Introdução a Teoria dos Códigos

Marcelo Muniz Alves - UFPR

VIII Semana da Matemática UNIFAL

6, 8 e 9 de maio de 2025

Os problemas motivadores da teoria

- Informação original: sequência

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ a_{j+4}\ldots$$

de elementos de um “alfabeto” A.

Os problemas motivadores da teoria

- Informação original: sequência

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ a_{j+4}\ldots$$

de elementos de um “alfabeto” A.

- Erros:

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ a_{j+4}\ldots$$

$$\Downarrow$$

$$\ldots a_{j-1}\ b_j\ a_{j+1}\ a_{j+2} b_{j+3}\ a_{j+4}\ldots$$

Os problemas motivadores da teoria

- Informação original: sequência

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ a_{j+4}\ldots$$

de elementos de um “alfabeto” A.

- Erros:

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ a_{j+4}\ldots$$

$$\Downarrow$$

$$\ldots a_{j-1}\ X\ a_{j+1}\ a_{j+2} X\ a_{j+4}\ldots$$

Os problemas motivadores da teoria

Problemas: sem ter acesso à informação original,

Os problemas motivadores da teoria

Problemas: sem ter acesso à informação original,

- como detectar a ocorrência de erros ?

Os problemas motivadores da teoria

Problemas: sem ter acesso à informação original,

- como detectar a ocorrência de erros ?
- havendo erros, como corrigir e obter a informação original?

Alguns precursores

- Richard Hamming, “Error Detecting and Error Correcting Codes” ((1947) 1950)
- Marcel Golay, “Notes on Digital Coding” (1949)
- Claude Shannon, “A Mathematical Theory of Communication” (1948)
- Irving Reed, Gustave Solomon “Polynomial Codes over Certain Finite Fields” (1960)
- Raj Chandra Bose, D. K. Ray-Chaudhuri “On A Class of Error Correcting Binary Group Codes”(1960);
  Alexis Hocquenghem (1959), “Codes correcteurs d’erreurs”.

Codificação por blocos

A “alfabeto”.
Mensagem é dividida em sequências de comprimento k.

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3}\ldots$$

Codificação por blocos

A “alfabeto”.
Mensagem é dividida em sequências de comprimento k.

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3} \ldots$$

$$a_1 a_2 \cdots a_{k-1} a_k$$
$$b_1 b_2 \cdots b_{k-1} b_k$$
$$c_1 c_2 \cdots c_{k-1} c_k$$

Codificação por blocos

A alfabeto.
Codificação: acrescentamos redundâncias para “afastar” as sequências entre si.

$$\ldots a_{j-1}\ a_j\ a_{j+1}\ a_{j+2}\ a_{j+3} \ldots$$

$$\begin{aligned}
a_1 a_2 \cdots a_{k-1} a_k &\longmapsto a_1 a_2 \cdots a_{k-1} a_k a_{k+1} a_{k+2} \cdots a_n \\
b_1 b_2 \cdots b_{k-1} b_k &\longmapsto b_1 b_2 \cdots b_{k-1} b_k a_{k+1} a_{k+2} \cdots a_n \\
c_1 c_2 \cdots c_{k-1} c_k &\longmapsto c_1 c_2 \cdots c_{k-1} c_k a_{k+1} a_{k+2} \cdots a_n
\end{aligned}$$

Codificação por blocos

Codificação: $n > k$, função injetora

$$A^k \xrightarrow{\quad f \quad} A^n$$

Codificação por blocos

Codificação: $n > k$, função injetora

$$A^k \xrightarrow{\quad f \quad} A^n$$

Detecção de erros:

$$\mathbf{v} \in A^n \text{ está em } f(A^k)?$$

Codificação por blocos

Codificação: $n > k$, função injetora

$$A^k \xrightarrow{\quad f \quad} A^n$$

Detecção de erros:

$$\mathbf{v} \in A^n \text{ está em } f(A^k)?$$

Correção de erros:

$$g : A^n \dashrightarrow f(A^k), \quad g \circ f = f.$$

Codificação por blocos

Codificação: $n > k$, função injetora

$$A^k \xrightarrow{\quad f \quad} A^n$$

Detecção de erros:

$$\mathbf{v} \in A^n \text{ está em } f(A^k)?$$

Correção de erros:

$$g : A^n \dashrightarrow f(A^k), \quad g \circ f = f.$$

Correção e decodificação:

$$A^n \dashrightarrow f(A^k) \xrightarrow{f^{-1}} A^k$$

Detecção de erros: ISBN-10

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$a_j \in \{0, 1, 2, \ldots, 9\}$.

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

  $a_j \in \{0, 1, 2, \ldots, 9\}$.
- Cálculo de a_{10}:

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

  $a_j \in \{0, 1, 2, \ldots, 9\}$.
- Cálculo de a_{10}:

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$a_j \in \{0, 1, 2, \ldots, 9\}$.

- Cálculo de a_{10}:

$$D = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11 - j)a_j$$

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$a_j \in \{0, 1, 2, \ldots, 9\}$.

- Cálculo de a_{10}:

$$D = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11 - j)a_j$$

r = resto da divisão de D por 11

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$a_j \in \{0, 1, 2, \ldots, 9\}$.

- Cálculo de a_{10}:

$$D = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11 - j)a_j$$

r = resto da divisão de D por 11

$$a_{10} = 11 - r \quad \text{se } r > 1;$$

Codificação para o ISBN-10

- **I**nternational **S**tandard **B**ook **N**umber (ISBN)

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$a_j \in \{0, 1, 2, \ldots, 9\}$.

- Cálculo de a_{10}:

$$D = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11 - j)a_j$$

r = resto da divisão de D por 11

$$a_{10} = 11 - r \quad \text{se } r > 1;$$

$$r = 0 \implies a_{10} = 0,$$
$$r = 1 \implies a_{10} = X.$$

Codificação para o ISBN-10

- Ideia: o número a_{10} é escolhido de modo que

$$\sum_{j=1}^{10}(11-j)a_j$$

seja múltiplo de 11 (interpretando “X” como 10).

Codificação para o ISBN-10

- Ideia: o número a_{10} é escolhido de modo que

$$\sum_{j=1}^{10}(11-j)a_j$$

seja múltiplo de 11 (interpretando "X" como 10).

-

$$\sum_{j=1}^{10}(11-j)a_j \quad = \quad \sum_{j=1}^{9}(11-j)a_j + a_{10}$$

Codificação para o ISBN-10

- Ideia: o número a_{10} é escolhido de modo que

$$\sum_{j=1}^{10}(11-j)a_j$$

seja múltiplo de 11 (interpretando "X" como 10).

-

$$\sum_{j=1}^{10}(11-j)a_j \quad = \quad \sum_{j=1}^{9}(11-j)a_j + a_{10}$$

Codificação para o ISBN-10

- Ideia: o número a_{10} é escolhido de modo que

$$\sum_{j=1}^{10}(11-j)a_j$$

seja múltiplo de 11 (interpretando "X" como 10).

-

$$\begin{aligned}\sum_{j=1}^{10}(11-j)a_j &= \sum_{j=1}^{9}(11-j)a_j + a_{10} \\ &= 11q + r + a_{10}\end{aligned}$$

Codificação para o ISBN-10

- Ideia: o número a_{10} é escolhido de modo que

$$\sum_{j=1}^{10}(11-j)a_j$$

seja múltiplo de 11 (interpretando "X" como 10).

-

$$\begin{aligned}\sum_{j=1}^{10}(11-j)a_j &= \sum_{j=1}^{9}(11-j)a_j + a_{10} \\ &= 11q + r + a_{10} \\ &= 11q + r + (11-r) = 11(q+1).\end{aligned}$$

Codificação para o ISBN-10

- A “mensagem” do ISBN-10 é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

Codificação para o ISBN-10

- A “mensagem” do ISBN-10 é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

Codificação para o ISBN-10

- A “mensagem” do ISBN-10 é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

A **mensagem codificada** é o ISBN completo

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}.$$

O dígito a_{10} é o **dígito verificador**.

- Alfabeto: $A = \{0, 1, 2, \ldots, 9\} \cup \{X\}$

Codificação para o ISBN-10

- A “mensagem” do ISBN-10 é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

A **mensagem codificada** é o ISBN completo

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}.$$

O dígito a_{10} é o **dígito verificador**.

- Alfabeto: $A = \{0, 1, 2, \ldots, 9\} \cup \{X\}$
- $k = 9$; $n = 10$.

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

$$D_1 = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11-j)a_j$$

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

$$D_1 = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11-j)a_j$$

r = resto da divisão de D_1 por 11

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

$$D_1 = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11-j)a_j$$

$$r = \text{resto da divisão de } D_1 \text{ por } 11$$

$$r > 1 \implies a_{10} = 11 - r$$

Codificação para o CPF

- **C**adastro de **P**essoa **F**ísica

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}$$

- Cálculo de a_{10}:

$$D_1 = 10a_1 + 9a_2 + \cdots + 2a_9 = \sum_{j=1}^{9}(11-j)a_j$$

$$r = \text{resto da divisão de } D_1 \text{ por } 11$$

$$r > 1 \implies a_{10} = 11 - r$$

$$r = 0 \text{ ou } 1 \implies a_{10} = 0.$$

- Para a_{11} faz-se o mesmo, mas começando de a_2, não de a_1 (e portanto usa-se a_{10}):
  $D_2 = 10a_2 + 9a_3 + \cdots + 2a_{10}$, etc.

Codificação para o CPF

- A “mensagem” do CPF é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

Codificação para o CPF

- A “mensagem” do CPF é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

Codificação para o CPF

- A “mensagem” do CPF é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

A **mensagem codificada** é o CPF completo

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}.$$

Os dígitos a_{10} e a_{11} são os **dígitos verificadores**.

Codificação para o CPF

- A “mensagem” do CPF é a sequência

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9.$$

A **mensagem codificada** é o CPF completo

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}\ a_{11}.$$

Os dígitos a_{10} e a_{11} são os **dígitos verificadores**.

- Alfabeto: $\{0, 1, \ldots, 9\}$.
- $k = 9$, $n = 11$.

Detecção de erros

Em ambos os casos, a codificação

- Detecta 1 erro em qualquer posição;

Detecção de erros

Em ambos os casos, a codificação

- Detecta 1 erro em qualquer posição;
- Detecta permuta de dois dígitos.

Detecção de erros

Em ambos os casos, a codificação

- Detecta 1 erro em qualquer posição;
- Detecta permuta de dois dígitos.

Detecção de erros

Em ambos os casos, a codificação

- Detecta 1 erro em qualquer posição;
- Detecta permuta de dois dígitos.

Não é possível corrigir erros.

Detecção de erros

Em ambos os casos, a codificação

- Detecta 1 erro em qualquer posição;
- Detecta permuta de dois dígitos.

Não é possível corrigir erros.

Questão: por que divisão por 11 e não por 10?

Congruência módulo m

Congruências

Dado m natural positivo e a, b inteiros, dizemos que a e b são **congruentes** módulo m se

$$a - b \text{ é múltiplo de } m.$$

Congruências

Dado m natural positivo e a, b inteiros, dizemos que a e b são **congruentes** módulo m se

$$a - b \text{ é múltiplo de } m.$$

De modo equivalente,

$$a \text{ e } b \text{ deixam o mesmo resto na divisão por } m.$$

Congruências

Dado m natural positivo e a, b inteiros, dizemos que a e b são **congruentes** módulo m se

$$a - b \text{ é múltiplo de } m.$$

De modo equivalente,

$$a \text{ e } b \text{ deixam o mesmo resto na divisão por } m.$$

Notação:

$$a \equiv b \mod m.$$

Congruências

Resto na divisão e congruência:

Congruências

Resto na divisão e congruência:

$$\begin{aligned} a &= qm + r \\ b &= q'm + r' \end{aligned}$$

com $r \geq r'$. Então

Congruências

Resto na divisão e congruência:

$$\begin{aligned} a &= qm + r \\ b &= q'm + r' \end{aligned}$$

com $r \geq r'$. Então

$$a - b = (q - q')m + (r - r')$$

com $m \geq r - r' \geq 0$.

Congruências

Resto na divisão e congruência:

$$\begin{aligned} a &= qm + r \\ b &= q'm + r' \end{aligned}$$

com $r \geq r'$. Então

$$a - b = (q - q')m + (r - r')$$

com $m \geq r - r' \geq 0$.

Pela unicidade do quociente e do resto,

$$a \equiv b \mod m \iff m \mid (a - b) \iff r = r'.$$

Congruências

Prova-se que se

$$a \equiv a' \mod m, \quad b \equiv b' \mod m$$

então

Congruências

Prova-se que se

$$a \equiv a' \mod m, \quad b \equiv b' \mod m$$

então

- $a + b \equiv a' + b' \mod m$

Congruências

Prova-se que se

$$a \equiv a' \mod m, \quad b \equiv b' \mod m$$

então

- $a + b \equiv a' + b' \mod m$
- $a \cdot b \equiv a' \cdot b' \mod m$

Congruências

Prova-se que se

$$a \equiv a' \mod m, \quad b \equiv b' \mod m$$

então

- $a + b \equiv a' + b' \mod m$
- $a \cdot b \equiv a' \cdot b' \mod m$

Congruências

Prova-se que se

$$a \equiv a' \mod m, \quad b \equiv b' \mod m$$

então

- $a + b \equiv a' + b' \mod m$
- $a \cdot b \equiv a' \cdot b' \mod m$

Em particular, se $r \in \mathbb{Z}$ então

- $a \equiv a' \mod m \implies ra \equiv ra' \mod m.$

Congruências

Reduzindo módulo m: podemos "zerar" múltiplos de m em qualquer cálculo módulo m, já que $km \equiv 0 \mod m$.

Congruências

Reduzindo módulo m: podemos "zerar" múltiplos de m em qualquer cálculo módulo m, já que $km \equiv 0 \mod m$.

Exemplo: qual o resto de $234 \cdot 457$ na divisão por 4?

Congruências

Reduzindo módulo m: podemos "zerar" múltiplos de m em qualquer cálculo módulo m, já que $km \equiv 0 \mod m$.

Exemplo: qual o resto de $234 \cdot 457$ na divisão por 4?

$$234 \cdot 456 = (200 + 32 + 2) \cdot (400 + 40 + 16 + 1) \equiv$$

$$\equiv (0 + 0 + 2) \cdot (0 + 0 + 0 + 1) \equiv 2 \mod 4.$$

Congruências

Reduzindo módulo m: podemos "zerar" múltiplos de m em qualquer cálculo módulo m, já que $km \equiv 0 \mod m$.

Exemplo: qual o resto de $234 \cdot 457$ na divisão por 4?

$$234 \cdot 456 = (200 + 32 + 2) \cdot (400 + 40 + 16 + 1) \equiv$$

$$\equiv (0 + 0 + 2) \cdot (0 + 0 + 0 + 1) \equiv 2 \mod 4.$$

Portanto o resto é 2.

Congruências

O **Teorema de Bezout** diz que se $mdc(a, b) = d$ então existem $r, s \in \mathbb{Z}$ tais que

$$ar + bs = d.$$

Congruências

Seja p primo e a inteiro tal que $p \nmid a$. Então existe r tal que

$$ar \equiv 1 \mod p.$$

Congruências

Seja p primo e a inteiro tal que $p \nmid a$. Então existe r tal que

$$ar \equiv 1 \mod p.$$

- p é primo, $p \nmid a \implies mdc(a, p) = 1$.

Congruências

Seja p primo e a inteiro tal que $p \nmid a$. Então existe r tal que

$$ar \equiv 1 \mod p.$$

- p é primo, $p \nmid a \implies mdc(a,p) = 1$.
- Bezout: existem r, s tais que

$$ar + ps = 1.$$

Congruências

Seja p primo e a inteiro tal que $p \nmid a$. Então existe r tal que

$$ar \equiv 1 \mod p.$$

- p é primo, $p \nmid a \implies mdc(a,p) = 1$.
- Bezout: existem r, s tais que

$$ar + ps = 1.$$

- $ar - 1 = p(-s) \implies ar \equiv 1 \mod p$.

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9}(11 - j)a_j$

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9} (11 - j)a_j$
- $a_{10} = 11 - r \equiv 11 - D \equiv -D \mod 11$

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9} (11 - j)a_j$
- $a_{10} = 11 - r \equiv 11 - D \equiv -D \mod 11$

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9} (11 - j)a_j$
- $a_{10} = 11 - r \equiv 11 - D \equiv -D \mod 11$
  $\Longrightarrow \; a_{10} + D \equiv 0 \mod 11$

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9}(11-j)a_j$
- $a_{10} = 11 - r \equiv 11 - D \equiv -D \mod 11$

  $\implies a_{10} + D \equiv 0 \mod 11$

  $$\implies \sum_{j=1}^{10}(11-j)a_j = \sum_{j=1}^{9}(11-j)a_j + a_{10} \equiv 0 \mod 11$$

Detecção de 1 erro no ISBN-10

- $D = \sum_{j=1}^{9}(11-j)a_j$
- $a_{10} = 11 - r \equiv 11 - D \equiv -D \mod 11$

  $\implies a_{10} + D \equiv 0 \mod 11$

  $$\implies \sum_{j=1}^{10}(11-j)a_j = \sum_{j=1}^{9}(11-j)a_j + a_{10} \equiv 0 \mod 11$$

a_{10} foi escolhido de modo a obtermos

$$\sum_{j=1}^{10}(11-j)a_j \equiv 0 \mod 11.$$

Detecção de 1 erro no ISBN-10

$$a_1\ a_2\ a_3.\ a_4\ a_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

$$\Downarrow$$

$$a_1\ a_2\ a_3.\ a_4\ b_5\ a_6.\ a_7\ a_8\ a_9 - a_{10}$$

Detecção de 1 erro no ISBN-10

- Suponha que a_i foi trocado por b_i e que $b_i > a_i$. Vamos denotar a nova sequência por

$$a'_1\ a'_2\ a'_3.\ a'_4\ a'_5\ a'_6.\ a'_7\ a'_8\ a'_9 - a'_{10}$$

Detecção de 1 erro no ISBN-10

- Suponha que a_i foi trocado por b_i e que $b_i > a_i$. Vamos denotar a nova sequência por

$$a'_1\ a'_2\ a'_3.\ a'_4\ a'_5\ a'_6.\ a'_7\ a'_8\ a'_9 - a'_{10}$$

- Vamos mostrar que o erro é detectado, provando que

$$\sum_{j=1}^{10}(11-j)a'_j \not\equiv 0 \mod 11,$$

o que diz que a sequência acima não corresponde a um ISBN.

Detecção de 1 erro no ISBN-10

$$\sum_{j=1}^{10}(11-j)a'_j =$$

Detecção de 1 erro no ISBN-10

$$\sum_{j=1}^{10}(11-j)a'_j =$$

$$= \sum_{j=1}^{10}(11-j)a'_j + (11-i)a_i - (11-i)a_i$$

Detecção de 1 erro no ISBN-10

$$\begin{aligned} & \sum_{j=1}^{10}(11-j)a'_j = \\ = & \sum_{j=1}^{10}(11-j)a'_j + (11-i)a_i - (11-i)a_i \\ = & \sum_{j=1}^{10}(11-j)a_j + (11-i)(b_i - a_i) \end{aligned}$$

Detecção de 1 erro no ISBN-10

$$
\begin{aligned}
& \sum_{j=1}^{10}(11-j)a'_j = \\
= & \sum_{j=1}^{10}(11-j)a'_j + (11-i)a_i - (11-i)a_i \\
= & \sum_{j=1}^{10}(11-j)a_j + (11-i)(b_i - a_i) \\
\equiv & \ (11-i)(b_i - a_i) \mod 11.
\end{aligned}
$$

Detecção de 1 erro no ISBN-10

- $\text{mdc}(11 - i, 11) = 1$

Detecção de 1 erro no ISBN-10

- $\text{mdc}(11 - i, 11) = 1$
- existe r tal que $r(11 - i) \equiv 1 \mod 11$

Detecção de 1 erro no ISBN-10

- $\text{mdc}(11 - i, 11) = 1$
- existe r tal que $r(11 - i) \equiv 1 \mod 11$
- se $(11 - i)(b_i - a_i) \equiv 0 \mod 11$ então, multiplicando por r,

$$b_i - a_i \equiv 0 \mod 11$$

ou seja, 11 divide $b_i - a_i$, o que é impossível.

Detecção de 1 erro no ISBN-10

- $\text{mdc}(11 - i, 11) = 1$
- existe r tal que $r(11 - i) \equiv 1 \mod 11$
- se $(11 - i)(b_i - a_i) \equiv 0 \mod 11$ então, multiplicando por r,

$$b_i - a_i \equiv 0 \mod 11$$

ou seja, 11 divide $b_i - a_i$, o que é impossível.

- Portanto

$$\sum_{j=1}^{10}(11 - j)a'_j \equiv (11 - i)(a_j - a'_j) \not\equiv 0 \mod 11.$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados.

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\sum_{l=1}^{10}(11-l)a'_l \quad =$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\sum_{l=1}^{10}(11-l)a'_l = \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i)$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\begin{aligned}\sum_{l=1}^{10}(11-l)a'_l &= \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i) \\ &\equiv ((11-j)-(11-i))(a_j-a_i) \mod 11\end{aligned}$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\begin{aligned}\sum_{l=1}^{10}(11-l)a'_l &= \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i)\\ &\equiv ((11-j)-(11-i))(a_j-a_i) \mod 11\\ &\equiv (i-j)(a_j-a_i) \mod 11.\end{aligned}$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\begin{aligned}\sum_{l=1}^{10}(11-l)a'_l &= \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i)\\ &\equiv ((11-j)-(11-i))(a_j-a_i) \mod 11\\ &\equiv (i-j)(a_j-a_i) \mod 11.\end{aligned}$$

11 não divide $i-j$, logo

$$mdc(11, i-j) = 1.$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\begin{aligned}\sum_{l=1}^{10}(11-l)a'_l &= \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i)\\ &\equiv ((11-j)-(11-i))(a_j-a_i) \mod 11\\ &\equiv (i-j)(a_j-a_i) \mod 11.\end{aligned}$$

11 não divide $i-j$, logo

$$mdc(11, i-j) = 1.$$

Assim, se

$$(i-j)(a_j-a_i) \equiv 0 \mod 11$$

Detecção de 2 erros no ISBN-10: permutação

Suponha que a_j e a_i foram permutados. Para a nova sequência obtida, temos

$$\begin{aligned}\sum_{l=1}^{10}(11-l)a'_l &= \sum_{l=1}^{10}(11-l)a_l + (11-j)(a_i-a_j) + (11-i)(a_j-a_i)\\ &\equiv ((11-j)-(11-i))(a_j-a_i) \mod 11\\ &\equiv (i-j)(a_j-a_i) \mod 11.\end{aligned}$$

11 não divide $i-j$, logo

$$mdc(11, i-j) = 1.$$

Assim, se

$$(i-j)(a_j-a_i) \equiv 0 \mod 11$$

obtemos

$$a_j - a_i \equiv 0 \mod 11,$$

contradição.

Detecção de 2 erros no ISBN-10

O ISBN-10 detecta 90% dos casos em que ocorrem dois erros [Egghe, Rousseau 2000]

Detecção de 2 erros no ISBN-10

O ISBN-10 detecta 90% dos casos em que ocorrem dois erros [Egghe, Rousseau 2000]

$$010.000.000 - 2$$

$$\Downarrow$$

$$000.000.030 - 2$$

Detecção de 2 erros no ISBN-10

O ISBN-10 detecta 90% dos casos em que ocorrem dois erros [Egghe, Rousseau 2000]

$$010.000.000 - 2$$

$$\Downarrow$$

$$000.000.030 - 2$$

Neste caso, dois erros levaram a um novo ISBN válido (e com mesmo dígito verificador).

Correção de erros no ISBN-10

Veremos mais adiante, que se um código corrige 1 erro então ele necessariamente detecta pelo menos 2 erros (sempre).

Portanto, não há correção de erros para o ISBN-10.

Correção de erros: código de repetição

Código de Repetição

Alfabeto:

$$A = \{0, 1\}$$

Código:

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

Codificação:

$$A \to A^3$$

$$0 \mapsto 000$$
$$1 \mapsto 111$$

Código de Repetição

$$01001\ldots \mapsto 000|111|000|000|111|\ldots$$

Código de Repetição

- Detecção de erros: as mensagens codificadas são 000 e 111.
  Se a mensagem recebida tiver tanto 0 quanto 1, sabemos que há erros.
  Detectam-se até 2 erros.

Código de Repetição

- Detecção de erros: as mensagens codificadas são 000 e 111. Se a mensagem recebida tiver tanto 0 quanto 1, sabemos que há erros.
  Detectam-se até 2 erros.
- Correção de erros: substitui-se pelo "mais próximo" (= o que envolve o menor número de erros em coordenadas).

$$000, 100, 010, 001 \quad \mapsto \quad 000$$

$$111, 110, 101, 011 \quad \mapsto \quad 111$$

Código de Repetição

- Exemplo...

Enviado	000|111|000|000|111
	↓
Recebido	0**1**0|**0**11|000|0**1**0|111
	↓
Corrigido	0**0**0|**1**11|000|0**0**0|111

Código de Repetição

Este código

- tem $k = 1$, $n = 3$;
- detecta até 2 erros;
- corrige 1 erro.

Código de Repetição, $n = 4$

Código:

$$C_4 = \{0000, 1111\}$$

Codificação:

$$A \to A^4$$

$$0 \mapsto 0000$$
$$1 \mapsto 1111$$

Código de Repetição, $n = 4$

- Correção:

$$0000 \qquad 1111$$
$$1000 \qquad 0111$$
$$0100 \mapsto 0000 \qquad 1011 \mapsto 1111$$
$$0010 \qquad 1101$$
$$0001 \qquad 1111$$

Código de Repetição, $n = 4$

- Correção:

0000	1111
1000	0111
0100 $\mapsto$ 0000	1011 $\mapsto$ 1111
0010	1101
0001	1111

- 0011 $\mapsto$?

Código de Repetição, $n = 4$

- $k = 1$, $n = 4$;
- detecta até 3 erros;
- corrige 1 erro.

Código de Repetição, $n = 5$

- $k = 1$, $n = 5$;
- detecta até 4 erros;
- corrige 2 erros.

Código de Repetição, $n = 2m + 1$

- $k = 1$, $n = 2m + 1$;
- detecta até $2m$ erros;
- corrige m erros.

Códigos Corretores de Erros: fundamentos

Formalizando & Generalizando

- **Alfabeto**: corpo finito $\mathbb{F}$.

Formalizando & Generalizando

- **Alfabeto**: corpo finito $\mathbb{F}$.
- **Código**: subespaço vetorial C de $\mathbb{F}^n$.

Formalizando & Generalizando

- **Alfabeto**: corpo finito $\mathbb{F}$.
- **Código**: subespaço vetorial C de $\mathbb{F}^n$.
- Correção pelo “mais próximo”: **métrica de Hamming** em $\mathbb{F}^n$.

Corpos Finitos

- Um **anel** (comutativo) é um conjunto com soma e produto que tem propriedades análogas às de $\mathbb{Z}$ ou $\mathbb{R}[x]$

Corpos Finitos

- Um **anel** (comutativo) é um conjunto com soma e produto que tem propriedades análogas às de $\mathbb{Z}$ ou $\mathbb{R}[x]$
- Um **corpo** é um conjunto com soma e produto com as "melhores propriedades possíveis"; tem propriedades básicas análogas às de $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$.

Corpos Finitos

- Um **anel** (comutativo) é um conjunto com soma e produto que tem propriedades análogas às de $\mathbb{Z}$ ou $\mathbb{R}[x]$
- Um **corpo** é um conjunto com soma e produto com as "melhores propriedades possíveis"; tem propriedades básicas análogas às de $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$.
  - Em especial, todo elemento não-nulo em um corpo tem inverso.

Corpos Finitos

- Dado $m \in \mathbb{N}, m \geq 2$, o anel dos inteiros módulo m é

$$\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$$

com operações

$$a +_m b = \text{ resto de } a + b \text{ na divisão por } m$$

$$a \cdot_m b = \text{ resto de } ab \text{ na divisão por } m$$

Corpos Finitos

- Equivalente: dados $a, b, r, s \in \mathbb{Z}_m$,

$$a +_m b = r \iff a + b \equiv r \mod m$$

$$a \cdot_m b = s \iff ab \equiv s \mod m$$

Corpos Finitos

- Equivalente: dados $a, b, r, s \in \mathbb{Z}_m$,

$$a +_m b = r \iff a + b \equiv r \mod m$$

$$a \cdot_m b = s \iff ab \equiv s \mod m$$

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Sejam p primo, $0 \neq a \in \mathbb{Z}_p$.

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Sejam p primo, $0 \neq a \in \mathbb{Z}_p$.

$p \nmid a \implies \text{mdc}(a, p) = 1$

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Sejam p primo, $0 \neq a \in \mathbb{Z}_p$.

$p \nmid a \implies \text{mdc}(a, p) = 1$

(Bezout) Existem $r, s \in \mathbb{Z}$ tais que $ar + ps = 1$

Corpos Finitos

- $\mathbb{Z}_m$ é corpo $\iff$ m é primo.

Sejam p primo, $0 \neq a \in \mathbb{Z}_p$.

$p \nmid a \implies \text{mdc}(a, p) = 1$

(Bezout) Existem $r, s \in \mathbb{Z}$ tais que $ar + ps = 1$

$$ar \equiv 1 \mod p \implies a \cdot_p r = 1.$$

Corpos Finitos

- tabelas de $\mathbb{Z}_2$

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Corpos Finitos

- tabelas de $\mathbb{Z}_2$

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

- tabelas de $\mathbb{Z}_3$

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

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Corpos Finitos

- multiplicação em $\mathbb{Z}_4$ - **não** é corpo!

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	**0**	2
3	0	3	2	1

Corpos Finitos

Inversos em $\mathbb{Z}_{11}$

- $\text{mdc}(2, 11) = 1$

$$2 \cdot 6 - 1 \cdot 11 = 1 \implies 2 \cdot_{(11)} 6 = 1 \text{ em } \mathbb{Z}_{11}$$

Corpos Finitos

Inversos em $\mathbb{Z}_{11}$

- $\text{mdc}(2, 11) = 1$

$$2 \cdot 6 - 1 \cdot 11 = 1 \implies 2 \cdot_{(11)} 6 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(3, 11) = 1$

$$3 \cdot 4 - 11 = 1 \implies 3 \cdot_{(11)} 4 = 1 \text{ em } \mathbb{Z}_{11}$$

Corpos Finitos

Inversos em $\mathbb{Z}_{11}$

- $\text{mdc}(2, 11) = 1$

$$2 \cdot 6 - 1 \cdot 11 = 1 \implies 2 \cdot_{(11)} 6 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(3, 11) = 1$

$$3 \cdot 4 - 11 = 1 \implies 3 \cdot_{(11)} 4 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(5, 11) = 1$

$$5 \cdot 9 - 4 \cdot 11 = 1 \implies 5 \cdot_{(11)} 9 = 1 \text{ em } \mathbb{Z}_{11}$$

Inversos em $\mathbb{Z}_{11}$

- $\text{mdc}(2, 11) = 1$

$$2 \cdot 6 - 1 \cdot 11 = 1 \implies 2 \cdot_{(11)} 6 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(3, 11) = 1$

$$3 \cdot 4 - 11 = 1 \implies 3 \cdot_{(11)} 4 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(5, 11) = 1$

$$5 \cdot 9 - 4 \cdot 11 = 1 \implies 5 \cdot_{(11)} 9 = 1 \text{ em } \mathbb{Z}_{11}$$

- $\text{mdc}(10, 11) = 1$

$$10 \cdot 10 - 9 \cdot 11 = 1 \implies 10 \cdot_{(11)} 10 = 1 \text{ em } \mathbb{Z}_{11}$$

Corpos Finitos

- Teorema de Galois : para todo primo p e todo natural $n \geq 1$ existe essencialmente um único corpo com p^n elementos.

Corpos Finitos

- Teorema de Galois : para todo primo p e todo natural $n \geq 1$ existe essencialmente um único corpo com p^n elementos.
- Um corpo com 4 elementos: $\mathbb{F}_4 = \{0, 1, \alpha, 1+\alpha\}$, com $\alpha^2 = \alpha + 1$.

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

Distância de Hamming

- $\mathbb{F}$ corpo finito de q elementos.

Distância de Hamming

- $\mathbb{F}$ corpo finito de q elementos.
- Distância de Hamming entre dois elementos em $\mathbb{F}^n$ = número de coordenadas distintas.

$$d(\mathbf{u}, \mathbf{v}) = |\{i; u_i \neq v_i\}|$$

Distância de Hamming

- $\mathbb{F}$ corpo finito de q elementos.
- Distância de Hamming entre dois elementos em $\mathbb{F}^n$ = número de coordenadas distintas.

$$d(\mathbf{u}, \mathbf{v}) = |\{i; u_i \neq v_i\}|$$

- Peso de $\mathbf{u} \in \mathbb{F}^n$ = número de coordenadas não-nulas de $\mathbf{u}$

$$\omega(\mathbf{u}) = |\{i; u_i \neq 0\}|$$

Distância de Hamming

- $\mathbb{F}$ corpo finito de q elementos.
- Distância de Hamming entre dois elementos em $\mathbb{F}^n$ = número de coordenadas distintas.

$$d(\mathbf{u}, \mathbf{v}) = |\{i; u_i \neq v_i\}|$$

- Peso de $\mathbf{u} \in \mathbb{F}^n$ = número de coordenadas não-nulas de $\mathbf{u}$

$$\omega(\mathbf{u}) = |\{i; u_i \neq 0\}|$$

- $d(\mathbf{u}, \mathbf{v}) = \omega(\mathbf{u} - \mathbf{v})$

Distância de Hamming

- $\mathbb{F}$ corpo finito de q elementos.
- Distância de Hamming entre dois elementos em $\mathbb{F}^n$ = número de coordenadas distintas.

$$d(\mathbf{u}, \mathbf{v}) = |\{i; u_i \neq v_i\}|$$

- Peso de $\mathbf{u} \in \mathbb{F}^n$ = número de coordenadas não-nulas de $\mathbf{u}$

$$\omega(\mathbf{u}) = |\{i; u_i \neq 0\}|$$

- $d(\mathbf{u}, \mathbf{v}) = \omega(\mathbf{u} - \mathbf{v})$
- Número de erros = distância de Hamming.

Códigos em $\mathbb{F}^n$

- Um código de comprimento n é um subconjunto não vazio de $\mathbb{F}^n$.

Códigos em $\mathbb{F}^n$

- Um código de comprimento n é um subconjunto não vazio de $\mathbb{F}^n$.
- Os elementos de C são as palavras de C, ou as palavras-código ("*codewords*").

Códigos em $\mathbb{F}^n$

- Um código de comprimento n é um subconjunto não vazio de $\mathbb{F}^n$.
- Os elementos de C são as palavras de C, ou as palavras-código ("*codewords*").
- A distância mínima de um código é

$$d(C) = \min\{d(\mathbf{c}, \mathbf{c}'); \mathbf{c}, \mathbf{c}' \in C, \mathbf{c} \neq \mathbf{c}'\}.$$

Códigos em $\mathbb{F}^n$

- Um código de comprimento n é um subconjunto não vazio de $\mathbb{F}^n$.
- Os elementos de C são as palavras de C, ou as palavras-código ("*codewords*").
- A distância mínima de um código é

$$d(C) = \min\{d(\mathbf{c}, \mathbf{c}'); \mathbf{c}, \mathbf{c}' \in C, \mathbf{c} \neq \mathbf{c}'\}.$$

- A capacidade de correção de C é

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor$$

(*maior inteiro menor ou igual a* $(d-1)/2$)

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

com $d(\mathbf{c}, \mathbf{r}) \leq t$.

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

$$d(\mathbf{c}', \mathbf{r}) \leq t.$$

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

$$d(\mathbf{c}', \mathbf{r}) \leq t.$$

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

  $$d(\mathbf{c}', \mathbf{r}) \leq t.$$

  Então

  $$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{r}) + d(\mathbf{c}', \mathbf{r})$$

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

  $$d(\mathbf{c}', \mathbf{r}) \leq t.$$

  Então

  $$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{r}) + d(\mathbf{c}', \mathbf{r}) \leq 2t$$

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

  $$d(\mathbf{c}', \mathbf{r}) \leq t.$$

  Então

  $$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{r}) + d(\mathbf{c}', \mathbf{r}) \leq 2t < d = d(C),$$

  contradição.

Parâmetros - capacidade de correção

- $$\mathbf{c} \in C \longmapsto \mathbf{r} \in \mathbb{F}^n$$

  com $d(\mathbf{c}, \mathbf{r}) \leq t$.

- Suponha que $\mathbf{c}' \in C$, $\mathbf{c}' \neq \mathbf{c}$ e

  $$d(\mathbf{c}', \mathbf{r}) \leq t.$$

  Então

  $$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{r}) + d(\mathbf{c}', \mathbf{r}) \leq 2t < d = d(C),$$

  contradição.

- $\mathbf{c}$ é a única palavra em C tal que $d(\mathbf{c}, \mathbf{r}) \leq t$.

Códigos Lineares

- Operações em $\mathbb{F}^n$:

  soma

  $$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$

  produto por escalar

  $$\alpha \cdot (u_1, \ldots, u_n) = (\alpha u_1, \ldots, \alpha u_n)$$

  Com estas operações, $\mathbb{F}^n$ é um espaço vetorial sobre $\mathbb{F}$.

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;
  - $u \in C, \alpha \in \mathbb{F} \implies \alpha u \in C$.

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;
  - $u \in C, \alpha \in \mathbb{F} \implies \alpha u \in C$.
- Os códigos de repetição são códigos lineares sobre $\mathbb{Z}_2$.

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;
  - $u \in C, \alpha \in \mathbb{F} \implies \alpha u \in C$.
- Os códigos de repetição são códigos lineares sobre $\mathbb{Z}_2$.
- O ISBN-10 não é linear,

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;
  - $u \in C, \alpha \in \mathbb{F} \implies \alpha u \in C$.
- Os códigos de repetição são códigos lineares sobre $\mathbb{Z}_2$.
- O ISBN-10 não é linear,

Códigos Lineares

- Código **linear** em $\mathbb{F}^n$:
  - $u, v \in C \implies u + v \in C$;
  - $u \in C, \alpha \in \mathbb{F} \implies \alpha u \in C$.
- Os códigos de repetição são códigos lineares sobre $\mathbb{Z}_2$.
- O ISBN-10 não é linear, mas está contido em um código linear sobre $\mathbb{Z}_{11}$: as soluções da equação

$$\sum_{j=1}^{10} (11 - j)x_j = 0.$$

Códigos Lineares

- Todo subespaço de $\mathbb{F}^n$ é imagem de uma função **injetora**

$$f : \mathbb{F}^k \to \mathbb{F}^n$$

que é **linear**, isto é,
dados $u, v \in \mathbb{F}^k, \alpha \in \mathbb{F}$,

$$f(u+v) = f(u) + f(v)$$

e

$$f(\alpha u) = \alpha \cdot f(u)$$

Códigos Lineares

- Todo subespaço de $\mathbb{F}^n$ é imagem de uma função **injetora**

$$f : \mathbb{F}^k \to \mathbb{F}^n$$

que é **linear**, isto é,
dados $u, v \in \mathbb{F}^k, \alpha \in \mathbb{F}$,

$$f(u+v) = f(u) + f(v)$$

e

$$f(\alpha u) = \alpha \cdot f(u)$$

- A **dimensão** de C é k.

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código
  - k = dimensão do código (subespaço vetorial)

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código
  - k = dimensão do código (subespaço vetorial)
  - $d = d(C)$ = distância mínima = $\min\{\omega(\mathbf{c}); \mathbf{c} \in C, \mathbf{c} \neq 0\}$.

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código
  - k = dimensão do código (subespaço vetorial)
  - $d = d(C)$ = distância mínima = $\min\{\omega(\mathbf{c}); \mathbf{c} \in C, \mathbf{c} \neq 0\}$.
- Parâmetros “derivados”:

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código
  - k = dimensão do código (subespaço vetorial)
  - $d = d(C)$ = distância mínima = $\min\{\omega(\mathbf{c}); \mathbf{c} \in C, \mathbf{c} \neq 0\}$.
- Parâmetros “derivados”:
  - k/n = taxa (de informação) do código

Parâmetros

- Parâmetros fundamentais de um código linear $C \subset \mathbb{F}^n$:
  - n = comprimento do código
  - k = dimensão do código (subespaço vetorial)
  - $d = d(C)$ = distância mínima = $\min\{\omega(\mathbf{c}); \mathbf{c} \in C, \mathbf{c} \neq 0\}$.
- Parâmetros "derivados":
  - k/n = taxa (de informação) do código
  - $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ = capacidade de correção do código

Parâmetros

No caso do código de repetição temos

- comprimento $n = 3$

Parâmetros

No caso do código de repetição temos

- comprimento $n = 3$
- dimensão $k = 1$

Parâmetros

No caso do código de repetição temos

- comprimento $n = 3$
- dimensão $k = 1$
- distância mínima $d = 3$

Parâmetros

No caso do código de repetição temos

- comprimento $n = 3$
- dimensão $k = 1$
- distância mínima $d = 3$
- taxa $k/n = 1/3$

Parâmetros

No caso do código de repetição temos

- comprimento $n = 3$
- dimensão $k = 1$
- distância mínima $d = 3$
- taxa $k/n = 1/3$
- capacidade de correção $t = 1$

Parâmetros

O código de repetição de comprimento $2m+1$ tem parâmetros:

- comprimento $2m+1$
- dimensão $k=1$
- distância mínima $d=2m+1$
- taxa $1/(2m+1)$
- capacidade de correção $t=m$

Parâmetros

O código linear associado ao ISBN-10 tem parâmetros:

- comprimento 10
- dimensão $k = 9$
- distância mínima $d = 2$
- taxa $9/10$
- capacidade de correção $t = 0$.

Códigos de Hamming

Códigos Lineares - matriz de paridade

- Todo subespaço C de $\mathbb{F}^n$ com dimensão k pode ser escrito como o conjunto de soluções de um sistema linear

$$Hv = 0$$

em que H é matriz de ordem $n - k \times n$, de **posto máximo**: há um subdeterminante não-nulo de ordem $n - k$.

Códigos Lineares - matriz de paridade

- Todo subespaço C de $\mathbb{F}^n$ com dimensão k pode ser escrito como o conjunto de soluções de um sistema linear

$$Hv = 0$$

em que H é matriz de ordem $n - k \times n$, de **posto máximo**: há um subdeterminante não-nulo de ordem $n - k$.

Códigos Lineares - matriz de paridade

- Todo subespaço C de $\mathbb{F}^n$ com dimensão k pode ser escrito como o conjunto de soluções de um sistema linear

$$Hv = 0$$

em que H é matriz de ordem $n - k \times n$, de **posto máximo**: há um subdeterminante não-nulo de ordem $n - k$.

Dizemos que C é o núcleo da matriz H.

Códigos Lineares - matriz de paridade

- Todo subespaço C de $\mathbb{F}^n$ com dimensão k pode ser escrito como o conjunto de soluções de um sistema linear

$$Hv = 0$$

em que H é matriz de ordem $n - k \times n$, de **posto máximo**: há um subdeterminante não-nulo de ordem $n - k$.

Dizemos que C é o núcleo da matriz H.

H é uma matriz de paridade de C ("parity check matrix").

Código de repetição

- O código de repetição de comprimento 3 é o núcleo da matriz

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

Código de repetição

- O código de repetição de comprimento 3 é o núcleo da matriz

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

- O código de repetição de comprimento $2m+1$ é o núcleo da matriz cujas colunas são todos os elementos de $\mathbb{F}_2^{2m+1}$ de peso par.

Códigos de Hamming (R. Hamming, 1947)

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

cujas colunas são todos os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.

Códigos de Hamming (R. Hamming, 1947)

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todos os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $\text{posto}(H) = 3 \implies k = 4$.

Códigos de Hamming

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todas os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $d = 3$.

Códigos de Hamming

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todas os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $d = 3$.
  - O elemento $(1, 1, 1, 0, 0, 0, 0)$ está no código $\Longrightarrow$ $d \leq 3$

Códigos de Hamming

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todas os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $d = 3$.
  - O elemento $(1, 1, 1, 0, 0, 0, 0)$ está no código $\Longrightarrow$ $d \leq 3$
  - Não há duas colunas iguais $\Longrightarrow$ não há elemento de peso 2

Códigos de Hamming

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todas os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $d = 3$.
  - O elemento $(1, 1, 1, 0, 0, 0, 0)$ está no código $\Longrightarrow$ $d \leq 3$
  - Não há duas colunas iguais $\Longrightarrow$ não há elemento de peso 2
  - Não há coluna nula $\Longrightarrow$ não há elemento de peso 1.

Códigos de Hamming

- Seja $\mathcal{H}_3 \subset (\mathbb{F}_2)^7$ o núcleo da matriz

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

  cujas colunas são todas os elementos não-nulos de $(\mathbb{F}_2)^3$, sem repetições.
- $d = 3$.
  - O elemento $(1,1,1,0,0,0,0)$ está no código $\Longrightarrow$ $d \leq 3$
  - Não há duas colunas iguais $\Longrightarrow$ não há elemento de peso 2
  - Não há coluna nula $\Longrightarrow$ não há elemento de peso 1.
- Este é o código de Hamming $\mathcal{H}_3$ de parâmetros [7,4,3].

Códigos de Hamming

Processo de decodificação/correção:

- vetor recebido com 1 erro: $\mathbf{r} = \mathbf{c} + \mathbf{e}_j$

Códigos de Hamming

Processo de decodificação/correção:

- vetor recebido com 1 erro: $\mathbf{r} = \mathbf{c} + \mathbf{e}_j$
- $$H\mathbf{r} = H\mathbf{c} + H\mathbf{e}_j = \mathbf{h}_j$$

Códigos de Hamming

Processo de decodificação/correção:

- vetor recebido com 1 erro: $\mathbf{r} = \mathbf{c} + \mathbf{e}_j$
- $$H\mathbf{r} = H\mathbf{c} + H\mathbf{e}_j = \mathbf{h}_j$$
- a j-ésima coluna é a expansão binária de j

Códigos de Hamming

Processo de decodificação/correção:

- vetor recebido com 1 erro: $\mathbf{r} = \mathbf{c} + \mathbf{e}_j$
- $$H\mathbf{r} = H\mathbf{c} + H\mathbf{e}_j = \mathbf{h}_j$$
- a j-ésima coluna é a expansão binária de j
- correção: $\mathbf{r} \mapsto \mathbf{r} - \mathbf{e}_j = \mathbf{c}$.

Códigos de Hamming de comprimento maior (M.Golay, 1949)

- H_m é uma matriz cujas colunas são os elementos não-nulos de $(\mathbb{F}_2)^m$
- $d = 3$, logo $t = 1$
- $n = 2^m - 1$
- $k = 2^m - m - 1$.

Códigos de Hamming de comprimento maior (M.Golay, 1949)

- H_m é uma matriz cujas colunas são os elementos não-nulos de $(\mathbb{F}_2)^m$
- $d = 3$, logo $t = 1$
- $n = 2^m - 1$
- $k = 2^m - m - 1$.
- taxa: $\dfrac{2^m - m - 1}{2^m - 1} = 1 - \dfrac{m}{2^m - 1}$.

Códigos de Hamming não-binários (M.Golay, 1949)

- Corpo $\mathbb{F}_q$
- Matriz $H_{q,m}$: bijeção

$$\text{colunas de } H_{q,m} \iff \text{retas em } (\mathbb{F}_q)^m$$

- $d = 3$, logo $t = 1$
- $n = \dfrac{q^m - 1}{q - 1}$
- $k = \dfrac{q^m - 1}{q - 1} - m.$

Códigos de Hamming não-binários (M.Golay, 1949)

- Corpo $\mathbb{F}_q$
- Matriz $H_{q,m}$: bijeção

$$\text{colunas de } H_{q,m} \iff \text{retas em } (\mathbb{F}_q)^m$$

- $d = 3$, logo $t = 1$
- $n = \dfrac{q^m - 1}{q - 1}$
- $k = \dfrac{q^m - 1}{q - 1} - m.$
- taxa: $1 - \dfrac{m}{q^m - 1}.$

references

- L. Egghe, R. Rousseau, *The Detection of Double Errors in ISBN- and ISSN-like Codes*. Mathematical and Computer Modelling 33 (2001) 943-955.
- Abramo Hefez e Maria L.T. Vilela, Introdução aos Códigos Corretores de Erros. IMPA.
- R. Lidl e H. Niederreiter, Introduction to Finite Fields and their Applications (revised edition), caps 2, 3, 8. Cambridge.
- W.C. Huffman, Vera Pless, Fundamentals of Error-Correcting Codes. Cambridge.