CAPÍTULO 1

Introdução

O objectivo destas notas é agrupar num único texto toda a matéria dada na cadeira de Combinatória e Teoria de Códigos. O livro *A Fisrt Course in Coding Theroty* de R. Hill [**2**] continua a ser uma referência para esta cadeira, embora não cubra toda a matéria dada, assim como as fichas de exercícios usadas nos anos anteriores.

1. Primeiros exemplos e definições

Consideremos a seguinte situação: fulano X está perdido no meio de uma floresta mas está em contacto com fulano Y que consegue saber onde está X e qual o caminho que este deve tomar. A mensagem que Y gostaria de transmitir a X consiste numa sequência dos símbolos N (Norte), S (Sul), E (Este) e W (Oeste), no entanto o *canal de transmissão* entre Y e X apenas permite usar dois símbolos. Trata-se portanto de *codificar* os quatro pontos cardeais através de um *código binário*. Podemos escolher vários tipos de código.

Exemplo 1.1. Seja $C_1 = \{0, 1, 00, 11\}$ e consideremos a correspondência

$$N \to 0 \qquad S \to 1 \qquad E \to 00 \qquad W \to 11 \tag{1.1}$$

O conjunto C_1 diz-se um *código* binário (em dois símbolos) e a aplicação entre $\{N, S, E, W\}$ e C_1 definida por (1.1) diz-se uma *função de codificação*. Neste exemplo o código não é *unicamente decifrável* pois a mensagem 00 tanto pode significar NN ou E.

Exemplo 1.2. Consideremos agora o código $C_2 = \{0, 01, 011, 0111\}$ e a correspondência

$$N \to 0 \qquad S \to 01 \qquad E \to 011 \qquad W \to 0111 \tag{1.2}$$

Neste caso o código é unicamente decifrável, mas não é *instantâneo* pois é preciso esperar pela próxima palavra, ou pelo fim da mensagem, para se conseguir interpretar cada palavra.

Exemplo 1.3. Consideremos ainda um terceiro código $C_3 = \{0, 10, 110, 1110\}$ e a correspondência

$$N \to 0 \qquad S \to 10 \qquad E \to 110 \qquad W \to 1110 \tag{1.3}$$

Neste caso o código é unicamente decifrável e instantâneo – uma palavra acaba quando se recebe o símbolo 0.

Exemplo 1.4. Consideremos ainda um quarto código $C_4 = \{00, 01, 10, 11\}$ e a correspondência

$$N \to 00 \qquad S \to 01 \qquad E \to 10 \qquad W \to 11 \tag{1.4}$$

Trata-se de um código unicamente decifrável e instantâneo, pois todas as palavras têm o mesmo comprimento. Neste caso C_4 diz-se um *código uniforme*.

Exemplo 1.5. Para finalizar estes exemplos, consideremos o último código $C_5 = \{000, 011, 101, 110\}$ e a correspondência

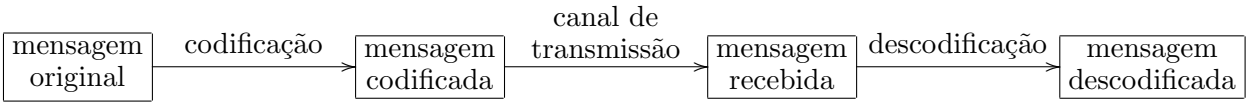
$$N \to 000 \qquad S \to 011 \qquad E \to 101 \qquad W \to 110 \tag{1.5}$$

Tal como no exemplo anterior, C_5 é um código uniforme.

Nesta cadeira iremos considerar apenas códigos uniformes, como os dos Exemplos 1.4 e 1.5. Entre estes, qual o melhor código, C_4 ou C_5? A resposta depende naturalmente do sentido que se der a "melhor". Mas, mesmo sem especificar esse sentido, já podemos comprar C_4 e C_5 nos seguintes aspectos:

- C_4 é um código de comprimento menor do que C_5, portanto é mais rápido transmitir uma mensagem usando C_4.
- C_4 é o conjunto de todas as palavras binárias de comprimento 2 (i.e., $C_4 = (\mathbb{Z}_2)^2$), portanto qualquer palavra recebida é uma palavra de código e, por isso, C_4 não permite detectar erros que ocorram durante a transmissão. Por outro lado, $C_5 \neq (\mathbb{Z}_2)^3$ e portanto C_5 vai permitir detectar alguns erros. Mas sera possível corrigi-los?

A situação geral considerada em Teoria de Códigos pode ser esquematizada na seguinte figura:



As mensagens codificada e recebida são ambas formadas por sequências de palavras do mesmo código. O canal de transmissão poderá ter ruído, de modo que a mensagem recebida poderá conter erros ou símbolos apagados e não será igual à mensagem enviada. O objectivo é estudar códigos tendo em conta certas características como a rapidez de transmissão, facilidade e eficiência de codificar e descodificar, capacidades detectoras e correctoras de erros, etc.

Comecemos então por definir os termos já usados na discussão anterior.

Definição 1.6.
- Um *alfabeto* é um conjunto finito de símbolos $\mathcal{A}_q = \{a_1, \ldots, a_q\}$.
- Uma *palavra* é uma sequência finita de elementos do alfabeto $\mathcal{A}_q$.
- Um *código q-ário* é um conjunto finito de palavras sobre um alfabeto de q elementos.
- Se todas as palavras do código C têm o mesmo comprimento n, i.e. se $C \subseteq \mathcal{A}_q^n$, então C diz-se um *código uniforme*.

Notação 1.7. Um código $(n, M)_q$ significa um código uniforme q-ário com M palavras de comprimento n. Também usamos (n, M) para denotar o mesmo tipo de códigos quando o número de símbolos q está subentendido.

Definição 1.8. Um *esquema de codificação* é um par (C, f) onde

- C é um código,
- $f : \mathcal{S} \to C$ é uma aplicação injectiva, chamada *função de codificação*,
- $\mathcal{S}$ diz-se o *alfabeto fonte*.

O alfabeto fonte pode ou não ser o mesmo do código C. Em todos os exemplos anteriores, o conjunto $\{N, S, E, W\}$ é o alfabeto fonte e o alfabeto do código é $\{0, 1\}$. As correspondências (1.1) a (1.5) definem funções de codificação.

Um alfabeto pode ser qualquer conjunto finito de símbolos à nossa escolha. O conjunto das letras $\{a, b, c, \ldots, x, y, z\}$ é naturalmente um alfabeto, e o conjunto de todas as palavras portuguesas formam um código que não é uniforme.

Os anéis $\mathbb{Z}_m = \mathbb{Z}/\langle m\rangle$ (com $m \geq 2$ um número inteiro) são também alfabetos. No caso particular de $\mathbb{Z}_2 = \{0, 1\}$, o código diz-se binário, e se o alfabeto é $\mathbb{Z}_3 = \{0, 1, 2\}$, o código diz-se ternário. Note-se que $\mathbb{Z}_2$ e $\mathbb{Z}_3$ têm uma estrutura de corpo. Os códigos lineares (Capítulo 4) são uma classe de códigos cujos alfabetos são corpos finitos (estes serão definidos/revistos no Capítulo 3).

A partir de agora, iremos condiderar apenas códigos uniformes, assim "código" significará sempre "código uniforme".

Exemplo 1.9. Fixemos um alfabeto $\mathcal{A}_q$ de q elementos, por exemplo, $\mathcal{A}_q = \mathbb{Z}_q$. O *código de repetição q-ário de comprimento n* é o conjunto formado por q palavras em que os símbolos de cada palavra são todos iguais. Concretamente, $\{0000, 1111\}$ é o código de repetição binário de comprimento 4 e tem parâmetros $(4, 2)$, $\{000, 111, 222, 333, 444\}$ é o código de repetição quinquenário de comprimento 3 e tem parâmetros $(3, 5)$, etc. Em geral, os parâmetros (ver Notação 1.7) de um código de repetição q-ário de comprimento n são (n, q).

Exemplo 1.10. Os parâmetros de um código não o definem univocamente. Seja $C_1 = \{0000, 1111\}$ o código de repetição binário e seja $C_2 = \{1010, 0101\}$. Estes dois códigos têm parâmetros $(4, 2)$, mas $C_1 \neq C_2$.

2. Canal de transmissão

Definição 1.11. Um *canal de transmissão* consiste num alfabeto $\mathcal{A}_q = \{a_1, a_2, \ldots, a_q\}$ e nas probabilidades de canal $P(\text{recebido } a_j \mid \text{enviado } a_i)$, para $i, j \in \{1, \ldots, q\}$, verificando a seguinte condição

$$\sum_{j=1}^{q} P(\text{recebido } a_j \mid \text{enviado } a_i) = 1 \quad \text{, para cada } i \text{ fixo.}$$

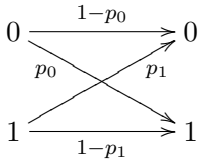
Para simplificar a notação, por vezes escrevemos $P(a_j|a_i)$ para denotar a probabilidade condicionada $P(\text{recebido } a_j \mid \text{enviado } a_i)$, e indicamos as probabilidades do canal através de um grafo onde cada

seta representa uma das probabilidades condicionais da definição

$$a_i \xrightarrow{P(a_j|a_i)} a_j \quad .$$

Vamos agora considerar vários exemplos.

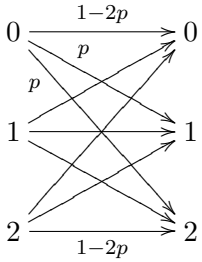
Um *canal de transmissão binário* ($q = 2$) é definido pelos dois valores $p_0 = P(1|0)$ (a probabilidade de troca do símbolo 0) e $p_1 = P(0|1)$ (a probabilidade de troca do símbolo 1), e pode ser representado pelo seguinte esquema



onde o número em cada seta é a probabilidade do símbolo da ponta da seta ser recebido dado que o símbolo da cauda da seta foi enviado. Portanto, neste exemplo, $P(0|0) = 1 - p_0$, $P(1|0) = p_0$, $P(0|1) = p_1$ e $P(1|1) = 1 - p_1$.

Se $p_0 = p_1$, obtém-se um *canal binário simétrico*, um caso particular que iremos usar bastante no resto destas notas. Neste caso, o número $p := p_0 = p_1$ diz-se a *probabilidade de troca de símbolos*.

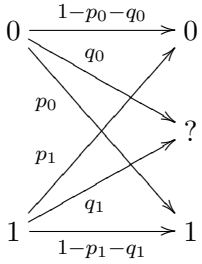
Para $q = 3$, temos o caso particular de um *canal símetrico ternário* com probabilidade de troca $p \in]0, 1[$ definido pelo esquema



onde as setas diagonais têm todas probabilidade p e, portanto, as setas horizontais têm probabilidade $1 - 2p$ (a figura acima está incompleta), ou seja, $P(a_j|a_i) = p$ se $j \neq i$, e $P(a_i|a_i) = 1 - 2p$.

Observação 1.12. Uma vez que, para cada símbolo $a_i \in \mathcal{A}_q$ enviado, se tem $\sum_{j=1}^{q} P(a_j|a_i) = 1$, basta definir as probabilidades $P(a_j|a_i)$ para $i \neq j$. Ou seja, na representação esquemática, basta definir as probabilidades das setas diagonais.

Outro exemplo interessante é o *canal binário de apagamento* definido por



i.e., para além de cada símbolo do alfabeto $\mathcal{A}_2 = \{0, 1\}$ poder ser trocado durante a transmissão, pode ainda ser apagado, o que corresponde a ser enviado para o novo símbolo '?'. É equivalente a usar o alfabeto $\{0, 1, ?\}$ em que o símbolo de apagamento '?' não é usado em nenhuma palavra de código.

Observação 1.13. Nestas notas assumimos sempre que o canal de transmissão é *sem memória*. Por definição, isto quer dizer que a transmissão de cada símbolo é independente das transmissões anteriores, de modo a se verificar a seguinte igualdade

$$P(\text{recebido } y \mid \text{enviado } x) = \prod_{i=1}^{n} P(y_i|x_i) \ , \tag{1.6}$$

onde $x = (x_1, x_2, \ldots, x_n) \in C$ é uma palavra de código e $y = (y_1, y_2, \ldots, y_n)$ é uma palavras arbitrária.

3. Descodificação

Fixemos um código q-ário C de comprimento n, isto é, $C \subseteq \mathcal{A}_q^n$ onde $\mathcal{A}_q$ é um alfabeto com q símbolos.

Um *método de descodificação* é uma correspondência entre palavras de $\mathcal{A}_q^n$ (vistas como as palavras recebidas) e palavras do código C. Caso esta correspondência não esteja definida em todas as palavras de $\mathcal{A}_q^n$, a descodificação diz-se *incompleta*. Nesta secção vamos considerar dois métodos de descodificação.

Definição 1.14. *Descodificação por máxima verosimilhança:* recebido $y \in \mathcal{A}_q^n$, procurar $x' \in C$ tal que

$$P(\text{recebido } y \mid \text{enviado } x') = \max_{x \in C} \{P(\text{recebido } y \mid \text{enviado } x)\} \ .$$

Como C é finito, o conjunto $\{P(\text{recebido } y \,|\, \text{enviado } x) \ : \ x \in C\}$ também é finito e, portanto, o máximo na definição anterior existe sempre, embora possa não ser único.

Exemplo 1.15. Seja $C = \{110, 111\}$ e considere-se um canal binário simétrico com probabilidade de troca $p = 0,03$. Suponhamos que recebemos a palavra 011. Como $011 \notin C$, sabemos que ocorreram erros durante a transmissão. Vamos usar o método de descodificação por máxima verosimilhança.

$$\begin{aligned} P(011 \text{ recebida} \,|\, 110 \text{ enviada}) &= P(0|\,1)P(1|\,1)P(1|\,0) \\ &= p(1-p)p = (0,03)^2 \times 0,97 = 0,000873 \\ P(011 \text{ recebida} \,|\, 111 \text{ enviada}) &= P(0|\,1)P(1|\,0)^2 \\ &= p(1-p)^2 = 0,03 \times (0,97)^2 = 0,028227 \end{aligned}$$

Como a última probabilidade é maior, concluimos que 111 é a palavra de código que provavelmente foi enviada, portanto descodificamos 011 por 111. Note-se que, no primeiro passo no cálculo de cada uma das probabilidades, usou-se a igualdade (1.6).

Exemplo 1.16. Consideremos a mesma situação do exemplo anterior, mudando apenas o código para $C = \{010, 111\}$. Continuamos a ter um canal simétrico binário e a mesma palavra recebida 011.

$$\begin{aligned} P(011 \text{ recebida } | 010 \text{ enviada}) &= P(0|\,0)P(1|\,1)P(1|\,0) \\ &= (1-p)^2 p \\ P(011 \text{ recebida } | 111 \text{ enviada}) &= P(0|\,1)P(1|\,1)^2 \\ &= p(1-p)^2 \end{aligned}$$

Como as duas probabilidades são iguais (e nem dependem do valor de p), o método de descodificação por máxima verosimilhança não nos permite tirar conclusões acerca de qual a palavra enviada com maior probabilidade. Temos então duas alternativas. Ou optamos por uma descodificação incompleta, o que quer dizer que não descoficamos a palavra recebida 011; ou escolhemos uma das palavra de código para descodificar 011 sempre que esta seja recebida. Neste último caso, se decidirmos descodificar 011 por 010, por exemplo, da próxima vez que 011 for recebida, teremos que descodificá-la novamente pela mesma palavra $010 \in C$.

Há esquemas de decisão ou descodificação que não envolvem probabilidades, mas usam uma noção de proximidade.

Definição 1.17. Sejam $x, y \in \mathcal{A}_q^n$. Define-se a *distância de Hamming* entre as palavras x e y por

$$\mathrm{d}(x, y) = \#\{i \,:\, x_i \neq y_i\} \ .$$

Ou seja, $\mathrm{d}(x, y)$ é o número de coordenadas em que x e y diferem, ou ainda, $\mathrm{d}(x, y)$ é o número mínimo de trocas de símbolos necessárias para obter y a partir de x. Por exemplo, $\mathrm{d}(00, 01) = 1$ e $\mathrm{d}(111000, 112012) = 3$.

Exemplo 1.18. Considere-se o alfabeto $\mathcal{A}_4 = \{1, 2, 3, 4\}$ e sejam $x = 1234$, $y = 2341$ e $z = 1243$. Então

$$\begin{aligned} \mathrm{d}(x, y) &= 4 \ , \\ \mathrm{d}(x, z) &= 2 \ , \\ \mathrm{d}(y, z) &= 3 \ . \end{aligned}$$

Definição 1.19. Seja C um código contendo pelo menos duas palavras. Define-se a *distância mínima de* C por

$$\mathrm{d}(C) = \min\{\mathrm{d}(x, y) \,:\, x, y \in C, x \neq y\} \ .$$

Este parâmetro $\mathrm{d}(C)$ vai ter bastante importância quando discutirmos as capacidades de detecção e correcção de um código C.

Notação 1.20. Se C é um código q-ário com M palavras de comprimento n e distância mínima $\mathrm{d}(C) = d$, dizemos que C é um código $(n, M, d)_q$. Os números n, M e d dizem-se os *parâmetros* de C.

Exemplo 1.21. Consideremos o código $C_5 = \{000, 011, 101, 110\}$ definido no Exemplo 1.5. A distância entre $000 \in C_5$ e qualquer outra plavara (de comprimento 3, claro) é o número de símbolos

não nulos nessa palavra, portanto $\mathrm{d}(000, x) = 2$ para qualquer $x \in C \setminus \{000\}$. Calculando a distância entre os restantes pares de palavras de código:

$$\mathrm{d}(011, 101) = 2 \ , \qquad \mathrm{d}(011, 110) = 2 \ , \qquad \mathrm{d}(101, 110) = 2 \ ,$$

conclui-se que $\mathrm{d}(C_5) = 2$ e portanto $(3, 4, 2)_2$ são os parâmetros deste código.

Exemplo 1.22. A distância mímina de um código de repetição q-ário C (definido no Exemplo 1.9) é o comprimento n das palavras, portanto $(n, q, n)_q$ são os parâmetros de C.

Proposição 1.23. *A distância de Hamming é uma métrica, i.e., verifica as seguintes propriedades:*

(i) $\mathrm{d}(x, y) \geq 0 \quad \forall x, y \in \mathcal{A}_q^n$,

(ii) $\mathrm{d}(x, y) = 0 \Leftrightarrow x = y$,

(iii) simetria: $\mathrm{d}(x, y) = \mathrm{d}(y, x) \quad \forall x, y \in \mathcal{A}_q^n$,

(iv) desigualdade triangular: $\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y) \quad \forall x, y, z \in \mathcal{A}_q^n$.

Estas propriedades são consequência directa da definição de distância de Hamming, por isso deixamos a sua demonstração como exercício.

Definição 1.24. *Descodificação por distância mínima:* recebida a palavra $y \in \mathcal{A}_q^n$, procurar $x' \in C$ tal que

$$\mathrm{d}(x', y) = \min\{\mathrm{d}(x, y) \ : \ x \in C\} \ ,$$

ou seja, descodificamos y pela palavra de código mais próxima.

Tal como no caso da descodificação por máxima verosimilhança, por C ser finito, o conjunto $\{\mathrm{d}(x, y) \ : \ x \in C\}$ também é finito e o mínimo na definição anterior existe sempre, embora possa não ser único.

Exemplo 1.25. Consideremos o código binário $C = \{0010, 0101, 1010, 1110\}$ e suponhamos que recebemos a palavra 0100. Como

$$\begin{aligned} &\mathrm{d}(0100, 0010) = 2 \ , \\ &\mathrm{d}(0100, 0101) = 1 \ , \\ &\mathrm{d}(0100, 1010) = 3 \ , \\ &\mathrm{d}(0100, 1110) = 3 \ , \end{aligned}$$

usando o método de descodificação por distância mínima, descodificamos 0100 por 0101.

Exemplo 1.26. Seja $C = \{0000, 1111\}$ o código de repetição de comprimento 4 e consideremos um canal de transmissão binario simétrico com propabilidade de troca $p = \frac{1}{4}$. Pretende-se descodificar a palavra recebida $y = 0010$ pelo dois métodos definidos.

Descodificação por máxima verosimilhança: Temos de calcular as probabilidades condicionadas $P(\text{recebido } y \,|\, \text{enviado } x)$ para $x \in C$. Otém-se

$$\begin{aligned} &P(\text{recebido } y \,|\, \text{enviado } 0000) = (1-p)^3 p = \tfrac{3^3}{4^4} \quad \text{e} \\ &P(\text{recebido } y \,|\, \text{enviado } 1111) = p^3(1-p) = \tfrac{3}{4^4} \ , \end{aligned} \tag{1.7}$$

pois y difere de 0000 em apenas um símbolos e difere de 1111 em três. Como $\frac{3^3}{4^4} > \frac{3}{4^4}$, descodificamos y por 0000.

Descodificação por distância mínima: Temos de calcular as distâncias entre y e cada uma das palavras do código C. Obtêm-se

$$\mathrm{d}(y, 0000) = 1 \qquad \text{e} \qquad \mathrm{d}(y, 1111) = 3 \ ,$$

portanto descodificamos y por 0000, a mesma que se obteve pelo outro método. Não se trata de uma coincidência uma vez que as probabilidades calculadas em (1.7) apenas dependem no número de coordenadas em que x e y diferem, i.e., da distância $\mathrm{d}(x, y)$.

Teorema 1.27. *Para um canal simétrico binário com probabilidade de troca $p < \frac{1}{2}$ os esquemas de descodificação por máxima verosimilhança e por distância mínima coincidem.*

4. Correcção e detecção de erros

Seja $C = \{000, 111\}$ o código de repetição binário de comprimento 3. Se usarmos a descodificação por distância mínima, cada palavra em $\mathcal{A}_2^3$ é descodificada de acordo com a seguinte tabela

recebido	descodificado por
000	000
100, 010, 001	000
011, 101, 110	111
111	111

Caso 1: Se 000 (ou 111) é a palavra enviada e ocorrem erros de transmissão em uma ou duas coordenadas, a palavra recebida y contém exactamente um ou dois símbolos 1. Embora não tenhamos informação para corrigir o erro (admitindo que não conhecemos a palavra enviada), podemos ainda concluir que ocorreram erros pois y não pertence ao código. Dizemos que C detecta até dois erros.

Caso 2: Se a palavra enviada foi 000 e ocorreu um erro na transmissão de um dos símbolos, a palavra recebida foi uma das da segunda linha da tabela, portanto é descodificada correctamente por ela própria. Ou seja, o erro foi corrigido. Analogamente para o caso de ocorrer um erro numa das coordenadas de 111. Caso ocorram dois erros na transmissão de 000, a palavra recebida é descodificada por 111 (terceira linha da tabela). Dizemos que C corrige um erro, mas não corrige dois.

Definição 1.28. Seja C um código e sejam s e t números inteiros positivos.

- Diz-se que C *detecta s erros* sse, quando ocorrem s erros ou menos, a palavra obtida não pertence ao código C.
- Diz-se que C *corrige t erros* sse o método de descodificação por distância mínima corrige t, ou menos, erros.

Em particular, "corrigir" quer dizer que há unicidade de mínimo na definição de descodificação, i.e., está-se a usar um método de descodificação incompleta em que não se descofdifica a palavra recebida em caso de "empate".

Teorema 1.29. *Seja C um código com distância mínima* $\mathrm{d}(C)$*. Então*

(a) C detecta s erros sse $\mathrm{d}(C) \geq s+1$*;*

(b) C corrige t erros sse $\mathrm{d}(C) \geq 2t+1$*.*

Dem. (a) Suponhamos que $\mathrm{d}(C) \geq s+1$. Seja $x \in C$ a palavra enviada e suponhamos que ocorrem no máximo s erros na transmissão e $y \neq x$ é a palavra recebida. Portanto $0 < d(x,y) \leq s$. Como $0 < \mathrm{d}(x,y) < \mathrm{d}(C)$, conclui-se que $y \notin C$ e os s erros são detectados.

Reciprocamente, se $\mathrm{d}(C) \leq s$, então existem palavras $x, y \in C$ tais que $\mathrm{d}(x,y) = \mathrm{d}(C) \leq s$. Logo é possível x ser a palavra enviada, ocorrerem $\mathrm{d}(C)$ erros e recebermos a palavra y. Como $y \in C$, estes erros não são detectados.

(b)($\Longleftarrow$) Suponhamos que $\mathrm{d}(C) \geq 2t+1$. Seja $x \in C$ a palavra enviada e suponhamos que ocorrem t erros na transmissão e $y \neq x$ é a palavra recebida. Portanto $0 < \mathrm{d}(x,y) \leq t$. Para qualquer $c \in C$, com $c \neq x$, temos

$$\mathrm{d}(x,c) \leq \mathrm{d}(x,y) + \mathrm{d}(y,c)$$

logo

$$\mathrm{d}(y,c) \geq \mathrm{d}(x,c) - \mathrm{d}(x,y) \geq \mathrm{d}(C) - t \geq 2t+1-t = t+1 > \mathrm{d}(x,y) \ ,$$

e assim, usando o método de descodificação por distância mínima, y é descodificada correctamente por x.

(b)($\Longrightarrow$) Seja C um código que corrige t erros e suponhamos que $\mathrm{d}(C) \leq 2t$. Então existem $x, x' \in C$ tais que $\mathrm{d}(x,x') = \mathrm{d}(C) \leq 2t$. Seja x a palavra enviada e seja y a palavra recebida com t erros, ou menos, durante a transmissão. Queremos ver que ou y é descodificada erradamente por x', ou existe outra palavra de código $z \in C$, $z \neq x$, tal que $\mathrm{d}(y,x) = \mathrm{d}(y,z)$ (i.e. não há unicidade de mínimo).

Se $\mathrm{d}(x,x') < t+1$, então podiamos ter $y = x'$ pois ocorreriam t erros no máximo, e estes erros nem seriam detectados porque $x' \in C$. Isto contradiz a hipótese de C corrigir t erros, portanto podemos assumir que $\mathrm{d}(x,x') \geq t+1$.

Sem perda de generalidade, podemos também assumir que x e x' diferem precisamente nas primeiras $d = \mathrm{d}(C)$ coordenadas, com $t+1 \leq d \leq 2t$. Seja

$$y = \underbrace{x_1 \cdots x_t}_{\text{como } x'} \underbrace{x_{t+1} \cdots x_d}_{\text{como } x} \underbrace{x_{d+1} \cdots x_n}_{\text{como } x \text{ e } x'} \ .$$

Então as três chavetas contêm t, $d-t$ e $n-t$ coordenadas, repectivamente, e

$$\mathrm{d}(y,x') = d - t \leq t = \mathrm{d}(y,x) \ .$$

Há dois casos a considerar. Ou $\mathrm{d}(y,x') < \mathrm{d}(y,)$ e y é descodificada incorrectamente por x'. Ou $\mathrm{d}(y,x') = \mathrm{d}(y,x)$ e não podemos decidir entre x e x' na descodificação por distância mínima. □

Corolário 1.30. *Seja C um código de distância mínima* $\mathrm{d}(C) = d$*. Então C detecta precisamente* $d-1$ *erros, ou corrige precisamente* $\left\lfloor \dfrac{d-1}{2} \right\rfloor$ *erros.*

Uma vez que a distância de Hamming é uma métrica, podemos definir bolas em $\mathcal{A}_q^n$. A bola de centro x e raio t é o conjunto

$$B_t(x) = \{y \in \mathcal{A}_q^n \ : \ \mathrm{d}(y,x) \leq t\} \subseteq \mathcal{A}_q^n \ .$$

No Teorema 1.29 provámos que, se $\mathrm{d}(C) = 2t + 1$, então quaisquer duas bolas de raio t e centro em palavras de código são disjuntas duas a duas. Assim, se soubermos que ocorrem no máximo t erros de transmissão e y é a palavra recebida, então existe um único $x \in C$ tal que $y \in B_t(x)$, nomeadamente, a palavra enviada.

Iremos voltar a usar esta noção de bola no Capítulo 2.

5. Probabilidade de descodificação (in)correcta

A probabilidade de erro na descodificação associada a um código C, de parâmetros $(n, M, d)_q$, é definida por

$$P_{err}(C) := \sum_{c \in C} P(\text{erro} \mid c \text{ enviado}) P(c \text{ enviado}) \ . \tag{1.8}$$

Naturalmente, precisamos das probabilidades $P(c \text{ enviado})$. Estas probabilidades definem a distribuição de entrada e não dependem de C, mas sim da situação concreta em que o código é usado. Iremos assumir sempre uma *distribuição uniforme*, i.e., $P(c \text{ enviado}) = \frac{1}{M}$.

A probabilidade condicionada $P(\text{erro} \mid c \text{ enviado})$ que ocorre na definição (1.8) denota a probabilidade da palavra enviada c ser descodificada com erro, ou seja, descodificada por uma outra palavra de código qualquer diferente de c. Estas probabilidades condicionadas dependem do canal de tranmissão usado e também podem depender da palavra $c \in C$.

Exemplo 1.31. Consideremos novamente o código binário de repetição de comprimento três, $C = \{000, 111\}$, e um canal de transmissão simétrico binário com probabilidade de troca p. Assumindo uma distribuição de entrada uniforme, tem-se

$$P(000 \text{ enviado}) = P(111 \text{ enviado}) = \frac{1}{2} \ .$$

Calculemos agora as probabilidades de erro condicionadas. Se 000 é a palavra enviada, a descodificação é incorrecta se ocorrerem erros de transmissão em pelo menos dois símbolos, portanto

$$\begin{aligned} P(\text{erro} \mid 000 \text{ enviado}) &= P(\text{recebido } 110, 101, 110 \text{ ou } 111 \mid 000 \text{ enviado}) \\ &= 3p^2(1-p) + p^3 \ . \end{aligned}$$

Analogamente

$$\begin{aligned} P(\text{erro} \mid 111 \text{ enviado}) &= P(\text{recebido } 001, 010, 100 \text{ ou } 000 \mid 111 \text{ enviado}) \\ &= 3p^2(1-p) + p^3 \ . \end{aligned}$$

Neste caso (e não é por acaso) as probabilidades $P(\text{erro} \mid c \text{ enviado})$ não dependem de $c \in C$. A probabilidade de descodificação incorrecta é então dada por

$$P_{err}(C) = \frac{1}{2}(3p^2(1-p) + p^3) + \frac{1}{2}(3p^2(1-p) + p^3) = 3p^2(1-p) + p^3 = 3p^2 - 2p^3 \ .$$

Quando $P(\text{erro} \mid c \text{ enviado})$ não depende de $c \in C$, e a distribuição de entrada é uniforme, a fórmula (1.8) simplifica-se para

$$P_{err}(C) = P(\text{erro} \mid c \text{ enviado})$$

escolhendo uma palavra de código c qualquer. No âmbito desta cadeira, assumir esta condição não é uma grande restrição, pois os códigos lineares satisfazem-na e são estes os códigos que iremos estudar a partir do Capítulo 4.

A probabilidade de descodificação correcta é o "complementar" de $P_{err}(C)$, ou seja,

$$P_{corr}(C) := 1 - P_{err}(C) \ .$$

Esta probabilidade também pode ser definida directamente (e por analogia com (1.8)) por

$$P_{corr}(C) := \sum_{c \in C} P(\text{descodificação correcta} \,|\, c \text{ enviado}) P(c \text{ enviado}) \ . \tag{1.9}$$

Exemplo 1.32. Considere a mesma situação do Exemplo 1.31: código de repetição $C = \{000, 111\}$ e canal de transmição binário simétrico com probabilidade de troca p.

Vamos calcular $P_{err}(C)$ calculando primeiro $P_{corr}(C)$ através de (1.9).

Se 000 é a palavra enviada, a descodificação é correcta se ocorrer no máximo um erro de transmissão, portanto

$$\begin{aligned} P(\text{descodificação correcta} \,|\, 000 \text{ enviado}) &= P(\text{recebido } 100, 010, 001 \text{ ou } 000 \,|\, 000 \text{ enviado}) \\ &= 3p(1-p)^2 + (1-p)^3 \ . \end{aligned}$$

Analogamente

$$\begin{aligned} P(\text{descodificação correcta} \,|\, 111 \text{ enviado}) &= P(\text{recebido } 110, 101, 011 \text{ ou } 111 \,|\, 111 \text{ enviado}) \\ &= 3p(1-p)^2 + (1-p)^3 \ . \end{aligned}$$

A probabilidade de descodificação correcta é então dada por

$$P_{corr}(C) = \frac{1}{2}(3p(1-p)^2+(1-p)^3)+\frac{1}{2}(3p(1-p)^2+(1-p)^3) = 3p(1-p)^2+(1-p)^3 = (1-p)^2(2p+1) \ ,$$

e portanto

$$P_{err}(C) = 1 - P_{corr}(C) = 1 - (1-p)^2(2p+1) = 3p^2 - 2p^3 \ ,$$

que coincide com o resultado obtido no Exemplo 1.31.

Exercícios

1.1. Resolver a Ficha 1.

1.2. Prove o Teorema 1.27, ou seja, prove que, para um canal de transmissão binário e simétrico, com probabilidade de troca $p < \frac{1}{2}$, os métodos de descodificação por distância mínima e por máxima verosimilhança coincidem.

CAPÍTULO 2

O Problema Principal da Teoria de Códigos

1. Enunciado do problema e alguns resultados

Seja C um código q-ário (n, M, d). Define-se *taxa de transmissão* de C por

$$R(C) = \frac{\log_q(M)}{n} \tag{2.1}$$

e defini-se *taxa de correcção de erros*[1] por

$$\delta(C) = \frac{\lfloor \frac{d-1}{2} \rfloor}{n} \ .$$

Exemplo 2.1. Dois casos extremos.

- Para o código binário de repetição de comprimento n, que tem parâmetros $(n, 2, n)$,

  $$R(C) = \frac{\log_2(2)}{n} = \frac{1}{n} \ .$$

  Se $n = 2t + 1$, o código corrige t erros (pelo Teorema 1.29) e

  $$\delta(C) = \frac{t}{n} = \frac{t}{2t+1} \longrightarrow \frac{1}{2} \qquad \text{quando } t \longrightarrow \infty \ .$$

  Por palavras, com n grande, o C corrige "quase" metade dos erros. No entanto, a taxa de transmissão é muito baixa – C apenas contém duas plavaras!
- Com $C = \mathbb{Z}_2^n$, um código de parâmetros $(n, 2^n, 1)$,

  $$R(C) = \frac{\log_2(2^n)}{n} = \frac{n}{n} = 1$$

[1] Esta é uma noção que varia de autor para autor. Nestas notas optou-se por esta definição, mas será muito pouco usada.

é a máxima taxa de transmissão possível mas, como $d = 1$, $\delta(C) = 0$ é mínima!

Os três parâmetros n, M e d de um código estão relacinados. Não é possível ter um código "ideal" com M grande (mais mensagens) e d grande (correcção de mais erros) e n pequeno (taxas de transmissão maiores).

Problema Principal na Teoria de Códigos: Para q, n e d fixos, determinar

$$A_q(n,d) := \max\{M : \exists \text{ código } q\text{-ário } (n,M,d)\} .$$

Ou seja, trata-se de determinar o maior número de palavras possível que um código q-ário de comprimento n e distância mínima d pode conter.

Na continuação do exemplo anterior, podemos deduzir o seguinte resultado.

Proposição 2.2. *Para $n \geq 1$ verifica-se*

- $A_q(n,1) = q^n$,
- $A_q(n,n) = q$ *e*
- $A_q(n,d) \leq q^n$ *para* $1 \leq d \leq n$.

Dem. No primeiro caso, como $d = 1$, todas as palavras são diferentes. O código $C = \mathcal{A}_q^n$ tem q^n palavras e distância mínima 1, logo $A_q(n,1) \geq q^n$. Qualquer outro código de comprimento n é subconjunto deste, logo $A_q(n,1) \leq q^n$.

No segundo caso, como $d = n$, cada palavra tem um símbolo diferente em cada posição (ou coordenada) fixa, logo $A_q(n,n) \leq \#\mathcal{A}_q = q$. Por outro lado, o código de repetição q-ário de comprimento n tem q palavars, logo $A_q(n,n) \geq q$.

No terceiro caso, basta notar que qualquer código C de comprimento n contendo pelo menos duas palavras[2] tem distância mínima $1 \leq d \leq n$ e é subconjunto de $\mathcal{A}_q^n$. Portanto $\#C \leq q^n$, tal como no primeiro caso. □

Para parâmetros n e d arbitrários, determinar $A_q(n,d)$ é um problema extremamamente difícil, e conhecem-se poucos resultados concretos. Para sistematizar a procura e construção de códigos, introduz-se uma noção de equivalência.

Definição 2.3. Seja C um código q-ário (n,M,d). C' diz-se um *código equivalente* a C se é obtido de C através da aplicação sucessiva das seguintes operações:

(i) permutar a ordem das coordenadas de todas as palavros do código, i.e., substituir todo o $c = c_1c_2\cdots c_n \in C$ por $c_{\sigma(1)}c_{\sigma(2)}\cdots c_{\sigma(n)}$, onde σ é uma permutação dos índices $\{1,2,\ldots,n\}$

(ii) permutar os símbolos de todas as palavras na coordenada i (fixa), mais precisamente, substituir todo o $c = c_1c_2\cdots c_n \in C$ por $\pi_1(c_1)\pi_2(c_2)\cdots\pi_n(c_n)$, onde π_1, π_2, ..., π_n são permutações do alfabeto $\mathcal{A}_q$.

[2]Recorde que a distância mínima de um código C só foi definida se $\#C \geq 2$.

Recorde que uma permutação de um conjunto finito X é apenas uma aplicação bijectiva de X em X. Assim, as permutações σ e π_1, ..., π_n na definição anterior são aplicações bijectivas da forma

$$\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\} \qquad \text{ou} \qquad \pi_i : \mathcal{A}_q \longrightarrow \mathcal{A}_q \ .$$

No caso de uma permutação σ do conjunto $\{1, 2, \ldots, n\}$, também escremos

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \ .$$

Por exemplo, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 6 & 5 \end{pmatrix}$ denota a permutação definida por $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$, $\sigma(4) = 4$, $\sigma(5) = 6$ e $\sigma(6) = 5$. Para uma revisão mais aprofundada, recomenda-se a consulta de [**1**].

Exemplo 2.4. Os códigos binários $C_1 = \{000, 111\}$, $C_2 = \{001, 110\}$ e $C_3 = \{100, 011\}$ são todos equivalentes porque:

- C_2 é obtido de C_1 trocando os símbolos 0 e 1 do alfabeto na terceira coordenada, i.e., na notação da Definição 2.3, aplicou-se a operação (ii) com π_3 dada por $\pi_3(0) = 1$ e $\pi_3(1) = 0$;
- C_3 é obtido de C_2 trocando a primeira e a terceira coordenadas das palavras de código, i.e., aplicou-se a operação (i) da Definição 2.3 com $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$.

Lema 2.5. *1. Qualquer código de comprimento n com alfabeto $\mathbb{Z}_q$ é equivalente a um código contendo a palavra $\vec{0} = 00\cdots0 \in \mathbb{Z}_q^n$.*

2. Dados dois códigos C e C' de parâmetros $(n, M, d)_q$ e $(n', M', d')_{q'}$, respectivamente, se C e C' são equivalentes, então $q = q'$, $n = n'$, $M = M'$ e $d = d'$.

Dem. 1. Aplicar permutações de símbolos de modo a uma palavra do código inicial previamente fixa se transformar em $\vec{0}$. Rigorosamente, fixar $c = c_1c_2\cdots c_n \in C$. Escolher permutações π_1, ...,π_n do alfabeto $\mathbb{Z}_q$ tais que $\pi_i(c_i) = 0$. Definir $C' = \{\pi_1(x_1)\cdots\pi_n(x_n) \ : \ x_1\cdots x_n \in C\}$. Portanto C' é equivalente a C e, por construção, $\vec{0} = \pi_1(c_1)\cdots\pi_n(c_n) \in C'$.

2. Directamente da Definição 2.3, tem-se $q = q'$, $n = n'$ e $M = M'$. Só falta ver que $d = d'$.

Trocar a ordem das coordenadas (operação (i)) não altera a distância entre palavras. Analisemos agora a operação (ii). Sejam $x = x_1\cdots x_n$ e $y = y_1\cdots y_n$ duas palavras do código C. Se $x_i = y_i$ então $\pi_i(x_i) = \pi_i(y_i)$ e se $x_i \neq y_i$ então $\pi_i(x_i) \neq \pi_i(y_i)$, porque π_i é uma aplicação bijectiva. Portanto

$$\mathrm{d}(x, y) = \mathrm{d}(\pi_1(x_1)\cdots\pi_n(x_n), \pi_1(y_1)\cdots\pi_n(y_n))$$

e conclui-se que $d = d'$. □

Exemplo 2.6. Vamos provar que $A_2(5, 3) = 4$. (Em [**2**] prova-se também que, a menos de equivalência, existe um único código binário $(5, 4, 3)$.)

1º passo: Mostrar que $A_2(4, 3) = 2$.

Seja C um código $(4, M, 3)$ binário. Sem perda de generalidade, como consequência do Lema 2.5, podemos assumir que $\vec{0} \in C$. Como $\mathrm{d}(C) = 3$ então $\mathrm{d}(x, \vec{0}) \geq 3$ para todo o $x \in C \setminus \{\vec{0}\}$, ou

seja, qualquer palavra de código x não nula tem pelo menos três símbolos 1, ou seja, $x \in X := \{1110, 1101, 1011, 0111, 1111\}$. Para quaisquer duas palavras distintas $y, z \in X$, tem-se

$$\mathrm{d}(y,z) = \begin{cases} 1, & \text{se } y = \vec{1} \text{ ou } z = \vec{1} \\ 2, & \text{se } y \neq \vec{1} \text{ e } z \neq \vec{1} \end{cases} ,$$

em ambos os casos verifica-se $\mathrm{d}(y,z) < 3 = \mathrm{d}(C)$, portanto C contém no máximo uma palavra de X, ou seja, C tem no máximo duas palavras. Como C é um código $(5, M, 3)_2$ arbitrário, provou-se que $A_2(5,3) \leq 2$.

Por outro lado $C = \{0000, 1110\}$ é um código binário de parâmetros $(4, 2, 3)$.

2º passo: Mostar que $A_2(5,3) = 4$.

Seja C um código binário $(5, M, 3)$. Seja[3]

$$C_1 = \{x = x_1x_2x_3x_4x_4 \in C \,:\, x_1 = 1\} \quad \text{e} \quad C_0 = \{x = x_1x_2x_3x_4x_4 \in C \,:\, x_1 = 0\} .$$

O código C_0 tem parâmetros $(5, M_0, d_0)$, com distância mínima $d_o = \mathrm{d}(C_0) \geq \mathrm{d}(C) = 3$ e $M_0 \leq \min\{A_2(4,3), A_2(4,4)\} = 2$ (justifique). Por simetria, também temos $M_1 \leq 2$. Como $C = C_1 \cup C_0$ e $C_1 \cap C_0 = \varnothing$, ou seja, C_1 e C_0 formam uma partição de C, então $M = M_1 + M_2$ e, portanto, $A_2(5,3) \leq 4$.

Por outro lado, $C = \{00000, 01101, 10110, 11011\}$ é um código binário $(5, 4, 3)$.

No resto desta secção, vamos considerar apenas códigos binários, ou seja, o alfabeto é $\mathbb{Z}_2 = \{0, 1\}$. Este conjunto tem uma estrutura de corpo, com as operações definidas pelas seguintes tabelas:

$+$	0	1
0	0	1
1	1	0

e

$\times$	0	1
0	0	0
1	0	1

.

$\mathbb{Z}_2^n$ tem então uma estrutura de espaço vectorial sobre $\mathbb{Z}_2$, com a soma de vectores e produto por um escalar em $\mathbb{Z}_2$ definidos da maneira habitual, coordenada a coordenada. Nomeadamente, se $x = (x_1, x_2, \ldots, x_n) \in \mathbb{Z}_2^n$, $y = (y_1, y_2, \ldots, y_n) \in \mathbb{Z}_2^n$ e $\lambda \in \mathbb{Z}_2$, então

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \qquad \text{e} \qquad \lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) . \tag{2.2}$$

Note que, uma vez que $-1 = 1$ em $\mathbb{Z}_2$, se verifica $x - y = x + y$ para quaisquer vectores $x, y \in \mathbb{Z}_2^n$.

Definição 2.7. Para $x, y \in \mathbb{Z}_2^n$, define-se

- *intersecção:* $x \cap y = (x_1y_1, x_2y_2, \ldots, x_ny_n) \in \mathbb{Z}_2^n$
- *peso:* $\mathrm{w}(x) = \#\{i \,:\, x_i \neq 0\} \in \mathbb{N}_0$

onde x_i e y_i são as coordenadas de x e y, respectivamente.

Por exemplo, se $x = (0,1,1,0,1)$ e $y = (0,0,1,1,0)$ ou, abreviadamente, $x = 01101$ e $y = 00110$, a intersecção é o vector $x \cap y = 00100$ e os pesos destes vectores são $\mathrm{w}(x) = 3$, $\mathrm{w}(y) = 2$ e $\mathrm{w}(x \cap y) = 1$. Note que $x \cap y = y \cap x$, pois a multiplicação em $\mathbb{Z}_2$ é uma operação comutativa.

[3] Estes códigos C_1 e C_0 dizem-se secções de C – ver Capítulo 5.

A noção de peso faz sentido para $\mathbb{Z}_q$ com q arbitrário e iremos considerar também estes casos mais tarde. Note-se que, para o alfabeto binário $\mathbb{Z}_2$, o peso de um vector x é também igual ao número de coordenadas iguais a 1.

Directamente das definições, vemos que

$$\mathrm{d}(x,\vec{0}) = \mathrm{w}(x) \qquad \forall\, x \in \mathbb{Z}_2^n \qquad \text{e} \tag{2.3}$$

$$\mathrm{w}(x \cap y) = \#\{i \,:\, x_i = y_i = 1\} \qquad \forall\, x, y \in \mathbb{Z}_2^n \ . \tag{2.4}$$

Para a última igualdade, convém observar que $ab = 1$ em $\mathbb{Z}_2$ sse $a = b = 1$.

Proposição 2.8. *Para quaisquer vectores $x, y \in \mathbb{Z}_2^n$*

(i) $\mathrm{d}(x,y) = \mathrm{w}(x-y)$,

(ii) $\mathrm{d}(x,y) = \mathrm{w}(x) + \mathrm{w}(y) - 2\,\mathrm{w}(x \cap y)$.

Deixa-se a demonstração desta proposição como exercício. Apenas se observa que a igualdade (ii) é falsa caso usássemos um outro alfabeto $\mathbb{Z}_q$ com $q \neq 2$.

Teorema 2.9. *Seja d um número inteiro positivo ímpar.*
Existe um código binário (n, M, d) sse existe um código binário $(n+1, M, d+1)$.

Dem. ($\Longrightarrow$) Seja C um código binário (n, M, d) e, para cada palavra de código $x = x_1 x_2 \cdots x_n$, defina-se

$$\widehat{x} = \begin{cases} x_1 \cdots x_n 0 & \text{se } w(x) \text{ é par} \\ x_1 \cdots x_n 1 & \text{se } w(x) \text{ é ímpar} \end{cases}$$

Seja $\widehat{C} = \{\widehat{x} \,:\, x \in C\}$. Por construção, $\widehat{C}$ é um código $(n+1, M, \widehat{d})$ com $d \leq \widehat{d} \leq d+1$ — justifique! Além disso $\mathrm{w}(\widehat{x})$ é sempre par, por definição de $\widehat{x}$, e, portanto, $\mathrm{d}(\widehat{x}, \widehat{y})$ também é par para qualquer $\widehat{x}, \widehat{y} \in \widehat{C}$ – nesta última afirmação aplicou-se a igualdade (ii) da Proposição 2.8. Donde se conclui que a distância mínima $d(\widehat{C}) = \widehat{d}$ é par. Atendendo a que $d \leq \widehat{d} \leq d+1$ com d ímpar e $\widehat{d}$ par, concluímos finalmente que $\widehat{d} = d+1$.

($\Longleftarrow$) Seja agora $\widehat{C}$ um código $(n+1, M, d+1)$ e fixemos $\widehat{x}, \widehat{y} \in \widehat{C}$ tais que $\mathrm{d}(\widehat{x}, \widehat{y}) = d+1 = \mathrm{d}(\widehat{C})$. Como esta distância é positiva, podemos escolher uma coordenada i tal que $\widehat{x}_i \neq \widehat{y}_i$. Seja C o código obtido apagando a coordenada i a todas as palavras de $\widehat{C}$, ou seja,

$$C = \{\widehat{z}_1 \cdots \widehat{z}_{i-1} \widehat{z}_{i+1} \cdots \widehat{z}_n \,:\, \widehat{z} \in \widehat{C}\} \ .$$

Deixamos como exercício justificar que o código C contém exactamente M palavras. Quanto à distância mínima $\mathrm{d}(C)$, basta observar que as palavras de C obtidas de $\widehat{x}$ e $\widehat{y}$ estão a uma distância d e usar a definição de $\mathrm{d}(C)$ e $\mathrm{d}(\widehat{C})$. □

As construções de códigos usadas na demonstração anterior são importantes. A primeira é um caso particular de uma *extensão de códigos* chamada *extensão por paridade*, a segunda chama-se *pontuação* – ver Capítulo 5.

Corolário 2.10. *Para d ímpar, $A_2(n,d) = A_2(n+1, d+1)$ ou, equivalentemente, para $d > 0$ par, $A_2(n,d) = A_2(n-1, d-1)$ ou, equivalentemente, para $t \in \mathbb{N}_0$,*

$$A_2(n, 2t+1) = A_2(n+1, 2t+2) \ .$$

2. Estimativas

Nesta secção apresentamos algumas desigualdades envolvendo $A_q(n,d)$ que, recordando da definição dada na página 14, designa o número máximo de palavras que um código q-ário de comprimento n e distância mínima d pode ter. O alfabeto dos códigos será sempre um conjunto arbitrário $\mathcal{A}_q$ de q elementos, sem qualquer estrutura adicional.

2.1. Estimativa de Singleton

Proposição 2.11. *Para q,n e $d \geq 1$ fixos, tem-se*

$$A_q(n,d) \leq q^{n-d+1} \ .$$

Dem. Fixemos um código arbitrário C de parâmetros $(n,M,d)_q$. Queremos mostar que $M \leq q^{n-d+1}$. Apagando as últimas $d-1$ coordenadas (ou outras $d-1$ coordenadas fixas à nossa escolha) de todas as palavras de C, obtém-se um código C' com M palavras de comprimento $n-d+1$ todas distintas entre si porque $d-1 < d = \mathrm{d}(C)$. Portanto $M \leq q^{n-d+1} = \#(\mathcal{A}_q)^{n-d+1}$, pois C' é um subconjunto de $\mathcal{A}_q^{n-d+1}$. □

Os códigos $(n,M,d)_q$ que satisfazem a igualdade $M = q^{n-d+1}$ dizem-se *códigos de distância máxima de separação* (ou simplesmente códigos MDS) e iremos estudar alguns exemplos mais tarde.

2.2. Empacotamento de esferas

Recorde que, usando a distância de Hamming d, se $c \in \mathcal{A}_q^n$ e r é um inteiro não negativo, a bola (ou esfera) de centro c e raio r é o subconjunto de $\mathcal{A}_q^n$ definido por

$$B_r(c) = \{x \in \mathcal{A}_q^n \ : \ \mathrm{d}(x,c) \leq r\} \ .$$

Sendo $\mathcal{A}_q$ um conjunto finito, $\mathcal{A}_q^n$ e qualquer seu subconjunto também o são. Define-se *volume de um subconjunto* S de $\mathcal{A}_q^n$ por

$$\mathrm{vol}(S) = \#S$$

ou seja, o volume de S é o seu cardinal.

Lema 2.12. *O volume da bola $B_r(c)$ é*

$$\mathrm{vol}(B_r(c)) = \sum_{j=0}^{r} \binom{n}{j} (q-1)^j \ ,$$

onde $0 \leq r \leq n$ e $x \in \mathcal{A}_q^n$.

Dem. A bola $B_r(c)$ é a união disjunta dos conjuntos $\{x \in \mathcal{A}_q^n \ : \ \mathrm{d}(x,c) = j\}$ com $j = 0,1,\ldots,r$. Portanto

$$\mathrm{vol}(B_r(c)) = \sum_{j=0}^{r} \#\{x \in \mathcal{A}_q^n \ : \ \mathrm{d}(x,c) = j\} \ .$$

Como

- $\mathrm{d}(x,c) = j$ se e só se x e c diferem exactamente em j coordenadas,

- $\binom{n}{j}$ é o número de maneiras diferentes de escolher j coordenadas em n e
- $q-1$ é o número de símbolos em $\mathcal{A}_q \setminus \{c_i\}$, i.e., o número de escolhas para a coordenada $x_i \neq c_i$,

conclui-se que

$$\#\{x \in \mathcal{A}_q^n \,:\, \mathrm{d}(x,c) = j\} = \binom{n}{j}(q-1)^j \ . \qquad \square$$

Caso $r \geq n$, tem-se obviamente que $B_r(c) = \mathcal{A}_q^n$, cujo volume é q^n.

Exemplo 2.13. Fixemos $\mathcal{A}_2 = \mathbb{Z}_2 = \{0,1\}$. Em $\mathbb{Z}_2^3$, a bola $B_4(001)$ tem volume 8, pois o raio é $r = 4 > 3 = n$, donde $B_4(001) = \mathbb{Z}_2^3$.
A bola de raio 1 e centro na origem em $\mathbb{Z}_2^3$ é o conjunto

$$B_1(000) = \{000, 001, 010, 100\} \ ,$$

sendo 000 o único vector à distância 0 do centro da bola, claro!, e sendo os restantes três elementos 001, 010 e 100 os vectores de $\mathbb{Z}_2^3$ à distância 1 do centro. Portanto $\mathrm{vol}(B_1(000)) = 4$. Também podemos aplicar o Lema 2.12 para o cálculo do volume.

Exemplo 2.14. Calcular o volume de $B_3(1100) \subseteq \mathbb{Z}_2^4$. Aplicando directamente o lema anterior e notando que $q - 1 = 1$ neste caso, fica

$$\mathrm{vol}(B_3(1100)) = \binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} = 1 + 4 + 6 + 4 = 15 \ .$$

Teorema 2.15 (Estimativa de Gilbert-Varshamov ou Minorante de Cobertura de Esferas)**.** *Para $q \geq 2$ e $1 \leq d \leq n$, temos*

$$\boxed{A_q(n,d) \geq \frac{q^n}{\mathrm{vol}(B_{d-1}(c))}} \ . \tag{2.5}$$

Dem. Seja C um código $(n, M, d)_q$ com $M = A_q(n,d)$. Vamos primeiro provar que

$$\forall x \in \mathcal{A}_q^n \quad \exists c \in C \quad \text{tal que} \quad \mathrm{d}(x,c) \leq d-1 \ . \tag{2.6}$$

Suponhamos que não. Nesse caso seja $y \in \mathcal{A}_q^n$ tal que $\mathrm{d}(y,c) \geq d$ para todo o $c \in C$. Em particular $y \notin C$. Então o conjunto $C' = C \cup \{y\}$ é um código $(n, M+1, d)$ [justifique que $\mathrm{d}(C') = d$] o que contradiz a hipótese $M = A_q(n,d)$. Provámos assim (2.6).
Em termos de conjuntos, (2.6) pode-se escrever na forma

$$\mathcal{A}_q^n = \bigcup_{c \in C} B_{d-1}(c)$$

Como $\mathrm{vol}(\mathcal{A}_q^n) = q^n$ e $\mathrm{vol}(\bigcup_{c\in C} B_{d-1}(c)) \leq M\,\mathrm{vol}(B_{d-1}(c))$ [porque é que não se tem necessarimante a igualdade?], obtém-se a desigualdade do enunciado do teorema. $\square$

Teorema 2.16 (Estimativa de Hamming ou Majorante de Empacotamento de Esferas)**.** *Para $q \geq 2$ e $2t+1 \leq d \leq n$, temos*

$$\boxed{A_q(n,d) \leq \frac{q^n}{\mathrm{vol}(B_t(c))}} \ . \tag{2.7}$$

Dem. Seja C um código $(n, M, d)_q$ com $M = A_q(n, d)$ e $d \geq 2t + 1$. Então, pelo Teorema 1.29,

$$B_t(c) \cap B_t(c') = \varnothing \qquad \forall\, c, c' \in C, \text{ com } c \neq c' \ .$$

Ou seja, as M bolas de raio t e centro nas M palavras do código C são disjuntas duas a duas, donde

$$\mathrm{vol}\Big(\bigcup_{c\in C} B_t(c)\Big) = \sum_{c\in C} \mathrm{vol}(B_t(c)) = M\,\mathrm{vol}(B_t(c)) \ , \tag{2.8}$$

uma vez que as bolas com o mesmo raio têm volumes iguais. Como $\mathrm{vol}(\mathcal{A}_q^n) = q^n$, a igualdade (2.8) implica que $M\,\mathrm{vol}(B_t(c)) \leq q^n$. □

Exemplo 2.17. Será que existe um código binário (8,29,3)?

A Estimativa de Singleton dá

$$A_2(8, 3) \leq 2^{8-3+1} = 64 \ .$$

É inconclusivo.

Como

$$\mathrm{vol}(B_2(c)) = \binom{8}{0} + \binom{8}{1} + \binom{8}{2} = 1 + 8 + 28 = 37 \ ,$$

o Minorante de Cobertura de Gilbert-Varshamov dá

$$A_2(8, 3) \geq \frac{2^8}{\mathrm{vol}(B_2(c))} = \frac{256}{37} = 6,9\ldots$$

logo $A_2(8, 3) \geq 7$. Também é inconclusivo.

O Majorante de Empacotamento de Hamming dá

$$A_2(8, 3) \leq \frac{2^8}{\mathrm{vol}(B_1(c))} = \frac{256}{9} = 28,4\ldots$$

logo $A_2(8, 3) \leq 28$ e, portanto, não existem códigos $(8, 29, 3)_2$.

E o que é que acontece quando se verifica a igualdade nas estimativas de Gilbert-Varshamov ou de Hamming?

Seja C um código q-ário de comprimento n qualquer. Define-se *raio de empacotamento* por

$$\rho_e(C) := \max\{r \in \mathbb{N}_0 \,:\, B_r(c) \cap B_r(c') = \varnothing \quad \forall c, c' \in C, \text{ com } x \neq x'\}$$

e *raio de cobertura* por

$$\rho_c(C) := \min\{r \in \mathbb{N}_0 \,:\, \bigcup_{c\in C} B_r(c) = \mathcal{A}_q^n\} \ .$$

Assim, o raio de empacotamento $\rho_e(C)$ é o maior raio possível de modo a todas as bolas de centro em palavras do código serem disjuntas duas a duas. Como não há sobreposições, é possível "empacotar" estas bolas no espaço $\mathcal{A}_q^n$. E o raio de cobertura $\rho_c(C)$ é o menor raio r de modo as bolas de raio r e centro nas palavras de código formarem uma cobertura do espaço $\mathcal{A}_q^n$.

Na demosntração do Teorema 2.15 provou-se que $\rho_c(C) \leq d - 1 = \mathrm{d}(C) - 1$ e na do Teorema 2.16 provou-se que $\rho_e(C) \geq t = \left\lfloor \dfrac{\mathrm{d}(C) - 1}{2} \right\rfloor$. Compare também com o Teorema 1.29 ou o Corolário 1.30.

Definição 2.18. Um código C de parâmetros $(n, M, 2t + 1)_q$ diz-se *perfeito* sse $\rho_e(C) = \rho_c(C)$.

Isto é, as bolas de raio $\rho = \rho_c(C) = \rho_e(C)$ e centro em $c \in C$ são disjuntas duas as duas e formam uma cobertura de $\mathcal{A}_q^n$. Diz-se também que estas bolas constituem um *empacotamento perfeito*, sem "sobreposições" nem "espaços vazios".

Exemplo 2.19. Seja $C = \{000, 111\}$. Uma vez que

$$B_1(000) = \{000, 100, 010, 001\} \qquad \text{e} \qquad B_1(111) = \{111, 011, 010, 110\} \ ,$$

verefica-se directamente que

$$\mathbb{Z}_2^3 = B_1(000) \cup B_1(111) \qquad \text{e} \qquad B_1(000) \cap B_1(111) = \varnothing \ ,$$

donde se conclui que $\rho_c(C) = \rho_e(C) = 1$ e, portanto, C é um código perfeito.

Exemplo 2.20. Códigos perfeitos triviais:

(a) Seja C um código contendo uma palavra apenas, de comprimento n. Neste caso, a distância mínima $\mathrm{d}(C)$ não foi definida, mas como C corrige n erros, convencionamos que $\mathrm{d}(C) = 2n+1$. Deste modo, os parâmetros de C são $(n, 1, 2n+1)_q$ e C verifica a igualdade na estimativa de Hamming (2.7) sendo, portanto, um código perfeito.

(b) $C = \mathcal{A}_q^n$, com parâmetros $(n, q^n, 1)_q$, é um código perfeito, porque o raio de empacotamento $\rho_e(C)$ e o raio de cobertura $\rho_c(C)$ são ambos zero.

(c) Os código de repetição binários de comprimento n ímpar também são perfeitos.

Alguns exemplos de códigos perfeitos não triviais, a ver mais tarde, são os códigos de Hamming e os códigos de Golay.

Exemplo 2.21. Será que existe um código perfeito binário de parâmetros $(7, 16, 3)$?
Como

$$\frac{q^n}{\mathrm{vol}(B_1(c))} = \frac{2^7}{\binom{7}{0} + \binom{7}{1}} = \frac{2^7}{1+7} = \frac{2^7}{2^3} = 2^4 = M$$

os parâmetros $(7, 16, 3)$ satisfazem a igualdade na Estimativa de Hamming. Daqui apenas se pode existir que pode existir um tal código.

Neste caso existe mesmo: o código do Exercício 6 da Ficha 1, que é um exemplo de código de Hamming binário, tem parâmetros $(7, 16, 3)$. Deixamos como exercício verificar que a distância mínima deste código é de facto 3.

2.3. Estimativas de Plotkin

Terminamos esta secção com as estimativas de Plotkin, primeiro enunciadas no caso binário no Teorema 2.22, depois generalizadas para o caso q-ário, com q arbitrário, no Teorema 2.23. As demostrações são parte dos exercícios da Ficha 2.

Teorema 2.22. *Seja C um código binário (n, M, d) com $n < 2d$. Então*

$$M \leq \begin{cases} \dfrac{2d}{2d-n} & \text{se } M \text{ é par} \\ \dfrac{2d}{2d-n} - 1 & \text{se } M \text{ é ímpar} \end{cases} .$$

Teorema 2.23. *Seja* $\theta = \dfrac{q-1}{q}$. *Se* $d > \theta n$, *então* $A_q(n,d) \leq \dfrac{d}{d-\theta d}$.

Note que, pondo $q = 2$ no Teorema 2.23, obtém-se uma estimativa mais fraca do que no Teorema 2.22 no caso de M ímpar.

Exercícios

2.1. Resolver a Ficha 2.

2.2. Mostre que $A_q(n,d) < A_{q+1}(n,d)$.

2.3. Mostre que $A_2(5,4) = 2$.

2.4. (a) Demonstre a Proposição 2.8.
(b) Através de um contra-exemplo, mostre que a segunda alínea da Proposição 2.8 não é verdadeira para vectores em $\mathbb{Z}_3^n$, $n > 1$.

2.5. Usando o Lema 2.12, verifique que o volume das bolas de raio n em $\mathcal{A}_q^n$ é de facto q^n.

2.6. Mostre que, se existe um código perfeito C de parâmetros $(n, M, d)_q$, então $A_q(n,d) = M$ e verifica-se a igualdade nas estimativas de Gilbet-Varshamov e de Hamming.

2.7. Justifique as afirmações do Exemplo 2.20 resolvendo as seguintes alíneas:
(a) Verifique que um código contendo apenas uma palavra satisfaz a igualdade na Estimativa de Hamming.
(b) Para o código $C = \mathcal{A}_q^n$, calcule os raios de empacotamento $\rho_e(C)$ e de cobertura $\rho_c(C)$. Verifique também que C satisfaz a igualdade na Estimativa de Hamming.
(c) Repita a alínea anterior para os código de repetição binários de comprimento ímpar.

CAPÍTULO 3

Corpos Finitos e Espaços Vectoriais

1. Corpos finitos

Nesta secção começamos por rever a definição e algumas propriedades dos anéis $\mathbb{Z}_m$ e também de anéis quocientes de polinómios. Depois introduzimos uma construção dos corpos finitos. Os anéis quocientes de polinómios são úteis quer na construção de corpos finitos, que faremos de seguida, quer na descrição de códigos cíclicos no Capítulo 8. Alguns dos resultas não serão demonstrados, ou porque os alunos já estudaram as demonstrações numa cadeira de álgebra anterior, ou porque não fazem parte do âmbito desta cadeira. Mas, para os alunos interessados, sugere-se a consulta do livro [**1**].

Seja m um número positivo (bastava assumir que $m \neq 0$, mas com $m > 0$ não precisamos de nos preocupar tanto com os sinais). No anel dos número inteiros $\mathbb{Z}$, temos a seguinte *relação de congruência*:

$$a \equiv a' \pmod{m} \iff a - a' = km \quad \text{para algum } k \in \mathbb{Z}$$

i.e. $a, a' \in \mathbb{Z}$ dizem-se congruentes módulo m sse $a - a'$ é divisível por m.

Note-se que, como caso particular, qualquer inteiro a é congruente com o resto r da sua divisão por m. Recorde ainda que, para cada $a \in \mathbb{Z}$, o algoritmo da divisão em $\mathbb{Z}$ garante que o resto r e o quociente q são os *únicos* inteiros tais que

$$a = qm + r \qquad \text{com} \qquad r \in \{0, \ldots, m-1\} \, .$$

Podemos então identificar as classes de equivalência da relação de congruência com os restos da divisão por m. Assim, cada número inteiro pertence exactamente a uma única classe de equivalência e denotamos o conjunto de todas elas por $\mathbb{Z}_m$. Por abuso de linguagem, nem sempre distinguimos entre a classe de equivalência (um conjunto) e os seus representantes (os elementos do conjunto) e

escrevemos

$$\mathbb{Z}_m = \{0, 1, \ldots, m-1\} .$$

Assim, por exemplo, se $m = 3$

$$7 \equiv 4 \equiv -2 \equiv 1 \pmod 3 ,$$

o resto da divisão de 7, 4, -2 e 1 por 3 é sempre 1, e estes inteiros pertencem todos à mesma classe de equivalência módulo 3. A sua classe de equivalência é o conjunto

$$\{1 + 3k : k \in \mathbb{Z}\} = \{\ldots, -5, -2, 1, 4, 7, 11, \ldots\} .$$

Com $m = 3$, há mais duas classes de equivalência, nomeadamente

$$\{0 + 3k : k \in \mathbb{Z}\} = \{\ldots, -6, -3, 0, 3, 6, 12, \ldots\} \quad \text{e}$$
$$\{2 + 3k : k \in \mathbb{Z}\} = \{\ldots, -7, -4, -1, 2, 5, 8, \ldots\} ,$$

e identificamos $\mathbb{Z}_3$ com $\{0, 1, 2\}$.

No caso de $m = 2$, há duas classes de equivalência, uma formada pelos números pares, a outra pelos ímpares, que identificamos com os restos 0 e 1, respectivamente, e escrevemos $\mathbb{Z}_2 = \{0, 1\}$, como já temos feito nos capítulos anteriores.

Proposição 3.1. *Se $a \equiv a' \pmod m$ e $b \equiv b' \pmod m$, então*

(i) $a + b \equiv a' + b' \pmod m$ *e*

(ii) $ab \equiv a'b' \pmod m$.

A proposição anterior permite definir as operações soma e produto em $\mathbb{Z}_m$ à custa das operações respectivas em $\mathbb{Z}$.

Teorema 3.2. *O conjunto $\mathbb{Z}_m$ com a soma e produto definidos pelo Proposição 3.1 é um anel comutativo com identidade, i.e., satisfaz as seguintes propriedades:*

(i) $a + b = b + a$ *e* $ab = ba$ *(comutatividade da soma e produto)*

(ii) $(a + b) + c = a + (b + c)$ *e* $(ab)c = a(bc)$ *(associetividade da soma e do produto)*

(iii) $(a + b)c = ac + bc$ *(distributividade da soma em relação ao produto)*

(iv) $a + 0 = a$ *(existência de elemento neutro da soma, ou zero)*

(v) $a \cdot 1 = a$ *(existência de elemento neutro do produto, ou identidade)*

(vi) $\forall\, a \in \mathbb{Z}_m \quad \exists\, -a \in \mathbb{Z}_m \quad$ *tal que* $\quad a + (-a) = 0$ *(existência de simétrico)*

para quaisquer $a, b, c \in \mathbb{Z}_m$.

Definição 3.3. Um *corpo* $\mathbb{F}$ é um anel comutativo com identidade que satisfaz a seguinte condição

$$\forall\, a \in \mathbb{F} \setminus \{0\} \quad \exists\, a^{-1} \in \mathbb{F} \quad \text{tal que} \quad a \cdot (a^{-1}) = 1 \qquad \text{(existência de inverso)}$$

Exemplo 3.4.

- $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são corpos.
- $\mathbb{Z}$ é um anel, mas não é um corpo.
- O conjunto das matrizes 2×2 de entradas reais, $M_2(\mathbb{R})$, é um anel com identidade mas não é comutativo.
- $< 2 >:=$ {inteiros pares} é um anel comutativo sem identidade.

Um corpo verifica ainda as seguintes propriedades.

Proposição 3.5. *Seja $\mathbb{F}$ um corpo. Então*

(1) $a \cdot 0 = 0$ *para qualquer* $a \in \mathbb{F}$,

(2) $a \cdot b = 0 \implies a = 0$ *ou* $b = 0$ *(lei do corte).*

Exemplo 3.6.

- $\mathbb{Z}_2 = \{0, 1\}$ é um corpo. O único elemento não nulo, o 1, é o seu próprio inverso.
- $\mathbb{Z}_3 = \{0, 1, 2\}$ é um corpo e tem as seguintes tabelas de operações (ou "tabuadas")

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

e

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

.

- $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ não é um corpo porque não satisfaz a lei do corte pois $2 \times 2 \equiv 0 \pmod 4$ mas $2 \not\equiv 0 \pmod 4$. Em particular, 2 não é invertível. No entanto 3 é invertível e o seu inverso, em $\mathbb{Z}_4$, é o próprio 3 pois

$$3 \times 3 = 9 = 1 + 2 \times 4 \equiv 1 \pmod 4 \ .$$

Teorema 3.7. *$\mathbb{Z}_m$ é um corpo sse m é um número primo.*

Nesta secção iremos ver que, embora 4 não seja um número primo, existe um corpo com quatro elementos. E outro com 8, e outro com 9, e muitos mais. Mas não existe nenhum corpo com 6 elementos, nem com 10.

Definição 3.8. Seja $\mathbb{F}$ um corpo finito.

- A *ordem* de $\mathbb{F}$ é o seu cardinal, que passamos a denotar por $|\mathbb{F}|$:

$$|\mathbb{F}| = \#\mathbb{F} = \text{ordem do corpo } \mathbb{F} \ .$$

- Dizemos que $a \in \mathbb{F} \setminus \{0\}$ tem *ordem* $n > 0$, que denotamos por $|a| = n$ ou $\text{ord}(a) = n$, se $a^n = 1$ e $a^k \neq 1$ para $0 < k < n$ ou, equivalentemente,

$$|a| = \min\{n \in \mathbb{N} \,:\, a^n = 1\} \ .$$

- A *característica*[1] de $\mathbb{F}$ é definida por

$$\text{car}(\mathbb{F}) = \min\{n \in \mathbb{N} \,:\, n \cdot 1 := \sum_{i=1}^{n} 1 = 0\} \ ,$$

se este mínimo existe, ou $\text{car}(\mathbb{F}) = 0$ caso contrário.

- $\alpha \in \mathbb{F}$ diz-se um *elemento primitivo*[2] se

$$\mathbb{F} \setminus \{0\} = \{\alpha^i \,:\, i \geq 0\} \ .$$

Notação 3.9. $\mathbb{F}_q$ ou $GF(q)$ (de "Galois Field") designa um corpo de ordem q.

[1] Esta definição faz sentido para qualquer corpo $\mathbb{F}$, não necessáriamente finito.

[2] Também existe a noção de elemento primitivo para corpos não necessarimente finitos, mas a definição aqui apresentada apenas se aplica ao caso finito.

Exemplo 3.10. Os corpos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ têm característica zero.

Exemplo 3.11. Consideremos o corpo $\mathbb{Z}_2 = \{0, 1\}$, ou $\mathbb{F}_2$. Obviamente tem-se que a sua ordem é $|\mathbb{Z}_2| = \#\mathbb{Z}_2 = 2$. Quando à característica, como $1 \neq 0$ e $2 \cdot 1 = 1 + 1 = 0$ em $\mathbb{Z}_2$, conclui-se que $\text{car}(\mathbb{Z}_2) = 2$. Neste caso só existe um elemento não nulo, a identidade, que é também um elemento primitivo de $\mathbb{Z}_2$.

Exemplo 3.12. Consideremos o corpo $\mathbb{Z}_3 = \{0, 1, 2\}$, ou $\mathbb{F}_3$. Tal como no exemplo anterior, $|\mathbb{Z}_3| = 3$. Também temos que $\text{car}(\mathbb{Z}_3) = 3$, pois

$$1 \neq 0\ , \qquad 2 \cdot 1 = 1 + 1 = 2 \neq 0 \qquad \text{e} \qquad 3 \cdot 1 = 1 + 1 + 1 = 3 = 0 \qquad \text{em } \mathbb{Z}_3\ .$$

Como

$$2^2 = 4 \equiv 1 \pmod 3\ ,$$

então $\mathbb{Z}_3 \setminus \{0\} = \{1, 2\} = \{2, 2^2\}$, donde se conclui que 2 é um elemento primitivo de $\mathbb{Z}_3$.

Exemplo 3.13. Seja p um número primo. Pelo Teorema 3.7 sabemos que $\mathbb{Z}_p$ é um corpo. Da construção de $\mathbb{Z}_p$ tem-se directamente que a sua ordem é $|\mathbb{Z}_p| = p$. Por isso também escrevemos $\mathbb{F}_p$ para designar este corpo.

Vamos agora calcular a característica de $\mathbb{Z}_p$. Seja n um inteiro tal que $0 < n < p$. Identificando n com a sua classe que equivalência módulo p, ou seja, pondo $\mathbb{Z}_p = \{0, 1, 2 \ldots, p-1\}$ como temos feito para $p = 2, 3$, tem-se que

$$\sum_{i=1}^{n} 1 = \underbrace{1 + \cdots + 1}_{n \text{ vezes}} = n \not\equiv 0 \pmod p\ ,$$

e

$$\sum_{i=1}^{p} 1 = \underbrace{1 + \cdots + 1}_{p \text{ vezes}} = p \equiv 0 \pmod p\ ,$$

logo $\text{car}(\mathbb{Z}_p) = p$.

Teorema 3.14. *Seja $\mathbb{F}$ um corpo qualquer. Então ou $\text{car}(\mathbb{F}) = 0$ ou $\text{car}(\mathbb{F}) = p$ para algum número primo p.*

Dem. Suponhamos que $\text{car}(\mathbb{F}) = n \neq 0$. Suponhamos também que n não é um número primo. Então existem inteiros a e b tais que $n = ab$ e $1 < a, b < n$. Da definição de característica, tem-se

$$0 = \sum_{i=1}^{n} 1 = n \cdot 1 = (ab) \cdot 1 = (a \cdot 1)(b \cdot 1)$$

em $\mathbb{F}$. Pela Lei do Corte, conclui-se que $a \cdot 1 = 0$ ou $b \cdot = 0$, o que contradiz o facto de n ser o menor inteiro positivo tal que $n \cdot 1 = 0$. □

Não é difícil de ver que, se $\text{car}(\mathbb{F}) = 0$, então $\mathbb{F}$ é um corpo infinito. Portanto, pelo teorema anterior, a característica de qualquer corpo finito é um número primo.

Embora não o vamos demonstrar nestas notas, é verdade que qualquer corpo finito contém um elemento primitivo. A seguinte proposição é bastante útil para averiguar se um dado elemento é primitivo ou não, sem calcular todas as suas potências.

Proposição 3.15. *Seja $\mathbb{F}$ um corpo finito de ordem q. Então*

(i) para todo o $a \in \mathbb{F}$, $a^q = a$,

(ii) para todo o $a \in \mathbb{F} \setminus \{0\}$, $|a|$ divide $q-1$,

(iii) $a \in \mathbb{F} \setminus \{0\}$ é um elemento primitivo de $\mathbb{F}$ sse $|a| = q-1$.

Dem. (i) O resultado é trivial se $a = 0$. Seja então $a \neq 0$, e sejam b_1,..., b_{q-1} os elementos não nulos de $\mathbb{F}$, i.e., seja $\mathbb{F} \setminus \{0\} = \{b_1, \ldots, b_{q-1}\}$. Como $a \neq 0$, também temos que $\mathbb{F} \setminus \{0\} = \{ab_1, \ldots, ab_{q-1}\}$. Assim, multiplicando todos os elementos não nulos de $\mathbb{F}$, fica

$$b_1 \cdots b_{q-1} = (ab_1) \cdots (ab_{q-1})$$

ou ainda

$$b_1 \cdots b_{q-1} = (a^{q-1})(b_1) \cdots b_{q-1}) \ ,$$

donde se obtém $a^{q-1} = 1$.

(ii) Fixemos $a \in \mathbb{F} \setminus \{0\}$ e seja m um intero positivo tal que $a^m = 1$. Sejam r, s inteiros tais que $m = |a|s + r$ e $0 \leq r < |a|$ (i.e., aplicámos o algoritmo da divisão aos inteiros m e $|a|$). Então

$$1 = a^m = a^{|a|s+r} = (a^{|a|})^s a^r = a^r \ ,$$

portanto $r = 0$, por definição de $|a|$, ou seja, $|a|$ divide m. O resultado agora segue da alínea (i), pois podemos escolher $m = q-1$.

(iii) Deixamos como exercício justificar que $a \in \mathbb{F} \setminus \{0\}$ tem ordem $q-1$ sse a, a^2, ..., a^{q-1} são todos distintos. □

Exemplo 3.16. Consideremos o corpo $\mathbb{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$, ou $\mathbb{F}_7$. Vamos determinar todos os seus elementos primitivos, calculando primeiro a ordem dos seus elementos não nulos.

Como $|\mathbb{Z}_7| = 7$ e, pela Proposição 3.15(iii), $|a|$ divide $|\mathbb{Z}_7| - 1 = 6$, então a ordem de qualquer $a \in \mathbb{Z}_7$, com $a \neq 0$, é 1,2,3 ou 6. Portanto basta calcular as potência a^2 e a^3 para decidir se a é primitivo ou não (porquê?). E basta fazê-lo para $a \neq 1$, uma vez que a identidade tem sempre ordem 1. Assim, temos

$$\begin{aligned}
&2^2 = 4 \ , \quad 2^3 = 8 = 1 && \Longrightarrow \quad |2| = 3 \\
&3^2 = 9 = 2 \ , \quad 3^3 = 3^2 \cdot 3 = 2 \cdot 3 = 6 && \Longrightarrow \quad |3| = 6 \\
&4^2 = (-3)^2 = 3^2 = 2 \ , \quad 4^3 = (-3)^3 = -6 = 1 && \Longrightarrow \quad |4| = 3 \\
&5^2 = (-2)^2 = 4 \ , \quad 5^3 = (-2)^3 = -1 = 6 && \Longrightarrow \quad |5| = 6 \\
&6^2 = (-1)^2 = 1 && \Longrightarrow \quad |6| = 2
\end{aligned}$$

donde se conclui que os elementos primitivos de $\mathbb{Z}_7$ são precisamente o 3 e o 5.

Exemplo 3.17. $\mathbb{Z}_4$ não é um corpo porque 4 não é um número primo. Mas será que existe um corpo de ordem 4?

Vamos considerar um conjunto de 4 elementos, $\mathbb{F}_4 = \{0, 1, a, b\}$, com 0 o zero e 1 a identidade de $\mathbb{F}_4$, e tentemos escrever as tabelas para a soma e para o produto de modo a satisfazer as propriedade de corpo.

As propriedades dos elementos zero e identidade forçam a primeira linha da tabela da soma e as duas primeiras linhas da tabela do produto e, por comutatividade das operações, as correspondentes colunas também ficam preenchidas. A lei do corte para o produto força as restantes 4 entradas da tabela do produto. Portanto já se calculou:

$+$	0	1	a	b
0	0	1	a	b
1	1			
a	a			
b	b			

e

$\times$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

A lei do corte para a soma implica que $1+a=0$ ou $1+a=b$. Suponhamos que $1+a=0$. Então, muliplicando por b obtém-se $b+1=0$ e, comprando com $1+a=0$ novamente, ficava $a=b$ o que é falso. Logo tem de ser $1+a=b$. Esta condição, juntamente com as propriedade de corpo e com o produto já definido, permite completar o resto da tabela da soma. Obtém-se, finalmente!, as seguintes tabuadas

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

e

$\times$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

(3.1)

Estas tabelas definem de facto uma estrutura de corpo em $\mathbb{F}_4$. Além disso, da tabela da soma conclui-se que a $\mathbb{F}_4$ tem característica 2 e, da tabela do produto, conclui-se que a e b são elementos primitivos.

Como se viu neste último exemplo, determinar as tabuadas apenas com base nas propriedades de corpo não é tarefa simples, mesmo se já soubermos que existe um corpo de determinarda ordem. Os alunos que já estudaram um pouco de teoria de grupos podiam ter chegado mais rapidamente à tabuada da soma, uma vez que $(\mathbb{F}_4,+)$ é um grupo abeliano e apenas há dois grupos com 4 elementos: a outra escolha para a soma leva às operações de $\mathbb{Z}_4$. Mesmo assim, estas observações nada simplificam no caso de corpos finitos de ordens maiores. Interessa portanto construir os corpos $\mathbb{F}_q$ de uma maneira mais eficiente. Para isso, precisamos de rever primeiro algumas definições e propriedades dos anéis polinómios.

Seja $\mathbb{F}$ um corpo qualquer. Seja $\mathbb{F}[t]$ o conjunto dos polinómios com coeficientes em $\mathbb{F}$, i.e.

$$\mathbb{F}[t] := \{a(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \,:\, n \in \mathbb{N}_0,\, a_i \in \mathbb{F}\}$$

Este conjunto com a soma e produto usuais de polinómios é um anel comutativo com identidade.

Define-se o *grau* de um polinómio $a(t)$ por

$$\deg(a(t)) = \begin{cases} -\infty & \text{se } a(t) \text{ é o polinómio nulo,} \\ \max\{i \in \mathbb{N}_0 \,:\, a_i \neq 0\} & \text{caso contrário.} \end{cases}$$

À semelhança de $\mathbb{Z}$, o anel $\mathbb{F}[t]$ também tem definido um *algoritmo de divisão*, ou seja,

$$\forall\, a(t), b(t) \in \mathbb{F}[b], b(t) \neq 0 \quad \exists\, q(t), r(t) \in \mathbb{F}[t] \quad \text{tais que} \quad a(t) = q(t)b(t) + r(t)$$

onde $\deg(r(t)) < \deg(b(t))$. Naturalmente chamamos quociente a $q(t)$ e resto a $r(t)$.

Fixemos $f(t) \in \mathbb{F}[t] \setminus \{0\}$. Analogamente a $\mathbb{Z}$, definimos uma *relação de congruência* em $\mathbb{F}[t]$ por

$$a(t) \equiv a'(t) \pmod{f(t)} \iff a(t) - a'(t) = k(t)f(t) \text{ para algum } k(t) \in \mathbb{F}[t] \ .$$

e definimos ainda o quociente $\mathbb{F}[t]/f(t)$, ou $\mathbb{F}[t]/\langle f(t)\rangle$, como o conjunto das classes de equivalência que, continuando com a analogia a $\mathbb{Z}$, identificamos com os restos $r(t)$ da divisão por $f(t)$:

$$\begin{aligned}\mathbb{F}[t]/f(t) &= \text{conjunto dos restos da divisão por } f(t) \in \mathbb{F}[t] \\ &= \text{conjunto dos polinómios em } \mathbb{F}[t] \text{ de grau estritamente menor que } \deg(f(t)) \ .\end{aligned}$$

Tal como fizemos para $\mathbb{Z}_m$, sempre que não haja ambiguidades, não distinguimos uma classe de equivalência dos seus representantes. E é isso que fizemos nas igualdades anteriores.

Proposição 3.18. *Seja $f(t) \in \mathbb{F}[t]$ um polinómio não nulo. O quociente $\mathbb{F}[t]/\langle f(t)\rangle$, com a soma e o produto definidos módulo $f(t)$, é um anel comutativo com identidade. Os elementos zero e identidade são representados pelo polinómio nulo $0 \in \mathbb{F}[t]$ e pelo polinómio constante $1 \in \mathbb{F}[t]$, respectivamente.*

Exemplo 3.19. Consideremos o anel dos polinómios com coeficientes em $\mathbb{F}_2$, $\mathbb{F}_2[t]$, e seja $f(t) = t^2 + t + 1$. Como $f(t)$ tem grau 2, temos

$$\mathbb{F}_2[t]/\langle f(t)\rangle = \{a + bt \ : \ a, b \in \mathbb{F}_2\} = \{0, 1, t, t+1\} \ .$$

Deixamos como exercício verificar que as tabelas da soma e produto deste anel são

$+$	0	1	t	$t+1$
0	0	1	t	$t+1$
1	1	0	$t+1$	t
t	t	$t+1$	0	1
$t+1$	$t+1$	t	1	0

e

$\times$	0	1	t	$t+1$
0	0	0	0	0
1	0	1	t	$t+1$
t	0	t	$t+1$	1
$t+1$	0	$t+1$	1	t

Portanto, $\mathbb{F}_2[t]/\langle t^2 + t + 1\rangle$ é um corpo. [Compare com as tabelas (3.1) do Exemplo 3.17.]

Exemplo 3.20. Seja $f(t) = t^2 + 1 \in \mathbb{F}_2[t]$. Tal como no exemplo anterior, tem-se que

$$\mathbb{F}_2[t]/\langle t^2 + 1\rangle = \{0, 1, t, t+1\} \ ,$$

porque $\deg(f(t)) = 2$. Mas as operações soma e produtos são agora dadas pelas tabelas

$+$	0	1	t	$t+1$
0	0	1	t	$t+1$
1	1	0	$t+1$	t
t	t	$t+1$	0	1
$t+1$	$t+1$	t	1	0

e

$\times$	0	1	t	$t+1$
0	0	0	0	0
1	0	1	t	$t+1$
t	0	t	1	$t+1$
$t+1$	0	$t+1$	$t+1$	0

Em particular, como $(t+1)(t+1) = t^2 + 2t + 1 = t^2 + 1 \equiv 0 \pmod{t^2+1}$, não se verifica a lei do corte e, portanto, $\mathbb{F}_2[t]/\langle t^2 + 1\rangle$ não é um corpo.

Pergunta: Quando é que o quociente $\mathbb{F}[t]/f(t)$ é um corpo?

Definição 3.21. Dizemos que $f(t) \in \mathbb{F}[t]$ é um *polinómio redutível* se é possível escrever $f(t) = a(t)b(t)$, em $\mathbb{F}[t]$, com $\deg(a(t)) < \deg(f(t))$ e $\deg(b(t)) < \deg(f(t))$. Caso contrário, dizemos que $f(t)$ é *irredutível*.

Exemplo 3.22. (a) Como $t^2+1=(t+1)(t+1)$ em $\mathbb{F}_2[t]$, o polinómio t^2+1 é redutível em $\mathbb{F}_2[t]$.

(b) t^2+1 é irredutível em $\mathbb{F}_3[t]$. Porquê?

(c) t^2+t+1 é irredutível em $\mathbb{F}_2[t]$. Porquê?

(d) $f(t)=t^4+t^2+1$ é redutível em $\mathbb{F}_2[t]$ porque $f(t)=(t^2+t+1)^2$, mas note que $f(t)$ não possui raízes em $\mathbb{F}_2$: $f(0)=f(1)=1\neq 0$.

Teorema 3.23. *Seja $\mathbb{F}$ um corpo e seja $f(t)\in\mathbb{F}[t]$ um polinómio não nulo. Então o quociente $\mathbb{F}[t]/\langle f(t)\rangle$ é um corpo se e só se $f(t)$ é irredutível em $\mathbb{F}[t]$.*

Portanto

- Se $\mathbb{F}$ é um corpo finito de ordem q e $f(t)\in\mathbb{F}[t]$ é um polinómio irredutível de grau $m>0$, então o quociente $\mathbb{F}[t]/\langle f(t)\rangle$ é um corpo de ordem q^m.
- O Exercício 2(c) da Ficha 3 implica que, para qualquer inteiro positivo m e para qualquer primo p, existe um polinómio irredutível em $\mathbb{F}_p[t]$ de grau m.

E fica assim justificado que exitem corpos de ordem p^m. O resultado geral acerca da existência de corpos finitos é o seguinte:

Teorema 3.24. *(i) Existe um corpo de ordem q se e só se $q=p^m$ para algum primo p e algum inteiro $m\geq 1$.*

(ii) Se E e F são corpos finitos com a mesma ordem, então E e F são isomorfos.

A segunda alínea deste teorema quer dizer que só há um corpo de ordem q, a menos de "mudar os nomes dos seus elementos".

Observação 3.25. Atenção! Se p é um número primo, tem-se que $\mathbb{F}_p=\mathbb{Z}_p$. Mas $\mathbb{F}_{p^m}\ncong\mathbb{Z}_{p^m}$, se $m>1$, porque $\mathbb{Z}_{p^m}$ não é um corpo.

Seja $f(t)$ um polinómio de grau m, irredutível em $\mathbb{F}_q[t]$. Seja α uma raiz de $f(t)$. Então o corpo $\mathbb{F}_{q^m}=\mathbb{F}_q[t]/\langle f(t)\rangle$ também pode ser representado por

$$\mathbb{F}_q[\alpha]=\{a_0+a_1\alpha+\cdots+a_{m-1}\alpha^{m-1}\,:\,a_i\in\mathbb{F}\}\,. \tag{3.2}$$

Se, além disso, α for também um elemento primitivo de $\mathbb{F}_{q^m}$, então

$$\mathbb{F}_{q^m}=\{0,\alpha,\alpha^2,\ldots,\alpha^{q^m-1}\}\,. \tag{3.3}$$

A descrição (3.2) é mais útil para determinar a soma de elementos, a descrição (3.3) é mais útil para calcular o produto de elementos.

Terminamos esta secção enunciando algumas propriedades úteis sobre polinómios (ir)redutíveis.

Proposição 3.26. *Seja $\mathbb{F}$ um corpo e fixemos $f(t)\in\mathbb{F}[t]$. Então*

(a) $t-a$ divide $f(t)$ sse $f(a)=0$ (i.e., se a é raíz de f),

(b) se $f(t)$ tem grau 2 ou 3, então $f(t)$ é irredutível sse não tem raízes em $\mathbb{F}$.

Dem. (a) Se $t-a$ divide $f(t)$, então $f(t)=(t-a)g(t)$ para algum polinómio $g(t)\in\mathbb{F}[t]$. Pondo $t=a$ fica $f(a)=0$.

Reciprocamente, o algoritmo de devisão implica que exitem polinómios $q(t), r(t)\in\mathbb{F}[t]$ tais que $f(t)=q(t)(t-a)+r(t)$ com $\deg(r(t))<\deg(t-a)=1$, ou seja, o resto é um polinómio constante $r(t)=r$. Como $q(a)(a-a)+r=f(a)=0$, obtêm-se $r=0$.

(b) $f(t)$ é redutível sse $f(t)=a(t)b(t)$ com $\deg(a(t))=1$ e $\deg(b(t))=1$ ou 2, porque $\deg(f(t))=2$ ou 3. Logo temos $a(t)=t-a$ para algum $a\in\mathbb{F}$ e, pela alínea (a), a é raíz de $f(t)$. □

A alínea (b) deste teorema permite responder às questões deixadas no Exemplo 3.22 sem recorrer a factorizações de polinómios.

2. Espaços vectoriais sobre corpos finitos

Seja $\mathbb{F}_q$ um corpo finito. A noção de espaço vectorial sobre $\mathbb{F}_q$ é em tudo análoga à estudada na cadeira de Álgebra Linear no caso do corpo dos escalares ser o $\mathbb{R}$. Em particular, não é difícil verificar que o conjunto $(\mathbb{F}_q)^n$ é um espaço vectorial com a soma de vectores e produto por um escalar definidos coordenada a coordenada respectivamente por

$$\vec{x}+\vec{y}=(x_1+y_1,\ldots,x_n+y_n)$$
$$a\vec{x}=(ax_1,\ldots,ax_n)$$

onde $\vec{x}=(x_1,\ldots,x_n)$ e $\vec{y}=(y_1,\ldots,y_n)$ são vectores em $\mathbb{F}_q^n$ e $a\in\mathbb{F}_q$ é um escalar.

O espaço vectorial $\mathbb{F}_q^n$ também costuma ser denotado por $V(n,q)$. Os vectores em $\mathbb{F}_q^n$ serão denotados por $\vec{x}$, $\vec{y}$, etc ou simplesmente por x, y, etc.

Se V é um espaço vectorial, um *subespaço vectorial* de V é um subconjunto $W\subseteq V$, não vazio, tal que W é ele prório um espaço vectorial.

Todos os espaços vectoriais que iremos considerar nesta cadeira serão subespaços de algum $\mathbb{F}_q^n$.

Tal como se verifica para espaços vectoriais reais, também temos o seguinte resultado quando o corpo dos esclares é $\mathbb{F}_q$ (ou mesmo qualquer corpo, finito ou infinito).

Teorema 3.27. *Seja V um espaço vectorial e $W\subseteq V$ com $W\neq\varnothing$. Então, W é um subespaço de V sse W é fechado para a soma e para o produto por um escalar.*

A demosntração é consequência directa dos axiomas de espaço vectorial.

Exemplo 3.28. Os seguintes conjuntos são espaços vectoriais:

(a) $\{\vec{0}\}$ e $\mathbb{F}_q^n$

(b) $V_1=\{(a,a,\ldots,a)\,:\,a\in\mathbb{F}_q\}\subset\mathbb{F}_q^n$

(c) $V_2=\{(0,0,0,0),(1,0,1,0),(0,1,0,1),(1,1,1,1)\}\subset\mathbb{F}_2^4$

(d) $V_3=\{(0,0,0),(0,1,2),(0,2,1)\}\subset\mathbb{F}_3^3$

De seguida iremos ver algumas definições e resultados para espaços vectoriais sobre o corpo finito $\mathbb{F}_q$. Muitos deles são completamente análogos aos já estudados na cadeira de Álgebra Linear, e as demonstrações serão omitidas, mas por vezes o caso finito tem um comportamento diferente.

Definição 3.29. Seja V um espaço vectorial sobre $\mathbb{F}_q$.

- Uma *combinação linear* de vectores $v_1, \ldots, v_r \in V$ é um vector da forma $a_1v_1 + \cdots + a_rv_r \in V$, com $a_i \in \mathbb{F}_q$, $i = 1, \ldots, r$.
- Um conjunto de vectores $\{v_1, \ldots, v_r\} \subseteq V$ diz-se *linearmente independente* se
$$a_1v_1 + \cdots + a_rv_r = \vec{0} \quad \Longrightarrow \quad a_1 = a_2 = \cdots = a_r = 0$$
- Diz-se que $\{v_1, \ldots, v_r\} \subseteq V$ é um *conjunto gerador de V* se qualquer vector em V é combinação linear de $v_1, \ldots, v_r$.
- O *espaço gerado* pelo conjunto $\{v_1, \ldots, v_r\} \subseteq V$ é
$$\langle v_1, \ldots, v_r \rangle = \{a_1v_1 + \cdots + a_rv_r \,:\, a_i \in \mathbb{F}_q\} \,,$$
portanto é o menor subespeço de V que contém os vectores $v_1, \ldots, v_r$.
- Uma *base de V* é um conjunto gerador linearmente independente.

Exemplo 3.30. Continuando o Exemplo 3.28:

- Qualquer conjunto que contem o vector nulo $\vec{0}$ é linearmente dependente.
- O espaço $\{\vec{0}\}$ não contém nenhuma base.
- $\{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\}$ é uma base de $\mathbb{F}_q^4$, é a chamada *base canónica*.
- $\{(1,1,\ldots,1)\}$ é uma base para o espaço V_1 do Exemplo 3.28(b).
- $\{(1,0,1,0), (0,1,0,1)\}$ é uma base para o espaço V_2 do Exemplo 3.28(c) e $\{(1,0,1,0), (1,1,1,1)\}$ é uma outra base para o mesmo espaço.
- $\{(0,1,2)\}$ é uma base para o espaço V_3 do Exemplo 3.28(d).

Teorema 3.31. *Seja $V \neq \{\vec{0}\}$ um espaço vectorial finito sobre $\mathbb{F}_q$. Então qualquer conjunto gerador de V contém uma base.*

Como consequência directa, tem-se que qualquer subespaço $V \neq \{\vec{0}\}$ de $\mathbb{F}_q^n$ tem uma base.

Exemplo 3.32. Seja $V \subseteq \mathbb{F}_3^4$ o espaço gerado pelo conjunto $\{(0,1,2,1), (1,0,2,2), (1,2,0,1)\}$. Vamos determinar uma base para V.

Uma vez que $\mathbb{F}_3$ é um corpo tal como $\mathbb{R}$, o método de eleminação de Gauss continua a ser válido. Aplicando então o método de eleminação de Gauss à matriz (de entradas em $\mathbb{F}_3$) cujas linhas são os vectores do conjunto dado, fica
$$\begin{bmatrix} 1 & 0 & 2 & 2 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 2 \\ 0 & 2 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 2 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
portanto, olhando para a última matriz, conclui-se que $(0,1,2,1)$ é combinação linear de $(1,0,2,2)$ e $(1,2,0,1)$. Conclui-se ainda que $\{(1,0,2,2), (1,2,0,1)\}$ é uma base para V.

Teorema 3.33. *Seja $V \neq \{\vec{0}\}$ um espaço vectorial finito sobre $\mathbb{F}_q$. Então*

(i) Fixada uma base $\mathcal{B} = \{v_1, \ldots, v_r\}$ de V, qualquer vector $v \in V$ se escreve de maneira <u>única</u> como combinação linear dos vectores da base $\mathcal{B}$, i.e.,
$$\forall\, v \in V \quad \exists!\, a_1, \ldots, a_r \in \mathbb{F}_q \quad \textit{tais que} \quad v = a_1v_1 + \cdots + a_rv_r \,.$$

(ii) Qualquer base de V tem o mesmo número de elementos.

Definição 3.34. A *dimensão* do espaço vectorial $V \subseteq \mathbb{F}_q^n$ é o número de elementos de uma base de V, se $V \neq \{\vec{0}\}$, e $\dim(V) = 0$ se V é o espaço nulo.

Note que o Teorema 3.33 garante que a definição anterior não depende da base escolhida para V.

Exemplo 3.35. Consideremos novamente os espaços V_1, V_2 e V_3 do Exemplo 3.28. Então, pelo que foi dito no Exemplo 3.30, $\dim(V_1) = 1$, $\dim(V_2) = 2$ e $\dim(V_3) = 1$.

Tratando-se de um espaço vectorial finito, não é sempre necessário determinar uma base para para calcular a sua dimensão.

Seja $V \neq \{\vec{0}\}$ um subespaço vectorial de $\mathbb{F}_q^n$ e seja k a sua dimensão, que queremos determinar. Também sabemos que V possui uma base com k vectores, e designemos por v_1,...,v_k os vectores dessa base. Quantos elementos é que V contém? Como qualquer $v \in V$ se escreve de maneira única como uma combinação linear de v_1,...,v_k (pelo Teorema 3.33 (i)), então escolhas diferentes de coeficientes escalares dão origem a vectores distintos de V, i.e,

$$a_1 v_1 + \cdots + a_k v_k = b_1 v_1 + \cdots + b_k v_k \ ,$$

com $a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{F}_q$, sse $a_i = b_i$ para $i = 1, \ldots, k$. Ou seja V contém exactamente q^k vectores, onde $k = \dim V$. Assim, provámos que o número de elementos em V é uma potência de $q = |\mathbb{F}_q|$ e que a sua dimensão é

$$\boxed{\dim V = \log_q(\#V)} \ . \tag{3.4}$$

Já agora repare que a fórmula (3.4) também é válida quando V é o espaço nulo.

Definição 3.36. Sejam $u, v \in \mathbb{F}_q^n$, de coordenadas $u = (u_1, \ldots, n_n)$ e $v = (v_1, \ldots, v_n)$.

(i) O *produto interno euclideano* dos vectores u e v é o escalar $u \cdot v = u_1 v_1 + \cdots u_n v_n \in \mathbb{F}_q$.

(ii) Dizemos que u é *ortogonal* a v, e escrevemos $u \perp v$, sse $u \cdot v = 0$.

Proposição 3.37. *Para quaisquer vectores $u, v, w \in \mathbb{F}_q^n$ e escalares $a, b \in \mathbb{F}_q$, verifica-se*

(i) $u \cdot v = v \cdot u$ *(simetria)*

(ii) $(au + bv) \cdot w = a(u \cdot w + b(v \cdot w)$ *(linearidade)*

(iii) $u \cdot v = 0$ *para todo o* $u \in \mathbb{F}_q^n$ *se e só se* $v = \vec{0}$ *(não degenerescência)*

Dem. As alíneas (i) e (ii) provam-se aplicando directamente a definição de produto interno e usando as propriedades de comutatividade e distributividade do corpo $\mathbb{F}_q$.

Na alínea (iii), se $v = \vec{0}$ então, usando outra vez a definição de produto interno, tem-se claramente que $u \cdot v = 0$. Reciprocamente, suponhamos que $v \neq \vec{0}$. Então existe uma coodenada i tal que $v_i \neq 0$. Seja u o vector com todas as coordenadas nulas excepto $u_i = 1$. Então $u \cdot v = u_i v_i = v_i \neq 0$. □

Atenção! Não é verdade, em geral, que $v \cdot v = 0$ apenas para o vector nulo $v = \vec{0}$. Por exemplo, em $\mathbb{F}_3^4$, se $v = (1, 0, 2, 1)$ então $v \cdot v = 1^2 + 0 + 2^2 + 1^2 = 1 + 0 + 1 + 1 = 0$ (as operações são feitas em $\mathbb{F}_3$). Este comportamento "estranho" deve-se ao facto dos corpos finitos $\mathbb{F}_q$ terem característica não nula.

Há outros produtos internos com interesse na Teoria de Códigos (ver o Exercício 3.5 para um exemplo), por isso damos aqui a definição geral.

Definição 3.38. Um *produto interno* em $\mathbb{F}_q^n$ é uma aplicação $\langle \cdot, \cdot \rangle : \mathbb{F}_q^n \times \mathbb{F}_q^n \longrightarrow \mathbb{F}_q$ tal que

- $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
- $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$,
- $\langle u, v \rangle = 0 \quad \forall\, u \in \mathbb{F}_q^n$ sse $v = \vec{0}$ e
- $\langle u, v \rangle = 0 \quad \forall\, v \in \mathbb{F}_q^n$ sse $u = \vec{0}$,

para quaisquer vectores $u, v, w \in \mathbb{F}_q^n$.

Mas, por defeito, usaremos o produto interno euclideano.

Definição 3.39. Seja V um subespaço vectorial de $\mathbb{F}_q^n$. O *complemento ortogonal* de V é o conjunto $V^\perp = \{w \in \mathbb{F}_q^n \,:\, w \cdot v = 0 \quad \forall\, v \in V\}$.

Note que, tal como no caso dos subespaços de $\mathbb{R}^n$, na definição de $V^\perp$ basta verificar a condição de ortogonalidade para os vectores v numa base de V, porque o produto interno é linear.

Teorema 3.40. *Seja V um subespaço de $\mathbb{F}_q^n$. Então o complemeto ortognal $V^\perp$ é também um subespaço de $\mathbb{F}_q^n$ e*

$$\dim V + \dim V^\perp = n \ . \tag{3.5}$$

Dem. Se V é o espaço nulo, então $V^\perp = \mathbb{F}_q^n$ e o teorema é válido. Suponhamos agora que $V \neq \{\vec{0}\}$. Seja $k = \dim V$ e seja $\{v_1, \ldots, v_k\}$ uma base de V. Como o produto interno é linear, se $x, y \in V^\perp$ e $a, b \in \mathbb{F}_q$, então

$$(ax + by) \cdot v_i = a(x \cdot v_i) + b(y \cdot v_i) = 0, \qquad \text{para} \quad i = 1, \ldots k \ ,$$

donde se conclui que $ax + by \in V^\perp$, logo $V^\perp$ é um subespaço vectorial de $\mathbb{F}_q^n$.

Se escrevermos as igualdades $v_1 \cdot x = \cdots = v_k \cdot x = 0$ em notação matricial, ficamos com $V^\perp = \mathcal{N}(A)$, onde A é a matriz $k \times n$ cuja linha i é formada pelas coordenadas do vector v_i:

$$A = \begin{bmatrix} \text{——} & v_1^T & \text{——} \\ & \vdots & \\ \text{——} & v_k^T & \text{——} \end{bmatrix}_{k \times n}$$

(v_1^T designa o transposto[3] de v_1.)

Portanto, a dimensão de $V^\perp$ é o número de variáveis livres no sistemas de equações $Ax = 0$, que é igual à diferença entre o número de colunas de A menos o número de linhas linearmente independentes. Neste caso fica $\dim V^\perp = n - k = n - \dim V$. □

[3] Em notação matricial, identificamos vectores com matrizes colunas. Assim, se v é um vector ou uma coluna, o seu transposto v^T é uma linha.

ATENÇÃO! Contrariamente ao que estamos habituados no caso real, não é verdade que $V \oplus V^\perp = \mathbb{F}_q^n$ para todo o subespaço $V \subseteq \mathbb{F}_q^n$.

Recorde que a *soma* de dois subespaços $V, W \subseteq \mathbb{F}_q^n$ é definida por

$$V + W = \{v + w \in \mathbb{F}_q^n \,:\, v \in V, w \in W\} \tag{3.6}$$

e é um subespaço de $\mathbb{F}_q^n$. Quando se verifica que $V \cap W = \{\vec{0}\}$, dizemos que o espaço soma é a *soma directa* de V e W e escrevemos $V \oplus W$. Assim, $\mathbb{F}_q^n = V \oplus W$ se e só se $\mathbb{F}_q^n = V + W$ e ainda $V \cap W = \{\vec{0}\}$. Recorde ainda que a intersecção $V \cap W$ dos subespaços V e W é sempre um espaço vectorial.

Exemplo 3.41. $\{\vec{0}\}^\perp = \mathbb{F}_q^n$ e, neste caso, tem-se trivialmente a decomposição em soma directa $\{\vec{0}\} \oplus \{\vec{0}\}^\perp = \mathbb{F}_q^n$.

Exemplo 3.42. Seja $V = \langle (1,1,1,1) \rangle \subseteq \mathbb{F}_2^4$. Então $V^\perp = \{x \in \mathbb{F}_2^4 \,:\, \mathrm{w}(x) \text{ é par}\}$, porque $x \in V^\perp$ se e só se $x \cdot (1,1,1,1) = 0$ se e só se $x_1 + x_2 + x_3 + x_4 = 0$ em $\mathbb{F}_2$, i.e., se e só se $\mathrm{w}(x) = x_1 + x_2 + x_3 + x_4 \equiv 0 \pmod 2$.

Da equação $x_1 + x_2 + x_3 + x_4 = 0$, ou $x_1 = x_2 + x_3 + x_4$, que descreve $V^\perp$, também se tira que $\{(1,1,0,0), (1,0,1,0), (1,0,0,1)\}$ é uma base de $V^\perp$ — o primeiro vector é obtido subtituindo $x_2 = 1$ e $x_3 = x_4 = 0$, o segundo substituindo $x_3 = 1$ e $x_2 = x_4 = 0$, etc.

Quanto às dimensões temos que $\dim V = 1$, pois V é gerado por um único vector, nomeadamente $(1,1,1,1)$, e $\dim V^\perp = 3$, aplicando o Teorema 3.40 ou directamente da definição de dimensão, uma vez que já temos uma basa para $V^\perp$.

Mas como $(1,1,1,1) \in V^\perp$, porque o seu peso é 4, conclui-se que $\boxed{V \subset V^\perp}$.

Exemplo 3.43. Seja V o subespaço de $\mathbb{F}_2^4$ gerado pelos vectores $u = (1,0,1,0)$ e $v = (0,1,0,1)$. Como u e v são linearmente independentes (já visto num exemplo anterior), conclui-se que V tem dimensão 2. Calculemos o complemento ortogonal $V^\perp$. Usando a definição

$$\begin{aligned} V^\perp &= \{x \in \mathbb{F}_2^4 \,:\, x \cdot u = 0,\, x \cdot v = 0\} \\ &= \{(x_1, x_2, x_3, x_4) \in \mathbb{F}_2^4 \,:\, x_1 = x_3,\, x_2 = x_4\}\ , \end{aligned}$$

donde se conclui que $V^\perp = \langle (1,0,1,0), (0,1,0,1) \rangle$, ou seja, $\boxed{V = V^\perp}$.

Teorema 3.44. *Seja V um subespaço vectorial de $\mathbb{F}_q^n$. Então $(V^\perp)^\perp = V$.*

Dem. 1º PASSO: Mostrar que $V \subseteq (V^\perp)^\perp$.

Seja $v \in V$. Então, para todo o $w \in V^\perp$ verifica-se $v \cdot w = 0$ (pela definição de $V^\perp$), mas isto também quer dizer que $v \in (V^\perp)^\perp$.

2º PASSO: Comparar dimensões.

Aplicando o Teorema 3.33 duas vezes, fica

$$\dim (V^\perp)^\perp = n - \dim V^\perp = n - (n - \dim V) = \dim V\ .$$

Logo, os espaços vectoriais V e $(V^\perp)^\perp$ têm o mesmo número de elementos (justifique), e como um é subconjunto do outro (pelo 1º passo), conclui-se que são iguais. □

Exercícios

3.1. Resolver a Ficha 3.

3.2. (a) Verifique que as tabelas dos Exemplos 3.19 e 3.20 estão correctas.
(b) Determine um isomorfismo (de anéis) entre $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ e $\mathbb{F}_2[t]/\langle t^2+1\rangle$.

3.3. Seja V um subespaço vectorial de $\mathbb{F}_q^n$ de dimensão k, com $1 \leq k \leq n$.
(a) Quantos vectores contém V?
(b) Quantas bases diferentes tem V?

3.4. (a) Mostre que $\mathbb{F}_{q^m}$ é um espaço vectorial sobre $\mathbb{F}_q$, com a soma e o produto por um escalar definidos à custa das operações em $\mathbb{F}_{q^m}$.
(b) Seja $f(x) \in \mathbb{F}_q[x]$ um polinómio de grau m, irredutível em $\mathbb{F}_q[x]$, e seja $\alpha \in \mathbb{F}_{q^m}$ uma raiz de $f(x)$. Mostre que $\{1, \alpha, \alpha^2, \ldots, \alpha^{m-1}\}$ é uma base de $\mathbb{F}_{q^m}$ sobre $\mathbb{F}_q$.

3.5. Considere a aplicação $< \cdot, \cdot >_H : \mathbb{F}_{q^2}^n \times \mathbb{F}_{q^2}^n \longrightarrow \mathbb{F}_{q^2}$ definida por

$$< u, v >_H = \sum_{i=1}^{n} u_i v_i^q \ ,$$

onde $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in \mathbb{F}_{q^2}^n$. Mostre que $< \cdot, \cdot >_H$ é um produto interno em $\mathbb{F}_{q^2}^n$. Nota: $< \cdot, \cdot >_H$ diz-se o *produto interno hermítico*. O *dual hermítico* do código linear C é definido por

$$C^{\perp_H} = \{v \in \mathbb{F}_{q^2}^n \ : < v, c >_H = 0 \quad \forall\, c \in C\} \ .$$

3.6. Recorde que $\mathbb{F}_4 = \mathbb{F}_2[x]/\langle x^2+x+1\rangle = \{0, 1, \alpha, \alpha^2\}$, onde α é uma raiz de $x^2+x+1 \in \mathbb{F}_2[x]$. Mostre que os seguintes códigos lineares sobre $\mathbb{F}_4$ são auto-duais em relação ao produto interno hermítico definido no problema anterior:
(a) $C_1 = \langle (1,1)\rangle \subset \mathbb{F}_4^2$,
(b) $C_2 = \langle (1,0,0,1,\alpha,\alpha), (0,1,0,\alpha,1,\alpha), (0,0,1,\alpha,\alpha,1)\rangle \subset \mathbb{F}_4^6$.
Serão estes códigos auto-duais em relação ao produto interno euclideano?

3.7. Sejam V e W subespaços de $\mathbb{F}_q^n$. Mostre que a soma $V+W$, definida em (3.6), e a intersecção $V \cap W$ são espaços vectoriais. Mostre ainda que a soma $V+W$ é o espaço vectorial gerado por V e W.

3.8. (a) Justifique que os polinómios t^3+t+1 e t^3+t^2+1 são irredutíveis em $\mathbb{F}_2[t]$.
(b) Justifique que os quocientes $A = \mathbb{F}_2[t]/\langle t^3+t+1\rangle$ e $B = \mathbb{F}_2[t]/\langle t^3+t^2+1\rangle$ são ambos isomorfos ao corpo $\mathbb{F}_8$, e determine um isomorfismo $\phi : A \longrightarrow B$.
[Sugestão: Seja $\alpha \in A$ uma raíz de $1+t+t^3$ e $\beta \in B$ uma raíz de $1+t^2+t^3$. Encontre uma relação entre α e β ou, mais precisamente, determine uma raíz de $1+t^2+t^3$ em A.]
(c) Para a descrição A de $\mathbb{F}_8$, determine um elemento primitivo e uma base de A como espaço vectorial sobre $\mathbb{F}_2$.

BIBLIOGRAFIA

[1] R.L. Fernandes, M. Ricou, *Introdução à Álgebra*, IST Press.

[2] R. Hill, *A Fisrt Course in Coding Thery*, Oxford Applied Mathematics and Computing Science Series, 1996, Oxford University Press.

[3] J.H. van Lint, R.M. Wilson, *A course in Combinatorics*, 2nd edition, Cambridge University Press, 2001.

[4] S. Roman, *Coding and Information Theory*, Graduate Texts in Mathematics, 134, Springer-Verlag, 1992.