $c(1) = 0$, perché $c(1) = \sum_i \gamma_i$ e γ_i assume solo i valori 1 e 0. Allora la relazione tra Z e Z_e è $Z_e = Z \cup \{0\}$, perché 1 è una radice del nuovo codice.

Nell'esempio precedente (codice di Gole), mostriamo che C_e **è autoortogonale**, ovvero mostriamo che $Z_e \supset -Z_e^c$.

$$Z_e = \{0, 1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$$

$$Z_e^c = \{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}$$

$$-Z_e^c = \{18, 16, 13, 12, 9, 8, 6, 4, 3, 2, 1\}$$

Proposizione 5.39. *Ogni elemento di C_E ha peso multiplo di 4.*

Dimostrazione. Sia

$$C_d = \{c \in C_e t.c. w(c) \equiv 0 \mod 4\}.$$

Dato un codice binario, $w_H(x + y) = w_H(x) + w_H(y) - 2w_H(x \cap y)$ Inoltre $w(x \cap y) \equiv 2 \mod x \cdot y$. C_e autoortogonale implica che $x \cap y \equiv 0$, allora se $x, y \in C_d$ segue che $x + y \in C_d$ (infatti $4 = W(x + y) = W(x) + W(y)$), quindi C_d è un realtà un sottospazio vettoriale. Se $C_d \neq C_e$, $\mathrm{Dim}_{F_2} C_d = \mathrm{Dim} C_e - 1$, ma poiché $C_d \subset C_e$, e poiché C_d è ciclico, gli zeri di C_d si ottengono aggiungendo un elemento a a Z_e. Mostriamo che questo non può avvenire.

Z_d e Z_e sono unione di due classi ciclotomiche. Allora a è una 2-classe ciclotomica, e questo significa che a viene fissato da 2, ma questo implica che $a \equiv 0 \mod 2$, e questo è impossibile perché si ha già $0 \in Z_e$, allora $Z_d = Z_e$. □

A priori, le parole di C_e devono avere peso maggiore di 5 e multiplo di 4, ovvero possono avere pesi 8,12,16,20, in realtà non tutte queste possibilità possono verificarsi. $\frac{x^{23}-1}{x-1} \equiv 0 \mod g_1(x)$. Il codice generato da $x - 1$ è il codice di ripetizione, quindi contiene $(1, \ldots, 1, 1)$. Se c è una parola di peso 20 $1 + c$ che sta in C ha peso 3, e questo è impossibile perché le parole del codice devono avere peso maggiore di 4. Segue allora che ci possono solo essere parole di peso 8,12,16.

Capitolo 6

FORMULA DI MACWILLIAMS

7-1-2016

6.1 Teorema di Macwilliams e sue generalizzazioni

Sia F un campo, allora il teorema di Macwilliams afferma che il problema della determinazione della distribuzione dei pesi delle parole di C è legato alla determinazione dei pesi delle parole di $C^\perp$ in un senso che preciseremo.

Definizione 6.1. Dato un codice lineare C, sia a_i il numero delle parole di peso i, chiamo *polinomio enumeratore* del codice

$$w_C(z) = \sum_{i=0}^{n} a_i z^i,$$

dove z è una variabile. $w_C(z)$ è detta *funzione generatrice* associata alla sequenza $\{a_i\}$.

Ci si chiede se, partendo da un codice C, è possibile ottenere informazioni sul suo polinomio enumeratore, o se si ottengono più facilmente informazioni analizzando il polinomio enumeratore del duale. Questo porterebbe vantaggi nel caso di codici di grandi dimensioni, perché se $\mathrm{Dim}C > n/2$ con n lunghezza del codice, segue che $\mathrm{Dim}C^\perp < n/2$, e quindi si lavora su polinomi di grado minore.

Si introduce un nuovo polinomio

$$w_C(x, y) := x^n W_c(y/x) = \sum_{i=0}^{n} a_i x^{n-i} y^i$$

che è detto *polinomio enumeratore omogeneo* (ovvero è tale che $p(tx) = t^n * p(x)$.

Teorema 6.2 (teorema di Macwilliams)**.** *Il polinomio enumeratore del codice duale è dato da*

$$1/o(C) * w_C(x + (s-1)y, x - y)$$

dove s è la cardinalità di F.

Chiamo T la matrice di trasformazione

$$\begin{pmatrix} 1 & 1 \\ s-1 & -1 \end{pmatrix}$$

e pongo $(x', y') = (x, y)T$.

Osservazione 6.3. Supponiamo che il teorema di Macwilliams sia vero, applicandolo a C segue che il polinomio enumeratore di C è dato da

$$w_C(x, y) = w_{(C^\perp)^\perp}(x, y) =$$

$$= 1/o(C^\perp) w_{C^\perp}((x, y)T)$$

$$= 1/o(C^\perp) 1/o(C) w_C(((x, y)T^2)$$

e $T^2 = \mathrm{diag}(s, s)$, quindi

$$= 1/o(C) * 1/o(C^{\perp}) w_C(s(x, y))$$

e passando al polinomio enumeratore non omogeneo:

$$= 1/o(C) * 1/o(C^{\perp}) s^n w_C(x, y)$$

Inoltre $o(C) * o(C^{\perp}) = s^{n-k} * s^k = s^n$. Allora

$$= 1/s^n * s^n w_C(x, y) = w_C(x, y)$$

come dovevamo ottenere, e quindi è ragionevole supporre che il teorema di Macwilliams valga.

6.1.1 Generalizzazioni

Teorema 6.4 (Versione non omogenea del teorema di Macwilliams)**.** *Sia b_i la cardinalità delle parole di $C^{\perp}$ con peso uguale a i. Allora*

$$w_{C^{\perp}}(z) = \sum_{m=0}^{n} b_m z^m = 1/o(C) \sum_i a_i (1 + (s-1)z)^{n-i} (1-z)^i.$$

Per un codice su un alfabeto generico vale la seguente versione del teorema:

Teorema 6.5. *Sia C un codice a blocchi definito su un alfabeto A di lunghezza n e cardinalità s, sia $w_C(z) = 1/o(C) \sum_i a_i z^i$ il polinomio enumeratore, con*

$$\bar{a}_i = |\{(c, d)\, t.c.\, d_H(c, d) = i\}|$$

e $\bar{b}_i$ definiti analogamente ma relativi a $C^{\perp}$ (non ha senso parlare di pesi perché non è detto che il codice sia lineare) Allora, ponendo

$$\bar{W}_{C^{\perp}}(z) := 1/o(C^{\perp}) \sum_{m=0}^{n} \bar{b}_m z^m$$

segue che

$$\bar{w}_{C^{\perp}}(z) = 1/o(C) \sum_{i=0}^{n} \bar{a}_i (1 + (s-1)z)^{n-i} (1-z)^i$$

e che $b_m \geq 0 \forall m$.

Osservazione 6.6. Se C è un codice lineare, il polinomio definito

$$\bar{w}_C(z) = 1/o(C) \sum_{i=0}^{n} \bar{a}_i z^i$$

coincide con $w_C(z)$. Infatti

$$\bar{a}_i = |\{(c, d)\, t.c.\, d_H(c, d) = i\}| = |C| * \{c \in C\, t.c.\, w_H(c) = i\}$$

ovvero esiste una corrispondenza "biunivoca" tra (c, d) con $d_H(c, d) = i$ e $c - d$ con $w_H(c - d) = i$. Infatti

$$d_H(c, d) = d_H(c - d, 0) = w_H(c - d)$$

Viceversa, dato t di peso i e $d \in C$, si ha $c := t + d \in C$ e $d_H(c, d) = w_H(t)$. Tale coppia può essere costruita in $|C|$ modi (uno per ogni scelta di $d \in C$). Quindi

$$\bar{a}_i = |C| * a_i$$

perciò

$$w_C(z) = \sum_i a_i z^i = 1/o(C) \sum_i \bar{a}_i z^i = \bar{w}_C(z).$$

6.1.2 Nozione di carattere

Per poter dare una dimostrazione del teorema di Macwilliams è necessario introdurre la nozione di carattere.

Definizione 6.7. Sia F un gruppo additivo (abeliano), si dice *carattere* di F un omomorfismo $\chi : F \to \mathbb{C}^*$, tale che dati $a, b \in F$, $\chi(a + b) = \chi(a) * \chi(b)$.

Se F è finito e ha cardinalità s, $\chi(a)^s = 1 \forall a \in F$ (ovvero l'immagine di F contiene radici dell'unità). Infatti, $sa = 0$, quindi $\chi(0) = \chi(sa) = \chi(a)^s$, e quindi $\chi(a)$ è una radice s-esima dell'unità.

Un caso particolare dell'osservazione precedente è il seguente: se F è un campo finito, e $o(F) = p^d$, allora per ogni $a \in F$, e per ogni carattere χ, $\chi(a)^p = 1$ (questo avviene perché in un campo di ordine p^d, $pa = 0$).

Si dà una struttura di gruppo all'insieme dei caratteri nel seguente modo: dati due omomorfismi χ, ψ, posso definire la loro composizione $\chi \circ \psi(a) := \chi(a) * \psi(a)$.

Proposizione 6.8. *$\chi \circ \psi$ è ancora un carattere.*

Dimostrazione.

$$\chi \circ \psi(a + b) = \chi(a + b) * \psi(a + b) = \chi(a) * \chi(b) * \psi(a) * \psi(b)$$

e per la commutatività in $\mathbb{C}$:b

$$= \chi(a) * \psi(a) * \chi(b) * \psi(b) = \chi \circ \psi(a) * \chi \circ \psi(b).$$

□

Si osserva inoltre che $\chi^{-1}(a) := (\chi(a))^{-1}$.

Teorema 6.9. *Se F è un campo, esiste un carattere χ tale che per ogni $a \neq 0 \in F$ esiste b tale che $\chi(ab) \neq id$.*

Osservazione 6.10. Si ha sempre $\chi(0) = 1$, e per ogni b, $\chi(0 * b) = 1$, quindi per $a = 0$ non esiste l'elemento b con le proprietà del teorema.

Dimostrazione. Considero il campo F di ordine p^d, allora esiste una base $\{v_1, \ldots, v_d\}$ di F su F_p, quindi $F \cong \otimes_{i=1}^d < v_i >$, con

$$< v_i >= \{\lambda v_i, \lambda \in F_p\}.$$

Dato $v = \sum_i \lambda_i v_i$ e un omomorfismo ϕ:

$$\phi(v) = \phi(\sum_i \lambda_i v_i) = \prod_i \phi(\lambda_i v_i) = \prod_i (\phi(v_i))^{\lambda_i}.$$

Se pongo $\omega_i = \phi(v_i)$, essa è una radice p-esima dell'unità, e posso definire un nuovo carattere χ_i tale che

$$\chi_i :< v_i > \to C^* : \chi_i(v_i) = \omega_i.$$

Estendendo χ_i a tutto F ponendo $\chi_i(v_j) = 1$ per $i \neq j$, si mostra che ϕ è il prodotto puntuale dei χ_i. Pongo $\chi = \chi_1$, con $\chi_1(v_1) = \omega$ con ω radice p-esima primitiva (ovvero diversa dall'unità).

Dato $a \neq 0$, considero $b = a^{-1}v_1$, allora $ab = v_1$, e

$$\chi(ab) = \chi_1(v_1) \neq 1$$

per ipotesi. □

Definizione 6.11. Sia $V = F^n$ uno spazio vettoriale, e sia χ un carattere arbitrario di F. Costruisco la *forma biadditiva* $\mathcal{B}_\chi : V \times V \to \mathbb{C}$ ponendo, per ogni v, w:

$$\mathcal{B}_\chi(v, w) = \chi(v \cdot w).$$

Data una forma biettiva, deve valere la proprietà

$$\mathcal{B}_\chi(u + u', v) = \mathcal{B}_\chi(u, v) * \mathcal{B}_\chi(u', v).$$

Definizione 6.12. Si chiama *radicale* di $\mathcal{B}_\chi$ l'insieme

$$\mathbf{Rad}\mathcal{B}_\chi = \{u \in V \, t.c. \, \mathcal{B}_\chi(u, v) = 1 \forall v \in V\}.$$

Proposizione 6.13. *Se χ è il carattere del teorema 6.9, allora $\mathbf{Rad}\mathcal{B}_\chi$ è ridotto al vettore nullo. (In tal caso si dice che la forma biadditiva è* non degenere*).*

Dimostrazione. Sia $u = (u_1, \ldots, u_n) \neq 0$, allora esiste i tale che $u_i \neq 0$. Sia $v = e_i$ l'i-esimo vettore della base standard, allora $u \cdot v = u_i \neq 0$, e per il teorema 6.9 esiste $b \in F$ tale che

$$1 \neq \chi(bu_i) = \mathcal{B}_\chi(u, be_i)$$

Allora l'unico vettore tale che $\mathcal{B}_\chi(u, v) = 1 \forall v$ è il vettore nullo. □

Osservazione 6.14. Se χ è il carattere banale, $\mathcal{B}_\chi$ è non degenere e $\mathbf{Rad}b_\chi = V$.

11-1-2016

6.1.3 Leggi di ortogonalità

Sia W un sottospazio di V, allora valgono le seguenti proprietà:

$$\sum_{u \in V} \mathcal{B}_\chi(u, v) = \begin{cases} 0 & \iff v \neq 0 \\ |V| & \iff v = 0 \end{cases} \tag{6.1}$$

$$\sum_{u\in W} \mathcal{B}_\chi(u,v) = \begin{cases} 0 & \iff v \in W \\ |V| & \iff v \in W^\perp \end{cases} \tag{6.2}$$

dove

$$W^\perp = \{v\, t.c. \mathcal{B}_\chi(v,w) = 1 \forall w \in W\}.$$

Notazione: Dati $A, B \subseteq V$, denoto con

$$\chi(A|B) = \sum_{a\in A, b\in B} \mathcal{B}_\chi(a,b)$$

e con questa notazione la relazione di ortogonalità si scrive in maniera compatta come:

$$\chi(W|v) == \begin{cases} 0 & \iff v \in W \\ |V| & \iff v \in W^\perp \end{cases}$$

Osservazione 6.15. Poiché χ è non degenere, $\mathbf{Rad}\mathcal{B}_\chi$ coincide con $W^\perp$. Se $W = V$, l'affermazione $v \notin W^\perp$ equivale a $v \neq 0$, quindi la prima formula è un caso particolare della seconda.

dimostrazione delle relazioni di ortogonalità. Per l'osservazione precedente basta dimostrare la formula (6.2).

Sia $v \in W^\perp$, allora

$$\mathcal{B}_\chi(w,v) = \chi(wv) = 1 \forall w \in W$$

allora

$$\sum_{w\in W} \mathcal{B}_\chi(w,v) = \sum_{w\in W} 1 = |W|.$$

Sia ora $v \notin W^\perp$, allora esisterà almeno un vettore $w \in W$ tale che $\mathcal{B}_\chi(w,v) \neq 1$. Si può scrivere

$$\chi(W|v) = \chi(W + w|v) = \sum_{u\in W} \mathcal{B}_\chi(u + w, v)$$

$$= \sum_{u\in W} \mathcal{B}_\chi(u,v) * \mathcal{B}_\chi(w,v)$$

$$= \mathcal{B}_\chi(w,v) * \chi(W|v)$$

allora $\chi(W|v) = \mathcal{B}_\chi(w,v) * \chi(W|v)$, e poiché abbiamo scelto w tale che $\mathcal{B}_\chi(v,w) \neq 1$, l'unica possibilità è che $\chi(W|v) = 0$. □

6.1.4 Lemma preliminare

Lemma 6.16. *Sia $u \in V$ tale che $w_H(u) = i$, allora*

$$\sum_{v\in V} z^{w_H(v)} \mathcal{B}_\chi(v,u) = (1 + (s-1)z)^{n-i} * (1-z)^i$$

Dimostrazione. Stimiamo la quantità da calcolare:

$$\sum_{v\in V} z^{w_H(v)}\chi(u\cdot v) =$$

ed espandendo il prodotto scalare

$$= \sum_{v\in V} z^{w_H(v)}\chi(\sum_j u_j v_j) =$$

$$= \sum_{v\in V} z^{w_H(v)} \prod_{j=1}^{n} \chi(u_j v_j)$$

e poiché $z^{w_H(v)} = \prod_{j=1}^{n} z^{w_H(v_j)}$:

$$= \sum_{v\in V} \prod_{j=1}^{n} z^{w_H(v_j)}\chi(u_j v_j)$$

e scambiando somma e prodotto:

$$= \prod_{j=1}^{n} \sum_{v\in V} z^{w_H(v_j)}\chi(v_j u_j)$$

$$= \prod_{j=1}^{n} \sum_{\mu\in F} z^{w_H(\mu)}\chi(u_j\mu)$$

Abbiamo quindi ottenuto

$$\sum_{v\in V} z^{w_H(v)}\chi(u\cdot v) = \prod_{j=1}^{n} \sum_{\mu\in F} z^{w_H(\mu)}\chi(u_j\mu) \tag{6.3}$$

Lavorando su

$$\sum_{\mu\in F} z^{w_H(\mu)}\chi(u_j\mu)$$

Se $\mu = 0$, si ottiene 1. Altrimenti $w_H(\mu) = 1$ e rimane

$$z \sum_{\mu\in F^*} \chi(u_j\mu).$$

Se $u_j = 0$, $\chi(0) = 1$ e la somma vale $s-1$. Se invece $u_j \neq 0$, quando μ varia in F^*, anche μu_j varia in F^*, quindi rimane

$$z \sum_{\nu\in F^*} \chi(\nu) = \sum_{\nu\in F} \chi(\nu) - 1$$
$$= \chi(F|1) - 1 = 0 - 1 = -1$$

per la formula 6.1 con $v = 1 \neq 0$.

Unendo queste informazioni segue che,

$$\sum_{\mu\in F} z^{w_H(\mu)}\chi(u_j\mu)$$

$$= \begin{cases} 1 + z(s-1) & \iff u_j = 0 \\ 1 - z & \iff u_j \neq 0 \end{cases}$$

allora tornando alla formula 6.3 nella produttoria $1-z$ compare $w_H(u) = i$ volte, mentre $1 + (s-1)z$ compare $n - w_H(u) = n - i$ volte, quindi vale l'asserto. □

6.1.5 Dimostrazione del teorema di Macwilliams

Possiamo ora dimostrare il teorema di Macwilliams:

Teorema 6.17. *Sia C un codice lineare su F^n, con polinomio enumeratore*

$$w_C(z) = \sum_{i=0}^{n} a_i z^i$$

allora il polinomio enumeratore del codice duale è

$$1/o(C) \sum_{i=0}^{n} (1 + (s-1)z)^{n-i}(1-z)^i.$$

Osservazione 6.18. Questo teorema ha anche una versione per codici non lineari: se C è un codice qualsiasi con una struttura di gruppo additivo, per il quale esiste un carattere non degenere, e poniamo

$$\bar{a}_i = \{(c,d) t.c. d_H(c,d) = i\}$$

allora i coefficienti b_m del polinomio

$$\bar{w}_{C^\perp}(z) = 1/o(C) \sum_i a_i (1 + (s-1)z)^{n-i}(1-z)^i$$

sono razionali e positivi.

dimostrazione del teorema di Macwilliams. Valutiamo l'espressione

$$\sum_{c,d \in C} \sum_{u \in V} z^{w_H(u)} \mathcal{B}_\chi(u, c-d) \tag{6.4}$$

per il lemma precedente

$$= \sum_{c,d \in C} (1 + (s-1)z)^{n - w_H(c-d)} * (1-z)^{w_H(c-d)}$$

$$= \sum_{c \in C} \sum_{d \in C \, t.c. \, w_H(c-d) = i} (1 + (s-1)z)^{n - w_H(c-d)} * (1-z)^{w_H(c-d)}$$

e se C è lineare, fissato un intero $i \in \{0, \ldots, n\}$, segue che le coppie $c - d$ con $w_H(c-d) = i$ sono

$$|C| * |\{c \, t.c. \, w_H(c) = i\}| = |C| * a_i,$$

quindi

$$= o(C) * \sum_{i=0}^{n} (1 + (s-1)z)^{n-i} * (1-z)^i \tag{6.5}$$

Scambiando le sommatorie in (6.4), valuto

$$\sum_{u \in V} \sum_{c,d \in C} z^{w_H(u)} \mathcal{B}_\chi(u, c-d)$$

$$= \sum_{u \in V} z^{w_H(u)} \sum_{c,d \in C} \mathcal{B}_\chi(u,c) \mathcal{B}_\chi(u,-d)$$

$$= \sum_{u \in V} z^{w_H(u)} \sum_{c,d \in C} \mathcal{B}_\chi(u,c)\overline{\mathcal{B}_\chi(u,d)}$$

$$= \sum_{u \in V} z^{w_H(u)} \sum_{c \in C} \mathcal{B}_\chi(u,c) \sum_{d \in C} \overline{\mathcal{B}_\chi(u,d)}$$

e applicando ripetutamente le relazioni di ortogonalità:

$$= \sum_{u \in V t.c. u \in C^\perp} z^{w_H(u)} \sum_{c \in C} \mathcal{B}_\chi(u,c) o(C)$$

$$= o(C)^2 \sum_{u \in V t.c. u \in C^\perp} z^{w_H(u)}$$

e detto b_m il numero di vettori di $C^\perp$ di peso m

$$= o(C)^2 \sum_m b_m z^m \tag{6.6}$$

Eguagliando le due espressioni equivalenti a (6.4), ovvero le formule (6.5) e (6.6) segue che

$$o(C)^2 \sum_m b_m z^m = o(C) * \sum_{i=0}^{n} (1 + (s-1)z)^{n-i} * (1-z)^i$$

$$\longrightarrow w_{C^\perp}(z) = \sum_m b_m z^m = 1/o(C) * \sum_{i=0}^{n} (1 + (s-1)z)^{n-i} * (1-z)^i$$

e segue la dimostrazione del teorema. □

Nel caso non lineare, svolgendo gli stessi conti segue che

$$b_m = 1/o(C)^2 \sum_{v \in V \, t.c. \, w_H(v) = m} |\chi(v \cdot c)|^2$$

6.1.6 Applicazioni

Esempio 6.19. Se considero il codice costituito da tutti gli elementi di $V = F_{q^n}$,

$$a_i = |\{v \in F_{q^n} \, t.c. w_H(v) = i\}| = (q-1)^i \binom{n}{i}$$

allora il polinomio enumeratore è

$$\sum_i a_i z^i = \sum_i \binom{n}{i} (q-1)^i z^i = (1 + (q-1)z)^n$$

Un altro modo per ricavare questo risultato applicando il teorema di Macwilliams è il seguente: se V è il codice nullo, che è il duale di F_{q^n}, il codice ha un'unica parola di peso 0, allora il suo polinomio enumeratore è costantemente uguale a 1, e si ha che il polinomio del duale è

$$1/o(V) \sum_i a_i (1-z)^i (1-(q-1)z)^{n-i}$$

e sopravvive solo il termine per $i = 0$, ovvero $(1-(q-1)z)^n$, come confermato dal calcolo precedente.

Esempio 6.20. Vogliamo determinare il polinomio enumeratore del duale del codice di ripetizione, che è il parity check code. Il codice di ripetizione ha un'unica parola di peso 0, e tutte le altre di peso n, quindi

$$a_0 = 1,\ a_i = 0 \forall i = 1, \dots, n-1,\ a_n = q-1$$

$$\longrightarrow\ w_C(z) = 1 + (q-1)z^n$$

$$w_{C^\perp} = 1/q * \sum_{i=0}^{n} a_i (1 + (s-1)z)^{n-i}(1-z)^i$$

e sopravvivono solo i termini per $i = 0$ e $i = n$:

$$= 1/q((1 + (q-1)z)^n + (q-1)(1-z)^n)$$

e sviluppando come binomio di Newton:

$$= 1/q * (\sum_{i=0}^{n} \binom{n}{i}((q-1)z)^i + (q-1)\sum_{i=0}^{n}\binom{n}{i} z^i(-1)^i)$$

allora

$$b_i = \binom{n}{i} z^i \frac{(q-1)^i + (-1)^i(q-1)}{q}$$

Questo polinomio è sempre a coefficienti interi nella variabile z.

Osservazione 6.21. Esistono polinomi in una variabile tali che $z \in \mathbb{Z} \mapsto f(z) \in \mathbb{Z}$ ma $f(x) \notin \mathbb{Z}[x]$. Considero ad esempio il polinomio

$$f(x) = q * (q-1)/2 = q^2/2 - q/2$$

esso non ha coefficienti interi, ma si ottiene sempre un valore intero quando lo si valuta in un intero, perché questo polinomio valutato in a è uguale a $\binom{a}{2}$ che è intero.

Se f è una funzione che soddisfa questa proprietà, allora è una $\mathbb{Z}$-combinazione di polinomi della forma indicata sopra (di coefficienti binomiali).

Esempio 6.22. I codici di Hamming hanno peso costante, e il loro peso è il numero di vettori che stanno nella versione proiettiva di un iperpiano. L'ordine del codice duale del codice di Hamming è q^r (ha dimensione r), allora $b_0 = 1$ e $b_{q^{r-1}} = q^r - 1$, quindi

$$w_{C^\perp} = 1 + q^{r-1}z^{q^r-1}$$

Il polinomio enumeratore del codice di Hamming originario è

$$1/q^r * \{(1 + (q-1)z)^n + (q^r-1)(1 + (q-1)z)^{n-q^{r-1}}) * (1-z)^{q^{r-1}})$$

6.1.7 Classificazione dei codici perfetti (lineari)

Definizione 6.23. Un codice si dice *e-perfetto* su un alfabeto di s elementi se

$$o(B_e(c)) = \sum_{i=0}^{e} \binom{n}{i}(s-1)^i \mid s^n$$

La capacità correttiva dei codici e-perfetti visti finora è 1 per i codici di Hamming, 2 per il codice di Golay G_{11} e 3 per il codice di Golay G_{23}. I codici di Golay sono gli unici codici non lineari che sono e-perfetti e correggono più di un errore.

Per i codici di Hamming

$$n = \frac{q^r - 1}{q - 1}$$

$$n - 1 = \frac{q^r - q}{q - 1} = q * \frac{q^{r-1} - 1}{q - 1}$$

quindi $\frac{n-1}{q} = \frac{q^{r-1}-1}{q-1}$. Il polinomio enumeratore di un codice e-perfetto ha grado $a_{2e+1} = 0$, e ha espressione

$$w_C(z) = (1 - z)^{n-(n-1)/q} * (1 + (q - 1)z)^{(n-1)/q}$$

Dato il polinomio

$$(1 + (s - 1)z))^{n-i}(1 - z)^i$$

$$= [\sum_{k=0}^{n-i}((s - 1)z)^{n-i-k}][\sum_{h=0}^{i}\binom{i}{h}z^h(-1)^h]$$

il coefficiente del termine di grado m è dato da

$$\sum_{k=0}^{m}\binom{n-i}{m-k}(q-1)^{m-k}(-1)^k\binom{i}{k}$$

Definizione 6.24. Definisco dei polinomi dipendenti da una serie di parametri: m, n, x.

$$K_m(x, n, q) = \sum_{k=0}^{m}\binom{n-x}{m-k}\binom{x}{k}(s-1)^{m-k}(-1)^k$$

detti *polinomi di Krawtchouk*.

Se il polinomio enumeratore di un codice è

$$\sum_i a_i z^i$$

e il polinomio del duale è

$$\sum_m b_m z^m$$

per le osservazioni precedenti posso esprimere il teorema di Macwilliams affermando che

$$b_i = \sum_i a_i K_m(i, n, s)$$

Teorema 6.25 (Teorema di Delsart)**.** *Se esiste un codice non necessariamente lineare, con distanza minima d su F_q di lunghezza n, i b_m devono essere positivi, ovvero*

$$\sum_i a_i K_m(i, s, n) \geq 0, \forall m = 0, \ldots, n$$

e gli a_i sono interi.

Osservazione 6.26. I polinomi di Krawtchouk entrano in gioco nel calcolo di $\chi(V|u)$ nel caso di codici e-perfetti. Dato un codice e-perfetto, lo spazio vettoriale F^n può essere visto come unione disgiunta di

$$V = B_e(0) \oplus C$$

Dato un vettore qualsiasi u con peso di Hamming i, segue che

$$\chi(V|u) = \chi(B_e(0) + C|u) = \chi(B|u) * \chi(C|u)$$

La bolla può essere considerata come unione disgiunta di gusci per $t = 0, \ldots, e$, ovvero

$$B_e(0) = \bigcup_{t=1}^{e} S_t(0)$$

quindi

$$\chi(V|u) = \sum_t \chi(S_t(0)|u) * \chi(C|u) = \sum_t K_t(i, n, s)$$

Pongo

$$l_e(x) = \sum_{t=0}^{e} K_t(x, n, s)$$

Il grado di ogni polinomio di Krawtschuk è t. Concludo che $\chi(V|u) = l_e(i)$.

Esercizio 6.27. Determinare il polinomio di Krawtschouk $K_1(x)$ e mostrare che ammette radici intere se e solo se $s \mid n$. Trovare delle terne (n, q, e) che soddisfano lo sphere packing condition ma tali che $q \nmid n - 1$.