Códigos Corretores de Erros - Notas de Aula

Marcelo Firer
Universidade Estadual de Campinas - UNICAMP

2^o semestre de 2007

Contents

Chapter 1

Códigos Corretores de Erros

O marco fundamental da chamada Teoria de Informações, que abrange a teoria de códigos, é o trabalho "A mathematical theory of communication", publicado em Julho de 1948 por Claude Shannon ([CS, CS]). Neste trabalho, Shannon define a *capacidade* c de um canal como

$$c = \lim_{t \to \infty} \frac{\log n\left(t\right)}{t}$$

onde $n\left(t\right)$ é o número de sináis possíveis de serem enviados através do canal em um intervalo de tempo t. O canal, a mensagem e os sinais podem variar de acordo com a stuação. No caso de um CD, a mensagem é a música e o canal o CD; no caso de transmissão de imagens captadas por naves espaciais o canal é o espaço sideral (junto com o hardware pertinente) e as mensagens são as imagens e os sinais são sequências de binárias com $8, 16$ ou 24 dígitos[1], dependendo da paleta escolhida ter $2^8, 2^{16}$ ou 2^{24} cores (os casos mais comuns, como você pode conststar ao tentar configurar o monitor de seu computador). Supondo que estamos em uma situação na qual todos os símbolos permitidos para as informações têm a mesma duração, ou seja, precisam do mesmo intervalo de tempo para serem transmitidos (como por exemplo assumindo que cada uma das 2^N cores da paleta escolhida seja transmitida com

[1]Um dígito binário é o que chamamos de Bit, abreviação para "binary digit" sugerida por J. W. Tukey.

exatamente N dígitos. Então, assumindo como unidade de tempo o intevalo necessário para transmissão de um sequência que represente uma única cor, temos que

$$c = \lim_{t\to\infty} \frac{\log\left(2^N\right)^t}{t} = N.$$

Nem sempre os sinais tem a mesma duração. Por exemplos, os telégrafos trabalhavam com as seguintes combinações de sinais básicos, em que a linha de transmissão era aberta (A) e ou fechada (F) conforme a tabela abaixo:

Ponto	F	A				
Traço	F	F	F	A		
Separação de Letras	A	A	A			
Separação de Palavras	A	A	A	A	A	A

Se considerarmos o caso geral que são permitidos os símbolos $S_1, ..., S_k$ e que cada um destes símbolos tem respectivamente a duração de $t_1, ..., t_k$, então, como a quantitade total de sequências possíveis com duração total igual a t é a soma daquelas terminadas em $S_1, S_2, ..., S_k$, temos que

$$n(t) = n(t - t_1) + n(t - t_2) + ... + n(t - t_k).$$

usando resultado conhecido a respeito de diferenças finitas, é possível concluir que $n(t)$ é assintótica a função X_0^t, onde X_0 é a maior raiz real da equação característica

$$X^{-t_1} + X^{-t_2} + ... + X^{-t_k} = 1$$

de modo que a capacidade é bem definida: $c = \lim_{t\to\infty} \log\left(X_0^t/t\right) = \log X_0$.

Em seu trabalho seminal de 1948, Shannon demonstrou que, usando qualquer taxa de transmissão inferior a capacidade do canal, é possível ter comunicação tão confiável quanto desejado.

A palavra confiabilidade é a chave para introduzir o conceito de código. No contexto de Teoria de Informações, é uma regra para converter algum dado em outra forma de informação, não necessariamente

da mesma natureza. É o que ocorre, por exemplo, quando digitalizamos uma fotografia. Se trabalharmos por exemplo com a resolução de um megapixel, estaremos escolhendo em cada unidade de área da fotografia 10^6 pontos que serão varridos pelo instrumento de leitura, sendo identificada a cor de cada um destes pontos. Por sua vez, cada uma destas é identificada com uma cor de uma paleta, que comumente tem 2^8, 2^{16} ou 2^{24} cores. Se considerarmos como exemplo o caso de $2^8 = 256$ cores, cada uma delas é associada a um número entre 1 e 256, que é representado em forma binária por 8 algarismos. É essa sequência de informações que é armazenada no computador. Este processo é chamado de codificação. Ao se acessar a informação guardada na memória, o computador usa um outro processo para transformar esta informação sobre pontos isolados em uma imagem, no processo chamado de decodificação.

Um dos principais problemas em Teoria de Códigos (e em Teoria de Informação de modo mais geral) é a existência de erros: ao se transmitir ou armazenar informações, ocorrem erros que podem comprometer a confiabilidade dos dados. Em outras palavras, assim como na brincadeira de "telefone sem fio", a mensagem recebida não é igual aquela transmitida. São diversas as fontes de erros, mas estes estão quase sempre presentes e o problema se resume em duas etapas: detectar a existência de erros e tentar corrigi-lo. A confiabilidade é adquirida através de algum tipo de redundância (como por exemplo repetir cada palavra), mas a redundância tem sempre um custo, o que nos leva a um dos grandes desafios da Teoria de Códigos: adquirir a confiabilidade desejada ao menor custo (taxa de redundância) possível.

Vamos considerar neste texto apenas os códigos feitos para os chamados *canais simétricos*, aqueles em todos os símbolos têm a mesma probabilidade de serem recebidos errados e caso o símbolo seja trocado, a probalidade de em seu lugar ser enviado qualquer um dos outros é a mesma.

A principal referência para estas notas de aula é o extenso e agradável livro de Huffman e Pless [HP]. Não podemos deixar de indicar também o clássico livro de MacWilliams e Sloane [MS], de caráter enciclopédico, reunindo referências importantes para parte significativa da teoria de códigos corrtores de erros. Ainda é possível sugerir o livro [HV] do

Abramo Hefez e Maria Villela é outra excelente sugestão de leitura para aqueles que não têm qualquer familiaridade com a Teoria dos Códigos.

1.1 Espaços Vetoriais sobre Corpos Finitos

O ambiente que iremos trabalhar é o de espaços vetoriais sobre corpos finitos.

Um corpo é um conjunto com duas operações que mimetizam as operações de soma e produto usuais de números racionais ou reais, no sentido de satisfazer as usuais propriedades associativa, comutativa, distributiva, existência de elementos neutros, aditivo (0) e multiplicativo (1) e existência de oposto aditivo e inverso multilicativo (para elementos diferentes de 0).

Para o propósito deste texto, meramente introdutório, será suficiente nos atermos aos corpos finitos $\mathbb{Z}_p$, com p primo. Como conjunto $\mathbb{Z}_p$ possui p elementos, que denotamos por $\mathbb{Z}_p = \left\{\bar{0}, \bar{1}, \ldots, \overline{p-1}\right\}$. As operações são definidas do seguinte modo: dados $\bar{x}, \bar{y} \in \mathbb{Z}_p$, consideramos a soma (produto) usual $x+y$, dividimos por p, obtendo um resto $0 \leq r \leq p-1$ e definimos $\bar{x}+\bar{y} = \bar{r}$ ($\bar{x} \cdot \bar{y} = \bar{r}$). Apresentamos abaixo a título de exemplo as tabelas de soma e produto de $\mathbb{Z}_2$ e $\mathbb{Z}_3$:

$+$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

$\cdot$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Como existem outros corpos finitos além dos $\mathbb{Z}_p$, vamos usar a notação mais genérica, comumente encontrada na bibliografia, $\mathbb{F}_q$, onde assumimos que $\mathbb{F}_p = \mathbb{Z}_p$ se p for primo. Mais ainda, sempre que não houver possibilidade de confusão, vamos omitir as barras, denotando $\bar{n} \in \mathbb{Z}_p$ simplesmente or n. Eventualmente vamos trabalhar com o corpo quaternário

$\mathbb{F}_4 = \{0, 1, \omega, \overline{\omega}\}$, com as operações definidas por:

$+$	0	1	ω	$\overline{\omega}$
0	0	1	ω	$\overline{\omega}$
1	1	0	$\overline{\omega}$	ω
ω	ω	$\overline{\omega}$	0	1
$\overline{\omega}$	$\overline{\omega}$	ω	1	0

$\times$	0	1	ω	$\overline{\omega}$
0	0	0	0	0
1	0	1	ω	$\overline{\omega}$
ω	0	ω	$\overline{\omega}$	1
$\overline{\omega}$	0	$\overline{\omega}$	1	ω

Mais adiante, quando necessário, falaremos um pouco mais sobre a estrutura de corpos genéricos, mas no momento temos instrumentos suficientes para desenvolvermos parte significativa da teoria.

Seja $\mathbb{F}_q^n$ o conjunto de todas as n-uplas de elementos de $\mathbb{F}_q$. O conjunto $\mathbb{F}_q^n$ possui uma estrutura de *espaço vetorial* sobre $\mathbb{F}_q$, onde a soma e o produto por escalar são defindas coordenada a coordenada: se $x, y \in \mathbb{F}_q^n$, $\lambda \in \mathbb{F}_q$ com $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$, então

$$x + y = (x_1 + y_1, \ldots, x_n + y_n)$$

e

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n)\,.$$

1.2 Códigos Lineares e Detecção de Erros

Um $(n; M)$ *código* $\mathcal{C}$ sobre $\mathbb{F}_q$ é um subconjunto de $\mathbb{F}_q^n$ com M elementos. Chamamos $\mathbb{F}_q$ de *alfabeto* (e seus elementos de *letras*) e os elementos de $\mathcal{C}$ são chamadas de *palavras*. A analogia com as linguagens usuais (português, por exemplo) é absolutamente pertinente, onde entendemos o conjunto $\mathcal{C}$ como o vocabulário da língua, ou seja, o conjunto de todas as palavras possíveis de serem escritas com o alfabeto estendido, que compreende as 27 letras, a cedilha, os acentos e um símbolo para indicar a inexistência de letras, de modo que usamos este símbolo para tornar todas as palavras com o mesmo comprimento.

É possível (e necessário, se quisermosalgum resultado que transceda as questões de contagem) enriquecer $\mathcal{C}$ através de diversas estruturas, a mais simples e mais imortante de todas é a de espaço vetorial. Dizemos que $\mathcal{C} \subset \mathbb{F}_q^n$ é um $[n; k]$ *código linear* sobre $\mathbb{F}_q$ se $\mathcal{C}$ for um sub-espaço vetorial

de dimensão k de $\mathbb{F}_q^n$. Dizer que $\mathcal{C}$ é subespaço linear de $\mathbb{F}_q^n$ significa que dados $u, v \in \mathcal{C}$ e $\lambda \in \mathbb{F}_q$ quaisquer, então $u + v$ e λu pertencem a $\mathcal{C}$ (em particular o vetor nulo $0 = (0, \ldots, 0) = 0 \cdot v \in \mathcal{C}$). A dimensão de $\mathcal{C}$ é definida da maneira usual, como o número de elementos de uma *base*, ou seja, a quantidade mínima de elementos $v_1, \ldots, v_l \in \mathcal{C}$ tal que todo elemento $v \in \mathcal{C}$ pode ser descrito como combinação linear $v = \lambda_1 v_1 + \ldots + \lambda_l v_l$, com $\lambda_1, \ldots, \lambda_l$ escalares em $\mathbb{F}_q$. Observe que cada um dos k escalares pode assumir q valores distintos e como estamos considerando uma base de $\mathcal{C}$, obtemos q^k combinações lineares distindas, ou seja, $\mathcal{C}$ tem q^k elementos e um $[n; k]$ código linear é um $\left(n; q^k\right)$ código sobre $\mathbb{F}_q$.

Apenas a estrutura de espaço vetorial já permite , sob certas circunstâncias, detectarmos erros. Seja $\mathcal{C}$ um $[n; k]$ código linear, com $k < n$ e seja $v \in \mathcal{C}$ uma palavra. Suponha que ao transmitirmos a palavra v a palavra recebida seja w (podendo ter eventualmente $v = w$). Ao recebermos a mensagem w, procedemos antes de tudo com a verificação de que esta mensagem é uma palavra do nosso vocabulário, ou seja, verificamos se $w \in \mathcal{C}$. Esta verificação é simples, se lembrarmos que um subespaço vetorial é definido por um sistema de equações lineares homogêneas. Se considerarmos a matriz H definida pelos coeficientes do sistema linear, podemos representar este sistema matricialmente pela equação

$$H \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

e temos que $\mathcal{C}$ é o conjunto de soluções deste sistema. Uma matriz H satisfazendo esta propriedade é chamada de *matriz de verificação de paridade*. Observemos que sendo $\mathcal{C}$ um código de dimensão k, o posto de H é $n - k$. Assim, a matriz H deve ter ao menos $n - k$ linhas linearmente independentes.

Vejamos agora como proceder para detectarmos a existência de erros. Se $w = (w_1, \ldots, w_n)$ for uma palavra recebida, e denotarmos por w^t a matriz transposta (na realidade um vetor coluna), basta efetuar o produto Hw^t para sabermos se w pertence a $\mathcal{C}$. Se $w \notin \mathcal{C}$, sabemos que a mensagem recebida é equivocada, ou seja, conseguimos detectar a ocorrência de erro.

Note, no entanto, que é possível termos $w \neq v$ mas com $w \in \mathcal{C}$. Observemos que $\mathcal{C}$ possui q^k elementos, enquanto $\mathbb{F}_q^n$ possui q^n elementos, dos quais exatamente $q^n - q^k = q^k\left(q^{n-k}-1\right)$ elementos não pertencem a $\mathcal{C}$. Se assumirmos a hipótese de que o ruído perturba a palavra transmitida v de modo que possamos receber qualquer elemento de $\mathbb{F}_q^n$ temos que a probabilidade de detectarmos o erro é dada por

$$\begin{aligned} P &= \frac{\#\left(\text{elementos de } \mathbb{F}_q^n \text{ não pertencentes a } \mathcal{C}\right)}{\#\left(\text{elementos de } \mathbb{F}_q^n\right)} \\ &= \frac{q^n - q^k}{q^n} \\ &= 1 - \frac{1}{q^{n-k}}. \end{aligned}$$

Como $q \geq 2$ e $n-k \geq 1$, temos que $P < 1$ e esta cresce conforme q ou $n-k$ crescem.

Assim, se quisermos detectar em média 999 erros a cada 1000 ocorrências, se tivermos $q = 2$, basta termos $n - k \geq 10$. Como

$$\lim_{q \to +\infty} \frac{1}{q^{n-k}} = \lim_{n-k \to +\infty} \frac{1}{q^{n-k}} = 0,$$

podemos detectar erros com a confiança tão grande quanto desejarmos.

Para detectarmos a existência de erros em um $[n;k]$ código linear precisamos efetuar o produto Hw^t. Se $H = (a_{ij})$, e $w = (w_1, \ldots, w_n)$ a j-ésima entrada de Hw^t é $a_{j1}w_1 + \ldots + a_{jn}w_n$, ou seja, precisamos realizar n produtos e n somas, devendo ainda verificar se $a_{j1}w_1 + \ldots + a_{jn}w_n = 0$, sendo portanto necessário $2n+1$ operações. Como o vetor Hw^t tem $n-k$ entradas, para verificar a ocorrência de erros executamos $(n-k)(2n+1)$ operações. Apesar deste número parecer relativamente grande (cresce quadraticamente com n) se não tivessemos a estrutura vetorial, teríamos de comparar w com todos os q^k elementos de $\mathcal{C}$ e, de modo geral, para $q > 1$, temos que

$$\lim_{n-k \to \infty} \frac{(n-k)(2n+1)}{q^{(n-k)}} = \lim_{n-k \to \infty} \frac{((n-k)(2n+1))^k}{q^{(n-k)}} = 0,$$

para todo $d \geq 0$, ou seja, para códigos com *codimensão* $n - d$ grande, a estrutura linear verifica-se comparativamente muito econômica para a verificação de erros.

Outra maneira de descrevermos um código é através de uma *matriz geradora* G, uma matriz $k \times n$ em que as k linhas formam uma base de $\mathcal{C}$. De modo geral existem muitas matrizes geradoras para um código. Para cada escolha de k colunas linearmente independentes de G, correspondem k coordenadas que são chamadas de coordenadas ou *conjunto de informação de* $\mathcal{C}$ e as $r = n-k$ coordenadas remanescentes são chamadas de *conjundo de redundância de* $\mathcal{C}$ e r de *redundancia*. Se as k primeiras coordenadas forem um conjunto de informação, então o código $\mathcal{C}$ possui uma única matriz geradora da forma $\left[Id_{k\times k}|A_{k\times(n-k)}\right]$, chamada de *forma padrão*.

Teorema 1.1 *Se $G = [I_k|A]$ é uma matriz geradora na forma padrão de um $[n,k]$-código $\mathcal{C}$, então $H = [-A^t|I_{n-k}]$ é uma matriz de verificação de paridade de $\mathcal{C}$.*

Demonstração Claramente temos que

$$HG^t = -A^t I_k + I_{n-k}A^t = -A^t + A^t = 0$$

donde segue que os vetores linhas de G, geradores de $\mathcal{C}$, estão contidos no núcleo da transformação linear $T_H(w) = Hw^t$. Mas H claramente possui posto $n - k$, de modo que o $\dim(\ker T_H) = k = \dim(\mathcal{C})$ e segue que $\mathcal{C} = \ker T_H$, ou seja, H é matriz de verificação de paridade de $\mathcal{C}$. □

1.3 Códigos Duais

Dados $\mathbf{x} = (x_1, ..., x_n), \mathbf{y} = (y_1, ..., y_n) \in \mathbb{F}_q^n$, o produto interno formal de $\mathbf{x}$ por $\mathbf{y}$ é definido por $\langle\rangle\ \rangle 2\ \rangle\ \ \rangle \Delta 0 \langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 + ... + x_ny_n$. Dado código $\mathcal{C}$ definimos o *código dual* de $\mathcal{C}$ como

$$\mathcal{C}^{\perp} = \left\{\mathbf{x} \in \mathbb{F}_q^n | \langle \mathbf{x}, \mathbf{c}\rangle = 0, \forall c \in \mathcal{C}\right\}.$$

Exercício 1.1 *Dado código $\mathcal{C}$, G e H são matrizes geradora e de verificação de paridade de $\mathcal{C}$ se e somente se H e G são matrizes geradora e de verificação de paridade de $\mathcal{C}^{\perp}$.*

Exemplo 1.1 *Considere o* $[k, r \cdot k]$código de repetição *em que cada letra da mensagem é repetida um número fixo de vezes (digamos r) e que cada letra é formada por k dígitos sobre $\mathbb{F}_2$. Ou seja, a letra* 100...00 *($k-1$ zeros) é transmitida como* 100...00 100...00 ...100...00 *(r vezes). Este é o tipo mais simples de código e ao receber uma mensagem, devemos receber sempre blocos de r letras repetidas. Se não for este o caso, corrigimos o código escolhendo a letra que aparece maior número de vezes no bloco. Considerando Id a matriz identidade $k \times k$ temos que*

$$G = [Id|Id \ \cdots \ Id] \ \text{ e } H = \begin{bmatrix} -Id & \\ \vdots & |Id_{k(r-1)} \\ -Id & \end{bmatrix}$$

são respectivamente as matrizes geradora e verificaçãod e paridade de $\mathcal{C}$. É fácil verificar que G é matriz de verificação de paridade de $\mathcal{C}^{\perp}$. Observe que $\mathcal{C}^{\perp}$ tem a propriedade que um único erro pode ser sempre detectado (pois se $\mathbf{x} \notin \mathcal{C}^{\perp}$ temos que $G\mathbf{x}^t \neq 0$, mas como estamos considerando o corpo $\mathbb{F}_2$, ao mudarmos qualquer coordenada de $\mathbf{x}$ obtemos vetor $\mathbf{x}'$ tal que $G\mathbf{x}'^t$.

Um código é dito *auto-ortogonal* se $\mathcal{C} \subset \mathcal{C}^{\perp}$ e *auto-dual* se $\mathcal{C} \subset \mathcal{C}^{\perp}$.

Exercício 1.2 *Prove que um código auto-dual tem comprimento n par e dimensão $n/2$.*

O $[7, 4]$ *código de Hamming* $\mathcal{H}_3$ é o código que tem como matriz verificadora de paridade uma matriz H cujas colunas são todos os $2^3 - 1$ vetores não nulos de $\mathbb{F}_2^3$:

$$H = \left[\begin{array}{cccc|ccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

e matriz geradora

$$G = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]$$

Acrescentando a cada vetor de G uma coordenada de controle, obtemos o $[8, 4]$-código $\widehat{\mathcal{H}}_3$ que tem como matriz geradora

$$\widehat{G} = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{array}\right]$$

É imediato verificar que $\widehat{\mathcal{H}}_3$ é auto dual (basta verificar que $\widehat{G} \cdot \widehat{G}^{\tau} = 0$).

1.4 Correção de Erros

Se apenas a estrutura de espaço vetorial basta para detectarmos a existência de erros, ela é insuficiente para nos ajudar a corrigi-los, de modo que precisamos introduzir em $\mathbb{F}_q^n$ uma estrutura adicional.

Do mesmo modo que a definição de corpo mimetiza a estrutura dos números reais, assumindo como definição as propriedades básicas das operações aritméticas, o conceito de métrica mimetiza a geometria (euclideana) clássica, adotando como definição as propriedades básicas satisfeitas pela função distância entre pontos do plano.

Definição 1.1 *Uma* métrica *em um conjunto X é uma função $d : X \times X \to \mathbb{R}$ satisfazendo as seguintes propriedades:*

(*i*) $d(x, y) > 0$ *se* $x \neq y$ *e* $d(x, x) = 0$, *para quaisquer* $x, y \in X$;

(*ii*) $d(x, y) = d(y, x)$, *para quaisquer* $x, y \in X$;

(*iii*) $d(x, z) \leq d(x, y) + d(y, z)$, *para quaisquer* $x, y, z \in X$.

Conforme dissemos anteriormente, o exemplo primitivo de métrica é a distância euclideana d_E, dada pelo Teorema de Pitágoras: para $x, y \in \mathbb{R}^n$, com $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$, $d_E(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$.

Dada uma métrica d em um conjunto X definimos os conceitos de esfera e bola como lugares geométricos de pontos equidistantes ou adistância majorada de seu centro:

$$S(x_0; r) := \{x \in X : d(x, x_0) = r\},$$

$$B(x_0; r) := \{x \in X : d(x, x_0) \leq r\}.$$

Quando estivermos trabalhando com métricas especificadas, usaremos sub-índices $d_*(\cdot, \cdot)$ para identificar a métrica. O mesmo sub-índice será adotado para designar todos os conceitos derivados da métrica, tais como $S_*(\cdot; \cdot)$, $B_*(\cdot; \cdot)$ e $\omega_*(\cdot)$.

Dada uma métrica d em $\mathbb{F}_q^n$, temos uma maneira razoável de tentarmos corrigir erros. Vamos supor como antes que a palavra transmitida é $v \in \mathcal{C}$ e a palavra recebida foi $w \notin \mathcal{C}$. Se a métrica em si for razoável (considerando-se as características físicas do canal de transmissão de informação), é razoável supormos que a probabilidade de w estar "longe" de v é menor que a probabilidade de estar "perto" de v, no sentido de que para quaisquer $\alpha, \beta, m \in \mathbb{R}$, com $\alpha \geq \beta \geq 0$ e $m > 0$, então a probabilidade de termos $\alpha \leq d(z, w) \leq \alpha + m$ é menor ou igual que a probabilidade de termos $\beta \leq d(z, w) \leq \beta + m$. Obviamente esta probabilidade é definida pelas características físicas do canal de transmissão, mas como neste texto não estamos tratando de canais específicos, vamos sempre assumir que esta hipótese é verdadeira.

Com isto, passamos a ter um mecanismo sistemático para corrigir erros: se a palavra recebida $x \notin \mathcal{C}$, escolhemos o ponto w de $\mathcal{C}$ mais próximo de x, ou seja, escolhemos $w \in \mathcal{C}$ tal que

$$d(x, w) = \inf \{d(x, u) : u \in \mathcal{C}\}$$

(observe que o ínfimo é atingido pois $\mathcal{C}$ é um conjunto finito).

Do mesmo modo que ocorre com a detecção de erro, a correção de erros pode falhar e temos na realidade que entender quais são suas

limitações. Assim, acrescentar às restrições já determinadas (a estrutura vetorial do código $\mathcal{C}$ e a estrutura métrica $d(\cdot,\cdot)$ definida em $\mathbb{F}_q^n$ as seguintes condições:

(a) A métrica definida é compatível com a estrutura vetorial, no sentido de ser invariante por translações: $d(x+y, x+z) = d(y,z)$, para quaisquer $x, y, z \in \mathbb{F}_q^n$.

(b) A métrica é compatível com a estrutura discreta de um espaço finito, no sentido que $d(x,y) \in \mathbb{N}$ para quaisquer $x, y \in \mathbb{F}_q^n$.

Consideremos os pontos do código $\mathcal{C}$ e a menor distância entre dois destes pontos:

$$d := d(\mathcal{C}) = \min\{d(v,w) : v, w \in \mathcal{C}, v \neq w\}.$$

Se lembrarmos que o código $\mathcal{C}$ é linear e que a distância é compatível com a métrica, temos que

$$d(x,y) = d(x-x, y-x) = d(0,z)$$

para quaisquer $x, y \in \mathbb{F}_q^n$, onde $z = x - y$. Assim, se definirmos o *peso do vetor* $x \in \mathbb{F}_q^n$ como $\omega(x) := d(x,0)$, como $v - w \in \mathcal{C}$ sempre que $v, w \in \mathcal{C}$, temos que

$$d = \min\{\omega(v) : 0 \neq v \in \mathcal{C}\}.$$

Denotando por $r := \left\lfloor \frac{d-1}{2} \right\rfloor$ (chamado de *raio do código*) onde $\lfloor t \rfloor$ denota a parte inteira do real t, temos que duas bolas $B(v;r)$ e $B(w;r)$, com $v, w \in \mathcal{C}$ são necessariamente disjuntas. De fato, supondo que u pertença a intersecção destas, temos pela desigualdade triângular que

$$\begin{aligned} d(v,w) &\leq d(v,u) + d(u,w) \\ &\leq \left\lfloor \frac{d-1}{2} \right\rfloor + \left\lfloor \frac{d-1}{2} \right\rfloor \\ &\leq d-1, \end{aligned}$$

um absurdo, pois por definição $d \leq d(v, w)$. De modo geral, dizemos que a *capacidade de correção* do código é R se podemos garantir que, caso a a palavra transmitida v e a recebida w distem no máximo R, então temos certeza de estar corrigindo o código. Assim, temos que se a distância entre a mensagem enviada v e a mensagem recebida w for menor ou igual a $\left\lfloor \frac{d-1}{2} \right\rfloor$, temos que v é o ponto de $\mathcal{C}$ mais próximo de w e estamos de fato corrigindo o erro ocorrido durante a transmissão da mensagem. Neste caso, dizemos que $\mathcal{C}$ é um $[n; k; d]$ *código*.

Geometricamente podemos pensar na capacitade de correção como sendo o maior natural R tal que $B(v; R) \cap B(w; R) = \varnothing$ para quaisquer $v, w \in \mathcal{C}$ distintos, ou seja, o raio máximo que nos permite empacotar bolas disjuntas centradas nas palavras do código e por isto chamamos esta grandeza de *raio de empacotamento*:

$$R_e(\mathcal{C}) = \max \{r \in \mathbb{N} : B(v; r) \cap B(w, r) = \varnothing, \forall v, w \in \mathcal{C}, v \neq w\}.$$

Observamos que o raio $\left\lfloor \frac{d-1}{2} \right\rfloor$ de um código $\mathcal{C}$ é apenas um limitante inferior para o raio de empacotamento, este sim definindo a capcidade de correção do código.

As k variáveis que definem a dimensão do código e a cardinalidade q do corpo definem o tamanho q^k do alfabeto, que por sua vez é determinado pela natureza da informação que desejamos transmitir e por isto, estas k variáveis são chamadas de variáveis de informação. Ao escolhermos um subespaço k-dimensional de um espaço sobre $\mathbb{F}_q$ de dimensão $n > k$, estamos introduzindo variáveis que não contém informação adicional, mas que são usadas para detectar e corrigir erros e por isto são chamadas de variáveis de controle. Assim como o tamanho do alfabeto, a capacidade de correção do código R_e também é determinada a *priori* pela natureza da informação (dados referentes a operações bancárias necessitam de confiabilidade maior que a transmissão de televisão). Assim, tendo q^k e R_e definidos pela natureza da informação, boa parte do desafio na Teoria de Códigos Corretores de Erros é buscar códigos em que a relação entre as variáveis de informação e as de controle seja a maior possível, ou seja, em que precisemos adicionar o mínimo de variáveis de controle. Esta relação $\frac{k}{n}$ entre as variáveis de informação e o total de variáveis é chamada de *taxa de informação* do código.

Obviamente, qualquer que seja a taxa de informação, não podemos ter a certeza de efetivamente corrigir todos os erros de transmissão, mas do mesmo modo que ocorre com a detecção, podemos corrigir erros com grau de segurança tão grande quanto desejado.

Para podermos trabalhar alguns exemplos, devemos definir algumas métricas específicas, que apresentaremos na próxima sessão.

Chapter 2

Métricas em Espaços de Códigos

Até o momento exploramos aspectos de códigos que dependem exclusivamente das propriedades intrínsecas aos conceitos genéricos de subespaço vetorial e métrica. Nesta seção apresentaremos algumas métricas possíveis de serem definidas nos espaços $\mathbb{F}_q^n$. Sem qualquer sombra de dúvidas, a métrica mais importante em termos de aplicações práticas é a métrica de Hamming, utilizada geralmente em códigos e espaços vetoriais sobre o corpo com dois elementos $\mathbb{F}_2$. No entanto, neste texto enfatizaremos as métricas ponderadas, definidas a partir de ordens parciais. A opção por esta ênfase se deve a um único motivo: as métricas de Hamming e de Lee foram definidas respectivamente em 1950 e 1958 e têm sido trabalhadas ao longo de décadas e temos uma variedade de livros textos, inclusive em português, que trabalham estas métricas de maneira detalhada e didática, enquanto as métricas ponderadas foram introduzidas por Brualdi em 1995 e tem sido desenvolvida com mais ênfase apenas a artir de 2003, de modo que estes conceitos podem ser encontrados exclusivamente em artigos de pesquisa.

Indicamos o livro de Schröder [SB] para os leitores que desejam obter mais informações sobre Teoria de Ordem. Para os leitores interessados em aprofundar o conhecimento em métricas ponderadas indicamos os artigos [BG], [LE] e [LY].

2.1 Métrica de Hamming

A distância de Hamming entre dois vetores $x, y \in \mathbb{F}_q^n$ é simplesmente o número de coordenadas distintas entre estes dois vetores:

Definição 2.1 *Dados $x, y \in \mathbb{F}_q^n$ com $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, a* distânCia de Hamming *é definida como*

$$d_H(x, y) = \#\{i : x_i - y_i \neq 0; i = 1, 2, \ldots, n\}.$$

Definimos o peso de Hamming *de x como $\omega_H(x) := d_H(x, 0)$.*

Começamos demonstrando que esta é de fato uma métrica:

Proposição 2.1 *d_H é uma métrica em $\mathbb{F}_q^n$.*

Demonstração Por definição $d_H(x, y)$ assume apenas valores naturais e $d_H(x, y) = 0$ se e somente se $x = y$. A condição de simetria também é trivialmente satisfeita. Observe ainda que se y e x têm n_1 coordenadas distintas e y e z possuem n_2 coordenadas distintas, então x e y possuem no máximo $n_1 + n_2$ (e no mínimo $|n_1 - n_2|$) coordenadas distintas, de modo que a desigualdade triangular também é satisfeita. □

Vimos anteriormente que se $d = d(\mathcal{C})$ for a distância mínima do código e $R_e(\mathcal{C})$ o seu raio de empacotamento, então

$$\left\lfloor \frac{d-1}{2} \right\rfloor \leq R_e(\mathcal{C}) < d,$$

independentemente da métrica em consideração. Vamos mostrar que, no caso de uma métrica de Hamming, a situação é bem melhor definida.

Proposição 2.2 *Considerando em $\mathbb{F}_q^n$ a métrica de Hamming, temos que para todo código linear $\mathcal{C} \subset \mathbb{F}_q^n$, $\left\lfloor \frac{d-1}{2} \right\rfloor = R_e(\mathcal{C})$.*

Demonstração Vamos mostrar que se $r > \left\lfloor \frac{d-1}{2} \right\rfloor$ então existem $u, v \in \mathcal{C}$ tais que $B_H(u;r) \cap B_H(v;r) \neq \varnothing$. Se $d = 2k+\varepsilon$, com $\varepsilon \in \{0,1\}$, então $k+\varepsilon = \left\lfloor \frac{d-1}{2} \right\rfloor + 1$. Vimos na seção 1.3 que existe $u \in \mathcal{C}$, $u = (u_1, \ldots, u_n)$ tal que $\omega_H(u) = d$. Temos então que u possui exatamente $2k+\varepsilon$ coordenadas não nulas. Suponhamos que estas sejam as coordenadas no conjunto $I \cup J \cup \{l_\varepsilon\}$ onde

$$I = \{i_1, \ldots, i_k\}, \; J = \{j_1, \ldots, j_k\}$$

e $\{l_\varepsilon\} = \varnothing$ se $\varepsilon = 0$. Seja $x = (x_1, \ldots, x_n)$ o vetor definido por

$$x_m = \begin{cases} u_m & \text{se} \quad m \notin I \cup \{l_\varepsilon\} \\ 0 & \text{se} \quad m \in I \cup \{l_\varepsilon\} \end{cases}.$$

Temos então que

$$d_H(x,u) = \#(I \cup \{l_\varepsilon\}) = k + \varepsilon$$

e

$$d_H(x,0) = \#J = k$$

e temos que $x \in B_H(0; k+\varepsilon) \cap B_H(u; k+\varepsilon)$, de modo que $R_e(\mathcal{C}) < k + \varepsilon = \left\lfloor \frac{d-1}{2} \right\rfloor + 1$ e concluímos que $R_e(\mathcal{C}) = \left\lfloor \frac{d-1}{2} \right\rfloor$. □

Exemplo 2.1 *Vamos considerar em* $\mathbb{F}_2^{2^2-1} = \mathbb{F}_2^3$ *o código* $\mathcal{C}$ *que é definido pelo sistema linear* $Hx^t = 0$*, onde*

$$H = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Como H *é uma matriz* 2×3*, temos que* $\mathcal{C}$ *é um código de dimensão* 1 *em* $\mathbb{F}_2^3$ *e é imediato que* $\mathcal{C} = \{(0,0,0), (1,1,1)\}$*, bastando para isto verificar que*

$$Hx^t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Note agora que

$$B_H((0,0,0);1) = \{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\}$$

e

$$B_H((1,1,1);1) = \{(1,1,1),(0,1,1),(1,0,1),(1,1,0)\}.$$

Observe ainda que

$$B_H((0,0,0);1) \cap B_H((1,1,1);1) = \varnothing$$

e que

$$\mathbb{F}_2^3 = B_H((0,0,0);1) \cup B_H((1,1,1);1).$$

Ou seja, não apenas podemos empacotar bolas unitárias em pontos do código, como estas recobrem todo o espaço ambiente $\mathbb{F}_2^3$.

Exemplo 2.2 *Vamos considerar em* $\mathbb{F}_2^{2^3-1} = \mathbb{F}_2^7$ *o código que é definido pelo sistema linear* $Hx^t = 0$*, onde*

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Para encurtar os símbolos, vamos denotar por $i_1 i_2 \ldots i_k$ *o vetor* $(i_1, \ldots, i_k)$ *e por verificação direta observamos que*

$$\beta = \{0010110, 1100110, 1101001, 1111111\}$$

é um conjunto linearmente independente de vetores satisfazendo a equação $Hx^t = 0$*, ou seja,* β *é uma base para* $\mathcal{C}$*. Como o primeiro vetor da base,* $u = 0010110$ *tem peso de Hamming* 3*, temos que* $d(\mathcal{C}) \leq 3$*. Supondo que* $d(\mathcal{C}) < 3$*, teríamos um vetor com exatamente uma ou duas coordenadas não nulas. É imediato verificar que um vetor nestas condições não pode ser solução do sistema proposto. Temos então que* $d(\mathcal{C}) = 3$ *e* $R_e(\mathcal{C}) = 1$*. Se denotarmos por* e_i *o vetor em* $\mathbb{F}_2^7$ *que tem a* i*-ésima coordenada igual a* 1 *e todas as demais nulas, é imediato verificar que para qualquer* $x \in \mathbb{F}_2^7$*,*

$$\begin{aligned} B_H(x;1) &= S_H(x;0) \cup S_H(x;1) \\ &= \{x\} \cup \{x + e_i : i = 1, \ldots, 7\} \end{aligned}$$

de modo que $\#\left(B_H\left(x;1\right)\right)=8$. *Observe que o código* $\mathcal{C}$ *tem dimensão* 4 *e portanto possui* 2^4 *elementos. As bolas de raio* 1 *centradas nestes elementos são disjuntas duas a duas e como cada uma destas tem* $8=2^3$ *elementos concluímos que*

$$\begin{aligned}\#\left(\bigcup_{u\in\mathcal{C}}B_H\left(u;1\right)\right) &= \sum_{u\in\mathcal{C}}\#\left(B_H\left(u;1\right)\right)\\ &= 2^4\cdot 2^3\\ &= \#\left(\mathbb{F}_2^7\right).\end{aligned}$$

Assim, temos que estas bolas não apenas são disjuntas, como sua união engloba todo o espaço, ou seja,

$$B_H\left(u;1\right)\cap B_H\left(v;1\right)=\varnothing \text{ se } u,v\in\mathcal{C},\ u\neq v$$

e

$$\bigcup_{u\in\mathcal{C}}B_H\left(u;1\right)=\mathbb{F}_2^7.$$

2.2 Métrica de Lee

Se considerarmos um vetor $x=(x_1,\ldots,x_n)\in\mathbb{F}_q^n$, o peso de Hamming $\omega_H\left(x\right)$ identifica o número de coordenadas não nulas de x mas ignora totalmente o valor assumido por estas coordenadas quando estes não se anulam. Se considerarmos que $0\leq x_i\leq p-1$ para cada $i=1,\ldots,n$ e que estamos identificando os inteiros com o resto de sua divisão por p, podemos identificar $\mathbb{F}_p=\mathbb{Z}_p$ com as i-ésimas raízes da unidade

$$\left\{e^0,e^{i2\pi/p},e^{i4\pi/p},...,e^{i2(p-1)\pi/p}\right\},$$

vértices de um polígono regular de p lados $\mathcal{P}_p$. Definimos a *distânCia de Lee* $\left|a-b\right|_L$ entre dois pontos $a,b\in\mathbb{Z}_p$ como sendo o menor número de arestas de $\mathcal{P}_p$ que precisamos percorrer para ligar os vértices $e^{ia2\pi/p}$ e $e^{ib2\pi/p}$:

$$\left|a-b\right|_L=\min\left\{\left|a-b\right|,p-\left|a-b\right|\right\}$$

onde $|\cdot|$ é o valor absoluto usual em $\mathbb{R}$. Definimos a *métrica de Lee* $d_L(x,y)$ entre dois pontos $x, y \in \mathbb{F}_p^n$ como

$$d_L(x,y) = d_L((x_1,\ldots,x_n),(y_1,\ldots,y_n)) := \sum_{i=1}^{n} |x_i - y_i|_L .$$

De modo análogo definimos o *peso de Lee* de um elemento $x \in \mathbb{F}_q^n$ como $\omega_L(x) = d_L(x,0)$. Antes de tudo, observemos que se $p = 2$ ou $p = 3$, a métrica de Lee coincide com a métrica de Hamming. De fato, nestes casos, com a, b podendo assumir apenas os valores $0, 1$ e 2 (este último se tivermos $p = 3$) temos que para $a \neq b$, $|a-b| = 1$, de modo que $|a-b|_L = 0$ se $a = b$ e $|a-b|_L = 1$ se $a \neq b$.

Exemplo 2.3 *Considere em* $\mathbb{F}_5^2$ *o código*

$$\mathcal{C} = \{\lambda(1,2) : \lambda = 0,1,2,3,4\} .$$

Observe que $\omega_L((1,2)) = \omega_L(\lambda(1,2)) = 3$*, para todo* $\lambda \neq 0$*, de modo que* $\left\lfloor \frac{d_L(\mathcal{C})-1}{2} \right\rfloor = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$*. Vamos descrever as bolas unitárias, relativas a métrica de Lee, centradas nas palavras do código. Começamos descrevendo a bola centrada na origem:*

$$B_L((0,0);1) = \{(0,0),(1,0),(4,0),(0,1),(0,4)\} .$$

As outras bolas são obtidas por translação: dado $\lambda(1,2) \in \mathcal{C}$*, temos que*

$$\begin{aligned} B_L(\lambda(1,2);1) &= \lambda(1,2) + B_L((0,0);1) \\ &= \{\lambda(1,2) + v : v \in B_L((0,0);1)\} . \end{aligned}$$

Por verificação direta podemos constatar que

$$B_L(\lambda(1,2);1) \cap B_L(\alpha(1,2);1) = \varnothing$$

se $\lambda \neq \alpha$*, de modo que ao tomarmos a união* $\bigcup_{\lambda=0}^{4} B_L(\lambda(1,2);1)$ *destas cinco bolas obtemos exatamente* 25 *elementos distintos e temos que*

$$\mathbb{F}_5^2 = \bigcup_{\lambda=0}^{4} B_L(\lambda(1,2);1) .$$

Observe ainda que a situação muda radicalmente se considerarmos a métrica de Hamming, ao invés da métrica de Lee. Antes de tudo, temos que

$$\omega_H((1,2)) = \omega_H(\lambda(1,2)) = 2,$$

para todo $\lambda \neq 0$*, de modo que* $\left\lfloor \frac{d_H(\mathcal{C})-1}{2} \right\rfloor = \left\lfloor \frac{2-1}{2} \right\rfloor = 0$ *e temos que as bolas unitárias devem obrigatoriamente se interceptar. De fato,*

$$B_H((0,0);1) = \left\{ \begin{array}{c} (0,0),(1,0),(2,0),(3,0),(4,0), \\ (0,1),(0,2),(0,3),(0,4) \end{array} \right\}$$

de modo que as bolas unitárias possuem 9 *elementos cada. Como* $\#(\mathbb{F}_5^2) = 5^2 = 25$ *e* 9 *não é divisor de* 25*, a união destas bolas não pode recobrir todo o espaço sem intersecção.*

2.3 Métricas Ponderadas

Nesta seção introduziremos não uma, mas na realidade uma família de métricas, que chamamos de métricas ponderadas (*poset-metric* em inglês). O conceito de métricas ponderadas por ordens parcias foi introduzido or Brualdi em 1995 e a partir de 2003, diversos trabalhos têm aprofundado o conhecimento que temos sobre estes espaços.

Considere um conjunto finito com n elementos. Sem perda de generalidade vamos assumir que este é o conjunto que contém os naturais $1, 2, \ldots, n$ e denotá-lo por $[n] := \{1, 2, \ldots, n\}$.

Definição 2.2 *Uma* ordem par$\mathcal{C}$ial P em um $\mathcal{C}$onjunto X *é um subconjunto* $R \subset X \times X$ *satisfazendo as seguintes condições:*

(*i*) $(x,x) \in R$, $\forall x \in X$;

(*ii*) *Dados* $x, y \in X$*, se* $(x,y) \in R$ *e* $(y,x) \in R$*, então* $x = y$;

(*iii*) *Se* $(x,y) \in R$ *e* $(y,z) \in R$*, então* $(x,z) \in R, \forall x, y, z \in X$.

Neste caso dizemos que X é ordenado por R e usamos a notação $P = (X, R)$.

Exemplo 2.4 *Seja $X \subset \mathbb{R}$ e*

$$R = \left\{(x, y) \in \mathbb{R}^2 : x \leq y\right\}$$

temos que $P = (X, R)$ é uma ordem parcial em X.

Assim, dada uma ordem R em um conjunto X, adotamos a notação $x \leq y$ para dizer que $(x, y) \in R$ e escrevemos $P = (X, \leq)$.

Se X for finito, podemos representar uma ordem (X, R) por um *diagrama de Hasse* no plano, construído do seguinte modo:

(a) Identificamos os elementos de X com os vértices do grafo.

(b) Estabelecemos uma aresta ligando os vértices x e y se tivermos que $x \leq y$ e não existir elemento não trivial entre eles, ou seja, se $x \leq z \leq y$ então $z = x$ ou $z = y$.

(c) Ao posicionarmos os vértices no plano $\mathbb{R}\overrightarrow{i} + \mathbb{R}\overrightarrow{j}$, o fazemos de modo que se $x \leq y$ então a coordenada de x na direção $\overrightarrow{j}$ é menor ou igual que a coordenada de y na direção $\overrightarrow{j}$ e neste caso, a igualdade vale se e somente se $x = y$.

Nos exemplos abaixo, vamos considerar as seguintes ordens em $X = [n]$:

Exemplo 2.5 (Ordem Total ou Cadeia) *Seja*

$$R = \{(x, y) \in [n] \times [n] : x \leq y\},$$

onde $\leq$ é a ordem usual dos reais.

$$itbpFU2.7017in2.0314in0inFigura1-$$

[3] *cadeiafig.1.wmf*

Exemplo 2.6 (Anti-Cadeia) $R = \{(x, y) \in [n] \times [n] : x = y\}$.

$$itbpFU2.5071in1.6639in0inFigura2-$$

[4] *anti-cadeiafig.2.wmf*

Exemplo 2.7 (Coroa) *Se* $n = 2k$*, definimos* R *a partir das seguintes desigualdades:*

$$\begin{aligned} x &\leq x \text{ para todo } x \in [n], \\ i &\leq k + i \text{ para todo } i = 1, 2, \ldots, k-1, \\ i+1 &\leq k + i \text{ para todo } i = 1, 2, \ldots, k-1, \\ 1 &\leq 2k \text{ e } k \leq 2k. \end{aligned}$$

$$itbpFU2.7985in2.1041in0inFigura3-$$

[6] *coroafig.3.wmf*

Exemplo 2.8 (Ordem Semi-Fraca) *Sejam* $0 = n_0 < n_1 < \ldots < n_{k-1} < n_k = n$*. Dados* $x, y \in [n]$*, existem* $i = i(x), j = j(y) \in \{0, 1, \ldots, k\}$ *únicos tais que* $n_i + 1 \leq x \leq n_{i+1}$ *e* $n_j + 1 \leq y \leq n_{j+1}$*. Se* $x \neq y$ *dizemos que* $x \leq y$ *se e somente se* $i < j$*. Denotamos esta ordem por* $P = [n_1, \ldots, n_k]$ *e a chamamos de ordem* $[n_1, ..., n_k]$-semi-fraca*.*

$$itbpFU2.7138in2.0401in0inFigura4 - Ordem$$

$[3, 6, 9]$*-semi-fracafig.4.wmf*

Um *ideal* em uma ordem $P = (X, \leq)$ é um subconjunto $I \subset X$ que contem todos os elementos de X menores que algum elemento de I, ou seja, se $x \in X$, $y \in I$ e $x \leq y$, então $x \in I$. Dado $J = \{x_1, \ldots, x_r\} \subset X$, chamamos de *ideal gerado por* J o menor ideal de X contendo J, usando qualquer uma das notações $\langle J \rangle$ ou $\langle x_1, \ldots, x_r \rangle$. Como X é um ideal

contendo J e a intersecção de ideais é um ideal, temos que o ideal gerado por um conjunto sempre existe e

$$\langle J \rangle = \bigcap_{\substack{I \text{ é ideal} \\ J \subset I}} I.$$

Usando o conceito de ideais em uma ordem, podemos induzir uma métrica nos espaços vetoriais $\mathbb{F}_q^n$. Dado $x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$, definimos o seu *suporte* como sendo o conjunto de coordenadas não nulas de x, ou seja, $\text{supp}(x) := \{i \in [n] : x_i \neq 0\}$. Dada uma ordem P em $[n]$ definimos o *P-peso ponderado* de x como sendo a cardinalidade do ideal gerado pelo suporte de x:

$$\omega_P(x) = \#\left(\langle \text{supp}(x) \rangle\right).$$

Usando-se o P-peso, definimos a *métrica ponderada* (por P) de modo similar a relação estabelecida entre peso e métrica nos casos de Hamming e Lee:

$$d_P(x, y) := \omega_p(x - y).$$

Proposição 2.3 *Se P é uma ordem em $[n]$, então d_P é uma métrica em $\mathbb{F}_q^n$.*

Demonstração A positividade e a simetria são claramente satisfeitas. Para demonstrarmos que $d_P(x, y) \leq d_P(x, z) + d_P(z, y)$ começamos observando que $d_P(x + z, y + z) = d_P(x, y)$ para quaisquer $x, y, z \in \mathbb{F}_q^n$ já que $(x + z) - (y + z) = x - y$, de modo que ambos os conjuntos tem o mesmo suporte. Assim, substituindo (x, y), (x, z) e (z, y) respectivamente por $(x - y, y - y)$, $(x - z, z - z)$ e $(z - y, y - y)$ vemos que basta demonstrar que

$$d_P(0, x - y) \leq d_P(0, x - z) + d_P(0, z - y),$$

e como $x - y = (x - z) + (z - y)$, basta mostrar que

$$\omega_P(u + v) \leq \omega_P(u) + \omega_P(v)$$

para quaisquer $u, v \in \mathbb{F}_q^n$. Mas $supp(x+y) \subset supp(x) \cup supp(y)$ e $\langle I \cup J \rangle = \langle I \rangle \cup \langle J \rangle$, donde temos que

$$\begin{aligned} \omega_P(u+v) &= \#(\langle supp(u+v) \rangle) \\ &\leq \#(\langle supp(u) \cup supp(v) \rangle) \\ &= \#(\langle supp(u) \rangle \cup \langle supp(v) \rangle) \\ &\leq \omega_P(u) + \omega_P(v). \end{aligned}$$

□

Sendo d_P uma métrica, todos os conceitos referentes a métrica, como bolas, esferas, distância mínima e outros, serão denotados com o uso do sub-índice P ($B_P(u;r)$, $S_P(u;r)$ e $d_P(\mathcal{C})$ respectivamente), a menos que a ausência do índice não cause confusão.

Exemplo 2.9 *Vamos considerar as métricas cadeia (P_1), anti-cadeia (P_2), coroa (P_3), a ordem $[3,4]$-fraca (P_4), a ordem $[1,4]$-fraca (P_5) no conjunto $[4]$ e descrever as bolas de $\mathbb{F}_2^4$ relativas as métricas dadas. Como métricas ponderadas são invariantes por translações (conforme vimos na demonstração da Proposição 2.3), basta descrever as bolas, ou esferas, centradas na origem $(0,0,0,0)$. Para encurtar a notação, assim como fizemos no Exemplo 2.2, vamos denotar o vetor (i_1, i_2, i_3, i_4) simplesmente por $i_1 i_2 i_3 i_4$.*

	P_1
$S_P(0;1)$	1000
$S_P(0;2)$	$0100, 1100$
$S_P(0;3)$	$0010, 1010, 0110, 1110$
$S_P(0;4)$	$0001, 1001, 0101, 0011, 0111, 1011, 1101, 1111$

	P_2
$S_P(0;1)$	$1000, 0100, 0010, 0001$
$S_P(0;2)$	$1100, 1010, 1001, 0110, 0101, 0011$
$S_P(0;3)$	$1110, 1101, 1011, 0111$
$S_P(0;4)$	1111

	P_3
$S_P(0;1)$	$1000, 0100$
$S_P(0;2)$	1100
$S_P(0;3)$	$0010, 1010, 0110, 1110, 0001, 1001, 0101, 1101$
$S_P(0;4)$	$0011, 1011, 0111, 1111$

	P_4
$S_P(0;1)$	$1000, 0100, 0100$
$S_P(0;2)$	$1100, 1010, 0110$
$S_P(0;3)$	1110
$S_P(0;4)$	$0001, 1001, 0101, 0011, 1101, 1010, 0111, 1111$

	P_5
$S_P(0;1)$	1000
$S_P(0;2)$	$0100, 1100, 0010, 1010, 0001, 1001$
$S_P(0;3)$	$0110, 0101, 0011, 1110, 1101, 1011$
$S_P(0;4)$	$0111, 1111$

Observação 2.1 *As métricas ponderadas abrangem a métrica de Hamming, pois quando P é uma ordem anti cadeia, ou seja, cada elemento é comparável apenas consigo próprio, temos que $d_P = d_H$.*

Chapter 3

Códigos Perfeitos e Raio de Empacotamento

Nosso objetivo neste capítulo é estudar os assim chamados $\mathcal{C}$*ódigos perfeitos*. O uso do adjetivo não é superlativo. Lembremos que dado um código $\mathcal{C}$ com raio de empacotamento $R_e(\mathcal{C})$, consideramos todas as bolas com raio R_e centradas nos pontos do código. Ao se receber uma mensagem e constatar-se a existência de erro, verifica-se em qual destas bolas a mensagem recebida se encontra e corrigimos (assim esperamos) o erro assumindo que a mensagem enviada é a mais próxima da recebida, ou seja, o centro da bola em questão. Uma das situações problemáticas para a correção de erros ocorre quando a mensagem recebida não pertence a nenhuma destas bolas. Um código é perfeito quando esta situação não pode ocorrer.

Definição 3.1 *Diremos que um código linear $\mathcal{C} \subset \mathbb{F}_p^n$ é um* $\mathcal{C}$ódigo perfeito *de existe $r \in \mathbb{N}$ tal que a união de todas as bolas de raio r centradas nos elementos de $\mathcal{C}$ é igual a $\mathbb{F}_p^n$ sendo esta união disjunta. Em outras palavras, $\mathcal{C}$ é perfeito se*

$$\bigcup_{u \in \mathcal{C}} B(u; r) = \mathbb{F}_p^n$$

e

$$B(u; r) \cap B(v; r) = \varnothing$$

qualquer que seja $u, v \in \mathcal{C}$ *com* $u \neq v$.

Temos neste caso que $r = R_e(\mathcal{C})$.

Nas próximas duas seções veremos duas famílias de códigos que são perfeitos, dependendo da métrica adotada.

3.1 Códigos de Hamming

Seja $n = 2^r - 1$, com $r \geq 2$, e H_r a matriz de ordem $r \times (2^r - 1)$ cujas colunas são todos os vetores não nulos de $\mathbb{F}_2^r$. Observe que H_r contém um máximo de r linhas linearmente independentes. De fato, o número de linhas e colunas linearmente independentes (LI) é sempre igual. As colunas de H_r contém uma base de $\mathbb{F}_2^r$, portanto tem ao menos r colunas LI e não pode conter mais do que r colunas LI pois todo subconjunto de um espaço vetorial contendo mais elementos do que a dimensão é linearmente dependente.

O código linear

$$\mathcal{H}_r = \left\{ x \in \mathbb{F}_2^{2^r-1} : H_r \cdot x = 0 \right\}$$

que tem H_r como matriz de verificação de paridade, é chamado de $\mathcal{C}$*ódigo de Hamming*. Temos que $\mathcal{H}_r$ é um $[2; 2^r - 1; 2^r - r - 1]$ código linear. É possível mostrar que, considerando a métrica de Hamming, a distância mínima do código de Hamming é igual a 3 ([HP, Corolário 1.4.14]), obtendo o seguinte:

Lema 3.1 *Um código de Hamming é um* $[2; 2^r - 1; 2^r - r - 1; 3]$ *código linear.*

Exemplo 3.1 *Temos que* $\mathcal{H}_2 = \{(0,0,0), (1,1,1)\}$ *é um* $[3; 1; 3]$ *código de Hamming com matriz de verificação de paridade*

$$H_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Este é o mesmo código apresentado no exemplo 2.1, onde podemos verificar que

$$B((0,0,0);1) \cap B((1,1,1);1) = \varnothing$$

e

$$B\left((0,0,0);1\right) \cup B\left((1,1,1);1\right) = \mathbb{F}_2^3,$$

de modo que o código é perfeito e seja qual for a mensagem $x \in \mathbb{F}_2^3$ *recebida saberemos sempre o que fazer com ela: trocamos* x *pelo centro da bola que* x *pertence. Observe que* $\mathcal{H}_2$ *tem a capacidade de corrigir* $R_e\left(\mathcal{H}_2\right) = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$ *erro.*

Vamos demonstrar que todo código de Hamming é perfeito se considerarmos a métrica de Hamming d_H.

Seja $x \in \mathbb{F}_p^n$ e $r \in \mathbb{N}$. Note inicialmente que $\#\left(B_H\left(x;r\right)\right)$ é igual a $\#\left(B_H\left(0;r\right)\right)$, pois a translação $y \longmapsto y - x$ estabelece uma bijeção entre $B\left(x;r\right)$ e $B\left(0;r\right)$. Se $y \in B\left(0;r\right)$ então y possui exatamente i coordenadas não nulas para algum $i \leq r$. Sendo assim temos $\binom{n}{i}$ possíveis escolhas para as coordenadas de y. Cada coordenada não nula de y pode assumir os valores $1, 2, \ldots, p-1$, de modo que temos um total de $\binom{n}{i}\left(p-1\right)^i$ vetores que distam i de um ponto. Portanto

$$\#\left(B_H\left(x;r\right)\right) = \sum_{i=0}^{r} \binom{n}{i} \left(p-1\right)^i.$$

Teorema 3.1 *O código de Hamming* $\mathcal{H}_r$ *é um código perfeito.*

Demonstração Sabemos pelo Lema 3.1 que a distância mínima de $\mathcal{H}_r$ é 3. Daí segue que o raio de empacotamento de $\mathcal{H}_r$ é igual a $\left\lfloor \frac{3-1}{2} \right\rfloor = 1$. Afirmamos que as bolas de raio 1 centradas nos elementos de $\mathcal{H}_r$ cobrem $\mathbb{F}_2^{2^r-1}$. De fato, como

$$\#\left(B_H\left(u;1\right)\right) = \binom{2^r-1}{0} + \binom{2^r-1}{1} = 1 + \left(2^r - 1\right) = 2^r$$

para todo $u \in \mathcal{H}_r$ e

$$2^r \cdot \#\mathcal{C} = 2^r \cdot \left(2^{2^r-r-1}\right) = 2^{2^r-1} = \#\left(\mathbb{F}_2^{2^r-1}\right)$$

concluímos que as bolas $B_H\left(u;1\right)$ cobrem $\mathbb{F}_2^{2^r-1}$. Portanto $\mathcal{H}_r$ é um código perfeito. □

3.2 Códigos de Hamming Estendidos

Dado o código de Hamming $\mathcal{H}_r$ com matriz de verificação de paridade H_r, acrescentamos a esta uma linha com todas as entradas iguais a 1 e completamos a última coluna com 0's:

$$\widehat{H}_r = \left(\begin{array}{ccc|c} 1 & \dots & 1 & 1 \\ \hline & & & 0 \\ & H_r & & \vdots \\ & & & 0 \end{array}\right).$$

O *Código de Hamming estendido* $\widehat{\mathcal{H}}_r$ é definido como o código linear que tem $\widehat{H}_r$ como matriz de paridade:

$$\widehat{\mathcal{H}}_r = \left\{x \in \mathbb{F}_2^{2^r} : \widehat{H}_r \cdot x = 0\right\}.$$

Se lembrarmos que H_r é uma matriz com $2^r - 1$ colunas e um máximo de r linhas linearmente independentes, é imediato verificar que $\widehat{H}_r$ tem 2^r colunas e $r+1$ linhas LI, de modo que $\widehat{\mathcal{H}}_r$ é um $[2; 2^r; 2^r - r - 1]$ código linear.

Considerando-se a métrica de Hamming, é possível demonstrar ([HP, Corolário 1.4.14]) que a distância mínima de $\widehat{\mathcal{H}}_r$ é $d\left(\widehat{\mathcal{H}}_r\right) = 4$, de modo que o raio de empacotamento é $R_e\left(\widehat{\mathcal{H}}_r\right) = \left\lfloor \frac{4-1}{2} \right\rfloor = 1$.

Conforme vimos anteriormente, em um $\mathbb{F}_q^n$, considerando-se a métrica de Hamming, temos que $\#\left(B_H\left(0;1\right)\right) = n\left(q-1\right)+1$. Assim, considerando $\widehat{\mathcal{H}}_r \subset \mathbb{F}_2^{2^r}$ temos que $\#\left(B_H\left(0;1\right)\right) = 2^r + 1$. Observe ainda que $\widehat{\mathcal{H}}_r$ tem dimensão $2^r - r - 1$, de modo que possui 2^{2^r-r-1} elementos. As bolas de raio centradas nestes pontos são disjuntas mas a união destas têm $(2^r+1)\cdot 2^{2^r-r-1}$ pontos e como 2^r+1 é ímpar, temos que

$$\#\left(B_H\left(u;1\right)\right) \cdot \#\left(\widehat{\mathcal{H}}_r\right) = (2^r+1)\cdot 2^{2^r-r-1} < 2^{2^r} = \#\left(\mathbb{F}_2^{2^r}\right),$$

ou seja, estas bolas não cobrem o espaço $\mathbb{F}_2^{2^r}$ e, considerando-se a métrica de Hamming, $\widehat{\mathcal{H}}_r$ não é perfeito.

Exemplo 3.2 *Se*

$$\widehat{H}_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

então $\widehat{\mathcal{H}}_2 = \{0000, 1111\}$. *Neste caso temos que*

$$B_H(0000; 1) = \{0000, 1000, 0100, 0010, 0001\}$$

e

$$B_H(1111; 1) = \{1111, 0111, 1011, 1101, 1110\}$$

de modo que

$$\mathbb{F}_2^4 \setminus (B_H(0000; 1) \cup B_H(1111; 1)) =$$
$$= \{1100, 1010, 1001, 0110, 0101, 0011\} \tag{3.1}$$

e o código não é perfeito. Mais ainda, caso a palavra recebida seja uma das palavras em 3.1, esta não pode ser decodificada.

Em contraste com a situação clássica descrita anteriormente, exibiremos uma métrica d_P ponderada por uma ordem parcial que torna $\widehat{\mathcal{H}}_r$ um código perfeito. Mais ainda, a capacidade de correção de $\widehat{\mathcal{H}}_r$ aumenta para 2. Vamos aos detalhes.

Seja $[2^r] = \{1, 2, \ldots, 2^r\}$ munido com a ordem $[1, 2^r]$-semi-fraca (Exemplo 2.8). Lembramos que nesta ordem as únicas relações existentes são $i \leq i$ e $1 \leq i$ para $i = 1, 2, 3, \ldots, 2^r$.

$$itbpFU2.8167in2.1179in0inFigura5 - Ordem$$

$[1, 4]$-semi-fracafig.5.wmf

Sendo assim, se pretendemos mostrar que $\widehat{\mathcal{H}}_r$ é perfeito temos que determinar inicialmente a cardinalidade das P-bolas em $\mathbb{F}_2^{2^r}$ e para isto basta determinar $\#(B_P(0; r))$.

Seja $x \in B_P(0; r)$. Então $\omega_P(x) = i$ com $i \leq r$. Se $i = 0$ ou 1 temos exatamente um vetor em $\mathbb{F}_2^{2^r}$ com peso 0 ou 1 respectivamente, a saber $(0, 0, \ldots, 0)$ e $(1, 0, \ldots, 0)$. Se $i > 1$ então temos em $[2^r]$ um

total de $\binom{2^r-1}{i-1}$ ideais com cardinalidade i, determinados pela escolha de $i-1$ elementos diferentes de 1, que são os elementos maximais. Em $\mathbb{F}_2^{2^r}$, cada coordenada associada a um dos $i-1$ elementos maximais escolhidos em $[2^r]$ deve ser igual a 1, enquanto as associadas aos outros elementos maximais devem ser nulas. Já a primeira coordenada, cujo suporte está contido no ideal gerado por qualquer elemento de $[2^r]$, esta pode assumir qualquer um dos dois valores possíveis, 0 ou 1. Assim, para $i > 1$, temos que $2 \cdot \#\left(S\left(0;i\right)\right) = 2 \cdot \binom{2^r-1}{i-1}$ possibilidades para x. Portanto

$$\#\left(B_P\left(0;r\right)\right) = 1 + 1 + \sum_{i=2}^{r} 2\binom{2^r-1}{i-1}.$$

Considerando $\widehat{\mathcal{H}}_r$ como sendo um P-código, temos que as bolas de raio 2 centradas nos pontos de $\widehat{\mathcal{H}}_r$ cobrem todo o espaço $\mathbb{F}_2^{2^r}$ e são duas a duas disjuntas, ou seja, com esta P-métrica, $\widehat{\mathcal{H}}_r$ é perfeito e o raio de empacotamento é 2 (lembre que no caso da métrica de Hamming o raio de empacotamento de $\widehat{\mathcal{H}}_r$ é 1).

Antes de tudo mostraremos que $B_P\left(u;2\right) \cap B_P\left(v;2\right) = \varnothing$ para todo $u, v \in \widehat{\mathcal{H}}_r$ com $u \neq v$. É suficiente provar que $B_P\left(0;2\right) \cap B_P\left(u;2\right) = \varnothing$ para todo $0 \neq u \in \widehat{\mathcal{H}}_r$. Seja $u \in \widehat{\mathcal{H}}_r$ e suponha que exista $x \in \mathbb{F}_2^{2^r}$ tal que

$$x \in B_P\left(0;2\right) \cap B_P\left(u;2\right).$$

Já sabemos que o único elemento de peso 1 é $(1, 0, \ldots, 0)$. Já os elementos de peso 2 são aqueles que têm exatamente uma coordenada não nula a partir da segunda posição. Como $\omega_P\left(u\right) \geq 4$ então u possui pelo menos três posições entre $2, 3, \ldots, 2^r$ iguais a 1. Como x tem no máximo uma destas posições não nulas, temos que a diferença $u - x$ tem no minímo duas dentre estas posições não nulas, de modo que $d_P\left(x,u\right) = \omega_P\left(u-x\right) \geq 3$, o que contradiz a hipótese de termos $x \in B_P\left(0;2\right)$ e temos que $B_P\left(0;2\right) \cap B_P\left(u;2\right) = \varnothing$ para todo $u \in \widehat{\mathcal{H}}_r \setminus \{0\}$.

Mas cada bola de raio 2 têm

$$\#\left(B_P\left(x;2\right)\right) = 1 + 1 + 2 \cdot \left(2^r - 1\right) = 2^{r+1}$$

elementos e $\widehat{\mathcal{H}}_r$ tem 2^{2^r-r-1} elementos temos que

$$\#\left(B\left(u;2\right)\right) \cdot \#\left(\widehat{\mathcal{H}}_r\right) = 2^{r+1}2^{2^r-r-1} = 2^{2^r} = \#\left(\mathbb{F}_2^{2^r}\right),$$

donde segue que $\widehat{\mathcal{H}}_r$ é um P-código perfeito.

3.3 Códigos sobre Ordens Totais

Vimos na proposição 2.2 que, considerando a métrica de Hamming temos que para qualquer código linear $\mathcal{C}$ o raio de empacotamento é determinado pela distância mínima do código: $R_e(\mathcal{C}) = \left\lfloor \frac{d_H - 1}{2} \right\rfloor$.

Em um espaço métrico (X, d), usando apenas a desigualdade triangular, e independentemente de qualquer estrutura adicional, dados $x, y \in X$ e $r = d(x, y)$, temos que $B(x, s) \cap B(y, s) = \varnothing$ se $s < r/2$ e $x, y \in B(x, s) \cap B(y, s)$ se $s \geq r$. Segue, no caso particular de trabalharmos com um código $\mathcal{C} \in \mathbb{F}_q^n$ que, independetemente da métrica em questão

$$\left\lfloor \frac{d_P - 1}{2} \right\rfloor \leq R_e(\mathcal{C}) \leq d_P - 1. \tag{3.2}$$

Nada podemos afirmar a priori sobre a intersecção destas bolas se $r/2 \leq s < r$, mas já sabemos que a primeira das desigualdades ($\left\lfloor \frac{d_P - 1}{2} \right\rfloor \leq R_e(\mathcal{C})$) não é necessariamente justa, pois mostramos na seção anterior que considerando P a ordem $[1; 2^r]$-semi-fraca, temos que o código de Hamming estendido têm distância mínima $d_P(\mathcal{C}) = 4$ e raio de empacotamento $R_e\left(\widehat{\mathcal{H}}_r\right) = 2$, e neste caso

$$\left\lfloor \frac{d_P - 1}{2} \right\rfloor < R_e\left(\widehat{\mathcal{H}}_r\right).$$

Vamos ver agora, através de um exemplo, que a segunda das desigualdades em 3.2 não é estrita, ou seja, vamos exibir um código $\mathcal{C}$ e uma ordem P tais que $R_e(\mathcal{C}) = d_P(\mathcal{C}) - 1$.

Exemplo 3.3 *Seja $P = ([3], R)$ a ordem total $1 \leq 2 \leq 3$. Afirmamos que o raio de empacotamento do código $\mathcal{C} = \{000, 101\}$ é 2, que coincide com $d_P(\mathcal{C}) - 1$ já que $\omega_P(101) = 3 = d_P(\mathcal{C})$. De fato, note inicialmente que as bolas de raio 2 centradas nos elementos de $\mathcal{C}$ são disjuntas:*

$$B_P(000; 2) = \{000, 010, 110\}$$

e

$$B_P(101;2) = \{101, 001, 011, 111\}.$$

O mesmo não acontece se aumentamos o raio das bolas: 001 *pertence a interseção* $B_P(000;3) \cap B_P(101;3)$. *Concluímos então que* $R_e(\mathcal{C}) = 2$.

O exemplo acima pode ser extendido para dimensões arbitrárias. Isto será feito com o auxílio do próximo lema.

Lema 3.2 *Seja* $[n] = \{1, 2, \ldots, n\}$ *munido com a ordem total* $1 \leq 2 \leq \ldots \leq n$. *Se* $u \in B_P(v;r)$ *e* $w_P(v) > r$, *então*

$$\langle supp(u) \rangle = \langle supp(v) \rangle,$$

ou seja, $\max\{supp(u)\} = \max\{supp(v)\}$.

Demonstração Seja $u \in B_P(v;r)$. Se $\langle supp(u) \rangle \subset \langle supp(v) \rangle$, então $\langle supp(v) \rangle = \langle supp(v-u) \rangle$. Daí segue que $d_P(u,v) = \omega_P(v) > r$, o que é um absurdo. Suponha agora que $\langle supp(u) \rangle \supset \langle supp(v) \rangle$. Então $\langle supp(u) \rangle = \langle supp(v-u) \rangle$ donde segue que $d_P(u,v) = \omega_P(u)$. Como $\langle supp(u) \rangle \supset \langle supp(v) \rangle$ e por hipótese $w_P(v) > r$, segue que $\omega_P(u) > r$. Logo $d_P(u,v) > r$ o que contraria o fato de que $u \in B_P(v;r)$. Como em P as únicas possibilidades são $\langle i \rangle \subset \langle j \rangle$, $\langle j \rangle \subset \langle i \rangle$ ou $\langle i \rangle = \langle j \rangle$, concluímos que $\langle supp(u) \rangle = \langle supp(v) \rangle$. □

Finalmente, podemos demonstrar que nas ordens totais o raio de empacotamento é sempre o máximo possível $(d_P - 1)$, propiciando fartura de códigos perfeitos.

Teorema 3.2 *Seja* $[n] = \{1, 2, \ldots, n\}$ *totalmente ordenado. Se* $\mathcal{C} \subset \mathbb{F}_q^n$ *é um código com distância mínima igual a* d_P, *então* $\mathcal{C}$ *tem a capacidade de corrigir* $d_P - 1$ *erros.*

Demonstração Queremos provar que

$$B_P(0; d_P - 1) \cap B_P(v; d_P - 1) = \varnothing$$

para todo $v \in \mathcal{C}\backslash\{0\}$. Se $u \in B_P(0; d_P - 1) \cap B_P(v; d_P - 1)$, como $\omega_P(v) \geq d_P$, segue do Lema que

$$\langle supp(v) \rangle = \langle supp(u) \rangle$$

já que $u \in B_P(v; d_P - 1)$. Daí tem-se que $\omega_P(u) = \omega_P(v) \geq d_P$, o que é um absurdo pois u também pertence a $B_P(0; d_P - 1)$, ou seja, $\omega_P(u) \leq d_P - 1$. Portanto $B_P(0; d_P - 1) \cap B_P(v; d_P - 1) = \varnothing$. □

Bibliography

[CS] C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal, vol. 27, pp. 379-423 and 623-656, July and October, 1948.

[BG] Brualdi, R., Graves, J. S. and Lawrence, M. - *Codes with a poset metric* - Discrete Mathematics 147 (1995) 57-72.

[LE] Lee, K. - *Automorphism group of the Rosenbloom-Tsfasman space* - Eur. J. Combin. 24 (2003) 607-612.

[LY] Lee, Y. - *Projective systems and perfect codes with a poset metric* - Finite Fields and Their Applications 10 (2004) 105-112.

[MS] MacWilliams, F. J. and Sloane, N. J. A. - *The Theory of Error-Correcting Codes* - North-Holland, 1996.

[HV] Hefez, A., Villela, M. L. T. - *Códigos Corretores de Erros* - Série de Computação e Matemática, 2002.

[HP] Huffman, W. C. and Pless, V. - *Fundamentals of Error-Correcting Codes* - Cambrige, 2003.

[SB] Schröder, B. S. W. - *Ordered Set. An Introduction* - Birkhäuser Boston, 2003.