Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Campus São Paulo

VALTER FÉLIX PEREIRA DA SILVA

APLICAÇÕES DOS CÓDIGOS CORRETORES DE ERROS COM FUNDAMENTAÇÃO TEÓRICA EM ÁLGEBRA

São Paulo

2020

VALTER FÉLIX PEREIRA DA SILVA

APLICAÇÕES DOS CÓDIGOS CORRETORES DE ERROS COM FUNDAMENTAÇÃO TEÓRICA EM ÁLGEBRA

Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, como parte dos requisitos para a obtenção do grau acadêmico de Licenciado em Matemática.

Orientadora: Profa. Dra. Valéria Ostete Jannis Luchetta

São Paulo

2020

Catalogação na fonte
Biblioteca Francisco Montojos - IFSP Campus São Paulo
Dados fornecidos pelo(a) autor(a)

s586a Silva, Valter Félix Pereira da
APLICAÇÕES DOS CÓDIGOS CORRETORES DE ERROS COM FUNDAMENTAÇÃO TEÓRICA EM ÁLGEBRA / Valter Félix Pereira da Silva. São Paulo: [s.n.], 2020.
125 f.

Orientadora: Valéria Ostete Jannis Luchetta

Trabalho de Conclusão de Curso (Licenciatura em Matemática) - Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, IFSP, 2020.

1. Códigos de Hamming. 2. Códigos de Reed-muller de Primeira Ordem. 3. Códigos Lineares. 4. Álgebra. 5. Álgebra Linear. I. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo II. Título.

CDD 510

SERVIÇO PÚBLICO FEDERAL
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SÃO PAULO
DIRETORIA GERAL/CAMPUS SAO PAULO
Câmpus São Paulo, (11) 2763-7520, Rua Pedro Vicente, 625, CEP 01109-010, São Paulo (SP)

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Na presente data realizou-se a sessão pública de defesa da Trabalho de Conclusão de Curso intitulada **APLICAÇÕES DOS CÓDIGOS CORRETORES DE ERROS COM FUNDAMENTAÇÃO TEÓRICA EM ÁLGEBRA** apresentada pelo aluno **Valter Felix Pereira da Silva (SP1673131)** do Curso **LICENCIATURA EM MATEMÁTICA (Câmpus São Paulo)**. Os trabalhos foram iniciados às **10:00** pela Professora presidente da banca examinadora, constituída pelos seguintes membros:

Membros	IES	Presença	Aprovação/Conceito *(quando exigido)*
Valeria Ostete Jannis Luchetta (Orientadora)			aprovado/ 10,0
Vania Batista Flose Jardim (Examinadora Interna)			aprovado/ 10,0
Douglas Silva Maioli (Examinador Externo)			aprovado/ 10,0

Observações:

A banca examinadora, tendo terminado a apresentação do conteúdo da monografia, passou à argüição do candidato. Em seguida, os examinadores reuniram-se para avaliação e deram o parecer final sobre o trabalho apresentado pelo aluno, tendo sido atribuído o seguinte resultado:

[X] Aprovado [ ] Reprovado Nota (quando exigido): 10,0 (dez)

Proclamados os resultados pelo presidente da banca examinadora, foram encerrados os trabalhos e, para constar, eu lavrei a presente ata que assino juntamente com os demais membros da banca examinadora.

SÃO PAULO / SP, 19/08/2020

Valéria O. Jannis Luchetta
Valeria Ostete Jannis Luchetta

Vania B. Flose Jardim
Vania Batista Flose Jardim

Douglas Silva Maioli
Douglas Silva Maioli

Aos meus pais Gilvânia e Valter
e à minha namorada Lucy

AGRADECIMENTOS

Agradeço primeiramente aos meus pais Gilvânia e Valter que sempre me deram ótimas condições e incentivo para estudar. Tudo que eu sou hoje, e todas as conquistas que eu obtive foram graças a vocês. Sempre estiveram ao meu lado, e sempre estarão.

A maior contribuição intelectual foi graças a minha orientadora Valéria, queria deixar aqui o enorme apreço que eu tenho por ela, desde as aulas de álgebra abstrata quando a conheci, sempre ensinando com felicidade e alegria, até as últimas correções deste trabalho que não foram poucas. Agradeço por nunca ter perdido a paciência comigo, e não faltaram oportunidades para isso. Indiretamente essa admiração que eu tenho sempre me motivava a continuar seguindo em frente e melhorar.

Agradeço à minha namorada Lucy que sempre tentou me animar durante o processo de escrita do trabalho, não me fazendo desistir do tema para buscar algo mais fácil. Ela continuou ao meu lado mesmo eu estando extremamente estressado e chato. Todos os dias eu enchia o saco dela com detalhes deste trabalho, muito obrigado pela paciência. Saiba que eu sou seu maior admirador e sei o quão esforçada e inteligente você é, basta confiar em si mesma. Meu coração não aguenta mais de saudade.

Queria agradecer a todo corpo docente do Instituto Federal que contribuiu de forma significativa com a minha formação, nesses quatro anos que foram momentos de bastante sofrimento e evolução. Agradeço ao professor Curvello que foi o primeiro professor do curso que me fez ter a certeza de estar seguindo o caminho certo. Em particular aos professores Granero, Lucas, Flávia, Cesar e Larissa que tiveram um grande impacto em minha formação, sendo excelentes profissionais e amigos, acho que eu não poderia ter professores melhores que vocês.

Também gostaria de agradecer a todos os meus amigos. Sempre compartilhavamos nossos sofrimentos e ajudávamos uns aos outros nos momentos mais difíceis. Tenho certeza que serão ótimos profissionais independente da carreira que seguirem, pois o principal vocês já têm que é o caráter. Queria agradecer em especial aos meus amigos Gustavo, Jéssica e Wemerson.

Quero deixar um agradecimento aos professores Douglas e Vania pelas contribuições feitas ao meu trabalho e principalmente pela mensagem deixada pela professora, esses pequenos detalhes me deixam muito feliz e realizado com trajetória que eu percorri até o momento.

E por fim, gostaria de agradecer ao leitor, espero que este trabalho possa contribuir de alguma forma. Boa leitura, ou melhor, bons estudos.

"*It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment*".

Carl Friedrich Gauss

RESUMO

Tendo em vista o caráter abstrato da matemática e a importância de aplicações para o desenvolvimento em outras ciências, este trabalho apresenta e descreve duas aplicações de códigos corretores de erros pertencentes à família de códigos lineares, chamadas de códigos de Hamming e de códigos de Reed-Muller de primeira ordem. O objetivo é entender como funciona essas aplicações de álgebra na teoria dos códigos corretores de erros. Para atingir o objetivo especificado, explica-se o que é um código corretor de erro, apresenta-se conceitos básicos da teoria de códigos, desenvolve-se conceitos de álgebra e álgebra linear e define-se o que é um código linear. O método utilizado foi a pesquisa bibliográfica e documental. Espera-se que esta pesquisa estimule o leitor a se aprofundar neste campo da Álgebra que é interessantíssimo e possui muitas aplicações em diversos ramos da ciência.

Palavras-chave: Códigos de Hamming. Códigos de Reed-Muller de Primeira Ordem. Códigos Lineares. Álgebra. Álgebra Linear.

ABSTRACT

In view of the abstract math feature and the importance of applications for development in other sciences, this study introduces and describes two applications of error-correcting codes belonging to the families of linear codes, called Hamming codes and first order Reed-Muller codes. The aim is to understand how these applications of algebra works in the theory of error-correcting codes. To achieve the specified objective, this study explains what a code, presents basic concepts of this theory, develops concepts of algebra, linear algebra and defines a linear code. This is an bibliographic and documentary research. It's expected that this research encourages the reader delve into the study of algebra which is very important and have many applications in several science branches.

Keywords: Hamming Codes. First Order Reed-Muller Codes. Linear Codes. Algebra. Linear Algebra.

LISTA DE ILUSTRAÇÕES

LISTA DE ABREVIATURAS E SIGLAS

assoc.	Associativa
comut.	Comutativa
dist.	Distributiva
ad.	Adição
mult.	Multiplicação

LISTA DE SÍMBOLOS

$\mathbb{N}$	Conjunto dos números Naturais
$\mathbb{Z}$	Conjunto dos números Inteiros
$\mathbb{Q}$	Conjunto dos números Racionais
$\mathbb{R}$	Conjunto dos números Reais
α	Letra grega minúscula alfa
β	Letra grega minúscula beta
γ	Letra grega minúscula gama
κ	Letra grega minúscula kappa
λ	Letra gerga minúscula lambda
π	Letra grega minúscula pi
σ	Letra grega minúscula sigma
τ	Letra grega minúscula tau
ϕ	Letra grega minúscula phi
ω	Letra grega minúscula omega
$\forall$	Para todo
$\exists$	Existe
$\exists!$	Existe e é único
$\in$	Pertence
$\subset$	Contido
$\cup$	União
$\cap$	Intersecção

SUMÁRIO

1 INTRODUÇÃO

A teoria dos códigos corretores de erros é um campo de pesquisa muito dinâmico e ativo atualmente, devido as suas aplicações em diversas áreas do conhecimento como a matemática, computação, engenharia elétrica, estatística, etc.

Segundo Milies (2009), o que motivou o desenvolvimento desta teoria foi uma consequência dos computadores serem muito caros durante a década de 1940. Assim, apenas grandes instituições como o governo ou as universidades possuíam essas máquinas. Uma das poucas instituições que disponha destes computadores era o Laboratório Bell de Tecnologia, e o matemático estadunidense Richard Wesley Hamming trabalhava nessas máquinas em 1947.

Nesta época, programas eram gravados em cartões perfurados, onde os computadores detectavam erros a partir da leitura destes cartões. Em caso de detecção de erros, a leitura era interrompida e o computador era programado para ler um próximo cartão. Por causa disto, Hamming questionou, se as máquinas podem detectar um erro, porque não conseguem localizar e corrigir este erro?

Esta questão foi crucial para o desenvolvimento desta teoria. Hamming conseguiu desenvolver códigos capazes de detectar até dois erros e corrigir um erro, este trabalho em específico foi publicado em 1950. Ele também desenvolveu trabalhos que questionavam sobre a possibilidade de criar códigos mais eficientes.

Em paralelo aos trabalhos de Hamming, o matemático estadunidense Claude Elwood Shannon publicou o artigo *A Mathematical Theory of Communication* no *The Bell System Technical Journal* em 1948. Este artigo respondeu indiretamente o questionamento feito por Hamming, dando assim início a dois novos campos de pesquisa em matemática: a teoria dos códigos corretores de erros e a teoria da informação.

De acordo com Hefez e Villela (2008), a teoria dos códigos corretores de erros foi estabelecida em 1948, pelo matemático estadunidense Claude Elwood Shannon. Este tema interessou primeiramente aos matemáticos das décadas de 50 e 60. Com o início da corrida espacial na década de 70, ela ganhou grande importância em aplicações espaciais entre os engenheiros, dando início ao desenvolvimento tecnológico, tornando-se essencial o estudo e desenvolvimento dessa Teoria.

Nos dias de hoje, os códigos corretores de erros são utilizados em qualquer processo que se utiliza transmissão e armazenamento de dados, o que garante uma comunicação satisfatória, permitindo identificar a presença de erros. Alguns exemplos desses códigos são a comunicação via satélite, as comunicações internas de dispositivos eletrônicos (com-

putadores, celulares, etc.), o armazenamento de dados em disco rígido (HD), *pen drive*, cartões de memória, entre outros.

A teoria dos códigos corretores de erros nos remete a algo interessante, pois contém diversas aplicações na vida real, constituindo-se essencialmente de tópicos da álgebra abstrata, o que aproxima a Matemática Pura da Matemática Aplicada.

Este tema propicia a utilização da álgebra e álgebra linear em aplicações práticas da vida cotidiana dos alunos, além de ser um ótimo motivador para a utilização destes conceitos no curso de Licenciatura em Matemática, ampliando as oportunidades para os graduandos em participações de Congressos, fóruns, SEDCITEC[1] e etc.

Ao trabalharmos com a teoria dos códigos corretores de erros, vemos que o potencial de inovação desse trabalho constitui-se em introduzir novidades e aperfeiçoamento no ambiente social, que resulte em novos estudos e pesquisas sobre aplicações da Álgebra.

Os Parâmetros Curriculares Nacionais (PCN) descrevem o papel da Matemática como decisivo, pois possibilita a resolução de problemas no cotidiano, apresentando diversas aplicações no mundo do trabalho e funciona como ferramenta para construção de outros conhecimentos. Interfere também na formação de capacidade intelectual, estruturação do pensamento e no aperfeiçoamento do raciocínio lógico.

A relação da Matemática Abstrata com aplicações práticas é abordada, segundo Brasil (1997), da seguinte forma:

> "Mas a vitalidade da Matemática deve-se também ao fato de que, apesar de seu caráter abstrato, seus conceitos e resultados têm origem no mundo real e encontram muitas aplicações em outras ciências e em inúmeros aspectos práticos da vida diária: na indústria, no comércio e na área tecnológica. Por outro lado, ciências como Física, Química e Astronomia têm na Matemática ferramenta essencial." (BRASIL, 1997, p. 23).

Nesse trecho, podemos perceber a real importância das aplicações matemáticas no desenvolvimento e construção de novos conhecimentos dos alunos na Educação Básica.

Além disso, segundo Brasil (1997), a Matemática tem uma contextualização sociocultural que é compreender o conhecimento científico e tecnológico como um resultado da construção humana, inseridos em um processo social e histórico. A Matemática também nos possibilita avaliar e reconhecer o desenvolvimento tecnológico atual, suas relações com as ciências, seu papel na vida humana, sua presença no mundo cotidiano e seus impactos na vida social.

[1] Semana de Ciência, Educação e Tecnologia do campus São Paulo. De acordo com Brasil (2019), é um evento que envolve diversas áreas acadêmicas e aborda temas pertinentes ao desenvolvimento profissional dos alunos, além de oferecer experiências culturais e oficinas práticas que aproximam os estudantes através de temáticas e vivências manuais.

O problema central desta pesquisa é como transformar o código da fonte em código de canal[2], de maneira com que se possa identificar e corrigir erros de comunicação ou transmissão, ou qualquer perda de informação em uma das partes do processo de transmissão.

Tendo em vista a importância do tema para aplicações matemáticas, este trabalho tem como objetivo geral entender como funcionam algumas aplicações de álgebra abstrata na teoria dos códigos corretores de erros que são os códigos de *Hamming* e os códigos de *Reed-Muller* de primeira ordem.

Com o intuito de atingir o objetivo geral, temos como objetivos específicos entender o que é um código corretor de erro e como funciona a teoria dos códigos, relembrar conceitos e as ferramentas necessárias de álgebra e álgebra linear, para que seja possível compreender a construção e a classificação das famílias dos códigos corretores de erros.

Em relação a metodologia, trata-se de uma pesquisa bibliográfica e documental. Documental pois foram utilizados documentos públicos nacionais. De acordo com Marconi e Lakatos (2003), a pesquisa bibliográfica

> "[...] abrange toda bibliografia já tornada pública em relação ao tema de estudo, desde publicações avulsas, boletins, jornais, revistas, livros, pesquisas, monografias, teses, material cartográfico etc., até meios de comunicação orais: rádio, gravações em fita magnética e audiovisuais: filmes e televisão."(MARCONI; LAKATOS, 2003, p. 183).

A bibliografia básica utilizada nesta pesquisa foi o livro *Códigos Corretores de Erros* escrito, planejado e revisado por Abramo Hefez e Maria Lúcia T. Villela. Publicado pelo Instituto Nacional de Matemática Pura e Aplicada (IMPA) na Série de Computação e Matemática no Rio de Janeiro, 2008.

Como é comum nas pesquisas na área da Matemática Pura e Aplicada, ter a parte teórica dos estudos conduzida de maneira conjunta com o orientador, consistindo na leitura crítica e cuidadosa dos trechos selecionados dos livros de Domingues e Iezzi (2018) e Hefez e Villela (2008), na resolução de exercícios e na confecção de exemplos. Em reuniões semanais, além de apontar as seções de exercícios e resolver as dúvidas levantadas, o progresso dos estudos foi continuamente avaliado. Breves exposições dirigidas pela orientadora foram efetuadas, especialmente a respeito dos conteúdos abordados em álgebra abstrata e códigos corretores de erros.

O processo de ensino–aprendizagem também foi dado por meio do desenvolvimento do pesquisador, na confecção de exercícios e exemplos representativos da matéria estudada por meio de exemplos e exercícios desenvolvidos e apresentados à orientadora, que por sua vez, revisou e indicou correções e aperfeiçoamentos ao texto.

[2] Estes conceitos serão definidos no Capítulo 2 Seção 2.1.

Este trabalho contém seis capítulos, referências bibliográficas, um apêndice e um anexo. A seguir, descreveremos sucintamente a sua organização.

O Capítulo 2 é dedicado ao estudo de conceitos básicos da teoria dos códigos corretores de erros. Nele, explicamos o que é um código, definimos os códigos corretores de erros, métrica de *Hamming*, parâmetros, isometria e equivalência de códigos.

Nos Capítulos 3 e 4 estudaremos os conteúdos de álgebra e álgebra linear, necessários para o desenvolvimento e o entendimento das aplicações de códigos lineares. No Capítulo 3 recordamos algumas estruturas algébricas básicas, apresentando conceitos de Anéis e Corpos com ênfase no anel dos números inteiros, onde veremos a divisibilidade em $\mathbb{Z}$ e o conceito de inteiros módulo m, também estudaremos a mudança de alfabeto de um código. No Capítulo 4 trabalhamos noções básicas de álgebra linear e estudamos corpos finitos, vendo algumas de suas propriedades, e aprendemos a classificá-los.

O Capítulo 5 é dedicado aos códigos lineares. Neste capítulo definimos o que são códigos lineares, determinamos seus parâmetros, algoritmos gerais de correção e apresentamos duas aplicações dessa classe de códigos.

O trabalho se encerra com o Apêndice A e o Anexo A dedicados a construção dos anéis de polinômios e a divisão entre polinômios, conteúdos necessários para outras aplicações de códigos lineares, como por exemplo a família dos códigos de *Reed-Solomon*.

Os pré-requisitos para ler este trabalho consistem em ter uma certa maturidade em matemática, sendo necessário que o leitor tenha noções básicas de álgebra abstrata e álgebra linear, pois todos os conteúdos discutidos nos Capítulos 3 e 4 foram apenas a título de recordação.

2 CÓDIGOS CORRETORES DE ERROS

Neste capítulo vamos introduzir conceitos básicos da Teoria dos Códigos Corretores de Erros, de modo que, o leitor compreenda por meio de exemplos intuitivos o que são códigos e códigos corretores de erros. Definiremos os códigos corretores de erros, a métrica de *Hamming*, os parâmetros de um código, a isometria e a equivalência de códigos. Para a composição deste capítulo foram utilizados os seguintes materiais: Hefez e Villela (2008); Luchetta (2005); Couto (2010); Milies (2009) e Lima (1970).

2.1 O que é um Código?

Para entendermos o que é um Código, vamos trabalhar com alguns exemplos:

Exemplo 2.1.1. Um código bastante utilizado é o idioma. Considere as palavras da língua Portuguesa como nosso conjunto $\mathscr{P}$, supondo que seja escrita a palavra "concentimento", comparando-a com o conjunto $\mathscr{P}$, vemos que ela não pertence ao conjunto, sendo assim, podemos inferir que ocorreu pelo menos um erro na gráfia da palavra. Nesse caso a correção é factível, pois a palavra que pertence a $\mathscr{P}$ que mais se assemelha a "concentimento" é consentimento, pois difere apenas em uma letra.

Porém esse tipo de **código de comparação** não é muito eficiente, pois, se a palavra bela fosse erroneamente escrita como vela, ou como cela, ou ainda como tela, o erro não seria identificado pois todas essas palavras pertencem a $\mathscr{P}$.

Um Código Corretor de Erros bastante utilizado é aquele que contêm dígitos verificadores, como por exemplo o Cadastro de Pessoa Física (CPF) e os números de contas bancárias que fornecem dígitos adicionais para detectar erros.

Exemplo 2.1.2. Verificando agora como funciona o código com dígitos verificadores no CPF. Considere o CPF: 510.702.040-21[1], neste CPF os dígitos verificadores são os últimos dois números. Para facilitar o entendimento do algoritmo, denota-se o primeiro dígito verificador 2 por v_1 e o segundo dígito verificador 1 por v_2. Além disso sabe-se que o CPF possui mais 9 números, então considere que o CPF tenha a forma genérica:

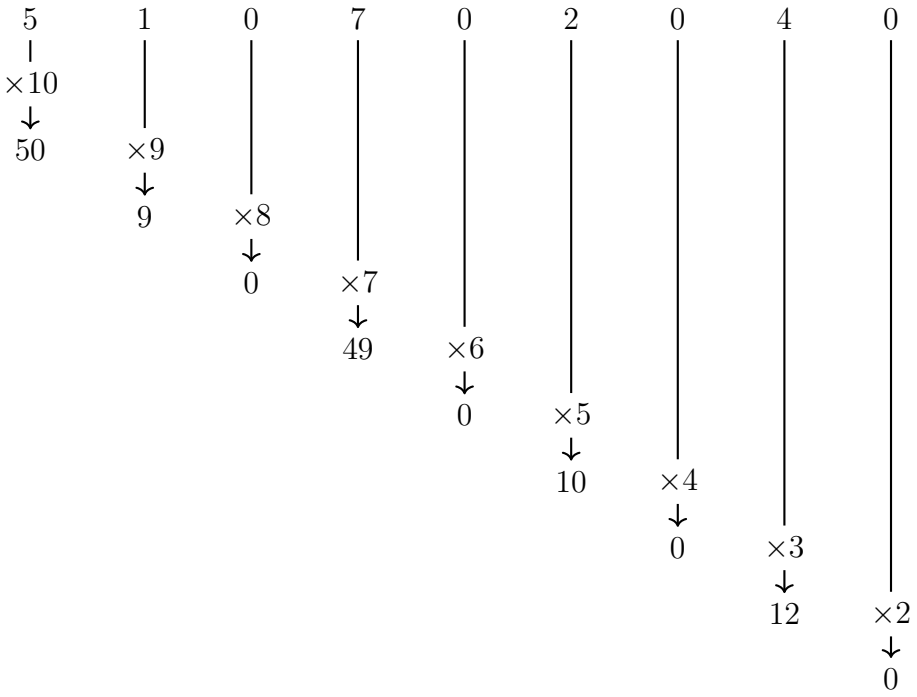
$$n_1n_2n_3.n_4n_5n_6.n_7n_8n_9 - v_1v_2$$

para quaisquer n_i e $v_i \in \mathbb{N}$ tais que $0 \leq n_i \leq 9$ e $0 \leq v_i \leq 9$.

[1] CPF gerado apenas para fins informativos, pelo *website* Devs, 4. (2012), "portal com ferrementas *online* e gratuitas com intuito de ajudar estudantes, programadores, analistas e testadores de *software*".

Vamos descrever o algoritmo para encontrar os dígitos verificadores.

Para encontrar v_1 devemos multiplicar n_1 por 10, e adicionar n_2 multiplicado por 9, e adicionar n_3 multiplicado por 8, e assim sucessivamente até o número n_9. Como esquematizado a seguir:



$$50 + 9 + 0 + 49 + 0 + 10 + 0 + 12 + 0 = 130$$

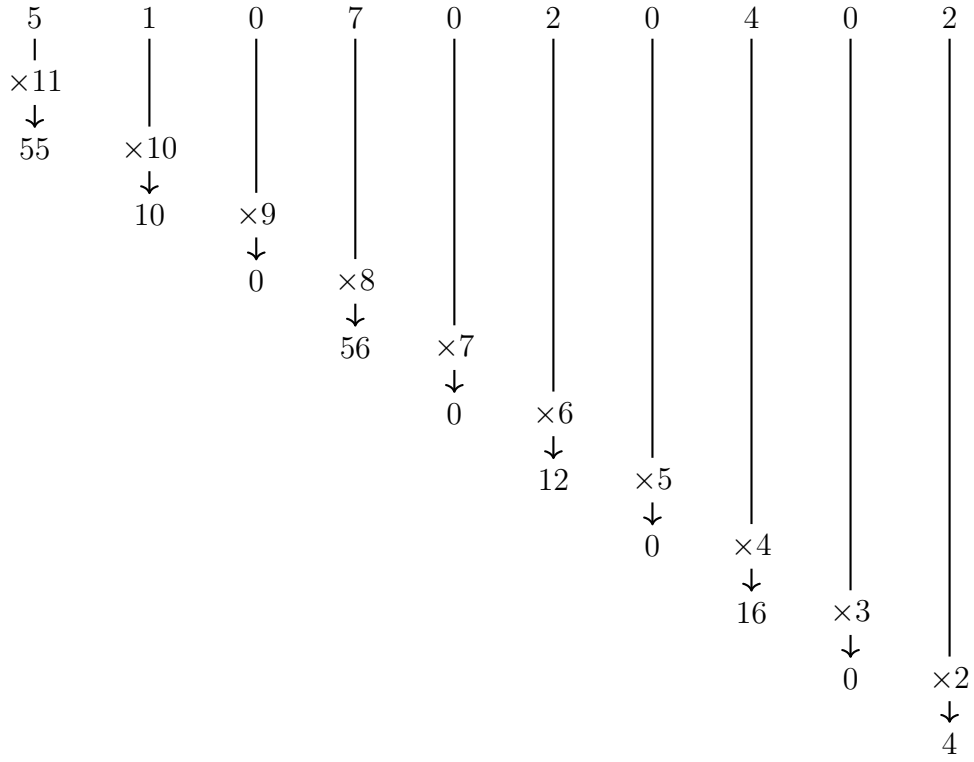
Se o resto da divisão entre a soma desses números e 11 for 0 ou 1, então o primeiro dígito verificador v_1 será 0, caso contrário $v_1 = 11 - r$, sendo r o resto da divisão.

Então, em nosso exemplo numérico, o resto da soma desses números dividido por 11 é:

$$\begin{array}{r|l} 130 & 11 \\ \hline 11 & 11 \\ \hline 20 & \\ 11 & \\ \hline 9 & \end{array}$$

Como o resto é 9, então $v_1 = 11 - 9 = 2$, corresponde exatamente ao primeiro dígito verificador.

Agora para encontrar v_2 deve-se multiplicar n_1 por 11, e adicionar n_2 multiplicado por 10, e adicionar n_3 multiplicado por 9, e assim sucessivamente até v_1. Como esquematizado a seguir:



$$55 + 10 + 0 + 56 + 0 + 12 + 0 + 16 + 0 + 4 = 153$$

Se o resto da divisão entre a soma desses números e 11 for 0 ou 1, então o segundo dígito verificador v_2 será 0, caso contrário $v_2 = 11 - r$, sendo r o resto da nova divisão.

Então, em nosso exemplo numérico, o resto da soma desses números dividido por 11 é:

$$\begin{array}{r|l} 153 & 11 \\ \hline 11 & 13 \\ \hline 43 & \\ 33 & \\ \hline 10 & \end{array}$$

Como o resto é 10, então $v_2 = 11 - 10 = 1$, corresponde exatamente ao segundo dígito verificador.

Tomemos um outro exemplo que ilustra bem o princípio desta teoria.

Exemplo 2.1.3. Considere um tabuleiro quadriculado. Suponha que uma peça se mova neste tabuleiro, de tal forma que, os comandos (*Leste*, *Oeste*, *Norte* e *Sul*), desloque a peça para posições vizinhas da posição atual.

Codificando os comandos dessa forma:

$$\begin{array}{cccccc} Leste & \longmapsto & 00 & Norte & \longmapsto & 10 \\ Oeste & \longmapsto & 01 & Sul & \longmapsto & 11 \end{array}$$

Assim, esses elementos fazem parte de um código, que é chamado de **Código da Fonte**. Suponha que por qualquer motivo ocorra uma interferência na transmissão do comando, a peça seguiria uma direção diferente da desejada. Por exemplo, caso o comando desejado seja 00 (*Leste*), sofra algum tipo de interferência, ou seja erroneamente escrito. Suponhamos então que, recebemos o código 01 (*Oeste*), neste caso o código pertence ao Código da Fonte, o que torna impossível a identificação do erro. Qualquer outro erro cometido pertenceria ao código, então para evitar este tipo de situação, acrescentamos algumas informações ao código, como pode ser observado a seguir:

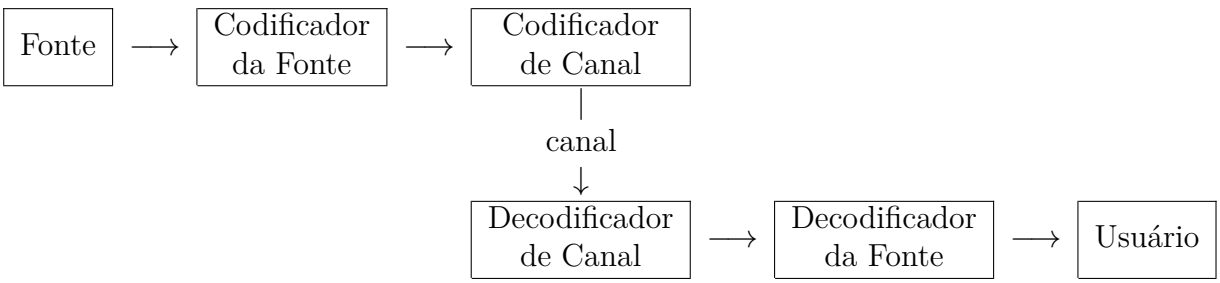
$$\begin{aligned} 00 &\longmapsto 00000 \\ 01 &\longmapsto 01011 \\ 10 &\longmapsto 10110 \\ 11 &\longmapsto 11101 \end{aligned}$$

Nessa recodificação as duas primeiras posições reproduzem o código da fonte, enquanto as restantes são redundâncias para que se identifique possíveis erros de transmissão, este novo código é chamado de **Código de Canal**. Por exemplo, se o comando 10110 (*Norte*) seja transmitido como 11110, nota-se que esse comando não pertence ao nosso Código da Fonte, portanto foi identificado um erro, porém ele se assemelha com *Norte*, que seria de fato o comando transmitido.

Esse tipo de Código Corretor de Erros organiza e acrescenta alguns dados na informação que está sendo transmitida ou armazenada, de tal forma que seja possível recuperar, detectar ou até mesmo corrigir alguns possíveis erros contidos na informação.

Esquematizando o procedimento descrito no exemplo da peça sobre um tabuleiro quadriculado, temos:

Figura 1 – Processo de Transmissão de um Código



Fonte: Hefez e Villela (2008).

O canal pode ser de radiofrequência, micro-ondas, cabo, circuito integrado digital, fita magnética, disco de armazenamento, etc.

Neste trabalho trataremos apenas dos códigos corretores de erros que consistem em transformar o código da fonte em código de canal. Vamos considerar apenas canais simétricos, isto é, todos os símbolos têm a mesma probabilidade de serem recebidos errados e, se recebido errado, a probabilidade de ser um dos outros símbolos restantes é a mesma.

2.2 Métrica de Hamming

Para que possamos começar a construir um código corretor de erros é necessário considerar um conjunto finito $\mathcal{A}$ que chamaremos de alfabeto. O número de elementos de $\mathcal{A}$ será denotado por $|\mathcal{A}|$, e simbolizado por q.

Os elementos de $\mathcal{A}$ serão indicados como **letras** ou **dígitos**. Uma sequência de elementos de $\mathcal{A}$ chamaremos de **palavra**, e o número de elementos desta sequência será denominado como **comprimento**.

Consideraremos $\mathcal{A}^n$ o conjunto de palavras de comprimento n sobre $\mathcal{A}$, ou seja, $\mathcal{A}^n = \{(a_1, \cdots, a_n) \mid a_i \in \mathcal{A},\ 1 \leq a_i \leq n\}$.

Definição 2.2.1. Um **código** C é um subconjunto próprio qualquer de $\mathcal{A}^n$, para algum número natural n.

A fim de tornar intuitiva a noção de proximidade entre palavras, que foi utilizada na Seção 2.1, definiremos a seguir uma forma de medir a distância entre palavras pertencentes ao conjunto $\mathcal{A}^n$.

Definição 2.2.2. Dados dois elementos $u, v \in \mathcal{A}^n$, a **distância de *Hamming*** entre u e v é o número de coordenadas em que u e v diferem, ou seja,

$$d(u, v) = |\{i \mid u_i \neq v_i,\ 1 \leq i \leq n\}|,$$

quando $u = v$ definiremos a distância como

$$d(u, v) = |\varnothing| = 0.$$

Exemplo 2.2.1. Considere $\mathcal{A} = \{0, 1\}$ e $\mathcal{A}^4$, sendo $u = 0111$ e $v = 0010$, temos que por definição a distância é obtida pela quantidade de componentes diferentes entre u e v, logo $d(0111, 0010) = |\{2, 4\}| = 2$ pois os componentes 2 e 4 dos códigos são os únicos diferentes.

Análogamente

$$d(0001, 0101) = |\{2\}| = 1,$$
$$d(0101, 1010) = |\{1, 2, 3, 4\}| = 4,$$

$$d(1100, 0010) = |\{1, 2, 3\}| = 3.$$

Proposição 2.2.1. *Dados $u, v, w \in \mathcal{A}^n$, então valem as seguintes propriedades para a distância de Hamming:*

i) Positividade: $d(u, v) \geq 0$, valendo a igualdade se, e somente se, $u = v$.

ii) Simetria: $d(u, v) = d(v, u)$.

iii) Desigualdade Triangular: $d(u, v) \leq d(u, w) + d(w, v)$.

Demonstração. i) Para $u \neq v$, existe $i \in \mathbb{N}^*$, tal que $u_i \neq v_i$. Desse modo, temos que $d(u, v) = |\{i \mid u_i \neq v_i,\ 1 \leq i \leq n\}|$. Logo, $d(u, v) > 0$.

Para $u = v$, nenhuma componente entre os códigos u e v são diferentes, então por definição $d(u, v) = |\varnothing| = 0$.

Suponhamos que $d(u, v) = 0$, por definição todas as componentes de u e v são iguais, logo $u = v$.

Portanto, $d(u, v) \geq 0$.

ii) Para $u \neq v$,

$$d(u, v) = |\{i \mid u_i \neq v_i,\ 1 \leq i \leq n\}| = |\{i \mid v_i \neq u_i,\ 1 \leq i \leq n\}| = d(v, u).$$

Para $u = v$,

$$d(u, v) = 0 = d(v, u).$$

iii) Sejam u, v e $w \in C$. Note que a contribuição da i-ésima componente de u, v e w será:

Para $u_i \neq v_i$, a contribuição da i-ésima componente de u e v para $d(u, v)$ é 1, nesse caso não podemos ter $u_i = w_i$ e $w_i = v_i$, pois desta forma teríamos $u_i = v_i$, o que contrária a nossa hipótese. Consequentemente a contribuição das i-ésimas componentes de u, v e w para $d(u, w) + d(w, v)$ é maior ou igual a 1, logo por hipótese $d(u_i, v_i) = 1 \leq d(u_i, w_i) + d(w_i, v_i)$.

Para $u_i = v_i$, a contribuição da i-ésima componente de u e v para $d(u, v)$ é 0, logo $d(u_i, v_i) \leq d(u_i, w_i) + d(w_i, v_i)$ pois $d(u_i, w_i) + d(w_i, v_i)$ pode ser igual a 0, 1 ou 2.

Portanto,

$$d(u, v) \leq d(u, w) + d(w, v).$$

■

Provada as três propriedades da Proposição 2.2.1, temos o que chamamos na Matemática de métrica. Por essa razão, a distância de *Hamming* entre elementos de $\mathcal{A}^n$ pode ser chamada de métrica de *Hamming*.

Definição 2.2.3. Dados $a \in \mathcal{A}^n$ e $r > 0$, tal que $r \in \mathbb{R}$, definiremos **disco** e **superfície esférica** de centro a e raio r como sendo os respectivos conjuntos:

$$D(a,r) = \{u \in \mathcal{A}^n \mid d(u,a) \leq r\},$$

$$S(a,r) = \{u \in \mathcal{A}^n \mid d(u,a) = r\}.$$

O Lema a seguir nos fornece as cardinalidades desses conjuntos, lembrando que utilizaremos a notação usual de Análise Combinatória,

$$C_{n,i} = \binom{n}{i} = \frac{n!}{i!(n-i)!}.$$

Lema 2.2.2. *Para todo $a \in \mathcal{A}^n$ e todo número natural $r > 0$, temos que*

$$|D(a,r)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i.$$

Demonstração. Sabemos pela Definição 2.2.3 que $S(a,i) = \{u \in \mathcal{A}^n \mid d(u,a) = i\}$ e $S(a,j) = \{v \in \mathcal{A}^n \mid d(v,a) = j\}$ assim, temos que se $i \neq j$ então $d(u,a) \neq d(v,a)$. Disso podemos concluir que $S(a,i) \cap S(a,j) = \varnothing$.

Como $D(a,r)$ trata-se do disco, temos que a união de todas as superfícies esféricas resultará no disco

$$\bigcup_{i=0}^{r} S(a,i) = D(a,r).$$

Podemos ver que

$$|S(a,i)| = \binom{n}{i}(q-1)^i,$$

pois temos que o número de palavras que diferem i componentes de a em $\mathcal{A}^n$ é $\binom{n}{i}$, e para cada componente temos $q-1$ possíveis escolhas de elementos de $\mathcal{A} - \{0\}$, desse modo, temos $(q-1)^i$ escolhas de componentes não nulas e diferentes de a. Observe que o número de palavras que coincidem com a é $\binom{n}{0}$.

De forma geral

$$\{u \in \mathcal{A}^n \mid d(u,a) = 0\} \Rightarrow |\{u \in \mathcal{A}^n \mid d(u,a) = 0\}| = \binom{n}{0}(q-1)^0 = 1;$$

$$S(a,1) = \{u \in \mathcal{A}^n \mid d(u,a) = 1\} \Rightarrow |S(a,1)| = \binom{n}{1}(q-1)^1 = n(q-1);$$

$$S(a,2) = \{u \in \mathcal{A}^n \mid d(u,a) = 2\} \Rightarrow |S(a,2)| = \binom{n}{2}(q-1)^2;$$

$$\vdots$$

$$S(a,r) = \{u \in \mathcal{A}^r \mid d(u,a) = r\} \Rightarrow |S(a,r)| = \binom{n}{r}(q-1)^r.$$

Portanto, se somarmos todas as cardinalidades das superfícies esféricas obtemos a cardinalidade do disco

$$|D(a,r)| = \sum_{i=0}^{r} |S(a,i)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i.$$

■

Note que $|D(a,r)|$ depende apenas de n, q e r. Agora faremos um exemplo que ilustre essa demonstração.

Exemplo 2.2.2. Seja $\mathcal{A} = \{0,1\}$ e considere $\mathcal{A} = \{000, 001, 010, 100, 011, 101, 110, 111\}^3$, sendo $a = (000)$ e $r = 2$ temos

$$|\{(000)\}| = \binom{3}{0}(2-1)^0 = 1,$$

$$|S(000,1)| = |\{(100),(010),(001)\}| = \binom{3}{1}(2-1)^1 = 3,$$

$$|S(000,2)| = |\{(110),(011),(101)\}| = \binom{3}{2}(2-1)^2 = 3,$$

então

$$|D(000,2)| = \sum_{i=0}^{2} |S(000,i)| = |S(000,0)| + |S(000,1)| + |S(000,2)| = 1+3+3 = 7.$$

Definição 2.2.4. Seja C um código. A **distância mínima** de C é o número

$$d = min\{d(u,v) \mid u, v \in C \ e \ u \neq v\}.$$

Exemplo 2.2.3. Seja C o código do Exemplo 2.1.3, calculemos as distâncias entre todas as palavras de C:

$$d(00000, 01011) = 3,$$
$$d(00000, 10110) = 3,$$
$$d(00000, 11101) = 4,$$
$$d(01011, 10110) = 4,$$
$$d(01011, 11101) = 3,$$
$$d(10110, 11101) = 3.$$

Podemos perceber que a menor das distâncias é 3, então por definição de distância mínima $d = 3$.

Para calcularmos d é necessário fazer a combinação de todos os elementos do código C dois a dois, isto é $\binom{M}{2}$, onde M é a quantidade de palavras de C. Porém considerando nosso código sendo o conjunto das palavras da língua portuguesa $\mathscr{P}$, então nesse caso M é um número relativamente grande, o que torna o custo computacional para calcular d muito elevado. Posteriormente veremos como calcular d de maneira mais econômica em códigos que possuam estruturas algébricas.

Definição 2.2.5. Seja C um código com distância mínima d, chamamos de **capacidade de correção** de C o número

$$\kappa = \left[\frac{d-1}{2}\right],$$

de modo que $[t]$ representa a parte inteira, para qualquer número real t.

Lema 2.2.3. *Seja C um código com distância mínima d. Se u e $v \in C$ e $u \neq v$, então*

$$D(u,\kappa) \cap D(v,\kappa) = \varnothing.$$

Demonstração. Sejam u e $v \in C$. Suponha, por absurdo, que existe $x \in D(u,\kappa) \cap D(v,\kappa)$, logo $x \in D(u,\kappa)$ e $x \in D(v,\kappa)$ então $d(u,x) \leq \kappa$ e $d(v,x) \leq \kappa$.

Por desigualdade triângular, simetria e pela definição de κ temos respectivamente

$$d(u,v) \leq d(u,x) + d(x,v) = d(u,x) + d(v,x) \leq 2\kappa \leq d-1,$$

que é um absurdo, pois pela Definição 2.2.4, $d(u,v) \geq d$, ou seja, a distância mínima d não pode ser maior do que a distância entre u e v.

■

Teorema 2.2.4. *Seja C um código com distância mínima d. Então C pode detectar até $d-1$ erros e corrigir até $\kappa = \left[\frac{d-1}{2}\right]$ erros.*

Demonstração. Suponha que ao transmitirmos uma palavra qualquer u de C ela sofra uma interferência de t erros, tal que $t \leq d-1$, desse modo não conseguimos encontrar outra palavra diferente de u que pertença a C, sendo assim, é possível detectar o erro. Caso contrário, se $t > d-1$ é possível encontrar outra palavra de C diferente de u, o que torna a detecção do erro impossível.

Em contrapartida, se transmitirmos uma palavra u do código C e ocorrer t erros, tal que $t \leq \kappa$, recebemos uma palavra e, então $d(e,u) = t \leq \kappa$, logo $e \in D(u,\kappa)$. Portanto $D(u,\kappa) \cap D(e,\kappa) \neq \varnothing$, pelo Lema 2.2.3 $e \notin C$. Isso nos garante uma correção de até κ erros e determina u univocamente a partir de c.

■

Exemplo 2.2.4. Considere o código $C = \{(000000), (010111), (111001)\}$, temos que $d = 4$, logo é possível corrigir até $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{4-1}{2}\right] = [1, 5] = 1$ erro e detectar $d - 1 = 4 - 1 = 3$ erros sem encontrar outra palavra do código. Se por qualquer motivo ocorra uma interferência, e a palavra recebida seja 000001, nesse caso é possível detectar o erro, pois ela não pertence a C, e ainda é possível corrigi-la para 000000, que é a palavra mais próxima. Agora suponha que recebemos a palavra 000011, note que é possível detectar o erro, mas não é possível corrigi-lo, pois $d(000011, 000000) = d(000011, 010111) = 2$.

Atente que, é essencial calcular ou determinar uma cota inferior para a distância mínima na teoria dos códigos, pois por consequência direta do Teorema 2.2.4 um código tem maior capacidade de correção quanto maior for sua distância mínima.

Definição 2.2.6. Sejam $C \subset \mathcal{A}^n$ um código com distância mínima d e $\kappa = \left[\frac{d-1}{2}\right]$. O código C será **perfeito** se

$$\bigcup_{a \in C} D(a, \kappa) = \mathcal{A}^n.$$

A partir do Teorema 2.2.4 é possível traçar uma estratégia para detectar e corrigir erros.

Sejam C um código com distância mínima d e $\kappa = \left[\frac{d-1}{2}\right]$. Quando um receptor recebe uma palavra x, pode ocorrer as seguintes situações:

i) A palavra x pertence a um disco D de raio κ, onde a é o centro de D e $a \in C$, pela demonstração do Teorema 2.2.4 a é única, diante disso substitui-se x por a.

ii) A palavra x não pertence a nenhum disco D de raio κ, onde a é o centro de D e $a \in C$, nesse caso não é possível decodificar.

Mesmo em i) não temos certeza de que a seja a palavra que pretendia-se transmitir, pois é possível ocorrer mais do que κ erros, se isso acontecer, x se aproxima mais de outra palavra do código C.

O caso ii) não ocorre se nosso código C for perfeito.

Um código $C \subset \mathcal{A}^n$ possui três parâmetros fundamentais $[n, M, d]$, nos quais, n é o comprimento do código, o que significa à dimensão do espaço $\mathcal{A}^n$; M é o número de elementos de C; e d é a distância mínima entre quaisquer dois elementos de C.

Nesse trabalho interessam os códigos nos quais M e d são relativamente grandes em relação a n. Dados três inteiros positivos arbitrários n, M e d, nem sempre existe um código com esses parâmetros, pois existe uma interdependência complexa entre eles, o que nos leva em um dos problemas fundamentais da teoria dos códigos.

2.3 Parâmetros de um Código

Nesta seção temos o objetivo de mostrar algumas relações entre os parâmetros $[n, M, d]$, vistos na seção anterior, de um código corretor de erro, trata-se de uma breve apresentação de algumas Cotas Assintóticas.

Considere nesta seção a quantidade de elementos de um um alfabeto sendo $q \geq 2$.

Teorema 2.3.1. *(Cota de Singleton) Seja C um código com parâmetros $[n, M, d]$, definido sobre um alfabeto $\mathcal{A}$ com q elementos. Temos que*

$$M \leq q^{n-d+1}.$$

Demonstração. Sejam $C \subset \mathcal{A}^n$ com parâmetros $[n, M, d]$, e $u, v \in C$. Considere a projeção

$$\begin{array}{rccc} Pr\Big|_C: & \mathcal{A}^n & \longrightarrow & \mathcal{A}^{n-(d-1)} \\ & (x_1, x_2, \cdots, x_d, x_{d+1}, \cdots, x_n) & \longmapsto & (x_d, x_{d+1}, \cdots, x_n). \end{array}$$

Logo, se

$$\begin{aligned} Pr(u) = Pr(v) &\Rightarrow Pr(u_1, u_2, \cdots, u_d, u_{d+1}, \cdots, u_n) = Pr(v_1, v_2, \cdots, v_d, v_{d+1}, \cdots, v_n) \\ &\Rightarrow \underbrace{(u_d, u_{d+1}, \cdots, u_n)}_{n-(d-1)} = \underbrace{(v_d, v_{d+1}, \cdots, v_n)}_{n-(d-1)}, \end{aligned}$$

percebemos que possuem a mesma quantidade $n - (d-1)$ de componentes, e só valerá a igualdade quando $u_i = v_i$, onde $d \leq i \leq n$.

Disso conclui-se que, se u e v possuem n componentes e $n - (d-1)$ dessas componentes são iguais, é possível que até $d-1$ componentes sejam diferentes, ou seja, $d(u, v) \leq d - 1$, e por definição de distância mínima, segue que $u = v$. Então $Pr(C)$ é injetora.

Como $Pr(C) \subset \mathcal{A}^{n-(d-1)}$, e possui M elementos, sendo assim $M \leq q^{n-d+1}$.

■

Definição 2.3.1. Para todo $n, r \in \mathbb{N}$ e $q \geq 2$, define-se

$$V_q(n, r) = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

Note que trata-se da cardinalidade do disco, já demonstrada na seção anterior no Lema 2.2.2, válido apenas para $q \geq 2$.

Teorema 2.3.2. *(Cota de Hamming) Se C é um código com parâmetros $[n, M, d]$, e se $\kappa = \left[\frac{d-1}{2}\right]$, então*

$$M \leq \frac{q^n}{V_q(n, \kappa)},$$

valendo a igualdade se, e somente se, C for perfeito.

Demonstração. Seja C um código com distância mínima d. Se u e $v \in C$ e $u \neq v$, então pelo Lema 2.2.3

$$D(u,\kappa) \cap D(v,\kappa) = \varnothing \Rightarrow \bigcup_{a \in C} D(a,\kappa) \subseteq \mathcal{A}^n$$

valendo a igualdade se, e somente se, o código for perfeito, disso podemos concluir que $\forall u \in C$

$$\begin{aligned} \sum_{u \in C} |D(u,\kappa)| \leq q^n &\Rightarrow M\left(\sum_{i=0}^{r} \binom{n}{i}(q-1)^i\right) \leq q^n && \text{(pelo Lema 2.2.2)} \\ &\Rightarrow M(V_q(n,r)) \leq q^n && \text{(pela Definição 2.3.1)} \\ &\Rightarrow M \leq \frac{q^n}{V_q(n,r)}, \end{aligned}$$

a existência de C nos garante que $V_q(n,r) \neq 0$.

■

Considere o seguinte número

$$A(n,d) = max\{M \mid \textit{existe um código com parâmetros } [n,M,d]\}.$$

A partir da Cota de *Hamming*, obtemos o resultado a seguir.

Corolário 1. *Para todos números naturais n e d, temos que*

$$A(n,d) \leq \frac{q^n}{V_q(n,r)}.$$

Note ainda que a partir da Cota de *Singleton* obtemos que

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

Definição 2.3.2. Seja C um código de comprimento n e distância mínima d, tal que $|C| = A(n,d)$, chamaremos de **código ótimo**.

Teorema 2.3.3. *(Cota de Gilbert-Varshamov) Para todos os números naturais n e d com $n \geq d$, temos*

$$A(n,d) \geq \frac{q^n}{V_q(n,d-1)}.$$

Demonstração. Seja C um código ótimo de parâmetros $[n,M,d]$, temos pela Definição 2.3.2 que $M = A(n,d)$. Considere, por absurdo, que existe $a' \in \mathcal{A}^n \setminus \bigcup_{a \in C} D(a,d-1)$, $\forall a \in C$, nesse caso teríamos um conjunto $C' = C \cup \{a'\}$ com $M+1$ elementos, o que é absurdo, pois C é um código ótimo. Então

$$\bigcup_{a \in C} D(a,d-1) = \mathcal{A}^n.$$

Portanto,

$$q^n = |\mathcal{A}^n| = |\bigcup_{a \in C} D(a, d-1)| \leq MV_q(n, d-1) = A(n,d)V_q(n, d-1)$$

então

$$A(n,d) \geq \frac{q^n}{V_q(n, d-1)}.$$

■

2.4 Equivalência de Códigos

Sempre que definimos uma classe de objetos matemáticos podemos inserir a noção de equivalência entre esses objetos, com os códigos não será diferente, nessa seção apresentaremos a noção de equivalência de códigos, que basea-se no conceito de isometria, ou seja, satisfaz uma bijeção de um espaço métrico[2] que preserva as distâncias de *Hamming*, como segue a definição.

Definição 2.4.1. Sejam $\mathcal{A}$ um alfabeto e $n \in \mathbb{N}$, dizemos que uma função $F : \mathcal{A}^n \longrightarrow \mathcal{A}^n$ é uma **isometria** de $\mathcal{A}^n$ se preservar as distâncias de *Hamming*, ou seja

$$d(F(x), F(y)) = d(x, y);\ \forall x, y \in \mathcal{A}^n.$$

Proposição 2.4.1. *Toda isometria de $\mathcal{A}^n$ é uma bijeção de $\mathcal{A}^n$.*

Demonstração. Seja $F : \mathcal{A}^n \longrightarrow \mathcal{A}^n$ uma isometria. Suponha que $x, y \in \mathcal{A}^n$ tal que $F(x) = F(y)$, pela Definição 2.2.2 $d(F(x), F(y)) = 0$, e ainda, pela Definição 2.4.1 $d(F(x), F(y)) = d(x, y) = 0$, portanto $x = y$. Desse modo provamos que F é uma aplicação injetora, e toda aplicação injetora de um conjunto finito nele próprio é sobrejetora, logo F é uma bijeção.

■

Proposição 2.4.2.

i) A função identidade de $\mathcal{A}^n$ é uma isometria de $\mathcal{A}^n$.

ii) Se F é uma isometria de $\mathcal{A}^n$, então F^{-1} é uma isometria de $\mathcal{A}^n$.

iii) Se F e G são isometrias de $\mathcal{A}^n$, então $F \circ G$ é uma isometria de $\mathcal{A}^n$.

Demonstração.

i) Seja

[2] Um espaço métrico é um par (M, d) formado por um conjunto M e uma métrica d em M. Uma métrica no conjunto M é definida pela função $d : M \times M \longrightarrow \mathbb{R}$ que satisfaz a Proposição 2.2.1.

$$\begin{array}{rccc} Id\colon & \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ & x & \longmapsto & x. \end{array}$$

Temos que, $d(Id(x), Id(y)) = d(x, y)$, portanto Id é isometria de $\mathcal{A}^n$ pois preserva a distância de *Hamming*.

ii) Como F é isometria por hipótese, pela Proposição 2.4.1 F é uma bijeção, logo admite inversa F^{-1}. Partindo da definição de isometria, $\forall x, y \in \mathcal{A}^n$ tais que

$$\begin{aligned} d(F(x), F(y)) &= d(x, y) \\ &= d(Id(x), Id(y)) \\ &= d(F(F^{-1}(x)), F(F^{-1}(x))) \\ &= d(F^{-1}(x), F^{-1}(y)), \end{aligned}$$

isto é, $d(x, y) = d(F^{-1}(x), F^{-1}(y))$, portanto F^{-1} também é uma isometria de $\mathcal{A}^n$.

iii) Sejam $x, y \in \mathcal{A}^n$, temos que

$$\begin{aligned} d(F(G(x)), F(G(y))) &= d(G(x), G(y)) && \text{(pois } F \text{ é isometria)} \\ &= d(x, y), && \text{(pois } G \text{ é isometria)} \end{aligned}$$

portanto $F \circ G$ é uma isometria de $\mathcal{A}^n$.

■

Definição 2.4.2. Dado dois códigos C e C' em $\mathcal{A}^n$, diremos que C' é **equivalente** a C se existir uma isometria F de $\mathcal{A}^n$ tal que $F(C) = C'$.

Por consequência da Proposição 2.4.2 a equivalência de códigos satisfaz as seguintes propriedades de uma relação de equivalência.

i) Reflexiva: Seja C em $\mathcal{A}^n$, considere

$$\begin{array}{rccc} F\colon & \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ & C & \longmapsto & C, \end{array}$$

pela Proposição 2.4.2 i) F é uma isometria, e pela Definição 2.4.2 C é equivalente C.

ii) Simétrica: Sejam C e C' em $\mathcal{A}^n$, se C' é equivalente a C existe uma isometria

$$\begin{array}{rccc} F\colon & \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ & C & \longmapsto & C', \end{array}$$

tal que $F(C) = C'$, pela Proposição 2.4.2 ii) existe uma isometria F^{-1} tal que $F^{-1}(C') = C$, então pela Definição 2.4.2 C é equivalente a C'.

iii) Transitiva: Sejam C, C' e C'' em $\mathcal{A}^n$, se C'' é equivalente a C' e C' é equivalente a C existem duas isometrias

$$F\colon \begin{array}{ccc} \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ C' & \longmapsto & C'' \end{array} \quad \text{e} \quad G\colon \begin{array}{ccc} \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ C & \longmapsto & C' \end{array}$$

respectivamente. Pela Proposição 2.4.2 iii) $F \circ G$ é uma isometria, então $F(G(C)) = F(C') = C''$, portanto C'' é equivalente a C.

■

Sucede de imediato da Definição 2.4.2 que dois códigos equivalentes possuem os mesmos parâmetros.

Agora iremos apresentar exemplos de famílias de isometrias que ajudarão a demonstrar o principal teorema desta seção.

Exemplo 2.4.1. Se $f : \mathcal{A} \longrightarrow \mathcal{A}$ é uma bijeção, e $i \in \mathbb{Z}$ tal que $1 \leq i \leq n$, a aplicação a seguir é uma isometria.

$$T_f^i\colon \begin{array}{ccc} \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ (a_1, \cdots, a_n) & \longmapsto & (a_1, \cdots, f(a_i), \cdots, a_n). \end{array}$$

De fato, dados $x, y \in \mathcal{A}^n$ quaisquer, por hipótese f é bijeção, então

$$f(x_i) = f(y_i) \Rightarrow x_i = y_i,$$

disso temos que

$$\begin{aligned} d(T_f^i(x), T_f^i(y)) &= d(T_f^i(x_1, \cdots, x_n), T_f^i(y_1, \cdots, y_n)) \\ &= d((x_1, \cdots, f(x_i), \cdots, x_n), (y_1, \cdots, f(y_i), \cdots, y_n)) \\ &= d((x_1, \cdots, x_i, \cdots, x_n), (y_1, \cdots, y_i, \cdots, y_n)) \\ &= d(x, y). \end{aligned}$$

Exemplo 2.4.2. Se π é uma bijeção do conjunto $\{1, \cdots, n\}$ nele próprio, também chamada de permutação de $\{1, \cdots, n\}$, a aplicação permutação de coordenadas a seguir é uma isometria.

$$T_\pi\colon \begin{array}{ccc} \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ (a_1, \cdots, a_n) & \longmapsto & (a_{\pi(1)}, \cdots, a_{\pi(n)}). \end{array}$$

Verificando se é uma isometria. Dados $x, y \in \mathcal{A}^n$ quaisquer, e sejam $i, j \in \{1, 2, \cdots, n\}$, por hipótese π é bijeção, logo as permutações dos subíndices $\pi(1), \pi(2), \cdots, \pi(n)$ nada mais são que uma reordenação de $1, 2, \cdots, n$. Assim, se $x_i \neq y_i$ então $x_{\pi(i)} \neq y_{\pi(i)}$, e se $x_j = y_j$ então $x_{\pi(j)} = y_{\pi(j)}$.

Logo,

$$\begin{aligned} d(T_\pi(x), T_\pi(y)) &= d(T_\pi(x_1, \cdots, x_n), T_\pi(y_1, \cdots, y_n)) \\ &= d((x_{\pi(1)}, \cdots, x_{\pi(n)}), (y_{\pi(1)}, \cdots, y_{\pi(n)})) \\ &= d(x, y). \end{aligned}$$

O principal teorema desta seção será uma consequência direta dos dois lemas que serão apresentados e demonstrados a seguir. A partir de agora considere $\mathcal{A}$ um alfabeto com q elementos.

Lema 2.4.3. *Dada uma isometria F de $\mathcal{A}^n$, com $n \geq 2$, e dados os elementos $a_1, \cdots, a_{n-1} \in \mathcal{A}$, existem $a'_1, \cdots, a'_{n-1} \in \mathcal{A}$, uma bijeção $f_n : \mathcal{A} \longrightarrow \mathcal{A}$ e uma permutação σ de $\{1, \cdots, n\}$ tais que*

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)), \forall x \in \mathcal{A}.$$

Demonstração. Para $q = 1$, temos que $\mathcal{A}^n = \{(a_1, \cdots, a_1)\}$ e $F(a_1, \cdots, a_1) = (a_1, \cdots, a_1)$, então

$$\begin{aligned} (T_\sigma \circ F)(a_1, \cdots, a_1) &= T_\sigma(F(a_1, \cdots, a_1)) \\ &= T_\sigma(a_1, \cdots, a_1) \\ &= (a_{\sigma(1)}, \cdots, a_{\sigma(1)}) \\ &= (a_1, \cdots, a_1). \end{aligned}$$

Portanto é fácil notarmos que nesse caso é válido o lema, pois o resultado independe da permutação e da bijeção f_n, ou seja, $(a_1, \cdots, a_1) = (a'_1, \cdots, a'_1, f_n(a_1))$.

Para $q \geq 2$, considere a_n e $b_n \in \mathcal{A}$ tais que $a_n \neq b_n$, de modo que

$$u = (a_1, \cdots, a_{n-1}, a_n) \text{ e } v = (a_1, \cdots, a_{n-1}, b_n) \in \mathcal{A}^n.$$

Temos que $d(u, v) = 1$ e F é uma isometria, então $d(F(u), F(v)) = d(u, v) = 1$. Portanto, $F(u)$ e $F(v)$ diferem apenas em uma componente.

Escolhendo convenientemente a permutação σ de $\{1, 2, \cdots, n\}$, que a princípio depende de u e v, podemos supor que

$$(T_\sigma \circ F)(u) = (a'_1, \cdots, a'_{n-1}, a'_n),$$

$$(T_\sigma \circ F)(v) = (a'_1, \cdots, a'_{n-1}, b'_n),$$

com $a'_n \neq b'_n$.

No caso de $q = 2$, vamos definir a bijeção f_n por $a_n \longmapsto a'_n$ e $b_n \longmapsto b'_n$, então temos que

$$(T_\sigma \circ F)(u) = (T_\sigma \circ F)(a_1, \cdots, a_{n-1}, a_n) = (a'_1, \cdots, a'_{n-1}, a'_n) = (a'_1, \cdots, a'_{n-1}, f_n(a_n)),$$

$$(T_\sigma \circ F)(v) = (T_\sigma \circ F)(a_1, \cdots, a_{n-1}, b_n) = (a'_1, \cdots, a'_{n-1}, b'_n) = (a'_1, \cdots, a'_{n-1}, f_n(b_n)),$$

logo nesse caso o lema está provado.

Agora para $q > 2$, considere

$$w = (a_1, \cdots, a_{n-1}, x), \forall x \in \mathcal{A} \text{ e } x \neq a_n,$$

pela Proposição 2.4.2 iii) $T_\sigma \circ F$ é uma isometria, então

$$d((T_\sigma \circ F)(w), (T_\sigma \circ F)(u)) = d(w, u) = 1.$$

Dessa forma podemos dizer que existe um único $y \in \mathcal{A}$, tal que

$$(T_\sigma \circ F)(w) = (a'_1, \cdots, a'_{i-1}, y, a'_{i+1}, \cdots, a'_n),$$

onde $y \neq a'_i$. Se $x = b_n$ teríamos que $w = v$, então

$$(T_\sigma \circ F)(w) = (a'_1. \cdots, a'_{n-1}, b'_n),$$

ou seja, $i = n$.

Por absurdo, suponha que $x \neq b_n$ e $i < n$, chegaríamos no seguinte resultado

$$d((T_\sigma \circ F)(w), (T_\sigma \circ F)(u)) = 2,$$

pois $y \neq a'_i$ e $a'_n \neq b'_n$ o que é absurdo, pois $d(w, u) = 1$ e $T_\sigma \circ F$ é isometria. Assim $n = i$.

Consequentemente,

$$(T_\sigma \circ F)(w) = (a'_1, \cdots, a'_{n-1}, y),$$

portanto podemos definir uma função $f_n : \mathcal{A} \longrightarrow \mathcal{A}$, de modo que

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)), \forall x \in \mathcal{A}.$$

Como $T_\sigma \circ F$ é uma isometria, pela Proposição 2.4.1 trata-se de uma bijeção, segue que f_n é injetora, e toda função injetora de um conjunto finito, no caso $\mathcal{A}$, nele próprio é também sobrejetora, portanto f_n é bijetora.

■

Lema 2.4.4. *Dados uma isometria G de $\mathcal{A}^n$ e $a_1, \cdots, a_{n-1}, a'_1, \cdots, a'_{n-1} \in \mathcal{A}$ fixos. Suponhamos que exista uma bijeção $f : \mathcal{A} \longrightarrow \mathcal{A}$, tal que*

$$G(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f(x)), \forall x \in \mathcal{A}.$$

Então, existe uma isometria H de $\mathcal{A}^{n-1}$, tal que

$$G(x_1, \cdots, x_{n-1}, x_n) = (H(x_1, \cdots, x_{n-1}), f(x_n)), \forall (x_1, \cdots, x_n) \in \mathcal{A}^n.$$

Demonstração. Seja $(b_1, \cdots, b_{n-1}) \in \mathcal{A}^{n-1}$, tal que

$$(b_1, \cdots, b_{n-1}) \neq (a_1, \cdots, a_{n-1})$$

e seja $a_n \in \mathcal{A}$. Suponha que

$$u = (a_1, \cdots, a_n) \text{ e } v = (b_1, \cdots, b_{n-1}, a_n).$$

Por hipótese

$$G(u) = (a'_1, \cdots, a'_{n-1}, f(a_n)),$$

agora cosidere

$$G(v) = (c_1, \cdots, c_n).$$

Suponha, por absurdo, que $c_n \neq f(a_n)$, temos que f é bijeção, então existe $b_n \in \mathcal{A}$ tal que $b_n \neq a_n$ e $c_n = f(b_n)$.

Considere

$$w = (a_1, \cdots, a_{n-1}, b_n)$$

logo, aplicando a G temos

$$G(w) = (a'_1, \cdots, a'_{n-1}, f(b_n)).$$

Seja $r \in \mathbb{N}^*$, tal que $d(u, v) = r$. Então

$$\begin{aligned} d(u, v) &= d((a_1, \cdots, a_n), (b_1, \cdots, b_{n-1}, a_n)) \\ &= d((a_1, \cdots, a_{n-1}), (b_1, \cdots, b_{n-1})) \\ &= r. \end{aligned} \tag{2.1}$$

Em contrapartida como $f(a_n) \neq c_n$ temos

$$\begin{aligned} d((a'_1, \cdots, a'_{n-1}), (c_1, \cdots, c_{n-1})) &= d(G(u), G(v)) - 1 \\ &= d(u, v) - 1 \\ &= r - 1. \end{aligned} \tag{2.2}$$

Como $a_n \neq b_n$, temos

$$\begin{aligned} d(w, v) &= d((a_1, \cdots, a_{n-1}, b_n), (b_1, \cdots, b_{n-1}, a_n)) \\ &= d((a_1, \cdots, a_{n-1}), (b_1, \cdots, b_{n-1})) + 1 \\ &= r + 1. \qquad \text{(de 2.1)} \end{aligned}$$

Por outro lado, como $f(b_n) = c_n$, temos que

$$d(G(w), G(v)) = d((a'_1, \cdots, a'_{n-1}, f(b_n)), (c_1, \cdots, c_n))$$

$$= d((a'_1, \cdots, a'_{n-1}), (c_1, \cdots, c_{n-1}))$$
$$= r - 1. \qquad \text{(de 2.2)}$$

O que é absurdo, pois G é uma isometria e $d(G(w), G(v)) \neq d(w, v)$, então $c_n = f(a_n)$.

Com isso, concluímos que para qualquer $(x_1, \cdots, x_n) \in \mathcal{A}^n$, existe $(y_1, \cdots, y_{n-1}) \in \mathcal{A}^{n-1}$, tal que

$$G(x_1, \cdots, x_n) = (y_1, \cdots, y_{n-1}, f(x_n)),$$

logo, $y_1, \cdots, y_{n-1}$ são univocamente determinados por $x_1, \cdots, x_{n-1}$.

Provemos agora a existência de H, é preciso mostrar que $y_1, \cdots, y_{n-1}$ depende apenas de $x_1, \cdots, x_{n-1}$ e não de x_n. Considere $z_n \neq x_n$ e suponha que

$$G(x_1, \cdots, x_{n-1}, z_n) = (y'_1, \cdots, y'_{n-1}, f(z_n)),$$

no entanto, como G é uma isometria, temos

$$d(G(x_1, \cdots, x_{n-1}, x_n), G(x_1, \cdots, x_{n-1}, z_n)) = d((x_1, \cdots, x_{n-1}, x_n), (x_1, \cdots, x_{n-1}, z_n)) = 1.$$

Além disso, G é uma bijeção, se $x_n \neq z_n$ então $f(x_n) \neq f(z_n)$, logo $y'_i = y_i, \forall i = 1, \cdots, n-1$.

Portanto, a função $H : \mathcal{A}^{n-1} \longrightarrow \mathcal{A}^{n-1}$ está bem definida e temos

$$G(x_1, \cdots, x_n) = (H(x_1, \cdots, x_{n-1}), f(x_n)).$$

Provemos que H é uma isometria. Sejam $(x_1, \cdots, x_{n-1}), (x'_1, \cdots, x'_{n-1}) \in \mathcal{A}^{n-1}$ e $x_n \in \mathcal{A}$, temos que

$$\begin{aligned} d((x_1, \cdots, x_{n-1}), (x'_1, \cdots, x'_{n-1})) &= d((x_1, \cdots, x_{n-1}, x_n), (x'_1, \cdots, x'_{n-1}, x_n)) \\ &= d(G(x_1, \cdots, x_{n-1}, x_n), G(x'_1, \cdots, x'_{n-1}, x_n)) \\ &= d((H(x_1, \cdots, x_{n-1}), f(x_n)), (H(x'_1, \cdots, x'_{n-1}), f(x_n))) \\ &= d(H(x_1, \cdots, x_{n-1}), H(x'_1, \cdots, x'_{n-1})). \end{aligned}$$

Portanto, H é uma isometria.

■

Proposição 2.4.5. *Sejam $\mathcal{A}$ um alfabeto e σ, σ' permutações de $\{1, \cdots, n\}$. Então para qualquer $u = (x_1, \cdots, x_n) \in \mathcal{A}^n$, temos*

i) $(T_\sigma \circ T_{\sigma'})(u) = T_{\sigma \circ \sigma'}(u)$

ii) $(T_\sigma)^{-1}(u) = T_{\sigma^{-1}}(u)$

Demonstração. Seja $u = (x_1, \cdots, x_n) \in \mathcal{A}^n$, então:

i) Temos que

$$\begin{aligned}(T_\sigma \circ T_{\sigma'})(u) &= T_\sigma(T_{\sigma'}(x_1, \cdots, x_n)) \\ &= T_\sigma(x_{\sigma'(1)}, \cdots, x_{\sigma'(n)}) \\ &= (x_{\sigma(\sigma'(1))}, \cdots, x_{\sigma(\sigma'(n))}) \\ &= (x_{\sigma\circ\sigma'(1)}, \cdots, x_{\sigma\circ\sigma'(n)}) \\ &= T_{\sigma\circ\sigma'}(u).\end{aligned}$$

ii) Considere

$$\begin{aligned}(T_\sigma \circ T_{\sigma^{-1}})(u) &= T_{\sigma\circ\sigma^{-1}}(u) \qquad\qquad \text{(pelo item } i\text{)} \\ &= T_{id}(u) \\ &= T_{id}(x_1, \cdots, x_n) \\ &= (x_{id(1)}, \cdots, x_{id(n)}) \\ &= (x_1, \cdots, x_n) \\ &= u.\end{aligned}$$

Portanto, $T_\sigma \circ T_{\sigma^{-1}} = Id$. Assim, $T_{\sigma^{-1}}$ é a inversa de T_σ, ou seja, $(T_\sigma)^{-1}(u) = T_{\sigma^{-1}}(u), \forall u \in \mathcal{A}^n$.

■

Teorema 2.4.6. *Seja $F : \mathcal{A}^n \longrightarrow \mathcal{A}^n$ uma isometria, então exitem uma permutação π de $\{1, \cdots, n\}$ e uma bijeção $f_i : \mathcal{A} \longrightarrow \mathcal{A},\ 1 \leq i \leq n$, tais que*

$$F = T_\pi \circ T^1_{f_1} \circ \cdots \circ T^n_{f_n}.$$

Demonstração. Considere $\mathcal{A} = \{a_1, \cdots, a_q\}$.

Por indução sobre n, temos:

Para n=1,

$$d(F(a_i), F(a_j)) = d(a_i, a_j), \forall i, j = 1, \cdots, q.$$

Se $i \neq j$ temos que $d(a_i, a_j) = 1$, logo $F(a_i) \neq F(a_j)$.

Sejam $F(a_i) = a_{i+1}$ e $F(a_j) = a_{j+1}$. Vamos definir f convenientemente por

$$f(a_k) = \begin{cases} a_{k+1}, & \text{se } 1 \leq k \leq q-1 \\ a_1, & \text{se } k = q, \end{cases}$$

se $f(a_i) \neq f(a_j)$ então $a_{i+1} \neq a_{j+1}$, logo f é injetora, toda função injetora de um conjunto finito nele próprio é também sobrejetora, portanto f é bijetora.

Logo, $F = I \circ T_{f_1}$

Agora considere como, hipótese de indução, que o resultado seja válido para $n-1$.

Sejam $a_1, \cdots, a_{n-1} \in \mathcal{A}$. Pelo Lema 2.4.3 existem $a'_1, \cdots, a'_{n-1} \in \mathcal{A}$, uma bijeção $f_n : \mathcal{A} \longrightarrow \mathcal{A}$ e uma permutação σ de $\{1, \cdots, n\}$, tais que

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)), \forall x \in \mathcal{A}.$$

Pelo Lema 2.4.4, existe uma isometria H de $\mathcal{A}^{n-1}$, tal que

$$(T_\sigma \circ F)(x_1, \cdots, x_n) = (H(x_1, \cdots, x_{n-1}), f(x_n)), \forall (x_1, \cdots, x_n) \in \mathcal{A}^n. \tag{2.3}$$

Por hipótese de indução, existem uma permutação τ' de $\{1, \cdots, n-1\}$ e bijeções $f_1, \cdots, f_{n-1}$ de $\mathcal{A}$, tais que

$$H = (T_{\tau'})' \circ (T^1_{f_1})' \circ \cdots \circ (T^{n-1}_{f_{n-1}})', \tag{2.4}$$

onde

$$\begin{array}{rccc} (T_{\tau'})': & \mathcal{A}^{n-1} & \longrightarrow & \mathcal{A}^{n-1} \\ & (x_1, \cdots, x_{n-1}) & \longmapsto & (x_{\tau'(1)}, \cdots, x_{\tau'(n-1)}) \end{array}$$

e, para $i \in \{1, \cdots, n-1\}$

$$\begin{array}{rccc} (T^i_{f_i})': & \mathcal{A}^{n-1} & \longrightarrow & \mathcal{A}^{n-1} \\ & (x_1, \cdots, x_{n-1}) & \longmapsto & (x_1, \cdots, f_i(x_i), \cdots, x_{n-1}). \end{array}$$

Agora definindo a permutação τ de $\{1, \cdots, n\}$ por

$$\tau(i) = \begin{cases} \tau'(i), & \text{se } 1 \leq i \leq n-1 \\ n, & \text{se } i = n \end{cases}$$

e considerando

$$\begin{array}{rccc} T_\tau: & \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ & (x_1, \cdots, x_n) & \longmapsto & (x_{\tau(1)}, \cdots, x_{\tau(n)}) \end{array}$$

e, para $i \in \{1, \cdots, n-1\}$ colocamos

$$\begin{array}{rccc} T^i_{f_i}: & \mathcal{A}^n & \longrightarrow & \mathcal{A}^n \\ & (x_1, \cdots, x_n) & \longmapsto & (x_1, \cdots, f_i(x_i), \cdots, x_n). \end{array}$$

De (2.3) e (2.4) concluímos que

$$T_\sigma \circ F = T_\tau \circ T^1_{f_1} \circ \cdots \circ T^n_{f_n}.$$

A partir da Proposição 2.4.5 e considerando $\pi = \sigma^{-1} \circ \tau$, obtemos que

$$\begin{aligned} T_\sigma \circ F = T_\tau \circ T_{f_1}^1 \circ \cdots \circ T_{f_n}^n &\Rightarrow T_{\sigma^{-1}} \circ T_\sigma \circ F = T_{\sigma^{-1}} \circ T_\tau \circ T_{f_1}^1 \circ \cdots \circ T_{f_n}^n \\ &\Rightarrow F = T_{\sigma^{-1} \circ \tau} \circ T_{f_1}^1 \circ \cdots \circ T_{f_n}^n \\ &\Rightarrow F = T_\pi \circ T_{f_1}^1 \circ \cdots \circ T_{f_n}^n, \end{aligned}$$

o que nos garante o resultado.

■

3 ANÉIS E CORPOS

Para que seja possível avançarmos nos estudos da teoria dos códigos corretores de erros, é preciso recordar algumas estruturas algébricas básicas. Neste capítulo será apresentado conceitos de anéis e corpos com ênfase no anel dos números inteiros, onde veremos a divisibilidade em $\mathbb{Z}$ e o conceito de inteiros módulo m, também estudaremos a mudança de alfabeto de um código. Para composição deste capítulo foram utilizados os seguintes materiais: Notas de aula Álgebra - ALGM5, 2º Semestre de 2018; Domingues e Iezzi (2018); Milies e Coelho (2013); Gonçalves (2017); Garcia e Lequain (2018); Tarcha (2019) e Hefez e Villela (2008).

3.1 Anéis

Definição 3.1.1. Seja A um conjunto não vazio onde estejam definidas duas operações, as quais chamaremos de **adição** e **multiplicação** em A e denotaremos por $+$ e $\cdot$ respectivamente.

Assim,

$$+: \begin{array}{ccc} A \times A & \longrightarrow & A \\ (a,b) & \longmapsto & a+b \end{array} \quad \text{e} \quad \cdot : \begin{array}{ccc} A \times A & \longrightarrow & A \\ (a,b) & \longmapsto & a \cdot b \end{array}$$

É dito que A é um **anel** se as propriedades a seguir são satisfeitas para quaisquer $a, b, c \in A$.

(A1) **Associativa da adição:** $a + (b + c) = (a + b) + c$.

(A2) **Comutativa da adição:** $a + b = b + a$.

(A3) **Existência do elemento neutro aditivo:** $\exists\ 0_A \in A$, tal que $a + 0_A = a$.

(A4) **Existência de simétricos aditivos:** $\forall x \in A$, $\exists\ y \in A$, denotado por $y = -x$, tal que $x + y = 0_A$.

(M1) **Associativa da multiplicação:** $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(D1) **Distributiva da multiplicação em relação a adição:** $a \cdot (b + c) = a \cdot b + a \cdot c$ e $(a + b) \cdot c = a \cdot c + b \cdot c$.

Utilizaremos a notação $(A, +, \cdot)$ para representar o anel A munido das operações de adição e subtração respectivamete

Exemplo 3.1.1. Os conjuntos dos números inteiros $\mathbb{Z}$, racionais $\mathbb{Q}$, reais $\mathbb{R}$ e complexos $\mathbb{C}$, com as operações de adição e multiplicação usuais, cujas propriedades cumprem os axiomas da definição acima, são anéis. Assim, podemos representá-los como os anéis numéricos $(\mathbb{Z}, +, \cdot), (\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ e $(\mathbb{C}, +, \cdot)$ respectivamente.

Decorre de imediato, do fato de que a adição é uma operação sobre A, as seguintes propriedades.

Propriedade 3.1.1. Seja $(A, +, \cdot)$ um anel.

i) O elemento neutro 0_A é único. Esse elemento é chamado de **zero do anel**.

ii) O simétrico $-a$ de um elemento $a \in A$ é único.

iii) Se $a_1, a_2, \cdots, a_n \in A$ então $-(a_1 + a_2 + \cdots + a_n) = (-a_1) + (-a_2) + \cdots + (-a_n)$.

iv) Se $a \in A$ então $-(-a) = a$.

v) Se $a + x = a + y$ então $x = y$, ou seja, todo elemento de A é regular para a adição. Em outras palavras, vale a **lei do cancelamento da adição**.

vi) A equação $a + x = b$ tem uma e uma só solução: o elemento $b + (-a)$.

vii) Se $a \in A$ então $a \cdot 0 = 0 \cdot a = 0$.

viii) Se $a, b \in A$ então $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$.

ix) Se $a, b \in A$ então $(-a) \cdot (-b) = a \cdot b$.

Demonstração. Consultar Tarcha (2019, p. 22-24).

■

Definição 3.1.2. Sejam $a, b \in A$. Chama-se **diferença** entre a e b, denotado por $a - b$ o elemento $a + (-b) \in A$. Portanto, $a - b = a + (-b)$.

Da definição acima, segue a seguinte propriedade.

Propriedade 3.1.2. Seja $(A, +, \cdot)$ um anel. Se $a, b, c \in A$ então $a \cdot (b - c) = a \cdot b - a \cdot c$ e $(a - b) \cdot c = a \cdot c - b \cdot c$.

Demonstração. Consultar Domingues e Iezzi (2018, p. 224-225).

■

Definição 3.1.3. Um anel $(A, +, \cdot)$ em que o conjunto A é finito chama-se **anel finito**.

Definição 3.1.4. Sejam $(A, +, \cdot)$ um anel e B um subconjunto não vazio de A. Seja B fechado para as operações de $+$ e $\cdot$ de A, ou seja,

i) $x, y \in B \Rightarrow x + y \in B$;

ii) $x, y \in B \Rightarrow x \cdot y \in B$.

Assim, podemos considerar a adição e a multiplicação de A como operações de B.

Se $(B, +, \cdot)$ for um anel com as operações de A dizemos que B é um **subanel** de A.

Definição 3.1.5. Seja $(A, +, \cdot)$ um anel. Se a multiplicação de A goza da propriedade comutativa, ou seja, se

$$a \cdot b = b \cdot a,$$

para $a, b \in A$ quaisquer, diz-se que A é um **anel comutativo**.

Definição 3.1.6. Seja $(A, +, \cdot)$ um anel. Se A possui elemento neutro multiplicativo, ou seja, se existe um elemento $1_A \in A$, $1_A \cdot a = a$, tal que

$$a \cdot 1_A = 1_A \cdot a = a,$$

para $a \in A$ qualquer, diz-se que 1_A é a **unidade** de A e que A é um **anel com unidade**.

Definição 3.1.7. (Potências num anel) Seja $(A, +, \cdot)$ um anel com unidade. Se $a \in A$ e $n \in \mathbb{N}$, define-se a^n, **potência n-ésima de** a, por recorrência da seguinte maneira

$$a^0 = 1_A \quad \text{e} \quad a^{n+1} = a^n \cdot a,$$

sempre que $n \geq 0$.

Proposição 3.1.1. *Seja $(A, +, \cdot)$ um anel com unidade. Se $a \in A$ e $m, n \in \mathbb{N}$ então*

i) $a^m \cdot a^n = a^{m+n}$;

ii) $(a^m)^n = a^{m \cdot n}$.

Demonstração. Consultar Domingues e Iezzi (2018, p. 232). ■

Definição 3.1.8. Seja $(A, +, \cdot)$ um anel comutativo com unidade. Se para esse anel vale a **lei do anulamento do produto**, ou seja, se uma igualdade do tipo

$$a \cdot b = 0_A,$$

para $a, b \in A$, só for possível se

$$a = 0_A \quad \text{e} \quad b = 0_A$$

então diz-se que A é um **anel de integridade** ou **domínio de integridade**.

3.2 Corpos

Definição 3.2.1. Um conjunto não vazio $\mathbb{K}$ é chamado de **corpo** se $\mathbb{K}$ é um domínio de integridade, tal que para qualquer elemento de $\mathbb{K}$, diferente do elemento nulo é inversível em relação a multiplicação, ou seja,

$$\forall\ x \in \mathbb{K}^*, \exists\ x^{-1} \in \mathbb{K} \text{ tal que } x \cdot x^{-1} = 1_{\mathbb{K}}.$$

Exemplo 3.2.1. Os conjuntos numéricos $\mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$ são alguns dos exemplos mais comuns de corpos.

Proposição 3.2.1. *Todo anel de integridade finito é um corpo.*

Demonstração. Consultar Domingues e Iezzi (2018, p. 263-237).

■

Definição 3.2.2. Seja $(\mathbb{K}, +, \cdot)$ um corpo. Um subconjunto não vazio $L \subset \mathbb{K}$ é chamado **subcorpo** de $\mathbb{K}$ se é fechado para a adição e multiplicação de $\mathbb{K}$ e se L também tem uma estrutura de corpo, para as operações de $\mathbb{K}$, restritas aos elementos de L.

Exemplo 3.2.2. O conjunto numérico $\mathbb{Q}$ é um subcorpo de $\mathbb{R}$ que, por sua vez, é um subcorpo de $\mathbb{C}$.

Proposição 3.2.2. *Sejam $\mathbb{K}$ um corpo e L um subconjunto não vazio de $\mathbb{K}$. Para que L seja um subcorpo de $\mathbb{K}$ é necessário e suficiente que*

i) $0_{\mathbb{K}}, 1_{\mathbb{K}} \in L$;

ii) Se $x, y \in L$ *então* $x - y \in L$;

iii) Se $x, y \in L$ *e* $y \neq 0_{\mathbb{K}}$ *então* $x \cdot y^{-1} \in L$.

Demonstração. Consultar Tarcha (2019, p. 30).

■

3.3 Divisibilidade em $\mathbb{Z}$ e Inteiros módulo m

Definição 3.3.1. Sejam $a, b \in \mathbb{Z}$. Diz-se que b divide a se existe um inteiro c tal que $b \cdot c = a$. Denotaremos por $b \mid a$. A negação desta afirmação será indicada por $b \nmid a$.

Exemplo 3.3.1. Sabemos que $3 \mid 12$, pois existe um inteiro c, tal que $3 \cdot c = 12$. De fato, $3 \cdot 4 = 12$.

Por outro lado, $3 \nmid 11$, pois não existe um inteiro c, tal que $3 \cdot c = 11$.

Definição 3.3.2. Um inteiro p diz-se primo se tem exatamente dois divisores positivos: 1 e $|p|$.

Exemplo 3.3.2. Note que, 11 é um número primo. De fato, 11 possui apenas dois divisores positivos: 1 e 11.

Definição 3.3.3. Seja $m \neq 0$ um inteiro fixo. Dois inteiros a e b dizem-se congruentes módulos m, se m divide a diferença $a - b$.

Exemplo 3.3.3. Seja $m = 2$. Os inteiros 15 e 7 são congruentes módulo 2, pois existe um inteiro c, tal que 2 divide a diferença $15 - 7$. De fato,

$$2 \cdot c = 15 - 7 \Rightarrow 2 \cdot c = 8 \Rightarrow c = 4.$$

Assim, pela Definição 3.3.1, temos que $2 \mid (15 - 7)$.

Definição 3.3.4. Sejam $a, m \in \mathbb{Z}$, tal que $m > 1$. Chama-se **classe de congruência de a módulo m** o conjunto formado por todos os inteiros que são congruentes a a módulo m. Denotaremos esse conjunto por $\bar{a}$. Em outras palavras

$$\bar{a} = \{x \in \mathbb{Z} \mid x \equiv a \pmod{m}\}.$$

Como $x \equiv a \pmod{m}$ se, e somente se, x é da forma $x = a + tm$, para algum $t \in \mathbb{Z}$. Assim, podemos escrever o conjunto $\bar{a}$ como

$$\bar{a} = \{a + tm \mid t \in \mathbb{Z}\}.$$

Agora vamos construir as classes residuais de $\mathbb{Z}$ módulo m.

Note que, as classes dos inteiros $0, 1, \cdots, m - 1$ são

$$\begin{aligned}
\bar{0} &= \{0, \pm m, \pm 2m, \cdots\} \\
\bar{1} &= \{1, 1 \pm m, 1 \pm 2m, \cdots\} \\
\bar{2} &= \{2, 2 \pm m, 2 \pm 2m, \cdots\} \\
&\vdots \\
\overline{m-1} &= \{m - 1, m - 1 \pm m, m - 1 \pm 2m, \cdots\}
\end{aligned}$$

Exemplo 3.3.4. Se $m = 3$, temos que todas as classes possíveis módulo 3 são

$$\begin{aligned}
\bar{0} &= \{0, 3, -3, 6, -6, \cdots\} \\
\bar{1} &= \{1, 4, -4, 7, -7, \cdots\} \\
\bar{2} &= \{2, 5, -5, 8, -8, \cdots\}.
\end{aligned}$$

Definição 3.3.5. O conjunto das classes de congruência módulo m, denotado por $\mathbb{Z}_m$, é chamado de conjunto dos **inteiros módulo** m, ou ainda,

$$\mathbb{Z}_m = \{\bar{0}, \bar{1}, \cdots, \overline{m-1}\}.$$

Dadas duas classes $\bar{a}, \bar{b} \in \mathbb{Z}_m$. Vamos definir as operações de adição e multiplicação sobre $\mathbb{Z}_m$, respectivamente, por

$$\bar{a} + \bar{b} = \overline{a+b} \qquad \text{e} \qquad \bar{a} \cdot \bar{b} = \overline{a \cdot b}.$$

Decorre da definição acima e dos axiomas para as operações com números inteiros as propriedades A1, A2, A3, A4, A5, M1 e D1 vistas na Seção 3.1, e ainda, $\mathbb{Z}_m$ possui unidade e a operação de multiplicação é comutativa.

Para as demonstrações destas propriedades consultar Domingues e Iezzi (2018, p. 137-138).

Segue destas propriedades que, $\mathbb{Z}_m$ é um anel comutativo e com unidade.

Proposição 3.3.1. *Um anel de classes de restos $\mathbb{Z}_m$ é um domínio de integridade se, e somente se, m é um número primo.*

Demonstração. Consultar Domingues e Iezzi (2018, p. 235). ■

3.4 Mudança de Alfabeto

É possível mudar o alfabeto $\mathcal{A}$ de um código C por outro, caso possua a mesma quantidade q de elementos e mantendo os mesmos parâmetros de C.

Sejam $\mathcal{A}$ e $\mathcal{B}$ dois conjuntos finitos, e a função $f : \mathcal{A} \longrightarrow \mathcal{B}$ uma bijeção. A partir de f podemos definir a função

$$\begin{array}{rccc} \phi\colon & \mathcal{A}^n & \longrightarrow & \mathcal{B}^n \\ & (x_1, \cdots, x_n) & \longmapsto & (f(x_1), \cdots, f(x_n)). \end{array}$$

Vamos verificar que ϕ é uma bijeção e preserva a distância de *Hamming*.

Sejam $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n) \in \mathcal{A}^n$, como f é bijeção temos

$$\begin{aligned} d(\phi(x), \phi(y)) &= d(\phi(x_1, \cdots, x_n), \phi(y_1, \cdots, y_n)) \\ &= d((f(x_1), \cdots, f(x_n)), (f(y_1), \cdots, f(y_n))) \\ &= d((x_1, \cdots, x_n), (y_1, \cdots, y_n)) \\ &= d(x, y). \end{aligned}$$

Logo, ϕ preserva a distância de *Hamming*, ou seja, ϕ é uma isometria, e pela Proposição 2.4.1, ϕ é uma bijeção.

■

Seja $C \subset \mathcal{A}^n$ um código com M elementos e distância mínima d, a sua imagem $C' = \phi(C) \subset \mathcal{B}^n$ é um código sobre o alfabeto $\mathcal{B}$ com os mesmos parâmetros de C.

Desse modo, dado um código C sobre um alfabeto $\mathcal{A}$ com m elementos, por meio de uma bijeção $f : \mathcal{A} \longrightarrow \mathbb{Z}_m$ podemos obter um código C' sobre o anel $\mathbb{Z}_m$ com os mesmos parâmetros de C.

Definição 3.4.1. Diremos que $\omega(u)$ é o **peso** do elemento $u \in \mathbb{Z}_m$, se

$$\omega(u) = |\{i \mid u_i \neq 0\}|,$$

ou seja, o número de coordenadas não nulas de u. Em outras palavras, temos que

$$\omega(u) = d(u, 0),$$

onde d é a distância de *Hamming* de C.

Se o código C' é fechado para a subtração, ou melhor, se

$$\forall u, v \in C', u - v \in C',$$

então vale a seguinte igualdade para a distância mínima d de C'

$$d = min\{\omega(u) \mid u \in C', u \neq 0\}.$$

Observe que em $(\mathbb{Z}_m)^n$ existe uma métrica diferente da métrica de *Hamming*, que é chamada de métrica de *Lee*[1], comumente utilizada em telefonia celular.

[1] Essa métrica foi definida por Lee (1958). Diferentemente da formulação que Hamming faz, onde as palavras de um código normalmente são **binárias**, ou seja, são constituídas apenas de zeros e uns, e a distância entre um par de palavras deste código é o número de coordenadas não correspondentes. Lee (1958) define esta métrica para códigos não-binários.

4 PRELIMINARES ALGÉBRICOS

Neste Capítulo vamos trabalhar com corpos finitos, estudar algumas de suas propriedades, e aprender a classificá-los. De modo geral, a teoria dos códigos corretores de erros baseia-se na álgebra linear sobre corpos finitos, por esse motivo também apresentaremos algumas noções básicas de álgebra linear. Para a composição deste capítulo foram utilizados os seguintes materiais: Notas de aula Álgebra Linear - Curso de Verão, IME-USP[1]; Coelho e Lourenço (2015); Callioli, Domingues e Costa (1990); Domingues e Iezzi (2018); McCoy (1973) Hefez e Villela (2008) Steinbruch e Winterle (1987); Boldrini et al. (1980); Hoffman e Kunze (1970) e Santos (2010).

4.1 Noções Básicas de Álgebra Linear

Nesta seção, vamos apresentar apenas as noções básicas de álgebra linear necessárias para o desenvolvimento de códigos lineares. Lembrando que, o conteúdo apresentado é um breve resumo a título de recordação.

Definição 4.1.1. Um conjunto não vazio V é um **espaço vetorial** sobre um corpo $\mathbb{K}$ qualquer, se em seus elementos, denominados **vetores**, estiverem definidas as duas operações a seguir:

(A) **Adição em** V

A cada par v, w de vetores de V corresponde um vetor $v + w$ pertencente a V, chamado soma de v e w, ou seja,

$$\begin{array}{rccc} +: & V \times V & \longrightarrow & V \\ & (v, w) & \longmapsto & v + w, \end{array}$$

com as seguintes propriedades:

(A1) **Comutativa:** $u + v = v + u,\ \forall u, v \in V$;

(A2) **Associativa:** $u + (v + w) = (u + v) + w,\ \forall u, v, w \in V$;

(A3) **Existência do Neutro:** Existe em V um vetor, denominado **vetor nulo** e denotado por 0, tal que $u + 0 = u,\ \forall u \in V$;

(A4) **Existência do Simétrico:** Para todo vetor $u \in V$, existe um vetor em V, denotado por $-u$, tal que $u + (-u) = 0$.

[1] Instituto de Matemática e Estatística - Universidade de São Paulo.

(M) **Multiplicação por um escalar em V**

A cada par $\alpha \in \mathbb{K}$ e $u \in V$, corresponde um vetor $\alpha \cdot u$ pertencente a V, chamado produto por escalar de α por u, ou seja,

$$\begin{array}{rccc} \cdot : & \mathbb{K} \times V & \longrightarrow & V \\ & (\alpha, u) & \longmapsto & \alpha \cdot u, \end{array}$$

com as seguintes propriendades:

(M1) **Associativa:** $\alpha(\beta \cdot u) = (\alpha\beta) \cdot u, \ \forall \alpha, \beta \in \mathbb{K}$ e $\forall u \in V$;

(M2) **Existência do Neutro:** $1 \cdot u = u, \ \forall u \in V$, onde 1 é o elemento neutro multiplicativo do corpo $\mathbb{K}$.

Além disso, vamos impor que as duas operações acima se distribuam, ou seja, que sejam válidas as seguintes propriedades:

(D1) **Distributiva da multiplicação de um escalar em relação à adição de vetores:** $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v, \ \forall \alpha \in \mathbb{K}$ e $\forall u, v \in V$;

(D2) **Distributiva da adição de escalares em relação a multiplicação de um vetor:** $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u, \ \forall \alpha, \beta \in \mathbb{K}$ e $\forall u \in V$.

Exemplo 4.1.1. Todo corpo é um espaço vetorial sobre si mesmo. Assim, são exemplos de espaços vetoriais os corpos $\mathbb{R}$ e $\mathbb{C}$, visto que as propriedades acima apresentadas são satisfeitas nesses corpos.

Podemos, desta forma, observar que os conjuntos a seguir são espaços vetoriais: $M_{m\times n}(\mathbb{R})$ sobre $\mathbb{R}$, $\mathbb{R}^n$ sobre $\mathbb{R}$, $\mathbb{C}^n$ sobre $\mathbb{C}$, e o conjunto dos polinômios $\mathbb{K}[X]$ sobre um corpo $\mathbb{K}$ qualquer. Esses são alguns exemplos dos espaços vetoriais mais estudados em álgebra linear.

Definição 4.1.2. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer. Um subconjunto W de V é um **subespaço vetorial** de V se a restrição das operações de V a W torna esse conjunto um espaço vetorial sobre o corpo $\mathbb{K}$.

Proposição 4.1.1. *Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer e $W \subseteq V$ um subconjunto não vazio. Então W é um subespaço de V se, e somente se, satisfaz as seguintes propriedades:*

i) $0 \in W$;

ii) Se $u, v \in W$ *então* $u + v \in W$;

iii) Se $\alpha \in \mathbb{K}$ *e* $u \in W$ *então* $\alpha \cdot u \in W$.

Demonstração. Consultar Santos (2010, p. 14-15).

■

Exemplo 4.1.2. $W = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0\}$ é subespaço vetorial de $V = \mathbb{R}^3$. De fato, verificando as condições da Proposição 4.1.1, temos que

i) $0_V = (0, 0, 0) \in W$, se $y = z = 0$;

ii) Se $u = (0, y_1, z_1), v = (0, y_2, z_2) \in W$ então

$$u + v = (0, y_1, z_1) + (0, y_2, z_2) = (0, y_1 + y_2, z_1 + z_2) \in W;$$

iii) Se $\alpha \in \mathbb{R}$ e $u = (0, y, z) \in W$ então

$$\alpha \cdot u = \alpha \cdot (0, y, z) = (0, \alpha \cdot y, \alpha \cdot z) \in W.$$

Portanto, W é um subespaço vetorial de $\mathbb{R}^3$.

Definição 4.1.3. Sejam U e V subespaços vetoriais de um espaço vetorial W sobre um corpo $\mathbb{K}$ qualquer. Indicaremos por $U + V$ e chamaremos de **soma** de U com V o seguinte subconjunto de W:

$$U + V = \{u + v \mid u \in U \text{ e } v \in V\}.$$

Proposição 4.1.2. *Se U e V são subespaços vetoriais de um espaço vetorial W sobre um corpo $\mathbb{K}$ qualquer então $U + V$ também é um subespaço vetorial de W.*

Demonstração. Consultar Callioli, Domingues e Costa (1990, p. 56).

■

Exemplo 4.1.3. Sejam $U = \{(x, 0, 0)\}$ e $V = \{(0, 0, z)\}$ subespaços vetoriais de um espaço vetorial $\mathbb{R}^3$ sobre $\mathbb{R}$. Temos que, a soma de U com V é

$$U + V = \{u + v \mid u \in U \text{ e } v \in V\} = \{(x, 0, 0) + (0, 0, z)\} = \{(x, 0, z)\}.$$

Além disso, $U + V$ é subespaço vetorial de $\mathbb{R}^3$. De fato, verificando as condições da Proposição 4.1.1, temos que

i) $0_{\mathbb{R}^3} = (0, 0, 0) \in U + V$, se $x = z = 0$;

ii) Se $u' = (x_1, 0, z_1),\ v' = (x_2, 0, z_2) \in U + V$ então

$$u' + v' = (x_1, 0, z_1) + (x_2, 0, z_2) = (x_1 + x_2, 0, z_1 + z_2) \in U + V;$$

iii) Se $\alpha \in \mathbb{R}$ e $u' = (x, 0, z) \in U + V$ então

$$\alpha \cdot u' = \alpha \cdot (x, 0, z) = (\alpha \cdot x, 0, \alpha \cdot z) \in U + V.$$

Portanto, $U + V$ é subespaço vetorial de $\mathbb{R}^3$.

Definição 4.1.4. Sejam U e V subespaços vetoriais de um espaço vetorial W sobre um corpo $\mathbb{K}$ qualquer, tais que $U \cap V = \{0_W\}$. Neste caso, é dito que $U + V$ é **soma direta** dos subespaços U e V, denotado por $U \oplus V$.

Proposição 4.1.3. *Sejam U e V subespaços vetoriais de um espaço vetorial W sobre um corpo $\mathbb{K}$ qualquer. Então $U \oplus V = W$ se, e somente se, cada vetor $w \in W$ admite uma única decomposição $w = u + v$, com $u \in U$ e $v \in V$.*

Demonstração. Consultar Callioli, Domingues e Costa (1990, p. 56-57).

■

Exemplo 4.1.4. O espaço vetorial $\mathbb{R}^2$ sobre $\mathbb{R}$ é uma soma direta dos subespaços $U = \{(x, 0) \mid x \in \mathbb{R}\}$ e $V = \{(0, y) \mid y \in \mathbb{R}\}$ de $\mathbb{R}^2$, pois $U \cap V = \{(0, 0)\}$. Além disso, cada vetor $(x, y) \in \mathbb{R}^2$ admite uma única decomposição $(x, y) = (x, 0) + (0, y)$, com $(x, 0) \in U$ e $(0, y) \in V$.

Definição 4.1.5. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer. Dizemos que um conjunto $L = \{u_1, u_2, \cdots, u_n\} \subset V$ é **linearmente independente (L.I.)** se, e somente se, uma igualdade

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0,$$

com os $\alpha_i \in \mathbb{K}$, só for possível para $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, chamaremos essa solução de **trivial**.

Exemplo 4.1.5. No espaço vetorial $V = \mathbb{R}^4$ sobre $\mathbb{R}$, o conjunto $L = \{(2, 5, 0, 1), (0, 0, 2, 0)\}$ é linearmente independente, pois existem α_1, $\alpha_2 \in \mathbb{R}$ tais que

$$\alpha_1 \cdot (2, 5, 0, 1) + \alpha_2 \cdot (0, 0, 2, 0) = (0, 0, 0, 0) \Rightarrow \begin{cases} 2\alpha_1 = 0 \\ 5\alpha_1 = 0 \\ 2\alpha_2 = 0 \\ \alpha_1 = 0 \end{cases} \Rightarrow \alpha_1 = \alpha_2 = 0.$$

Note que, o sistema admite apenas a solução trivial. Portanto, L é linearmente independente.

Definição 4.1.6. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer. Dizemos que um conjunto $L = \{u_1, u_2, \cdots, u_n\} \subset V$ é **linearmente dependente (L.D.)** se, e somente se, L não é L.I., isto é, seja possível uma igualdade

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0,$$

sem que os escalares $\alpha_i \in \mathbb{K}$ sejam todos iguais a 0.

Exemplo 4.1.6. No espaço vetorial $V = \mathbb{R}^2$ sobre $\mathbb{R}$, o conjunto $L = \{(2, -1), (4, -2)\}$ é linearmente dependente, pois existem $\alpha_1,\ \alpha_2 \in \mathbb{R}$ tais que

$$\alpha_1 \cdot (2, -1) + \alpha_2 \cdot (4, -2) = (0, 0) \Rightarrow \begin{cases} 2\alpha_1 + 4\alpha_2 = 0 \\ -\alpha_1 - 2\alpha_2 = 0 \end{cases} \Rightarrow \alpha_1 = -2\alpha_2.$$

Note que, α_1 depende de α_2, desse modo, o sistema admite outras soluções além da trivial. Portanto, L é linearmente dependente.

Definição 4.1.7. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer.

i) Um vetor $u \in V$ é uma **combinação linear** dos vetores $u_1, \cdots, u_n \in V$ se existirem escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$ tais que

$$u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = \sum_{i=1}^{n} \alpha_i u_i.$$

ii) Seja B um subconjunto de V. B é dito **gerador** de V, ou ainda, B **gera** V se todo elemento de V for uma combinação linear de um número finito de elementos de B.

Exemplo 4.1.7. Considere $\mathbb{R}^4$ como espaço vetorial sobre $\mathbb{R}$. Observe que, o conjunto $B = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ é um conjunto gerador de $\mathbb{R}^4$, pois podemos escrever qualquer vetor $(a, b, c, d) \in \mathbb{R}^4$ como combinação linear dos vetores de B, ou seja, existem $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ tais que

$$(a, b, c, d) = \alpha_1(1, 0, 0, 0) + \alpha_2(0, 1, 0, 0) + \alpha_3(0, 0, 1, 0) + \alpha_4(0, 0, 0, 1),$$

então $\alpha_1 = a,\ \alpha_2 = b,\ \alpha_3 = c,\ \alpha_2 = d$ com $a, b, c, d \in \mathbb{R}$.

Definição 4.1.8. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer. Dizemos que um subconjunto B de V é uma **base** de V se B for gerador de V e B for linearmente independente.

Exemplo 4.1.8. Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Observe que, o conjunto $B = \{(1, -1), (1, 0)\}$ é base de $\mathbb{R}^2$, pois

i) B é um conjunto gerador de $\mathbb{R}^2$, visto que podemos escrever qualquer vetor $(x, y) \in \mathbb{R}^2$ como combinação linear dos vetores de B, ou seja, existem $\alpha_1, \alpha_2 \in \mathbb{R}$ tais que

$$(x, y) = \alpha_1(1, -1) + \alpha_2(1, 0) \Rightarrow \begin{cases} \alpha_1 + \alpha_2 = x \\ -\alpha_1 \quad\ = y \end{cases} \Rightarrow \alpha_1 = -y \text{ e } \alpha_2 = x + y,$$

com $x, y \in \mathbb{R}$.

ii) B é um conjunto linearmente independente, uma vez que existem $\beta_1, \beta_2 \in \mathbb{R}$ tais que

$$\beta_1(1, -1) + \beta_2(1, 0) = (0, 0) \Rightarrow \begin{cases} \beta_1 + \beta_2 = 0 \\ -\beta_1 \quad\ = 0 \end{cases} \Rightarrow \beta_1 = \beta_2 = 0.$$

Exemplo 4.1.9. Considere $\mathbb{K}^n$ um espaço vetorial sobre um corpo $\mathbb{K}$ qualquer. Seja B o subconjunto constituído dos vetores

$$\begin{aligned}\varepsilon_1 &= (1, 0, 0, \cdots, 0)\\ \varepsilon_2 &= (0, 1, 0, \cdots, 0)\\ &\vdots\\ \varepsilon_n &= (0, 0, 0 \cdots, 1).\end{aligned}$$

Sejam $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$ e coloquemos $u = \alpha_1\varepsilon_1 + \alpha_2\varepsilon_2 + \cdots + \alpha_n\varepsilon_n \in \mathbb{K}^n$. Então

$$u = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Isto posto, temos que B é conjunto gerador de $\mathbb{K}^n$. Note que, $u = 0_{\mathbb{K}^n}$ se, e somente se $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Desse modo, os vetores de B são linearmente independentes.

Definição 4.1.9. O conjunto $B = \{(1, 0, \cdots, 0), (0, 1, 0, \cdots, 0), \cdots, (0, 0, \cdots, 1)\}$ construído acima é uma base do espaço vetorial $\mathbb{K}^n$ sobre $\mathbb{K}$. Denominaremos esta base particular a **base canônica** de $\mathbb{K}^n$.

Definição 4.1.10. Dizemos que um espaço vetorial V sobre um corpo $\mathbb{K}$ qualquer é **finitamente gerado** se possuir um conjunto gerador finito.

Exemplo 4.1.10. Considere $\mathbb{R}^3$ como um espaço vetorial sobre $\mathbb{R}$. Observe que, $\mathbb{R}^3$ é um espaço vetorial finitamente gerado, pois o conjunto $B = \{(-1, 0, 0), (0, 1, 0), (0, 0, -1)\}$ é gerador de $\mathbb{R}^3$ e B é um conjunto finito de três vetores.

Teorema 4.1.4. *Seja V um espaço vetorial finitamente gerado não nulo sobre um corpo $\mathbb{K}$ qualquer. Então duas bases quaisquer de V têm o mesmo número de vetores.*

Demonstração. Consultar Coelho e Lourenço (2015, p. 52).

■

Exemplo 4.1.11. Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Já vimos que, o conjunto $B = \{(1, -1), (1, 0)\}$ é uma base de $\mathbb{R}^2$. É fácil verificar que o conjunto $B' = \{(1, 0), (0, 1)\}$ também é uma base de $\mathbb{R}^2$. Note que, as duas bases de $\mathbb{R}^2$ têm o mesmo número de vetores.

Definição 4.1.11. Seja V um espaço vetorial finitamente gerado. Denomina-se **dimensão** de V, denotado por $dim\ V$, o número de vetores de uma base qualquer de V.

Exemplo 4.1.12. Considere $M_2(\mathbb{R})$ como um espaço vetorial sobre $\mathbb{R}$. Observe que, o conjunto $B = \left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}$ é conjunto gerador de $M_2(\mathbb{R})$, pois se $\begin{bmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{bmatrix} \in M_2(\mathbb{R})$ então

$$\begin{bmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{bmatrix} = \alpha_1 \cdot \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \alpha_2 \cdot \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} + \alpha_3 \cdot \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} + \alpha_4 \cdot \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix},$$

para $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$.

É fácil ver que, os vetores de B são linearmente independentes, portanto B é uma base de $M_2(\mathbb{R})$. B possui 4 vetores, logo $dim\ M_2(\mathbb{R}) = 4$, e como B é finito, $M_2(\mathbb{R})$ é um espaço vetorial finitamente gerado.

Corolário 2. *Seja V um espaço vetorial de dimensão $n \geq 1$ e B um subconjunto de V com n elementos. As seguintes afirmações são equivalentes:*

i) B é uma base de V.

ii) B é linearmente independente.

iii) B é um conjunto gerador de V.

Demonstração. Consultar Santos (2010, p. 75-76).

■

Definição 4.1.12. Sejam U e V espaços vetoriais sobre um corpo $\mathbb{K}$ qualquer. Uma aplicação $T : U \longrightarrow V$ é chamada **transformação linear** de U em V se, satisfaz as seguintes condições:

i) $T(u_1 + u_2) = T(u_1) + T(u_2),\ \forall u_1, u_2 \in U$;

ii) $T(\alpha \cdot u) = \alpha \cdot T(u),\ \forall \alpha \in \mathbb{K}$ e $\forall u \in U$.

No caso em que $U = V$, uma transformação linear de U em V é chamada de **operador linear**.

Exemplo 4.1.13. Considere $\mathbb{R}^3$ um espaço vetorial sobre $\mathbb{R}$ e a aplicação T definida por

$$\begin{array}{rccc} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x, y, z) & \longmapsto & (y, 2x + z, -z), \end{array}$$

para $x, y, z \in \mathbb{R}$.

Verifiquemos as condições para T ser uma transformação linear.

i) Sejam $u_1 = (x_1, y_1, z_1), u_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$ quaisquer, temos

$$\begin{aligned} T(u_1 + u_2) &= T((x_1, y_1, z_1) + (x_2, y_2, z_2)) \\ &= T(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (y_1 + y_2, 2(x_1 + x_2) + (z_1 + z_2), -(z_1 + z_2)) \\ &= (y_1 + y_2, (2x_1 + z_1) + (2x_2 + z_2), -z_1 + (-z_2)) \end{aligned}$$

$$\begin{aligned}
&= (y_1, 2x_1 + z_1, -z_1) + (y_2, 2x_2 + z_2, -z_2) \\
&= T((x_1, y_1, z_1)) + T((x_2, y_2, z_2)) \\
&= T(u_1) + T(u_2).
\end{aligned}$$

ii) Sejam $u = (x, y, z) \in \mathbb{R}^3$ e $\alpha \in \mathbb{R}$ quaisquer, temos

$$\begin{aligned}
T(\alpha \cdot u) &= T(\alpha \cdot (x, y, z)) \\
&= T(\alpha x, \alpha y, \alpha z) \\
&= (\alpha y, 2(\alpha x) + \alpha z, -\alpha z) \\
&= (\alpha y, \alpha(2x + z), \alpha(-z)) \\
&= \alpha \cdot (y, 2x + z, -z) \\
&= \alpha \cdot T(u).
\end{aligned}$$

Portanto, T é uma transformação linear de $\mathbb{R}^3$ em $\mathbb{R}^3$.

Definição 4.1.13. Sejam U e V espaços vetoriais sobre um corpo $\mathbb{K}$ qualquer e $T : U \longrightarrow V$ uma transformação linear.

i) O **núcleo** de T, denotado por $Ker\ T$, é o conjunto definido por

$$Ker\ T = \{u \in U \mid T(u) = 0\};$$

ii) A **imagem** de T, denotado por $Im\ T$, é o conjunto definido por

$$Im\ T = \{v \in V \mid \exists u \in U, \text{ com } T(u) = v\}.$$

Exemplo 4.1.14. Seja T a transformação linear definida por

$$\begin{array}{rccc}
T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^4 \\
 & (x, y) & \longmapsto & (0, 0, 0, x - y),
\end{array}$$

para $x, y \in \mathbb{R}$.

Achemos o núcleo de T, temos

$$(x, y) \in Ker\ T \Leftrightarrow (0, 0, 0, x - y) = (0, 0, 0, 0) \Leftrightarrow x = y.$$

Portanto, $Ker\ T = \{(x, x) \mid x \in \mathbb{R}\}$.

Achemos a imagem de T, temos $(a, b, c, d) \in Im\ T$, para $a, b, c, d \in \mathbb{R}$ se existe $(x, y) \in \mathbb{R}^2$ tal que

$$(0, 0, 0, x - y) = (a, b, c, d),$$

somente tem solução se $a = b = c = 0$ e $d = x - y$.

Portanto, $Im\ T = \{(a, b, c, d) \in \mathbb{R}^4 \mid a = b = c = 0 \text{ e } d = x - y\}$.

Proposição 4.1.5. *Sejam U e V dois espaços vetoriais sobre um corpo $\mathbb{K}$ qualquer e $T : U \longrightarrow V$ uma transformação linear. Então $Ker\ T$ é um subespaço vetorial de U e $Im\ T$ é um subespaço vetorial de V.*

Demonstração. Consultar Santos (2010, p. 221-222).

■

Proposição 4.1.6. *Sejam U e V dois espaços vetoriais sobre um corpo $\mathbb{K}$ qualquer e $T : U \longrightarrow V$ uma transformação linear. Então T é injetora se, e somente se, $Ker\ T = \{0\}$.*

Demonstração. Consultar Coelho e Lourenço (2015, p. 85).

■

Teorema 4.1.7. *Sejam U e V espaços vetoriais de dimensão finita sobre um corpo $\mathbb{K}$ qualquer e $T : U \longrightarrow V$ uma transformação linear. Então*

$$dim\ U = dim\ Ker(T) + dim\ Im(T).$$

Demonstração. Consultar Callioli, Domingues e Costa (1990, p. 112-113).

■

Podemos representar uma transformação linear por uma matriz. Sejam U e V espaços vetoriais de dimensão n e m, respectivamente, sobre um corpo $\mathbb{K}$ qualquer. Consideremos uma transformação linear $T : U \longrightarrow V$. Dadas as bases $B = \{u_1, \cdots, u_n\}$ de U e $B' = \{v_1, \cdots, v_m\}$ de V, temos que

$$\begin{aligned} T(u_1) &= \alpha_{11}v_1 + \alpha_{21}v_2 + \cdots + \alpha_{m1}v_m \\ T(u_2) &= \alpha_{12}v_1 + \alpha_{22}v_2 + \cdots + \alpha_{m2}v_m \\ &\vdots \\ T(u_n) &= \alpha_{1n}v_1 + \alpha_{2n}v_2 + \cdots + \alpha_{mn}v_m \end{aligned}$$

logo,

$$T(u_j) = \sum_{i=1}^{m} \alpha_{ij}v_i, j \in \{1, 2, \cdots, n\}$$

onde os $\alpha_{ij} \in \mathbb{K}$ e são univocamente determinados.

Definição 4.1.14. A matriz de ordem $m \times n$ sobre $\mathbb{K}$ é

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix},$$

que se obtém das considerações feitas acima é chamada **matriz** de T em relação as bases B e B', denotado por $[T]_{B,B'}$.

Exemplo 4.1.15. Seja T a transformação linear definida por

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & (x,y) & \longmapsto & (3x+y, x+3y, 0), \end{array}$$

para $x, y \in \mathbb{R}$ e considere as bases $B = \{(1,-1),(1,0)\}$ e $B' = \{(1,0,0),(0,1,0),(0,0,1)\}$ de $\mathbb{R}^2$ e $\mathbb{R}^3$ respectivamente. Com uma conta simples, temos

$$\begin{aligned} T((1,-1)) &= (2,-2,0) = 2(1,0,0) + (-2)(0,1,0) + 0(0,0,1) \\ T((1,0)) &= (3,1,0) = 3(1,0,0) + 1(0,1,0) + 0(0,0,1). \end{aligned}$$

Assim,

$$[T]_{B,B'} = \begin{bmatrix} 2 & 3 \\ -2 & 1 \\ 0 & 0 \end{bmatrix}.$$

Definição 4.1.15. Seja $M \in M_{m\times n}(\mathbb{K})$. O **posto** de M é o número máximo de linhas linearmente independentes de M.

Exemplo 4.1.16. Considere a matriz $A \in M_{4\times 3}(\mathbb{R})$, tal que

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Observe que, o número máximos de linhas linearmente independentes de A é 3. Portanto, o posto da matriz A é igual a 3.

Definição 4.1.16. Dada uma matriz $A = (a_{ij}) \in M_{m\times n}(\mathbb{K})$ denomina-se **transposta** de A e indica-se por A^t a matriz de ordem $n \times m$, tal que $A^t = (b_{ji})$, onde $b_{ji} = a_{ij}$, para $i \in \{1, \cdots, m\}$ e $j \in \{1, \cdots, n\}$.

Exemplo 4.1.17. Condisere a matriz $A \in M_{2\times 2}(\mathbb{R})$, tal que

$$A = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}.$$

Agora, trocando ordenadamente as linhas por colunas de A obtemos a matriz transposta de A. Portanto,

$$A^t = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix}.$$

Proposição 4.1.8. *Dadas as matrizes* $A = (a_{ij}), B = (c_{jk}) \in M_{m\times n}(\mathbb{K})$. *Vale a seguinte relação:*

$$(AB)^t = B^t A^t.$$

Demonstração. Consultar Callioli, Domingues e Costa (1990, p. 23).

■

Exemplo 4.1.18. Sejam $\mathbb{K}$ um corpo qualquer. O domínio de integridade $\mathbb{K}[X]$, é um espaço vetorial sobre $\mathbb{K}$.

De fato, seja $\mathbb{K}[X]_{n-1}$, com $n \in \mathbb{N}$, definido por

$$\mathbb{K}[X]_{n-1} = \{p(x) \in \mathbb{K}[X] \mid gr(p(x)) \leq n-1\} \cup \{0\}.$$

Note que, $\mathbb{K}[X]_{n-1} \subset \mathbb{K}[X]$ e as restrições das operações de $\mathbb{K}[X]$ a $\mathbb{K}[X]_{n-1}$ torna $\mathbb{K}[X]_{n-1}$ um espaço vetorial sobre $\mathbb{K}$, desse modo, $\mathbb{K}[X]_{n-1}$ é um subespaço vetorial de $\mathbb{K}[X]$ sobre $\mathbb{K}$ de dimensão n, com base $\{1, x, \cdots, x^{n-1}\}$.

De fato, os polinômios da base $\{1, x, \cdots, x^{n-1}\}$ são linearmente independentes sobre $\mathbb{K}$, dado que

$$p(x) = \alpha_1 1 + \alpha_2 x + \cdots + \alpha_n x^{n-1} = 0,$$

só é possível se $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, com os $\alpha_i \in \mathbb{K}$.

Portanto, temos um conjunto linearmente independente, cujo número de elementos é igual à dimensão de $\mathbb{K}[X]_{n-1}$, logo é uma base desse espaço vetorial.

4.2 Homomorfismos de Anéis

Definição 4.2.1. Sejam $(A, +, \cdot)$ e $(B, +, \cdot)$ dois anéis. Uma aplicação $f : A \longrightarrow B$ diz-se um **homomorfismo** de A em B se satisfaz as seguintes condições:

i) $f(x + y) = f(x) + f(y),\ \forall x, y \in A$;

ii) $f(x \cdot y) = f(x) \cdot f(y),\ \forall x, y \in A$.

Se a aplicação f for injetora chamaremos de **monomorfismo**. E, se for uma aplicação sobrejetora de **epimorfismo**. Caso a aplicação seja bijetora é chamado de **isomorfismo**.

Agora vamos verificar algumas propriedades fundamentais sobre homomorfismo.

Proposição 4.2.1. *Sejam* $(A, +, \cdot)$ *e* $(B, +, \cdot)$ *dois anéis com unidade e* $f : A \longrightarrow B$ *um homomorfismo. Temos que*

i) $f(0_A) = 0_B$;

ii) $f(-x) = -f(x),\ \forall x \in A$;

iii) $f(x-y) = f(x) - f(y),\ \forall x, y \in A$;

iv) Se f *é um epimorfismo então* $f(1_A) = 1_B$.

v) Se $x \in A$ *é invertível então* $f(x)$ *também é invertível e* $[f(x)]^{-1} = f(x^{-1})$;

vi) Se f *é um isomorfismo então a aplicação inversa* f^{-1} *de* f *é um homomorfismo;*

vii) Se A *e* B *são corpos então* f *é injetora e* $f(A)$ *é um subcorpo de* B.

Demonstração. Sejam A e B dois anéis com unidade.

i) Pela Definição 4.2.1 i), temos

$$f(0_A) = f(0_A + 0_A) = f(0_A) + f(0_A),$$

logo, somando o simétrico aditivo de $f(0_A)$ em ambos os lados da igualdade acima, obtemos

$$\begin{aligned} f(0_A) = f(0_A) + f(0_A) &\Rightarrow f(0_A) + (-f(0_A)) = [f(0_A) + f(0_A)] + (-f(0_A)) \\ &\Rightarrow 0_B = f(0_A). \end{aligned}$$

ii) Seja $x \in A$ qualquer. Temos que $x + (-x) = 0_A$. Assim, pela Definição 4.2.1 i) e pelo item i), temos

$$0_B = f(0_A) = f(x + (-x)) = f(x) + f(-x),$$

logo, somando o simétrico aditivo de $f(x)$ em ambos os lados da igualdade acima, obtemos

$$\begin{aligned} 0_B = f(x) + f(-x) &\Rightarrow 0_B + (-f(x)) = [f(x) + f(-x)] + (-f(x)) \\ &\Rightarrow -f(x) = f(-x). \end{aligned}$$

iii) Sejam $x, y \in A$ quaisquer. Temos que $x - y = x + (-y)$. Assim, pela Definição 4.2.1 i) e pelo item ii), temos

$$f(x-y) = f(x + (-y)) = f(x) + f(-y) = f(x) - f(y).$$

iv) Seja $y \in B$ qualquer. Como f é um epimorfismo, f é sobrejetora, desse modo existe $x \in A$ tal que $f(x) = y$. Assim, pela Definição 4.2.1 ii), temos

$$y \cdot f(1_A) = f(x) \cdot f(1_A) = f(x \cdot 1_A) = f(x) = y.$$

De modo análogo,

$$f(1_A) \cdot y = y.$$

Portanto, $f(1_A)$ é a unidade de B e denotamos por 1_B, ou seja, $f(1_A) = 1_B$.

v) Se $x \in A$ e x é invertível então $x \cdot x^{-1} = 1_A = x^{-1} \cdot x$. Assim, pela Definições 4.2.1 ii) e pelo item iv), temos

$$1_B = f(1_A) = f(x \cdot x^{-1}) = f(x) \cdot f(x^{-1}).$$

De modo análogo,

$$1_B = f(x^{-1}) \cdot f(x).$$

Assim, $f(x^{-1})$ é o inverso multiplicativo de $f(x)$. Portanto, podemos escreve-lo como $[f(x)]^{-1}$, ou seja, $f(x^{-1}) = [f(x)]^{-1}$.

vi) Seja f um isomorfismo. Dados $z, t \in B$ quaisquer. Como f é isomorfismo, f é bijetora, desse modo existem $x, y \in A$ tais que $z = f(x)$ e $t = f(y)$. Note que, pelo fato de f ser bijetora garante que exista $f^{-1} : B \longrightarrow A$ que também é uma aplicação bijetora. Observe que, as igualdades acima equivalem respectivamente a $f^{-1}(z) = x$ e $f^{-1}(t) = y$. Verificando as propriedades da Definição 4.2.1 para f^{-1}, temos

vi.i) $f^{-1}(z + t) = f^{-1}(f(x) + f(y)) = f^{-1}(f(x + y)) = x + y = f^{-1}(z) + f^{-1}(t)$;

vi.ii) $f^{-1}(z \cdot t) = f^{-1}(f(x) \cdot f(y)) = f^{-1}(f(x \cdot y)) = x \cdot y = f^{-1}(z) \cdot f^{-1}(t)$.

Portanto, a aplicação inversa f^{-1} de f é um homomorfismo.

vii) Sejam A e B dois corpos. Dados $x, y \in A$. Se $f(x) = f(y)$, do item iii), temos que $f(x - y) = f(x) - f(y) = 0_B$. Se $x \neq y$ então $x - y$ seria invertível, do item v), $f(x - y)$ seria invertível, consequentemente não nulo, o que é absurdo. Assim, $x = y$ e, portanto, f é um monomorfismo.

Para verificar que $f(A)$ é um subcorpo de B, suponhamos que $z, t \in f(A)$ com $t \neq 0$, então existem $x, y \in A$ tais que $z = f(x)$ e $t = f(y)$, e assim temos que

Por i),

$$f(0_A) = 0_B \in f(A) \text{ e } f(1_A) = 1_B \in f(A);$$

Por iii),

$$z - t = f(x) - f(y) = f(x - y) \in f(A);$$

Por v),

$$z \cdot t^{-1} = f(x) \cdot [f(y)]^{-1} = f(x) \cdot f(y^{-1}) = f(x \cdot y^{-1}) \in f(A).$$

Portanto, f é injetora e $f(A)$ é um subcorpo de B.

■

4.3 Múltiplos de um elemento de um Anel

Definição 4.3.1. Seja $(A, +, \cdot)$ um anel. Se $m \in \mathbb{Z}$ e $a \in A$ então $m \cdot a$ é assim definido:

i) Se $m = 0$,
$$m \cdot a = 0 \cdot a = 0_A;$$

ii) Se $m \geq 1$,
$$m \cdot a = (m-1) \cdot a + a;$$

iii) Se $m < 0$,
$$m \cdot a = (-m) \cdot (-a).$$

O elemento $m \cdot a$ é dito **múltiplo m-ésimo** de a.

Decorre desta definição as seguintes propriedades.

Propriedade 4.3.1. Sejam $a \in A$ qualquer e $m, n \in \mathbb{Z}$ quaisquer, temos que

i) $m \cdot a + n \cdot a = (m+n) \cdot a$;

ii) $(-m) \cdot a = -(m \cdot a)$;

iii) $n \cdot (m \cdot a) = (nm) \cdot a$.

Proposição 4.3.1. *Seja A um anel com unidade. Se 1_A indica a unidade e $m, n \in \mathbb{Z}$, então $(mn) \cdot 1_A = (m \cdot 1_A)(n \cdot 1_A)$.*

Demonstração. Consultar Domingues e Iezzi (2018, p. 261).

■

Corolário 3. *Seja A um anel com unidade. Então o conjunto $\mathbb{Z} \cdot 1_A = \{m \cdot 1_A \mid m \in \mathbb{Z}\}$ é um subanel unitário de A.*

Demonstração. Consultar Domingues e Iezzi (2018, p. 261).

■

4.4 Característica de um Corpo

Definição 4.4.1. Um domínio de integridade A é dito de **característica** 0 se $m = 0$ sempre que $m \cdot a = 0_A$ com $a \in A$, $a \neq 0_A$.

A diz-se de **característica finita** se existe $a \in A$, $a \neq 0_A$, tal que $m \cdot a = 0_A$ para algum inteiro $m \neq 0$. Nesse caso, definimos como a **característica** de A o menor inteiro positivo m, tal que $m \cdot a = 0_A$ para algum $a \in A$, $a \neq 0_A$.

Proposição 4.4.1. *Seja A um anel com unidade. Então a característica de A é um inteiro positivo $h > 0$ se, e somente se, h é o menor inteiro estritamente positivo tal que $h \cdot 1_A = 0_A$.*

Demonstração. ($\Rightarrow$) Por hipótese, $h > 0$ é característica de A, ou seja, $h \cdot a = 0_A$ para todo $a \in A$, em particular, $h \cdot 1_A = 0_A$.

Suponha, por absurdo, que existe $m \in \mathbb{Z}$, tal que $0 < m < h$ e $m \cdot 1_A = 0_A$. Seja $a \in A$, logo

$$\begin{aligned} m \cdot a &= \underbrace{a + a + \cdots + a}_{\text{m vezes}} \\ &= \underbrace{a1_A + a1_A + \cdots + a1_A}_{\text{m vezes}} \\ &= a\underbrace{(1_A + 1_A \cdots + 1_A)}_{\text{m vezes}} \\ &= a(m \cdot 1_A) \\ &= a(0_A) \\ &= 0_A, \end{aligned}$$

o que é absurdo, pois a característica de A é h.

($\Leftarrow$) Por hipótese h é o menor inteiro estritamente positivo tal que $h \cdot 1_A = 0_A$. Para todo $a \in A$, temos que

$$\begin{aligned} h \cdot a &= \underbrace{a + a + \cdots + a}_{\text{h vezes}} \\ &= \underbrace{a1_A + a1_A + \cdots + a1_A}_{\text{h vezes}} \\ &= a\underbrace{(1_A + 1_A \cdots + 1_A)}_{\text{h vezes}} \\ &= a(h \cdot 1_A) \\ &= a(0_A) \\ &= 0_A. \end{aligned}$$

Se existisse algum inteiro m tal que $0 < m < h$ e $m \cdot a = 0_A$ para todo $a \in A$ então $m \cdot 1_A = 0_A$, o que é absurdo segundo a hipótese.

■

Proposição 4.4.2. *Se a característica de um domínio de integridade A não é zero então é um número primo.*

Demonstração. Suponha, por absurdo, que a característica do domínio de integridade A é o inteiro h, tal que $h > 0$ e h não é um número primo.

Assim, $h = st$, onde $s, t \in \mathbb{Z}_+^*$ tais que $1 < s, t < h$. Como h é característica de A, pelas Proposições 4.4.1 e 4.3.1, temos

$$h \cdot 1_A = 0_A \Rightarrow (st) \cdot 1_A = 0_A \Rightarrow (s \cdot 1_A)(t \cdot 1_A) = 0_A.$$

Como A é domínio de integridade, temos que $(s \cdot 1_A) = 0_A$ ou $(t \cdot 1_A) = 0_A$, o que é absurdo, pois por hipótese h é a característica de A. Portanto, h é um número primo.

■

Proposição 4.4.3. *Seja A um anel com unidade finito. Então a característica de A é maior do que zero.*

Demonstração. Seja A um anel com unidade finito com n elementos. Consideremos a sequência

$$1_A,\ 2 \cdot 1_A,\ 3 \cdot 1_A, \cdots .$$

Note que, os elementos dessa sequência pertencem ao anel com unidade finito A. Logo, podemos afirmar que existem dois elementos $s \cdot 1_A, t \cdot 1_A \in A$, tais que $s > t$, de modo que

$$s \cdot 1_A = t \cdot 1_A \Rightarrow s \cdot 1_A - t \cdot 1_A = 0_A \Rightarrow (s - t) \cdot 1_A = 0_A,$$

para $s - t > 0$.

Observe que, o menor inteiro estritamente positivo h, tal que $h \cdot 1_A = 0_A$ é a característica do anel com unidade finito A.

■

Proposição 4.4.4. *Seja A um anel com unidade cuja característica é um número primo p, então a correspondência que associa a cada $\overline{s} \in \mathbb{Z}_p$ o elemento $\overline{s} \cdot 1_A \in \mathbb{Z} \cdot 1_A$ é um isomorfismo de anéis.*

Demonstração. Considere a aplicação f definida da seguinte forma

$$\begin{array}{rccc} f: & \mathbb{Z}_p & \longrightarrow & \mathbb{Z} \cdot 1_A \\ & \overline{s} & \longmapsto & s \cdot 1_A, \end{array}$$

para todo $\overline{s} \in \mathbb{Z}_p$.

Primeiro vamos mostrar que f é uma aplicação bem definida. Sejam $\overline{s}, \overline{t} \in \mathbb{Z}_p$, suponhamos que

$$\overline{s} = \overline{t} \Leftrightarrow p \mid (s-t) \Leftrightarrow s - t = ph,$$

para algum $h \in \mathbb{Z}$, pelo algoritmo da divisão.

Logo,

$$(s-t) \cdot 1_A = (ph) \cdot 1_A = (p \cdot 1_A)(h \cdot 1_A) = 0_A(h \cdot 1_A) = 0_A,$$

pela Proposição 4.3.1, temos

$$0_A = (s-t) \cdot 1_A = s \cdot 1_A - t \cdot 1_A \Rightarrow t \cdot 1_A = s \cdot 1_A.$$

Portanto, f está bem definida.

Agora vamos mostrar que f é um isomorfismo. Verificando as condições de homomorfismo, temos:

i) Sejam $\overline{s}, \overline{t} \in \mathbb{Z}_p$.

$$\begin{aligned} f(\overline{s} + \overline{t}) &= f(\overline{s+t}) \\ &= (s+t)1_A \\ &= s \cdot 1_A + t \cdot 1_A \\ &= f(\overline{s}) + f(\overline{t}). \end{aligned}$$

ii) Sejam $\overline{s}, \overline{t} \in \mathbb{Z}_p$.

$$\begin{aligned} f(\overline{s}\overline{t}) &= f(\overline{st}) \\ &= (st) \cdot 1_A \\ &= (s \cdot 1_A)(t \cdot 1_A) \\ &= f(\overline{s})f(\overline{t}). \end{aligned}$$

Portanto, f é um homomorfismo.

Verificando se f é injetora. Para todo $\overline{s_1}, \overline{s_2} \in \mathbb{Z}_p$, temos que

$$\begin{aligned} f(\overline{s_1}) = f(\overline{s_2}) &\Rightarrow s_1 \cdot 1_A = s_2 \cdot 1_A \\ &\Rightarrow s_1 \cdot 1_A - s_2 \cdot 1_A = 0_A \\ &\Rightarrow (s_1 - s_2) \cdot 1_A = 0_A. \end{aligned}$$

Como p é a característica de A, existe $h \in \mathbb{Z}$, tal que

$$s_1 - s_2 = ph \Rightarrow p \mid (s_1 - s_2) \Rightarrow s_1 \equiv s_2 \pmod{p} \Rightarrow \overline{s_1} = \overline{s_2}.$$

Portanto, f é monomorfismo.

Verificando se f é sobrejetora. Toda aplicação injetora de um conjunto finito nele próprio é sobrejetora, logo f é uma bijeção. Portanto, f é um isomorfismo de anéis.

■

Observação 1. Seja $\mathbb{K}$ um corpo. Uma vez que todo corpo é um anel de integridade, segue da Proposição 4.4.2 que a característica de $\mathbb{K}$ é um número primo p ou $0_{\mathbb{K}}$.

Definição 4.4.2. Seja $\mathbb{F}$ um corpo com quantidade finita de elementos, diremos que $\mathbb{F}$ é um **corpo finito**.

Teorema 4.4.5. *Um corpo finito $\mathbb{F}$ tem característica prima p.*

Demonstração. Segue da Observação 1 e da Proposição 4.4.3 que um corpo finito $\mathbb{F}$ possui característica maior do que zero. Assim, pela Proposição 4.4.2, temos que a característica de $\mathbb{F}$ é um número primo.

■

Proposição 4.4.6. *Seja $\mathbb{F}$ um corpo finito com característica p. Se para $m \in \mathbb{Z}$ e $a \in \mathbb{F}$ tem-se $m \cdot a = 0_{\mathbb{F}}$ então m é múltiplo de p ou $a = 0_{\mathbb{F}}$.*

Demonstração. Suponha que para todo $a \in \mathbb{F}$ e $m \in \mathbb{Z}$, temos que $m \cdot a = 0_{\mathbb{F}}$ então $(m \cdot 1_{\mathbb{F}})a = 0_{\mathbb{F}}$. Como $\mathbb{F}$ é um corpo, temos que $m \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$ ou $a = 0_{\mathbb{F}}$.

Se $a = 0_{\mathbb{F}}$ a proposição está demonstrada.

No caso de $m \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$. Como $m \in \mathbb{Z}$ e p é a característica de $\mathbb{F}$, pelo algoritmo da divisão sempre existem $q, r \in \mathbb{Z}_+$, tais que $m = qp + r$, com $0 \leq r < p$.

Logo,

$$0_{\mathbb{F}} = m \cdot 1_{\mathbb{F}} = (qp + r) \cdot 1_{\mathbb{F}} = (qp) \cdot 1_{\mathbb{F}} + r \cdot 1_{\mathbb{F}} = q(p \cdot 1_{\mathbb{F}}) + r \cdot 1_{\mathbb{F}} = q \cdot 0_{\mathbb{F}} + r \cdot 1_{\mathbb{F}} = r \cdot 1_{\mathbb{F}},$$

pela Proposição 4.4.1, p é o menor inteiro estritamente positivo, tal que $p \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$, consequentemente $r = 0$.

Portanto, p divide m, ou ainda, m é múltiplo de p.

■

Teorema 4.4.7. *Seja $\mathbb{F}$ um corpo finito. Então $\mathbb{F}$ tem p^n elementos, onde o primo p é a característica de $\mathbb{F}$ e n é o grau de $\mathbb{F}$ sobre seu subcorpo primo.*

Demonstração. Como $\mathbb{F}$ é um corpo finito pelo Teorema 4.4.5, temos que sua característica é um primo p. Além disso, pela Proposição 4.4.4 o subcorpo primo L de $\mathbb{F}$ é isomorfo a $\mathbb{Z}_p$ e portanto contém p elementos.

Note que, $\mathbb{F}$ é um espaço vetorial sobre L. Como $\mathbb{F}$ é um corpo finito, segue que tem dimensão finita sobre L. Seja $B = \{u_1, u_2, \cdots, u_n\}$ uma base de $\mathbb{F}$ sobre L, para algum número natural n, temos que $dim\ \mathbb{F} = n$.

Assim, cada vetor de $\mathbb{F}$ é escrito de maneira única na forma

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n,$$

com os $\alpha_i \in L$ e $i \in \{1, 2, \cdots, n\}$.

Como existem p possibilidades de valores para cada coeficiente α_i e, cada uma delas nos dá um elemento diferente em L, segue que $\mathbb{F}$ tem exatamente p^n elementos, ou seja, $|\mathbb{F}| = p^n$.

■

5 CÓDIGOS LINEARES

A proposta deste trabalho é estudar os códigos lineares. Neste capítulo vamos defini-los, determinar seus parâmetros e apresentar algumas aplicações dessa classe de códigos. Para a composição deste capítulo foram utilizados os seguintes materiais: Hefez e Villela (2008); Luchetta (2005); MacWillians e Sloane (1981); Dias (2005) e Roman (1992).

5.1 Códigos Lineares

Denotaremos por $\mathbb{F}$ um corpo finito com q elementos tomado como alfabeto. Temos, portanto, para cada número natural n, um espaço vetorial $\mathbb{F}^n$ sobre $\mathbb{F}$ de dimensão n.

Definição 5.1.1. Um código $C \subset \mathbb{F}^n$ será chamado de **código linear** se for um subespaço vetorial de $\mathbb{F}^n$.

Seja u um vetor do código linear C, u é dito **código nulo** se todas as suas componentes forem zero, denotado por 0_C.

Exemplo 5.1.1. Vamos mostrar que, no exemplo da peça sobre um tabuleiro quadriculado, descrito no Exemplo 2.1.3 do Capítulo 2, trata-se de um código linear. Verificando as condições para $C = \{(00000), (01011), (10110), (11101)\}$ ser subespaço vetorial de $\mathbb{F}_2^5$, temos que

i) $(00000) \in C$;

ii) Se $u, v \in C$ então $u + v \in C$. De fato,

$$
\begin{aligned}
(00000) + (00000) &= (00000) \in C \\
(00000) + (01011) &= (01011) \in C \\
(00000) + (10110) &= (10110) \in C \\
(00000) + (11101) &= (11101) \in C \\
(01011) + (01011) &= (00000) \in C \\
(01011) + (10110) &= (11101) \in C \\
(01011) + (11101) &= (10110) \in C \\
(10110) + (10110) &= (00000) \in C \\
(10110) + (11101) &= (01011) \in C \\
(11101) + (11101) &= (00000) \in C;
\end{aligned}
$$

iii) Se $\alpha \in \{0,1\}$ e $u \in C$ então $\alpha \cdot u \in C$. De fato, se $\alpha = 0$, temos que

$$\alpha \cdot u = 0 \cdot u = (00000) \in C.$$

Se $\alpha = 1$, temos que

$$\alpha \cdot u = 1 \cdot u = u \in C.$$

Portanto, C é um subespaço vetorial de $\mathbb{F}_2^5$ e pela definição acima C é um código linear.

Achemos a base e a dimensão desse código linear. Considere um subconjunto $B = \{(01011), (10110), (11101)\}$ de C. Verifiquemos a dependência linear deste conjunto. Suponhamos que existam $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_2$ tais que

$$\alpha_1(01011) + \alpha_2(10110) + \alpha_3(11101) = (00000) \Rightarrow \begin{cases} \quad\quad \alpha_2 + \alpha_3 = 0 \\ \alpha_1 \quad\quad + \alpha_3 = 0 \\ \quad\quad \alpha_2 + \alpha_3 = 0 \\ \alpha_1 + \alpha_2 \quad\quad = 0 \\ \alpha_1 \quad\quad + \alpha_3 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \quad\quad \alpha_2 = \alpha_3 \\ \alpha_1 \quad\quad = \alpha_3 \\ \alpha_1 + \alpha_2 = 0 \end{cases} \Rightarrow \alpha_1 = \alpha_2 \text{ e } \alpha_3 \text{ é qualquer.}$$

Logo, o conjunto B é lineramente dependente.

Consideremos $B' = \{(01011), (10110)\}$, verifiquemos a dependência linear deste conjunto. Suponhamos que existam $\beta_1, \beta_2 \in \mathbb{F}_2$ tais que

$$\beta_1(01011) + \beta_2(10110) = (00000) \Rightarrow \begin{cases} \quad\quad \beta_2 = 0 \\ \beta_1 \quad\quad = 0 \\ \quad\quad \beta_2 = 0 \\ \beta_1 + \beta_2 = 0 \\ \beta_1 \quad\quad = 0 \end{cases} \Rightarrow \beta_1 = \beta_2 = 0,$$

admitindo apenas a solução trivial. Logo, B' é linearmente independente. Note que, os vetores de B' geram C. De fato, dados $\gamma_1, \gamma_2 \in \mathbb{F}_2$, temos que

$$\gamma_1(01011) + \gamma_2(10110) = (00000), \text{ para } \gamma_1 = \gamma_2 = 0;$$
$$\gamma_1(01011) + \gamma_2(10110) = (01011), \text{ para } \gamma_1 = 1 \text{ e } \gamma_2 = 0;$$
$$\gamma_1(01011) + \gamma_2(10110) = (10110), \text{ para } \gamma_1 = 0 \text{ e } \gamma_2 = 1;$$
$$\gamma_1(01011) + \gamma_2(10110) = (11101), \text{ para } \gamma_1 = 1 \text{ e } \gamma_2 = 1.$$

Portanto, B' é uma base de C. Como B' possui dois vetores, temos que $dim\ C = 2$.

Definição 5.1.2. Diremos que $\omega(u)$ é o **peso** do elemento $u \in \mathbb{F}^n$, se

$$\omega(u) = |\{i \mid u_i \neq 0,\ 1 \leq i \leq n\}|,$$

ou seja, o número de coordenadas não nulas de u. Em outras palavras, temos que

$$\omega(u) = d(u, 0_C),$$

onde d é a distância de *Hamming* de C.

Exemplo 5.1.2. Se $u = (10110)$ então $\omega(u) = |\{1, 3, 4\}| = 3$.

Definição 5.1.3. O **peso de um código linear** C é o inteiro

$$\omega(C) = min\{\omega(u) \mid u \in C \setminus \{0_C\}\}.$$

Exemplo 5.1.3. Seja $C = \{(00), (01), (10), (11)\}$ um código linear. Temos

$$\omega(00) = 0$$
$$\omega(01) = 1$$
$$\omega(10) = 1$$
$$\omega(11) = 2.$$

Logo, o peso do código linear C é

$$\omega(C) = min\{\omega(u) \mid u \in C \setminus \{00\}\} = 1.$$

Proposição 5.1.1. *Seja $C \subset \mathbb{F}^n$ um código linear com distância mínima d. Temos que*

i) $\forall u, v \in \mathbb{F}^n$, $d(u, v) = \omega(u - v)$;

ii) $d = \omega(C)$.

Demonstração. i) Para qualquer $u, v \in \mathbb{F}^n$, temos, pelas definições de distância de *Hamming* 2.2.2 e peso 5.1.2, que

$$d(u, v) = |\{i \mid u_i \neq v_i,\ 1 \leq i \leq n\}| = |\{i \mid u_i - v_i \neq 0,\ 1 \leq i \leq n\}| = \omega(u - v).$$

ii) Sejam $u, v \in C$ quaisquer e $u \neq v$, temos que existe $z \in C \setminus \{0\}$, tal que $u - v = z$. Pelo item i), obtemos que $d(u, v) = \omega(u - v) = \omega(z)$. Desse modo, pela definição de distância mínima 2.2.4, temos que

$$d = min\{d(u, v) \mid u, v \in C \text{ e } u \neq v\} = min\{\omega(z) \mid z \in C \setminus \{0\}\} = \omega(C).$$

■

Note que, a proposição acima mostra que, em códigos lineares com M elementos, podemos calcular a distância mínima d a partir de $M-1$ cálculos de distâncias, em vez da maneira vista no Capítulo 2, onde para calcularmos d era necessário fazer a combinação de todos os elementos do código dois a dois, isto é $\binom{M}{2}$. Porém, na prática, para códigos grandes, esses métodos para calcular d são inviáveis pois possuem um custo computacional elevado.

Em razão da Proposição 5.1.1 ii), a distância mínima de um código linear C será chamada de peso do código C. Em Álgebra Linear, temos duas maneiras de descrever subespaços vetoriais C de um espaço vetorial $\mathbb{F}^n$, uma como imagem, e outra como núcleo de transformações lineares.

Exemplo 5.1.4. Observe que, o código linear C do Exemplo 5.1.1 pode ser representado pela imagem da seguinte transformação linear

$$\begin{array}{rccc} T\colon & \mathbb{F}_2^2 & \longrightarrow & \mathbb{F}_2^5 \\ & (x_1, x_2) & \longmapsto & (x_1, x_2, x_1, x_1 + x_2, x_2). \end{array}$$

De fato,

$$\begin{aligned} T(00) &= (0,0,0,0+0,0) = (00000); \\ T(01) &= (0,1,0,0+1,1) = (01011); \\ T(10) &= (1,0,1,1+0,0) = (10110); \\ T(11) &= (1,1,1,1+1,1) = (11101). \end{aligned}$$

Note que, $Im\ (T) = \{(00000), (01011), (10110), (11101)\} = C$.

Por definição, todo código linear é um espaço vetorial de dimensão finita. Sejam $n \in \mathbb{N}$ tal que n é a dimensão de C e $B = \{u_1, u_2, \cdots, u_n\}$ uma base de C. Assim, cada vetor de C é escrito de maneira única na forma

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n,$$

com os $\alpha_i \in \mathbb{F}$ e $i \in \{1, 2, \cdots, n\}$.

Logo, o número de elementos de C é

$$M = |C| = q^n,$$

e, consequentemente,

$$dim\ C = n = \log_q q^n = \log_q M.$$

Agora vamos mostrar as duas maneiras de descrever subespaços vetoriais C de um espaço vetorial $\mathbb{F}^n$.

Primeiro vamos obter a representação de C como imagem de uma transformação linear.

Seja $C \subset \mathbb{F}^n$ um código linear de dimensão k sobre $\mathbb{F}$. Então C é isomorfo a $\mathbb{F}^k$. Desse modo, vamos definir uma transformação linear injetora $T : \mathbb{F}^k \longrightarrow \mathbb{F}^n$, tal que $Im\ (T) = C$.

De fato, dado uma base $B' = \{c_1, c_2, \cdots, c_k\}$ de C e uma base $B'' = \{e_1, e_2, \cdots, e_k\}$ de $\mathbb{F}^k$, definimos T como

$$\begin{array}{rccc} T\colon & \mathbb{F}^k & \longrightarrow & \mathbb{F}^n \\ & e_j & \longmapsto & c_j, \end{array}$$

para $e_j \in \mathbb{F}^k, c_j \in C$ e $j \in \{1, \cdots, k\}$.

Vamos encontrar uma matriz G que representa a transformação linear T nas bases destes espaços. Considere $B = \{b_1, b_2, \cdots, b_n\}$ a base canônica de $\mathbb{F}^n$. Assim, qualquer vetor $u \in \mathbb{F}^n$ pode ser escrito de maneira única como

$$u = u_1 b_1 + \cdots + u_n b_n, \text{ para } u_i \in \mathbb{F},\ i \in \{0, \cdots, n\}.$$

Em particular, escrevendo todos os vetores de B' em função dos vetores de B, temos que

$$\begin{cases} c_1 &= a_{11}b_1 + a_{21}b_2 + \cdots + a_{n1}b_n \\ c_2 &= a_{12}b_1 + a_{22}b_2 + \cdots + a_{n2}b_n \\ &\vdots \\ c_k &= a_{1k}b_1 + a_{2k}b_2 + \cdots + a_{nk}b_n \end{cases}$$

para $a_{ij} \in \mathbb{F}, i \in \{1, \cdots, n\}$ e $j \in \{1, \cdots, k\}$.

Como $T(e_i) = c_j = a_{1j}b_1 + a_{2j}b_2 + \cdots + a_{nj}b_n$, para $j \in \{1, \cdots, k\}$ a matriz de T nas respectivas bases é

$$G = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}_{n \times k}.$$

Assim, considere $v = (v_1, \cdots, v_k) \in \mathbb{F}^k$, a matriz de T se escreve, em coordenadas, pela equação

$$Gv = w \Rightarrow \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Como $Im\ (T) = C$, temos que cada vetor coluna de G pertence ao código linear C, ou seja, C é o subespaço gerado pelas colunas de G, que formam uma base de C.

Os elementos de C são as palavras y da forma $Gx = y$, $\forall x \in \mathbb{F}^k$.

Definição 5.1.4. Uma matriz $G \in M_{n\times k}(\mathbb{F})$ cujas colunas formam uma base de C, chama-se **matriz de codificação** de C.

Recorde que, um espaço vetorial sobre um corpo pode ter várias bases, isto posto, podemos encontrar diferentes matrizes de codificação para um mesmo código.

Exemplo 5.1.5. Seja $\mathbb{F}_2$ um corpo finito com dois elementos. Considere o código linear $C \subset \mathbb{F}^5$ definido pela imagem da transformação linear injetora

$$\begin{array}{rccc} T\colon & \mathbb{F}_2^3 & \longrightarrow & \mathbb{F}_2^5 \\ & (u_1,u_2,u_3) & \longmapsto & (u_2, u_1+u_3, u_1, u_1+u_2, u_2+u_3). \end{array}$$

Verifiquemos as condições para T ser uma transformação linear injetora.

i) Sejam $u = (u_1,u_2,u_3), v = (v_1,v_2,v_3) \in \mathbb{F}_2^3$ quaisquer, temos que

$$\begin{aligned} T(u+v) &= T((u_1,u_2,u_3)+(v_1,v_2,v_3)) \\ &= T(u_1+v_1, u_2+v_2, u_3+v_3) \\ &= (u_2+v_2, (u_1+v_1)+(u_3+v_3), u_1+v_1, (u_1+v_1)+(u_2+v_2), (u_2+v_2) \\ &\quad +(u_3+v_3)) \\ &= (u_2+v_2, (u_1+u_3)+(v_1+v_3), u_1+v_1, (u_1+u_2)+(v_1+v_2), (u_2+u_3) \\ &\quad +(v_2+v_3)) \\ &= (u_2, u_1+u_3, u_1, u_1+u_2, u_2+u_3)+(v_2, v_1+v_3, v_1, v_1+v_2, v_2+v_3) \\ &= T(u)+T(v). \end{aligned}$$

ii) Sejam $\alpha \in \mathbb{F}_2$ e $u = (u_1,u_2,u_3) \in \mathbb{F}_2^3$ qualquer, temos que

$$\begin{aligned} T(\alpha\cdot u) &= T(\alpha\cdot(u_1,u_2,u_3)) \\ &= T(\alpha\cdot u_1, \alpha\cdot u_2, \alpha\cdot u_3) \\ &= (\alpha\cdot u_2, \alpha\cdot u_1+\alpha\cdot u_3, \alpha\cdot u_1, \alpha\cdot u_1+\alpha\cdot u_2, \alpha\cdot u_2+\alpha\cdot u_3) \\ &= (\alpha\cdot u_2, \alpha\cdot(u_1+u_3), \alpha\cdot u_1, \alpha\cdot(u_1+u_2), \alpha\cdot(u_2+u_3)) \\ &= \alpha\cdot(u_2, u_1+u_3, u_1, u_1+u_2, u_2+u_3) \\ &= \alpha\cdot T(u). \end{aligned}$$

iii) Sejam $u = (u_1,u_2,u_3), v = (v_1,v_2,v_3) \in \mathbb{F}_2^3$ quaisquer, temos que, se

$$T(u) = T(v) \Rightarrow (u_2, u_1+u_3, u_1, u_1+u_2, u_2+u_3) = (v_2, v_1+v_3, v_1, v_1+v_2, v_2+v_3)$$

$$\Rightarrow \begin{cases} u_2 = v_2 \\ u_1 + u_3 = v_1 + v_3 \\ u_1 = v_1 \\ u_1 + u_2 = v_1 + v_2 \\ u_2 + u_3 = v_2 + v_3 \end{cases} \Rightarrow \begin{cases} u_1 = v_1 \\ u_2 = v_2 \\ u_3 = v_3 \end{cases} \Rightarrow u = v.$$

Portanto, T é uma transformação linear injetora.

■

Vamos encontrar uma base de $Im\ (T) = C$. Todo elemento de C pode ser escrito como

$$(u_2, u_1 + u_3, u_1, u_1 + u_2, u_2 + u_3) = u_1(0,1,1,1,0) + u_2(1,0,0,1,1) + u_3(0,1,0,0,1).$$

Note que, esses vetores formam uma base de C.

Considere $B = \{b_1, \cdots, b_5\}$ uma base canônica de $\mathbb{F}_2^5$.

Vamos encontrar uma matriz G que representa a transformação linear T. Assim,

$$\begin{aligned} (0,1,1,1,0) &= 0b_1 + 1b_2 + 1b_3 + 1b_4 + 0b_5 \\ (1,0,0,1,1) &= 1b_1 + 0b_2 + 0b_3 + 1b_4 + 1b_5 \\ (0,1,0,0,1) &= 0b_1 + 1b_2 + 0b_3 + 0b_4 + 1b_5 \end{aligned}$$

Portanto, uma matriz G de codificação de C é da forma

$$G = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Exemplo 5.1.6. Seja $\mathbb{F}_2$ um corpo finito com dois elementos. Considere o código linear $C \subset \mathbb{F}_2^6$ definido pela imagem da transformação linear injetora

$$\begin{array}{rccc} T\colon & \mathbb{F}_2^3 & \longrightarrow & \mathbb{F}_2^6 \\ & (u_1, u_2, u_3) & \longmapsto & (u_1, u_2, u_3, u_1 + u_3, u_2 + u_3, u_1 + u_2 + u_3). \end{array}$$

Verifiquemos as condições para T ser uma transformação linear injetora.

i) Sejam $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in \mathbb{F}_2^3$ quaisquer, temos que

$$T(u + v) = T((u_1, u_2, u_3) + (v_1, v_2, v_3))$$

$$
\begin{aligned}
&= T(u_1+v_1, u_2+v_2, u_3+v_3)\\
&= (u_1+v_1, u_2+v_2, u_3+v_3, (u_1+v_1)+(u_3+v_3), (u_2+v_2)+(u_3+v_3),\\
&\quad (u_1+v_1)+(u_2+v_2)+(u_3+v_3))\\
&= (u_1+v_1, u_2+v_2, u_3+v_3, (u_1+u_3)+(v_1+v_3), (u_2+u_3)+(v_2+v_3),\\
&\quad (u_1+u_2+u_3)+(v_1+v_2+v_3))\\
&= (u_1, u_2, u_3, u_1+u_3, u_2+u_3, u_1+u_2+u_3)+(v_1, v_2, v_3, v_1+v_3, v_2+v_3,\\
&\quad v_1+v_2+v_3)\\
&= T(u)+T(v).
\end{aligned}
$$

ii) Sejam $\alpha \in \mathbb{F}_2$ e $u = (u_1, u_2, u_3) \in \mathbb{F}_2^3$ qualquer, temos que

$$
\begin{aligned}
T(\alpha \cdot u) &= T(\alpha \cdot (u_1, u_2, u_3))\\
&= T(\alpha \cdot u_1, \alpha \cdot u_2, \alpha \cdot u_3)\\
&= (\alpha \cdot u_1, \alpha \cdot u_2, \alpha \cdot u_3, \alpha \cdot u_1 + \alpha \cdot u_3, \alpha \cdot u_2 + \alpha \cdot u_3, \alpha \cdot u_1 + \alpha \cdot u_2 + \alpha \cdot u_3))\\
&= (\alpha \cdot u_1, \alpha \cdot u_2, \alpha \cdot u_3, \alpha \cdot (u_1+u_3), \alpha \cdot (u_2+u_3), \alpha \cdot (u_1+u_2+u_3))\\
&= \alpha \cdot (u_1, u_2, u_3, u_1+u_3, u_2+u_3, u_1+u_2+u_3)\\
&= \alpha \cdot T(u).
\end{aligned}
$$

iii) Sejam $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in \mathbb{F}_2^3$ quaisquer, temos que, se

$$
T(u) = T(v) \Rightarrow (u_1, u_2, u_3, u_1+u_3, u_2+u_3, u_1+u_2+u_3) = (v_1, v_2, v_3, v_1+v_3, v_2+
$$

$$
v_3, v_1+v_2+v_3) \Rightarrow \begin{cases} u_1 = v_1 \\ u_2 = v_2 \\ u_3 = v_3 \end{cases} \Rightarrow u = v.
$$

Portanto, T é uma transformação linear injetora.

Considere $B = \{b_1, \cdots, b_6\}$ e $B' = \{e_1, e_2, e_3\}$ bases canônicas de $\mathbb{F}_2^6$ e $\mathbb{F}_2^3$ respectivamente.

Vamos encontrar a matriz G que representa a transformação linear T. Assim,

$$
\begin{aligned}
T(e_1) &= (1,0,0,1,0,1) = 1b_1 + 0b_2 + 0b_3 + 1b_4 + 0b_5 + 1b_6\\
T(e_2) &= (0,1,0,0,1,1) = 0b_1 + 1b_2 + 0b_3 + 0b_4 + 1b_5 + 1b_6\\
T(e_3) &= (0,0,1,1,1,1) = 0b_1 + 0b_2 + 1b_3 + 1b_4 + 1b_5 + 1b_6
\end{aligned}
$$

Portanto, uma matriz de codificação G tem a forma

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

■

Observe que, assim como no código linear da peça sobre um tabuleiro quadriculado do Exemplo 5.1.1, as três primeiras coordenadas são dígitos de informação, logo possui uma decodificação simples, onde ao recebermos a palavra (001111) a mensagem enviada é (001), fazendo com que as últimas três coordenadas sejam redundâncias para que se indentifique possíveis erros de transmissão.

Uma codificação que aplica a mensagem $u = (u_1, \cdots, u_k)$ na palavra do código $c = (c_1, \cdots, c_n)$, tal que existem coordenada $i_1, \cdots, i_k$ onde $u_1 = c_{i_1}, \cdots, u_k = c_{ik}$, diz-se **codificação sistemática**.

Um código diz-se **separável** se admite uma codificação sistemática.

As matrizes de codificação com a forma do Exemplo 5.1.6 recebe um nome especial na teoria dos códigos.

Definição 5.1.5. Diz-se que uma matriz de codificação G de um código linear C está na **forma padrão**, ou ainda, na **forma canônica** se $G = \begin{pmatrix} Id_k \\ A \end{pmatrix}$, onde $Id_k \in M_{k\times k}(\mathbb{F})$ é a matriz identidade e $A \in M_{(n-k)\times k}(\mathbb{F})$ uma matriz qualquer.

Agora vamos descrever o código linear C através do núcleo de uma transformação linear sobrejetora. Definiremos uma transformação linear sobrejetora $\pi : \mathbb{F}^n \longrightarrow \mathbb{F}^{n-k}$, tal que $Ker\ (\pi) = C$.

De fato, dado uma base $B' = \{c_1, \cdots, c_k\}$ de C e a ampliando para uma base $\{c_1, \cdots, c_k, b_1, \cdots, b_{n-k}\}$ de $\mathbb{F}^n$, temos que qualquer vetor $u \in \mathbb{F}^n$ pode ser escrito de maneira única como

$$u = u_1c_1 + \cdots + u_kc_k + u_{k+1}b_1 + \cdots + u_nb_{n-k},$$

para $u_i \in \mathbb{F}$ e $i \in \{1, \cdots, n\}$.

Definimos π como

$$\begin{array}{rccc} \pi: & \mathbb{F}^n & \longrightarrow & \mathbb{F}^{n-k} \\ & u & \longmapsto & u' = u_{k+1}b_1 + \cdots + u_nb_{n-k}. \end{array}$$

Denotaremos por $H \in M_{(n-k)\times n}(\mathbb{F})$ a matriz de posto $n-k$ que respresenta a transformação linear sobrejetora π em suas respectivas bases canônicas. O conjunto das palavras $x \in \mathbb{F}^n$ que satisfazem $Hx = 0$ formam o código linear C, pois $Ker\ (\pi) = C$.

Definição 5.1.6. A matriz H construída acima é chamada de **matriz de teste** do código linear C.

Exemplo 5.1.7. Seja $\mathbb{F}_2$ um corpo finito com dois elementos. Considere a transformação linear sobrejetora

$$\begin{array}{rccc} \pi: & \mathbb{F}_2^6 & \longrightarrow & \mathbb{F}_2^3 \\ & (u_1, \cdots, u_6) & \longmapsto & (u_1+u_4, u_1+u_2+u_3+u_5, u_1+u_2+u_6). \end{array}$$

Verifiquemos as condições para π ser uma transformação linear.

i) Sejam $u = (u_1, \cdots, u_6), v = (v_1, \cdots, v_6) \in \mathbb{F}_2^6$ quaisquer, temos que

$$\begin{aligned} \pi(u+v) &= \pi((u_1, \cdots, u_6) + (v_1, \cdots, v_6)) \\ &= \pi(u_1+v_1, \cdots, u_6+v_6) \\ &= ((u_1+v_1)+(u_4+v_4), (u_1+v_1)+(u_2+v_2)+(u_3+v_3)+(u_5+v_5), \\ &\quad (u_1+v_1)+(u_2+v_2)+(u_6+v_6)) \\ &= ((u_1+u_4)+(v_1+v_4), (u_1+u_2+u_3+u_5)+(v_1+v_2+v_3+v_5), \\ &\quad (u_1+u_2+u_6)+(u_1+u_2+u_6)) \\ &= (u_1+u_4, u_1+u_2+u_3+u_5, u_1+u_2+u_6)+(v_1+v_4, v_1+v_2+v_3+v_5, \\ &\quad v_1+v_2+v_6) \\ &= \pi(u)+\pi(v). \end{aligned}$$

ii) Sejam $\alpha \in \mathbb{F}_2$ e $u = (u_1, \cdots, u_6) \in \mathbb{F}_2^6$ qualquer, temos que

$$\begin{aligned} \pi(\alpha \cdot u) &= \pi(\alpha \cdot (u_1, \cdots, u_6)) \\ &= \pi(\alpha \cdot u_1, \cdots, \alpha \cdot u_6) \\ &= (\alpha \cdot u_1 + \alpha \cdot u_4, \alpha \cdot u_1 + \alpha \cdot u_2 + \alpha \cdot u_3 + \alpha \cdot u_5, \alpha \cdot u_1 + \alpha \cdot u_2 + \alpha \cdot u_6) \\ &= (\alpha \cdot (u_1+u_4), \alpha \cdot (u_1+u_2+u_3+u_5), \alpha \cdot (u_1+u_2+u_6)) \\ &= \alpha \cdot (u_1+u_4, u_1+u_2+u_3+u_5, u_1+u_2+u_6) \\ &= \alpha \cdot \pi(u). \end{aligned}$$

Agora vamos encontrar o núcleo de π.

Por definição $Ker\ (\pi) = \{u \in \mathbb{F}_2^6 \mid \pi(u) = 0_{\mathbb{F}_2^3}\}$. Desse modo,

$$\pi(u) = \pi(u_1, \cdots, u_6) = (u_1+u_4, u_1+u_2+u_3+u_5, u_1+u_2+u_6) = (0,0,0)$$

$$\Rightarrow \begin{cases} u_1 + u_4 = 0 \\ u_1 + u_2 + u_3 + u_5 = 0 \\ u_1 + u_2 + u_6 = 0 \end{cases} \Rightarrow \begin{cases} u_1 = u_4 \\ u_1 + u_2 + u_3 = u_5 \\ u_1 + u_2 = u_6 \end{cases}.$$

Portanto,

$$Ker\ (\pi) = \{(u_1, u_2, u_3, u_1, u_1 + u_2 + u_3, u_1 + u_2) \mid u_1, u_2, u_3 \in \mathbb{F}_2\}.$$

Assim, definimos $C = Ker\ (\pi)$.

Podemos escrever os elementos de $Ker\ (\pi)$ como

$$(u_1, u_2, u_3, u_1, u_1+u_2+u_3, u_1+u_2) = u_1(1,0,0,1,1,1)+u_2(0,1,0,0,1,1)+u_3(0,0,1,0,1,0).$$

Assim, os vetores $(1,0,0,1,1,1), (0,1,0,0,1,1), (0,0,1,0,1,0)$ geram $Ker\ (\pi)$. E ainda, dados $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_2$, temos que

$$\alpha_1(1,0,0,1,1,1) + \alpha_2(0,1,0,0,1,1) + \alpha_3(0,0,1,0,1,0) = 0_{\mathbb{F}_2^6} \Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0.$$

Logo, são vetores linearmente independentes. Desse modo, formam uma base de $Ker\ (\pi)$. Então a dimensão de $Ker\ (\pi)$ é 3.

Pelo Teorema 4.1.7, temos que $dim\ \mathbb{F}_2^6 = dim\ Ker\ (\pi) + dim\ Im\ (\pi)$, ou seja,

$$dim\ Im\ (\pi) = 6 - 3 = 3.$$

Como a dimensão da $Im\ (\pi)$ é igual a dimensão de $\mathbb{F}_2^3$, temos que π é uma transformação linear sobrejetora.

Considere $B = \{e_1, \cdots, e_6\}$ e $B' = \{b_1, b_2, b_3\}$ as bases canônicas de $\mathbb{F}_2^6$ e $\mathbb{F}_2^3$ respectivamente.

Vamos encontrar a matriz H que representa a transformação linear π. Assim,

$$\begin{aligned} \pi(e_1) &= \pi(1,0,0,0,0,0) = 1b_1 + 1b_2 + 1b_3 \\ \pi(e_2) &= \pi(0,1,0,0,0,0) = 0b_1 + 1b_2 + 1b_3 \\ \pi(e_3) &= \pi(0,0,1,0,0,0) = 0b_1 + 1b_2 + 0b_3 \\ \pi(e_4) &= \pi(0,0,0,1,0,0) = 1b_1 + 0b_2 + 0b_3 \\ \pi(e_5) &= \pi(0,0,0,0,1,0) = 0b_1 + 1b_2 + 0b_3 \\ \pi(e_6) &= \pi(0,0,0,0,0,1) = 0b_1 + 0b_2 + 1b_3 \end{aligned}$$

Portanto, a matriz H tem a forma

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

Note que, ao escolhermos um vetor $y \in \mathbb{F}_2^6$ qualquer, para verificarmos se ele pertence ao código C, analisaremos se satisfaz a condição $Hy = 0$.

Dados os vetores $u = (100111) \in \mathbb{F}_2^6$ e $v = (010101) \in \mathbb{F}_2^6$. Como

$$Hu = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+0+0+1+0+0 \\ 1+0+0+0+1+0 \\ 1+0+0+0+0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

e

$$Hv = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+0+0+1+0+0 \\ 0+1+0+0+0+0 \\ 0+1+0+0+0+1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Concluímos que, $u \in C$ e $v \notin C$.

■

Exemplo 5.1.8. Considere o código linear C do Exemplo 5.1.5, onde $C = \{(u_2, u_1 + u_3, u_1, u_1 + u_2, u_2 + u_3) \mid u_1, u_2, u_3 \in \mathbb{F}_2\}$. Vamos encontrar uma matriz de teste desse código.

Sabemos que os vetores $c_1 = (01110), c_2 = (10011)$ e $c_3 = (01001)$ formam uma base B de C. Para ampliar essa base para uma base de $\mathbb{F}_2^5$ tomemos os vetores $c_4 = (00010), c_5 = (00001) \in \mathbb{F}_2^5$, desse modo o conjunto $B' = \{c_1, c_2, c_3, c_4, c_5\}$ é uma base de $\mathbb{F}_2^5$.

Definindo

$$\pi : \mathbb{F}_2^5 \longrightarrow \mathbb{F}_2^2 \text{ por } \begin{cases} \pi(c_1) = (00) \\ \pi(c_2) = (00) \\ \pi(c_3) = (00) \\ \pi(c_4) = (10) \\ \pi(c_5) = (01) \end{cases}$$

Vamos escrever os vetores da base canônica de $\mathbb{F}_2^5$ como combinação linear dos vetores da base B' de $\mathbb{F}_2^5$.

$$(10000) = 0(01110) + 1(10011) + 0(01001) + 1(00010) + 1(00001)$$
$$(01000) = 0(01110) + 0(10011) + 1(01001) + 0(00010) + 1(00001)$$

$$(00100) = 1(01110) + 0(10011) + 1(01001) + 1(00010) + 1(00001)$$
$$(00010) = 0(01110) + 0(10011) + 0(01001) + 1(00010) + 0(00001)$$
$$(00001) = 0(01110) + 0(10011) + 0(01001) + 0(00010) + 1(00001)$$

De modo resumido, temos

$$(10000) = 0c_1 + 1c_2 + 0c_3 + 1c_4 + 1c_5$$
$$(01000) = 0c_1 + 0c_2 + 1c_3 + 0c_4 + 1c_5$$
$$(00100) = 1c_1 + 0c_2 + 1c_3 + 1c_4 + 1c_5$$
$$(00010) = 0c_1 + 0c_2 + 0c_3 + 1c_4 + 0c_5$$
$$(00001) = 0c_1 + 0c_2 + 0c_3 + 0c_4 + 1c_5$$

Assim,

$$\begin{array}{lcllcll}
\pi(10000) & = & \pi(c_2) + \pi(c_4) + \pi(c_5) & = & (00) + (10) + (01) & = & (11) \\
\pi(01000) & = & \pi(c_3) + \pi(c_5) & = & (00) + (01) & = & (01) \\
\pi(00100) & = & \pi(c_1) + \pi(c_3) + \pi(c_4) + \pi(c_5) & = & (00) + (00) + (10) + (01) & = & (11) \\
\pi(00010) & = & \pi(c_4) & = & (10) & & \\
\pi(00001) & = & \pi(c_5) & = & (01) & &
\end{array}$$

Portanto, a matriz de teste do código linear C é

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

■

Observe que, existe uma relação entre as transformações lineares T e π vistas acima, onde $C = Im\ (T) = Ker\ (\pi)$. Dado $u \in \mathbb{F}^k$, temos que $\pi \circ T(u) = \pi(T(u)) = 0$, pois $T(u) \in Im\ (T) = C = Ker\ (\pi)$, em notação matricial $HG = 0$.

Definição 5.1.7. Sejam $u = (u_1, \cdots, u_n), v = (v_1 \cdots, v_n) \in \mathbb{F}^n$, chamaremos de **produto escalar** de u e v em $\mathbb{F}^n$ por

$$u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Caso $u \cdot v = 0$, dizemos que u e v são ortogonais. Assim os vetores linhas da matriz de teste H são ortogonais aos vetores colunas da matriz de codificação G.

O produto escalar satisfaz as seguintes propriedades.

Propriedade 5.1.1. i) $u \cdot v = v \cdot u,\ \forall u, v \in \mathbb{F}^n$;

ii) $(u + v) \cdot w = u \cdot w + v \cdot w,\ \forall u, v, w \in \mathbb{F}^n$;

iii) $(\alpha u)\cdot v = \alpha(u\cdot v),\ \forall \alpha \in \mathbb{F},\ \forall u, v \in \mathbb{F}^n$.

Demonstração. Sejam $u, v, w \in \mathbb{F}^n$ quaisquer e $\alpha \in \mathbb{F}$ qualquer, temos

i)

$$u \cdot v = u_1v_1 + \cdots + u_nv_n = v_1u_1 + \cdots + v_nu_n = v \cdot u.$$

ii)

$$\begin{aligned}(u+v)\cdot w &= (u_1+v_1, \cdots, u_n+v_n)\cdot(w_1, \cdots, w_n)\\ &= (u_1+v_1)w_1 + \cdots + (u_n+v_n)w_n\\ &= (u_1w_1+v_1w_1) + \cdots + (u_nw_n+v_nw_n)\\ &= (u_1w_1 + \cdots + u_nw_n) + (v_1w_1 + \cdots + v_nw_n)\\ &= u\cdot w + v\cdot w.\end{aligned}$$

iii)

$$\begin{aligned}(\alpha u)\cdot v &= (\alpha u_1, \cdots, \alpha u_n)\cdot(v_1, \cdots, v_n)\\ &= (\alpha u_1)v_1 + \cdots + (\alpha u_n)v_n\\ &= \alpha(u_1v_1) + \cdots + \alpha(u_nv_n)\\ &= \alpha(u_1v_1 + \cdots + u_nv_n)\\ &= \alpha(u\cdot v).\end{aligned}$$

■

Note que, se $u \cdot u = 0$ não necessariamente $u = 0_{\mathbb{F}^n}$. Por exemplo, seja $\mathbb{F}_2$ um corpo finito com dois elementos, dado o vetor $u = (100111) \in \mathbb{F}_2^6$, temos que

$$u\cdot u = (100111)\cdot(100111) = 1\cdot1+0\cdot0+0\cdot0+1\cdot1+1\cdot1+1\cdot1 = 1+0+0+1+1+1 = 0.$$

Definição 5.1.8. Seja C um código linear de comprimento n e dimensão k. É dito que o conjunto

$$C^{\perp} = \{u \in \mathbb{F}^n \mid u\cdot v = 0,\ \forall v \in C\}$$

é o **código dual** de C.

Lema 5.1.2. *Se $C \subset \mathbb{F}^n$ é um código linear, com matriz de codificação G, então*

i) $C^{\perp}$ é um subespaço vetorial de $\mathbb{F}^n$.

ii) $u \in C^{\perp} \Leftrightarrow uG = 0$.

Demonstração. Seja $C \subset \mathbb{F}^n$ um código linear, com matriz de codificação G e $C^\perp$ um código dual de C. Verifiquemos as condições para $C^\perp$ ser um subespaço vetorial de $\mathbb{F}^n$.

i) i.i) Se $u = (0, \cdots, 0) \in \mathbb{F}^n$ então $u \in C^\perp$ pois

$$u \cdot v = (0, \cdots, 0) \cdot v = 0, \ \forall v \in C.$$

i.ii) Sejam $u, v \in C^\perp$, $\forall w \in C$, temos que

$$(u + v) \cdot w = u \cdot w + v \cdot w = 0 + 0 = 0.$$

Logo, $u + v \in C^\perp$.

i.iii) Sejam $\alpha \in \mathbb{F}^n$, $u \in C^\perp$, $\forall v \in C$, temos que

$$(\alpha u) \cdot v = \alpha(u \cdot v) = \alpha 0 = 0.$$

Logo, $\alpha u \in C^\perp$.

Portanto, $C^\perp$ é subespaço vetorial de $\mathbb{F}^n$ e consequentemente um código linear.

ii) Pela Definição 5.1.8, temos que se $u \in C^\perp$ então ele é ortogonal a qualquer elemento de C, desse modo u também será ortogonal a qualquer vetor de uma base de C. Assim, como os vetores colunas de G formam uma base de C, concluímos que $uG = 0$.

■

Sejam C um código de comprimento n e dimensão k com matriz de codificação G e $u \in C^\perp$. Então $u \cdot v = 0$, $\forall v \in C$. Em particular, u é ortogonal a todos os vetores de uma base de C, ou seja, $uG = 0$. Assim, podemos definir o código dual $C^\perp$ a partir da matriz de codificação como

$$C^\perp = \{u \in \mathbb{F}^n \mid uG = 0\}.$$

Exemplo 5.1.9. Recorde que, do Exemplo 5.1.8 obtemos que $HG = 0$, onde H é a matriz de teste do código linear C. Assim, H é uma matriz de codificação de $C^\perp$.

Proposição 5.1.3. *Seja $C \subset \mathbb{F}^n$ um código linear de comprimento n e dimensão k com matriz de codificação $G = \begin{pmatrix} Id_k \\ A \end{pmatrix}$, na forma padrão. Então*

i) $dim\ C^\perp = n - k$;

ii) $H = (-A \mid Id_{n-k})$ é uma matriz de codificação de $C^\perp$;

iii) $(C^\perp)^\perp = C$.

Demonstração. i) Pelo Lema 5.1.2 item ii), temos que $u \in C^{\perp} \Leftrightarrow uG = 0$. E ainda, consideremos a transposta, temos que $(uG)^t = 0$. Pela Proposição 4.1.8, isso equivale a dizer que $G^t u^t = 0$, por esse motivo $dim\ C^{\perp} = dim\ Ker\ (G^t)$.

Pelo Teorema 4.1.7, temos que

$$dim\ \mathbb{F}^n = dim\ Ker\ (G^t) + dim\ Im\ (G^t) \Rightarrow n = dim\ Ker\ (G^t) + dim\ Im\ (G^t)$$

e $G^t = (Id_k \mid A^t)$ tem posto k, ou seja, k linhas linearmente independentes. Logo, $dim\ Im\ (G^t) = k$.

Assim,

$$\begin{aligned} dim\ Ker\ (G^t) + dim\ Im\ (G^t) = n &\Rightarrow dim\ Ker\ (G^t) + k = n \\ &\Rightarrow dim\ Ker\ (G^t) = n - k. \end{aligned}$$

Portanto, $dim\ C^{\perp} = dim\ Ker\ (G^t) = n - k$.

ii) Considere $H = (-A \mid Id_{n-k})$ uma matriz de posto $n - k$. Nessas condições, temos que

$$HG = (-A \mid Id_{n-k}) \begin{pmatrix} Id_k \\ A \end{pmatrix} = 0.$$

Desse modo, os vetores linhas de H são ortogonais aos vetores colunas de G. Logo, o espaço gerado pelos vetores linhas de H está contido em $C^{\perp}$ e como o posto de H é igual a dimensão de $C^{\perp}$, segue que H é uma matriz de codificação do código linear $C^{\perp}$.

iii) Para mostrar a igualdade entre conjuntos vamos verificar que $C \subset (C^{\perp})^{\perp}$ e $(C^{\perp})^{\perp} \subset C$.

($\Rightarrow$) Seja $u \in C$, uma palavra não nula. Por definição existe $v \in C^{\perp}$, tal que $v \cdot u = 0$ então pela Propriedade 5.1.1, temos que $u \cdot v = 0$. Concluímos que, $u \in (C^{\perp})^{\perp}$. Portanto, $C \subset (C^{\perp})^{\perp}$.

($\Leftarrow$) Seja $w \in (C^{\perp})^{\perp}$, uma palavra não nula. Por definição existe $y \in C^{\perp}$ tal que $w \cdot y = 0$ então pela Propriedade 5.1.1, temos que $y \cdot w = 0$. Concluímos que, $w \in C$. Portanto, $(C^{\perp})^{\perp} \subset C$.

■

Note que, espaços duais sobre um corpo finito tem propriedades diferentes das propriedades de espaços duais sobre o corpo dos números reais. Por exemplo, seja U um espaço vetorial sobre $\mathbb{R}$ de dimensão finita, e considere V um subespaço vetorial de U então $V \cap V^{\perp} = \{0\}$. Isto não ocorre sempre no caso de um subespaço W de U de dimensão finita sobre um corpo finito $\mathbb{F}$, pois é possível que $W \cap W^{\perp} \neq \{0\}$ e até $W = W^{\perp}$.

Exemplo 5.1.10. Seja $\mathbb{F}_2$ o corpo finito com dois elementos. Considere C um código linear de comprimento 4 e dimensão 2 definido por $C = \{(0000), (1100), (0011), (1111)\}$.

Vamos encontrar o código dual $C^{\perp}$ de C, ou seja, $C^{\perp}$ é formado por todas as palavras $u = (u_1, u_2, u_3, u_4) \in \mathbb{F}_2^4$, tais que $u \cdot v = 0$, $\forall v \in C$.

Assim,

$$\begin{aligned}
u \cdot (0000) &= (u_1, u_2, u_3, u_4) \cdot (0000) = u_1 0 + u_2 0 + u_3 0 + u_4 0 = 0 \\
u \cdot (1100) &= (u_1, u_2, u_3, u_4) \cdot (1100) = u_1 1 + u_2 1 + u_3 0 + u_4 0 = 0 \\
u \cdot (0011) &= (u_1, u_2, u_3, u_4) \cdot (0011) = u_1 0 + u_2 0 + u_3 1 + u_4 1 = 0 \\
u \cdot (1111) &= (u_1, u_2, u_3, u_4) \cdot (1111) = u_1 1 + u_2 1 + u_3 1 + u_4 1 = 0
\end{aligned}$$

$$\Rightarrow \begin{cases} u_1 + u_2 = 0 \\ u_3 + u_4 = 0 \\ u_1 + u_2 + u_3 + u_4 = 0 \end{cases} \Rightarrow \begin{cases} u_1 = u_2 \\ u_3 = u_4 \\ u_1 + u_1 + u_3 + u_3 = 0 \end{cases} \Rightarrow u = (u_1, u_1, u_3, u_3).$$

Desse modo, temos os seguintes casos:

i) Se $u_1 = u_3 = 0$, temos que $u = (0000)$;

ii) Se $u_1 = 1$ e $u_3 = 0$, temos que $u = (1100)$;

iii) Se $u_1 = 0$ e $u_3 = 1$, temos que $u = (0011)$;

iv) Se $u_1 = 1$ e $u_3 = 1$, temos que $u = (1111)$.

Portanto, o código dual $C^{\perp} = \{(0000), (1100), (0011), (1111)\}$ é igual ao código linear C.

5.2 Relação entre uma matriz de teste e o peso do código linear

Nesta seção, vamos mostrar que a matriz de teste H de um código linear C contém informações sobre o valor do peso $\omega(C)$.

Lema 5.2.1. *Seja H uma matriz de teste de um código linear C. Se existe $u \in C$, tal que $\omega(u) = t$ então existem t colunas de H linearmente dependentes.*

Demonstração. Seja H a matriz de teste do código linear C, tal que

$$H = \begin{matrix} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{H_1}{} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{H_2}{} \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{H_n}{} \\ \\ \\ \\ \end{matrix} \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n-k1} & h_{n-k2} & \cdots & h_{n-kn} \end{bmatrix}$$

e como $u \in C$, temos que

$$Hu = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n-k1} & h_{n-k2} & \cdots & h_{n-kn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} h_{11}u_1 + h_{12}u_2 + \cdots + h_{1n}u_n \\ h_{21}u_1 + h_{22}u_2 + \cdots + h_{2n}u_n \\ \vdots \\ h_{n-k1}u_1 + h_{n-k2}u_2 + \cdots + h_{n-kn}u_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Assim, temos

$$\begin{cases} h_{11}u_1 + h_{12}u_2 + \cdots + h_{1n}u_n = 0 \\ h_{21}u_1 + h_{22}u_2 + \cdots + h_{2n}u_n = 0 \\ \quad \vdots \\ h_{n-k1}u_1 + h_{n-k2}u_2 + \cdots + h_{n-kn}u_n = 0 \end{cases}$$

Note que, a componente u_1 multiplica todos os elementos da coluna H_1, isto ocorre com todas as componentes da palavra u, isto é, $H_1u_1, H_2u_2, \cdots, H_nu_n$.

Logo, podemos representar estas multiplicações por

$$Hu = \overset{H_1}{\begin{bmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{n-k1} \end{bmatrix}} u_1 + \overset{H_2}{\begin{bmatrix} h_{12} \\ h_{22} \\ \vdots \\ h_{n-k2} \end{bmatrix}} u_2 + \cdots + \overset{H_n}{\begin{bmatrix} h_{1n} \\ h_{2n} \\ \vdots \\ h_{n-kn} \end{bmatrix}} u_n = 0.$$

Por hipótese, $\omega(u) = t$. Logo, u tem t coordenandas não nulas então as colunas que multiplicam essas coordenadas são linearmente dependentes. Assim, H possui t colunas linearmente dependentes.

■

Lema 5.2.2. *Seja H uma matriz de teste de um código C. Se existem t colunas de H que são linearmente dependentes então $\omega(C) \leq t$.*

Demonstração. Suponhamos que existem t colunas de H linearmente dependentes, denotadas por $h_{i_1}, \cdots, h_{i_t}$. Então existem escalares $c_{i_1}, \cdots, c_{i_t} \in \mathbb{F}$, não todos nulos, tais que

$$c_{i_1}h_{i_1} + \cdots + c_{i_t}h_{i_t} = 0.$$

Tomando $c = (c_1, \cdots, c_n) \in \mathbb{F}^n$ definido por $\begin{cases} c_i = c_{i_j}, & \text{se } j \in \{1, \cdots, t\} \\ c_i = 0, & \text{caso contrário.} \end{cases}$

Então $Hc = 0$, portanto $c \in C$. Logo, como H possui t colunas linearmente dependentes, temos que c tem até t coordenadas não nulas. Portanto, $\omega(c) \leq t$. Assim, $\omega(C) \leq t$.

■

Teorema 5.2.3. *Seja H uma matriz de teste de um código C. Temos que o peso de C é igual a d se, e somente se, quaisquer $d-1$ colunas de H são linearmente independentes e existem d colunas de H linearmente dependentes.*

Demonstração. ($\Rightarrow$) Seja o peso de C igual a d e suponhamos, por absurdo, que a matriz H possui $d-1$ colunas linearmente dependentes segue, do Lema 5.2.2, que $\omega(C) \leq d-1 < d$, o que contradiz nossa hipótese.

Observe que, como $\omega(C) = d$, existe alguma palavra $u \in C$ não nula, tal que $\omega(u) = d$. Pelo Lema 5.2.1, temos que existem d colunas de H linearmente dependentes.

($\Leftarrow$) Considere, por hipótese, que $d-1$ colunas de H são linearmente independentes e d colunas de H são linearmente dependentes.

Denotaremos as colunas de H por $H_1, \cdots, H_n$. Seja $c \in C$ uma palavra qualquer e não nula. Então

$$Hc = H_1c_1 + \cdots + H_nc_n = 0.$$

Como quaisquer das $d-1$ colunas de H são linearmente independentes, temos que pelo menos d coordenadas de c são não nulas, ou seja, $\omega(c) \geq d$, $\forall c \in C$. Assim, $\omega(C) \geq d$.

Pelo Lema 5.2.2, como d colunas de H são linearmente dependentes, temos que $\omega(C) \leq d$.

Portanto, $\omega(C) = d$.

■

Lembre que, no Teorema 2.3.1 demonstramos a *Cota de Singleton*. Agora vamos reescrevê-lo a partir de outros parâmetros pois quanto maior o peso de um código linear melhor será sua capacidade de correção.

O Corolário a seguir é consequência imediata dos resultados anteriores demonstrados, porém como já o demonstramos no Capítulo 2, aqui apresentaremos uma demonstração que utiliza os conceitos trabalhados neste capítulo supracitado.

Corolário 4. *(Cota de Singleton) Seja C um código linear com parâmetros $[n, k, d]$, onde $d = \omega(C)$. Temos que*

$$d \leq n - k + 1.$$

Demonstração. Pelo Teorema 2.3.1, temos que

$$M \leq q^{n-d+1} \Rightarrow \log_q M \leq \log_q q^{n-d+1} \Rightarrow \log_q M \leq n - d + 1.$$

Já vimos neste capítulo que

$$dim\ C = \log_q M \Rightarrow dim\ C \leq n - d + 1.$$

Como a $dim\ C = k$, temos que

$$k \leq n - d + 1 \Rightarrow d \leq n - d + 1.$$

■

Observe que, a *Cota de Singleton* nos mostra que um código linear C de dimensão k em $\mathbb{F}^n$ tem o peso máximo de

$$\omega(C) = d = n - k + 1.$$

Definição 5.2.1. Um código diz-se **código separável pela distância máxima** se vale a igualdade $d = n - k + 1$.

Exemplo 5.2.1. Vamos construir um código linear C de comprimento 6 e dimensão 3, tal que uma palavra $c \in C$ qualquer possua os três primeiros dígitos de informação e os três últimos dígitos de redundância. Denotaremos por c_1, c_2, c_3 os dígitos de informação e c_4, c_5, c_6 os dígitos de redundância. Definiremos os dígitos de redundância por

$$\begin{aligned} c_4 &= c_1 + c_2 \\ c_5 &= c_1 + c_3 \\ c_6 &= c_2 + c_3 \end{aligned}$$

em notação matricial, temos que

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_1 + c_2 \\ c_1 + c_3 \\ c_2 + c_3 \end{bmatrix},$$

onde a matriz apresentada acima é a matriz de codificação G do código linear C.

Suponhamos que recebemos uma palavra $u = (u_1, \cdots, u_6) \in \mathbb{F}_2^6$. Para que possamos descobrir se essa palavra pertence ou não a C devemos verificar se as equações definidas

pelos dígitos de redundância são satisfeitas, ou seja,

$$\begin{cases} u_1 + u_2 \quad\quad = u_4 \\ u_1 \quad\quad + u_3 = u_5 \\ \quad\quad u_2 + u_3 = u_6 \end{cases} \Rightarrow \begin{cases} u_1 + u_2 \quad\quad + u_4 \quad\quad = 0 \\ u_1 \quad\quad + u_3 \quad\quad + u_5 \quad\quad = 0 \\ \quad\quad u_2 + u_3 \quad\quad + u_6 = 0 \end{cases}$$

em notação matricial, temos que

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

onde a matriz apresentada acima é a matriz de teste H do código linear C, em outras palavras

$$u \in C \Leftrightarrow Hu = 0.$$

Observe que, quaisquer duas colunas de H são linearmente independentes entre si e as três primeiras colunas são linearmente dependentes. De fato, sejam $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_2$, temos que

$$\alpha_1(110)+\alpha_2(101)+\alpha_3(011) = 0 \Rightarrow \begin{cases} \alpha_1 + \alpha_2 \quad\quad = 0 \\ \alpha_1 \quad\quad + \alpha_3 = 0 \\ \quad\quad \alpha_2 + \alpha_3 = 0 \end{cases} \Rightarrow \begin{cases} \alpha_1 = \alpha_2 \\ \alpha_1 = \alpha_3 \\ \alpha_2 = \alpha_3 \end{cases} \Rightarrow \alpha_1 = \alpha_2 = \alpha_3.$$

Admitindo assim outras soluções além da trivial.

Isto posto, pelo Teorema 5.2.3, temos que d é igual a 3. Como o comprimento n é igual a 6 e a dimensão k é igual a 3, temos, pela *Cota de Singleton*, que

$$d \leq n - k + 1 \Rightarrow 3 \leq 6 - 3 + 1 \Rightarrow 3 \leq 4.$$

Logo, C não se trata de um código separável pela distância máxima. Observe ainda que, G está na forma padrão e o número de dígitos de redundância é igual ao número de linhas de H.

É comumente visto em teoria dos códigos corretores de erros que a matriz transposta G^t da matriz de codificação G, cujas linhas geram o código é donotada por matriz geradora.

Logo, uma matriz geradora G^t está na forma padrão se $G^t = (Id_k \mid A)$, para $Id_k \in M_{k\times k}(\mathbb{F})$ é a matriz identidade e $A \in M_{k\times n-k}(\mathbb{F})$ uma matriz qualquer. Então a matriz de teste H fica da forma $H = (-A^t \mid Id_{n-k})$, para $Id_{n-k} \in M_{n-k\times n-k}(\mathbb{F})$ é a matriz identidade.

5.3 Aplicações

Nessa seção vamos apresentar algumas aplicações dos códigos lineares que desepenharam um papel importante no desenvolvimento da teoria dos códigos corretores de erros.

Aplicação 1. (*Códigos de Hamming)*

De acordo com Roman (1992, p. 253), os códigos de *Hamming* são possivelmente, o tipo de códigos corretores de erros mais famoso, devido a sua capacidade de corrigir erros simples. Esses códigos foram fundamentados por *Marcel Golay* em 1949 e *Richard Hamming* em 1950.

Seja C um código linear de parâmetros $[n, k]$, com uma matriz de teste H, sobre um corpo finito $\mathbb{F}_2$. Pelo Teorema 5.2.3, a distância mínima é o menor inteiro d, tal que existam d colunas de H linearmente dependentes.

Vamos construir um código capaz de corrigir um único erro, ou seja, a capacidade de correção κ de C é igual a 1. Desse modo, temos que

$$\kappa = \left[\frac{d-1}{2}\right] \Rightarrow 1 = \left[\frac{d-1}{2}\right] \Rightarrow 2 = d - 1 \Rightarrow d = 3.$$

Logo, consideremos o código C com os prâmetros $[n, k, 3]$ descrito acima. Sabemos que C pode ser definido a partir de uma matriz de teste H cuja ordem é $m \times n$, onde $m = n - k$.

Consequentemente, a matriz de teste de um código C de parâmetros $[n, k, 3]$ tem a propriedade que duas de suas colunas não são linearmente dependentes, isto é, nenhuma coluna é múltipla de outra coluna por um escalar, porém algum conjunto de três colunas são linearmente dependentes.

Recorde que, a matriz de teste representa uma transformação linear sobrejetora. Neste caso, $\pi : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$, onde $Ker\ (\pi) = C$.

Denotaremos o vetor nulo $(0, 0, \cdots, 0) \in \mathbb{F}_2^m$ por $\mathbf{0}$. Assim, construiremos uma matriz de teste H de C, tal que as colunas de H sejam os elementos de $\mathbb{F}_2^m \setminus \{\mathbf{0}\}$, independente da ordem.

Consideremos todos os vetores possíveis, não nulos, de comprimento m em $\mathbb{F}_2^m$. Logo, o número de palavras não nulas em $\mathbb{F}_2^m$ é 2^m menos o vetor nulo de $\mathbb{F}_2^m$, ou seja, $2^m - 1$. Assim, a matriz de teste H possui $2^m - 1$ colunas, o que equivale a dizer que o comprimento de C é $n = 2^m - 1$.

Segue que, para cada m dado, a matriz H construída acima determina um código linear C, sobre um corpo finito $\mathbb{F}_2$, de comprimento $n = 2^m - 1$.

Pelo Teorema 4.1.7, temos que

$$\begin{aligned} dim\ \mathbb{F}_2^n = dim\ Ker\ (\pi) + dim\ Im\ (\pi) &\Rightarrow n = k + m \\ &\Rightarrow k = n - m \\ &\Rightarrow k = 2^m - 1 - m. \end{aligned}$$

Definição 5.3.1. Um código linear C, com a matriz de teste H construída acima e parâmetros $[n, k, d]$, tais que $n = 2^m - 1,\ k = 2^m - 1 - m$ e $d = 3$, diz-se **código de *Hamming* de ordem m**.

Vamos escolher uma forma "padrão" de escrevermos a matriz de teste H de um código de *Hamming*, organizando a ordem das colunas desta matriz.

Como as colunas de H são elementos de $\mathbb{F}_2^m \setminus \{\mathbf{0}\}$, coloquemos os vetores colunas da base canônica de $\mathbb{F}_2^m$ no bloco direito da matriz H, ou seja, considere $B = \{b_1, \cdots, b_m\}$ a base canônica de $\mathbb{F}_2^m$, de modo que os b_i sejam vetores colunas, para $i \in \{1, \cdots, m\}$.

Seja U o conjunto de todos os vetores colunas, não nulos e diferentes dos vetores da base canônica de $\mathbb{F}_2^m$. Logo, U tem $2^m - 1 - m$ vetores.

Desta forma, construímos um conjunto cujos elementos são dois a dois linearmente independentes. Logo, os vetores $u_1, u_2, \cdots, u_{2^m-1-m} \in U$ formam as colunas do bloco esquerdo de H.

Assim, a matriz de teste tem a forma

$$H = \begin{bmatrix} u_1 & \cdots & u_{2^m-1-m} & b_1 & \cdots & b_m \end{bmatrix}.$$

De modo que, a ordem dos vetores colunas u_j é qualquer, onde $j \in \{1, \cdots, 2^m - 1 - m\}$.

Definição 5.3.2. A matriz H construída acima é dita como **matriz de *Hamming* de ordem m**.

Faremos um exemplo que ilustre bem esse algorítmo.

Exemplo 5.3.1. Sejam $\mathbb{F}_2$ um corpo finito e C um código de *Hamming* de ordem 4 sobre $\mathbb{F}_2$.

Temos que os parâmetros deste código são:

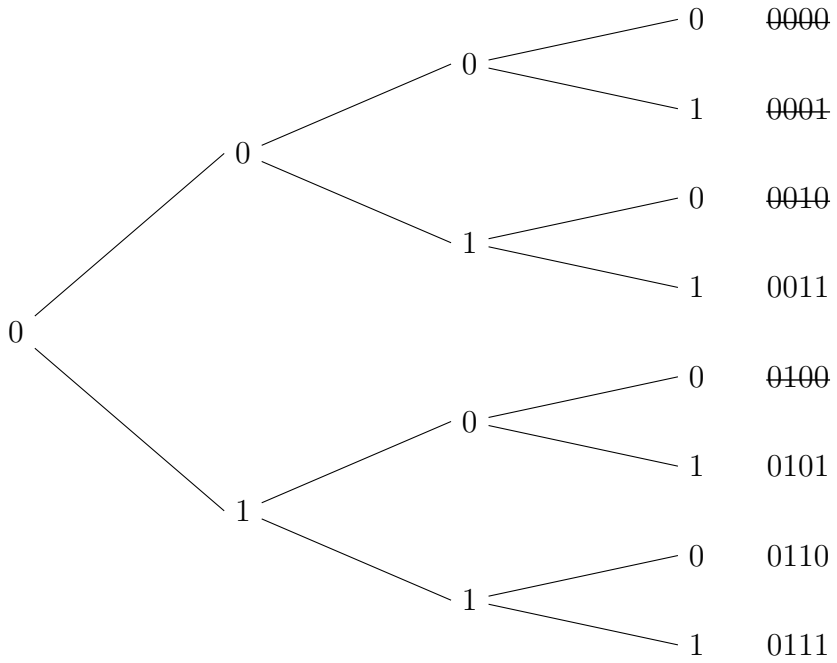
$$n = 2^m - 1 = 2^4 - 1 = 15, \quad k = n - m = 15 - 4 = 11 \quad \text{e} \quad d = 3.$$

Portanto, C é um código de *Hamming* com os parâmetros $[15, 11, 3]$.

Vamos contruir a matriz de *Hamming* H de ordem 4. Sabemos que, a base canônica B de $\mathbb{F}_2^4$ é igual a $\{(1000),(0100),(0010),(0001)\}$. Assim, os vetores transpostos de B formam as colunas do bloco direito de H.

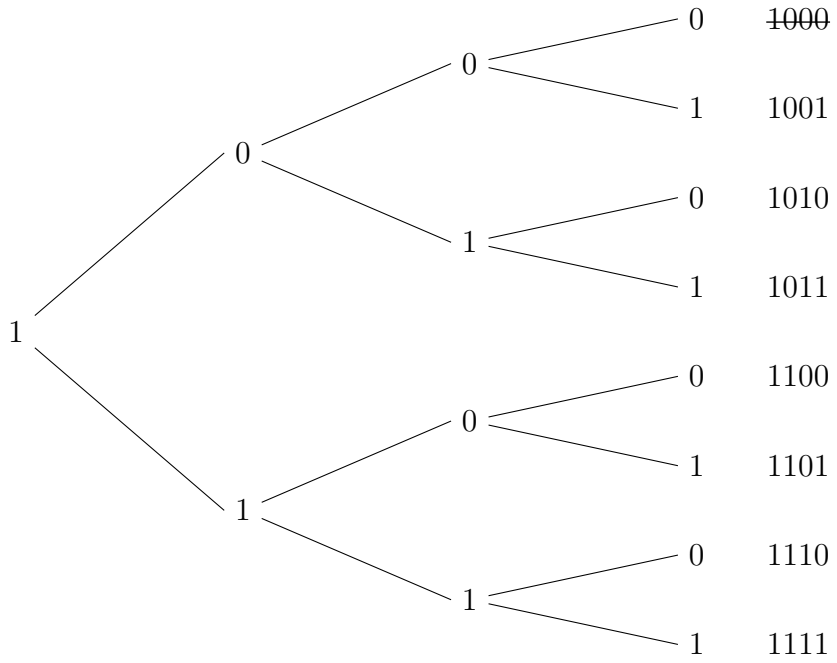
As palavras (vetores) de $\mathbb{F}_2^4$ são:

Figura 2 – Diagrama de árvore para os vetores de $\mathbb{F}_2^4$ que começam com o dígito 0



Fonte: Elaborado pelo autor.

Figura 3 – Diagrama de árvore para os vetores de $\mathbb{F}_2^4$ que começam com o dígito 1



Fonte: Elaborado pelo autor.

Note que, os vetores de $\mathbb{F}_2^4$ que foram tachados são os vetores de B e o vetor nulo. Assim, os vetores restantes pertencem a $\mathbb{F}_2^4 \setminus B \cup \{\mathbf{0}\}$. De modo que, esses vetores tanspostos formam o bloco esquerdo da matriz H.

Portanto, podemos escrever a matriz de *Hamming* de ordem 4 como

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

Proposição 5.3.1. *Todo código de Hamming é perfeito.*

Demonstração. Sejam C um código de *Hamming*, sobre um corpo finito $\mathbb{F}_2$, com parâmetros $[n, k, 3]$ e $u \in C$ uma palavra qualquer, segue do Lema 2.2.2, que o número de elementos do disco de centro u e raio $\kappa = 1$ é

$$\begin{aligned} |D(u,1)| &= \sum_{i=0}^{1} \binom{n}{i} (q-1)^i \\ &= \binom{n}{0}(q-1)^0 + \binom{n}{1}(q-1)^1 \\ &= \binom{2^m-1}{0}(2-1)^0 + \binom{2^m-1}{1}(2-1)^1 \\ &= \frac{(2^m-1)!}{(2^m-1)!} + \frac{(2^m-1)!}{(2^m-2)!} \\ &= 1 + 2^m - 1 \\ &= 2^m. \end{aligned}$$

Pelo Lema 2.2.3, temos que se u e v são palavras distintas de C então

$$D(u,1) \cap D(v,1) = \varnothing,$$

ou seja, os discos são disjuntos.

Assim, como o número de elementos de C é 2^k, temos

$$|\bigcup_{u \in C} D(u,1)| = (2^m)2^k = (2^m)2^{n-m} = 2^n = |\mathbb{F}_2^n|,$$

e consequentemente,

$$\bigcup_{u \in C} D(u,1) = \mathbb{F}_2^n.$$

Portanto, C é um código perfeito.

■

Note que, um código de *Hamming* de ordem m é um código separável pela distância máxima se, e somente se,

$$\begin{aligned} d = n - k + 1 \Rightarrow 3 &= 2^m - 1 - (2^m - 1 - m) + 1 \\ \Rightarrow 3 &= 2^m - 1 - 2^m + 1 + m + 1 \\ \Rightarrow 3 &= m + 1 \\ \Rightarrow m &= 2. \end{aligned}$$

Aplicação 2. (*Código da Mariner 9*)

Segundo Siddiqi (2018, p. 103), Mariner 9 foi uma missão com o objetivo de orbitar Marte. A espaçonave Mariner-71I / Mariner-I da missão se tornou o primeiro objeto feito pelo homem a entrar em órbita de outro planeta em 14 de Novembro de 1971. A sonda capturou 7329 imagens de Marte e mapeou cerca de 85% da superfície do planeta. Esta missão foi projetada e desenvolvida pela NASA[1]/JPL[2].

O código usado na nave espacial Mariner-71I / Mariner-I é um membro particular de uma família de códigos $R(1, m)$ definidos sobre um corpo finito $\mathbb{F}_2$, cujas matrizes de codificação sejam da seguinte forma:

Considere a matriz A cuja ordem é $m \times 2^m$, de modo que, as primeiras $2^m - 1$ colunas sejam a matriz de *Hamming* H de ordem m vista na Aplicação 1 e a última coluna de A seja o vetor nulo transposto de $\mathbb{F}_2^m$.

Logo,

$$A = (H \mid \mathbf{0}^t).$$

Construíremos a matriz de codificação G com 2^m linhas e $m + 1$ colunas, de modo que, os elementos da primeira coluna sejam todos 1 e o bloco direito da matriz G seja A^t. Denotaremos o vetor linha $(1, 1, \cdots, 1) \in \mathbb{F}_2^n$ por $\mathbf{1}$, e consequentemente, o vetor coluna por $\mathbf{1}^t$.

Assim,

$$G = \begin{bmatrix} \mathbf{1}^t & H^t \\ 1 & \mathbf{0} \end{bmatrix}.$$

Proposição 5.3.2. *Os parâmetros do código $R(1, m)$ são $[2^m, m + 1]$.*

Demonstração. Recorde que, as colunas de uma matriz de codificação G geram o código $R(1, m)$. Assim, como G tem $m + 1$ colunas e elas são linearmente independentes em virtude do bloco H, e do fato dos elementos da primeira coluna serem todos 1, nos garante que essa coluna é linearmente independente das demais. Isto posto, temos que a dimensão de $R(1, m)$ é $m + 1$ e G tem 2^m linhas, segue que o comprimento de $R(1, m)$ é igual 2^m.

[1] National Aeronautics and Space Administration.
[2] Jet Propulsion Laboratory.

■

Agora vamos verificar que a distância mínima de $R(1,m)$ é igual a 2^{m-1}. Para isso, vamos construir a família de códigos $R(1,m)$ de maneira recursiva.

i) Se $m=0$, temos que
$$R(1,0)=\{0,1\}=\mathbb{F}_2.$$

ii) Se $m=1$, temos que
$$R(1,1)=\mathbb{F}_2\times\mathbb{F}_2=\{00,01,10,11\}=\mathbb{F}_2^2$$

iii) Se $m>1$, temos que
$$R(1,m)=\left\{u\ u, u\ (u+\mathbf{1})\mid \forall u\in R(1,m-1) \text{ e } \mathbf{1}=\underbrace{11\cdots 1}_{2^{m-1}}\right\}.$$

Exemplo 5.3.2. Se $m=2$, temos um código $R(1,2)$ sobre um corpo finito $\mathbb{F}_2$, tal que
$$R(1,2)=\Big\{u\ u, u\ (u+11)\mid \forall u\in R(1,1)\Big\}.$$

Logo, para $u\ u$ temos

$$\begin{array}{cccc} 00 & 00 & \longrightarrow & 0000\\ 01 & 01 & \longrightarrow & 0101\\ 10 & 10 & \longrightarrow & 1010\\ 11 & 11 & \longrightarrow & 1111 \end{array}$$

e para $u\ (u+11)$ temos

$$\begin{array}{cccc} 00 & 00{+}11 & \longrightarrow & 0011\\ 01 & 01{+}11 & \longrightarrow & 0110\\ 10 & 10{+}11 & \longrightarrow & 1001\\ 11 & 11{+}11 & \longrightarrow & 1100 \end{array}$$

Portanto,
$$R(1,2)=\{0000,0101,1010,1111,0011,0110,1001,1100\}=\mathbb{F}_2^4.$$

Análogamente, a partir de $R(1,2)$, obtemos
$$R(1,3)=\begin{Bmatrix} 00000000 & 01010101 & 10101010 & 11111111\\ 00110011 & 01100110 & 10011001 & 11001100\\ 00001111 & 01011010 & 10100101 & 11110000\\ 00111100 & 01101001 & 10010110 & 11000011 \end{Bmatrix}.$$

Através de $R(1,3)$, obtemos $R(1,4)$ e assim sucessivamente.

Teorema 5.3.3. *A distância mínima do código $R(1,m)$ é 2^{m-1}.*

Demonstração. Vamos mostrar que o peso de qualquer palavra de $R(1,m)$, exceto as palavras $\mathbf{0} = \underbrace{00\cdots0}_{2^m}$ e $\mathbf{1} = \underbrace{11\cdots1}_{2^m}$ tem peso 2^{m-1}. Segue que, $\omega(\mathbf{0}) = 0$ e $\omega(\mathbf{1}) = 2^m$.

Por indução sobre m, temos:

Para $m = 1$, temos o código $R(1,1) = \{00, 01, 10, 11\}$, onde para qualquer palavra $c \in R(1,1)$, tal que $c \neq \mathbf{0} = 00$ e $c \neq \mathbf{1} = 11$, tem peso $2^{m-1} = 2^{1-1} = 1$. De fato, as palavras 01 e 10 possuem peso 1. Logo, o resultado é válido para $m = 1$.

Agora considere como, hipótese de indução, que o resultado seja válido para $m-1$. Assim, qualquer palavra $c \in R(1,m-1)$, tal que $c \neq \mathbf{0} = \underbrace{00\cdots0}_{2^{m-1}}$ e $c \neq \mathbf{1} = \underbrace{11\cdots1}_{2^{m-1}}$, tem peso $2^{(m-1)-1} = 2^{m-2}$.

Seja $c \in R(1,m)$, tal que $c \neq \mathbf{0} = \underbrace{00\cdots0}_{2^m}$ e $c \neq \mathbf{1} = \underbrace{11\cdots1}_{2^m}$. Temos as seguintes possibilidades:

i) Para $c = u\ u$, onde $u \in R(1,m-1)$. Temos que, como $c \neq \underbrace{00\cdots0}_{2^m}$ e $c \neq \underbrace{11\cdots1}_{2^m}$. Segue que, $u \neq \underbrace{00\cdots0}_{2^{m-1}}$ e $u \neq \underbrace{11\cdots1}_{2^{m-1}}$. Por hipótese de indução, $\omega(u) = 2^{m-2}$, ou seja, u tem 2^{m-2} componentes iguais a 1. Logo, $c = u\ u$ terá o dobro de componentes iguais a 1, isto é, $2(2^{m-2}) = 2^{m-1}$ componentes iguais a 1. Portanto,

$$\omega(c) = 2^{m-1}.$$

ii) Para $c = u\ (u + \mathbf{1})$, onde $u \in R(1,m-1)$ e $\mathbf{1} = \underbrace{11\cdots1}_{2^{m-1}}$. Temos que

ii.i) Se $u = \underbrace{00\cdots0}_{2^{m-1}}$ então $u + \mathbf{1} = \underbrace{11\cdots1}_{2^{m-1}}$. Logo,

$$c = \underbrace{00\cdots0}_{2^{m-1}}\underbrace{11\cdots1}_{2^{m-1}} \Rightarrow \omega(c) = 2^{m-1}.$$

ii.ii) Se $u = \underbrace{11\cdots1}_{2^{m-1}}$ então $u + \mathbf{1} = \underbrace{00\cdots0}_{2^{m-1}}$. Logo,

$$c = \underbrace{11\cdots1}_{2^{m-1}}\underbrace{00\cdots0}_{2^{m-1}} \Rightarrow \omega(c) = 2^{m-1}.$$

ii.iii) Por hipótese de indução, se $u \neq \underbrace{00\cdots0}_{2^{m-1}}$ e $u \neq \underbrace{11\cdots1}_{2^{m-1}}$ então $\omega(u) = 2^{m-2} = \frac{2^{m-1}}{2}$, isso equivale a dizer que, metade das componentes de u são iguais a zero e a outra metade iguais a 1.

Observe que, em $u+\mathbf{1}$ as componentes 0 viram 1, e as componentes 1 viram 0, ou seja, $\omega(u) = \omega(u+\mathbf{1}) = 2^{m-2}$. Logo, $c = u\ (u+\mathbf{1})$ terá o dobro de componentes iguais a 1, isto é, $2(2^{m-2}) = 2^{m-1}$ componentes iguais a 1. Portanto,

$$\omega(c) = 2^{m-1}.$$

■

Definição 5.3.3. Um código $R(1, m)$, com a matriz de codificação G construída acima e os parâmetros $[n, k, d]$, tais que $n = 2^m,\ k = m + 1$ e $d = 2^{m-1}$, é dito **código de *Reed-Muller* de primeria ordem**.

Exemplo 5.3.3. Considere o código $R(1, 4)$, sobre um corpo finito $\mathbb{F}_2$. Como já vimos no Exemplo 5.3.1 a forma da matriz de *Hamming* H de ordem 4, temos que

$$G = \begin{array}{c} \begin{matrix} u_1 & u_2 & u_3 & u_4 & u_5 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

Observe que, ao somarmos quaisquer duas colunas de G, temos que o peso é

$$2^{m-1} = 2^4 = 8.$$

Como exemplo numérico, considere a palavra $u = u_2 + u_4 \in R$, temos que

$$u = (0000111111110000) + (1011011001100100) = (1011100110010100).$$

Portanto,

$$\omega(u) = \omega(1011100110010100) = 8.$$

Assim, é fácil verificar que o peso será igual a 8 para qualquer palavra do código $R(1,4)$.

O código utilizado na Mariner 9 corresponde ao caso em que $m = 5$. Logo, é um código de *Reed-Muller* de Primeira Ordem $R(1,5)$ cujos parâmetros são $[32, 6, 16]$ e $\kappa = 7$.

6 CONSIDERAÇÕES FINAIS

É possível concluir a partir deste trabalho que a teoria dos códigos corretores de erros tem diversas aplicações em diferentes ramos da ciência, como por exemplo na matemática, computação, engenharia elétrica, estatística, etc. A nível elementar e introdutório, este trabalho depende apenas de conceitos básicos de álgebra, álgebra linear e técnicas de contagem, conceitos esses que podem ser aplicados por professores do ensino básico para ensinar matrizes e sistema binário. Deixo como proposta de trabalhos futuros a elaboração de planos de aula que utilizam códigos de Hamming ou códigos de Reed-Muller de Primeira Ordem. Um bom exemplo é Dias (2005, p. 20-22), ele elabora um plano de aula para 2º ano do Ensino Médio que utiliza o código da Mariner 9.

Em muitas disciplinas durante a graduação temos um olhar bastante abstrato sobre a matemática, diversas vezes não vemos aplicações e uma relação entre uma disciplina e outra. Por esse motivo, foram apresentadas duas aplicações de códigos lineares que utilizam conceitos de álgebra abstrata, teoria dos números e álgebra linear, criando assim uma relação entre esses conteúdos. É possível apresentar estas aplicações para alunos do curso de Licenciatura em Matemática, servindo como motivação aos estudos de anéis, corpos, corpos finitos, inteiros módulo m, transformações lineares, etc.

Em trabalhos futuros é possível utilizar-se dos conceitos apresentados no desenvolvimento desta pesquisa e no Apêndice A, que trata-se de polinômios, para estudar outras aplicações de códigos lineares, decodificação ou ainda estudar outras classes de códigos corretores de erros. Também é viável, a partir do Apêndice A, elaborar um trabalho sobre os códigos de Reed-Solomon, que utiliza conceitos de cálculo, cálculo numérico e anéis de polinômios.

Espero que, este trabalho sirva para estimular alunos e professores na elaboração de novas aulas e desenvolvimento de pesquisas, de modo que, aprofundem os estudos em diversos campos da Álgebra, contribuindo de maneira positiva com o desenvolvimento da educação e da ciência.

REFERÊNCIAS

BOLDRINI, J. L. et al. **Álgebra Linear**. 3. ed. São Paulo, SP: Harper & Row do Brasil, 1980.

BRASIL. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais: Matemática**. Brasília, DF: MEC/SEF, 1997.

BRASIL. Ministério da Educação. **SEDCITEC 2019: Semanas Acadêmicas**. 2019. Disponível em: <https://spo.ifsp.edu.br/destaques/1896-sedcitec-2019-9-a-13-de-setembro>. Acesso em: 08 de nov. de 2019.

CALLIOLI, C. A.; DOMINGUES, H. H.; COSTA, R. C. F. **Álgebra linear e Aplicações**. 6. ed. São Paulo: Atual, 1990.

COELHO, F. U.; LOURENÇO, M. L. **Um Curso de Álgebra Linear**. 1. ed. São Paulo: Edusp, 2015. v. 1.

COUTO, O. S. **Isometrias em Teoria de Códigos**. Dissertação (Mestrado em Matemática para Professores) — Faculdade de Ciência da Universidade do Porto, São Paulo, SP, 2010.

DEVS, 4. **Ferramentas Online: Gerador de CPF**. 2012. Disponível em: <https://www.4devs.com.br/gerador_de_cpf>. Acesso em: 29 de set. de 2019.

DIAS, J. S. **O Código da Mariner 9**. Dissertação (Programa de Mestrado Profissional em rede Nacional) — Universidade Federal de São João del-Rei, São Paulo, SP, 2005.

DOMINGUES, H. H.; IEZZI, G. **Álgebra Moderna**. 5. ed. São Paulo: Saraiva, 2018.

GARCIA, A.; LEQUAIN, Y. **Elementos de Álgebra**. 6. ed. Rio de Janeiro, RJ: IMPA, 2018. (Projeto Euclides).

GONÇALVES, A. **Introdução à Álgebra**. 6. ed. Rio de Janeiro, RJ: IMPA, 2017. (Projeto Euclides).

HEFEZ, A.; VILLELA, M. L. T. **Códigos Corretores de Erros**. 2. ed. Rio de Janeiro, RJ: IMPA, 2008. (Série de Computação e Matemática).

HOFFMAN, K.; KUNZE, R. **Álgebra Linear**. São Paulo, SP: Editora da Universidade de São Paulo e Polígono, 1970.

JACOBSON, N. **Basic algebra 1**. New York, NY: W. H. Freeman and Company, 1985.

LANG, S. **Undergraduate Algebra**. New Haven, CT: Springer, 2004.

LEE, C. Y. Some properties of nonbinary error-correcting codes*. **IRE Transactions on Information Theory**, v. 4, p. 77–82, 1958.

LIMA, E. L. **Elementos de Topologia Geral**. 1. ed. Rio de Janeiro, RJ: IMPA, 1970. (Elementos de Matemática).

LUCHETTA, V. O. J. **Códigos Cíclicos como Ideais em Álgebra de Grupos**. Dissertação (Mestrado em Matemática) — Instituto de Matemática e Estatística - Universidade de São Paulo, São Paulo, SP, 2005.

MACWILLIANS, F. J.; SLOANE, N. J. A. **The Theory of Error-Correcting Codes**. 3. ed. Amsterdam, NY. Oxford: North-Holland Publishing Company, 1981.

MARCONI, M. de A.; LAKATOS, E. M. **Fundamentos de Metodologia Científica**. 5. ed. São Paulo, SP: Editora Atlas, 2003.

MCCOY, N. H. **The Theory of Rings**. Bronx, NY: Chelsea Publishing Company, 1973.

MILIES, F. C. P. Breve introdução à teoria dos códigos corretores de erros. **Colóquio de Matemática da Região Centro-Oeste**, Campo Grande, MS, 2009.

MILIES, F. C. P.; COELHO, S. P. **Números: Uma Introdução à Matemática**. 3. ed. São Paulo, SP: Editora da Universidade de São Paulo, 2013.

ROMAN, S. **Coding and Information Theory**. 1. ed. Fullerton, CA: Springer, 1992. (Graduate Texts in Mathematics).

SANTOS, R. J. **Álgebra Linear e Aplicações**. 2010. Disponível em: <https://www.ime.unicamp.br/~deleo/MA327/ld2.pdf>. Acesso em: 26 de maio de 2020.

SIDDIQI, A. A. **Beyond Earth : a chronicle of deep space exploration, 1958–2016**. 2. ed. Washington, DC: NASA History Program Office, 2018.

STEINBRUCH, A.; WINTERLE, P. **Álgebra Linear**. 2. ed. São Paulo, SP: McGraw-Hill, 1987.

TARCHA, A. A. G. **Um estudo dos três problemas clássicos da geometria**. Monografia (Graduação em Licenciatura em Matemática) — Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, São Paulo, SP, 2019.

APÊNDICE A – POLINÔMIOS

Neste apêndice vamos construir os anéis de polinômios através de sequências, desenvolver a teoria da divisibilidade e apresentar o algoritmo de Euclides para polinômios. Para a composição deste apêndice foram utilizados os seguintes materiais: Garcia e Lequain (2018); Lang (2004); Jacobson (1985); Milies e Coelho (2013); Hefez e Villela (2008) e Tarcha (2019).

A.1 Anéis de Polinômios

Seja $(A, +, \cdot)$ um anel comutativo. Denotamos por $A[X]$ o conjunto formado pelas sequências infinitas $(a_0, a_1, \cdots)$, que tem apenas um número finito de termos a_i não nulos.

As sequências $(a_0, a_1, \cdots)$ e $(b_0, b_1, \cdots)$ são consideradas iguais se, e somente se, $a_i = b_i$ para todo $i \in \mathbb{N}$. Em outras palavras, $A[X]$ é o conjunto das aplicações $i \longmapsto a_i$ do conjunto dos naturais $\mathbb{N}$ no anel A, tal que $a_i = 0$ para algum i suficientemente grande.

Um **polinômio numa variável sobre** A é uma sequência $(a_0, a_1, \cdots, a_n, \cdots)$, onde $a_i \in A$ para todo índice e $a_i \neq 0$ somente para um número finito de índices.

Seja $A[X] = \{\text{polinômios numa variável sobre } A\}$. No conjunto $A[X]$ definimos as seguintes operações:

$$\begin{array}{rccc} +: & A[X] \times A[X] & \longrightarrow & A[X] \\ & (a_0, a_1, \cdots, a_n, \cdots), (b_0, b_1, \cdots, b_n, \cdots) & \longmapsto & (a_0 + b_0, a_1 + b_1, \cdots, a_n + b_n, \cdots) \end{array}$$

$$\begin{array}{rccc} \cdot: & A[X] \times A[X] & \longrightarrow & A[X] \\ & (a_0, a_1, \cdots, a_n, \cdots), (b_0, b_1, \cdots, b_n, \cdots) & \longmapsto & (\alpha_0, \alpha_1, \cdots, \alpha_n, \cdots), \end{array}$$

onde

$$\begin{aligned} \alpha_0 &= a_0 \cdot b_0 \\ \alpha_1 &= a_0 \cdot b_1 + a_1 \cdot b_0 \\ &\vdots \\ \alpha_n &= a_0 \cdot b_n + a_1 \cdot b_{n-1} + a_2 \cdot b_{n-2} + \cdots + a_{n-1} \cdot b_1 + a_n \cdot b_0 \\ &\vdots \end{aligned}$$

Como não há possibilidade de confusão, e por razões de ordem prática omitiremos a utilização do símbolo $\cdot$ para multiplicação. Por exemplo, para $a_0 \cdot b_0$ escreveremos como $a_0 b_0$.

Teorema A.1.1. *Com as duas operações definidas acima, $A[X]$ é um anel comutativo.*

Demonstração. Sejam $p = (a_0, a_1, \cdots), q = (b_0, b_1, \cdots), t = (c_0, c_1, \cdots) \in A[X]$ quaisquer, temos:

i) **Associativa da adição:** $p + (q + t) = (p + q) + t$.

$$\begin{aligned} p + (q + t) &= (a_0, a_1, \cdots) + [(b_0, b_1, \cdots) + (c_0, c_1, \cdots)] \\ &= (a_0, a_1, \cdots) + (b_0 + c_0, b_1 + c_1, \cdots) \\ &= (a_0 + (b_0 + c_0), a_1 + (b_1 + c_1), \cdots) \\ &= ((a_0 + b_0) + c_0, (a_1 + b_1) + c_1, \cdots) \quad \text{(assoc. da ad. de } A) \\ &= (a_0 + b_0, a_1 + b_1, \cdots) + (c_0, c_1, \cdots) \\ &= [(a_0, a_1, \cdots) + (b_0, b_1, \cdots)] + (c_0, c_1, \cdots) \\ &= (p + q) + t. \end{aligned}$$

ii) **Comutativa da adição:** $p + q = q + p$.

$$\begin{aligned} p + q &= (a_0, a_1, \cdots) + (b_0, b_1, \cdots) \\ &= (a_0 + b_0, a_1 + b_1, \cdots) \\ &= (b_0 + a_0, b_1 + a_1, \cdots) \qquad \text{(comut. da ad. de } A) \\ &= (b_0, b_1, \cdots) + (a_0, a_1, \cdots) \\ &= q + p. \end{aligned}$$

iii) **Existência do elemento neutro aditivo:** $\exists! g \in A[X]$ tal que $p + g = p$.

Seja $g = (e_0, e_1, \cdots) \in A[X]$. Suponha que a igualdade a seguir seja válida

$$\begin{aligned} p + g = p &\Rightarrow (a_0, a_1, \cdots) + (e_0, e_1, \cdots) = (a_0, a_1, \cdots) \\ &\Rightarrow (a_0 + e_0, a_1 + e_1, \cdots) = (a_0, a_1, \cdots). \end{aligned}$$

Logo,

$$\begin{aligned} a_0 + e_0 &= a_0 \\ a_1 + e_1 &= a_1 \\ &\vdots \end{aligned}$$

Como todos os elementos das sequências p e g pertencem ao anel A, e existe elemento neutro aditivo em A, temos

$$e_0 = e_1 = \cdots = 0_A.$$

Portanto, existe um único $g \in A[X]$ tal que $g = (0_A, 0_A, \cdots)$.

iv) **Existência de simétricos aditivos:** $\forall p \in A[X], \exists! p' \in A[X]$ tal que $p + p' = 0_{A[X]}$, onde $0_{A[X]} = (0_A, 0_A, \cdots)$.

Seja $p' = (a'_0, a'_1, \cdots) \in A[X]$. Suponha que a igualdade a seguir seja válida

$$\begin{aligned} p + p' = 0_{A[X]} &\Rightarrow (a_0, a_1, \cdots) + (a'_0, a'_1, \cdots) = (0_A, 0_A, \cdots) \\ &\Rightarrow (a_0 + a'_0, a_1 + a'_1, \cdots) = (0_A, 0_A, \cdots). \end{aligned}$$

Logo,

$$\begin{aligned} a_0 + a'_0 &= 0_A \\ a_1 + a'_1 &= 0_A \\ &\vdots \end{aligned}$$

Como todos os elementos das sequências p e p' pertencem ao anel A, e existem simétricos aditivos em A, segue que

$$\begin{aligned} a'_0 &= -\, a_0 \\ a'_1 &= -\, a_1 \\ &\vdots \end{aligned}$$

Portanto, para todo $p \in A[X]$, existe um $p' \in A[X]$ tal que $p' = (-a_0, -a_1, \cdots)$.

v) **Associativa da multiplicação:** $p(qt) = (pq)t$.

$$\begin{aligned} p(qt) &= (a_0, a_1, \cdots)[(b_0, b_1, \cdots)(c_0, c_1, \cdots)] \\ &= (a_0, a_1, \cdots)(b_0c_0, b_0c_1 + b_1c_0, \cdots) \\ &= (a_0(b_0c_0), a_0(b_0c_1 + b_1c_0) + a_1(b_0c_0), \cdots) \\ &= ((a_0b_0)c_0, (a_0b_0)c_1 + (a_0b_1)c_0 + (a_1b_0)c_0, \cdots) \qquad \text{(assoc. e dist. de } A) \\ &= ((a_0b_0)c_0, (a_0b_0)c_1 + (a_0b_1 + a_1b_0)c_0, \cdots) \qquad \text{(dist. de } A) \\ &= (a_0b_0, a_0b_1 + a_1b_0, \cdots)(c_0, c_1, \cdots) \\ &= [(a_0, a_1, \cdots)(b_0, b_1, \cdots)](c_0, c_1, \cdots) \\ &= (pq)t. \end{aligned}$$

vi) **Comutativa da multiplicação:** $pq = qp$.

$$\begin{aligned} pq &= (a_0, a_1, \cdots)(b_0, b_1, \cdots) \\ &= (a_0b_0, a_0b_1 + a_1b_0, \cdots) \\ &= (b_0a_0, b_1a_0 + b_0a_1, \cdots) \qquad \text{(comut. da mult. de } A) \\ &= (b_0a_0, b_0a_1 + b_1a_0, \cdots) \qquad \text{(comut. da ad. de } A) \\ &= (b_0, b_1, \cdots)(a_0, a_1, \cdots) \\ &= qp. \end{aligned}$$

vii) **Distributiva da multiplicação em relação a adição:** $p(q+t) = pq + pt$ e $(p+q)t = pt + qt$.

$$\begin{aligned}
p(q+t) &= (a_0, a_1, \cdots)[(b_0, b_1, \cdots) + (c_0, c_1, \cdots)] \\
&= (a_0, a_1, \cdots)(b_0 + c_0, b_1 + c_1, \cdots) \\
&= (a_0(b_0 + c_0), a_0(b_1 + c_1) + a_1(b_0 + c_0), \cdots) \\
&= (a_0b_0 + a_0c_0, [a_0b_1 + a_0c_1] + [a_1b_0 + a_1c_0], \cdots) \qquad \text{(A.1)} \\
&= (a_0b_0 + a_0c_0, [a_0b_1 + a_1b_0] + [a_0c_1 + a_1c_0], \cdots) \qquad \text{(A.2)} \\
&= (a_0b_0, a_0b_1 + a_1b_0, \cdots) + (a_0c_0, a_0c_1 + a_1c_0, \cdots) \\
&= [(a_0, a_1, \cdots)(b_0, b_1, \cdots)] + [(a_0, a_1, \cdots)(c_0, c_1, \cdots)] \\
&= pq + pt.
\end{aligned}$$

Note que, nas equações A.1 e A.2 todos os elementos pertencem ao anel A, então as propriedades distributiva, associativa e comutativa são válidas.

Como a comutativa das operações de adição e multiplicação são válidas em $A[X]$, temos que $(p+q)t = pt + qt$ também é válido.

Portanto, podemos concluir que $A[X]$ é um anel comutativo.

■

Proposição A.1.2. *Se A é domínio de integridade então $A[X]$ é um domínio de integridade.*

Demonstração. Note que no Teorema A.1.1 demonstramos que $A[X]$ é um anel comutativo. Sejam $p = (a_0, a_1, \cdots), q = (b_0, b_1, \cdots) \in A[X]$ quaisquer, desse modo, vamos mostrar as seguintes propriedades:

i) **Existência do elemento neutro multiplicativo:** $\exists! h \in A[X]$ tal que $ph = p$ e $h \neq 0_{A[X]}$.

Seja $h = (s_0, s_1, \cdots) \in A[X]$. Suponha que a igualdade a seguir é válida

$$\begin{aligned}
ph = p &\Rightarrow (a_0, a_1, \cdots)(s_0, s_1, \cdots) = (a_0, a_1, \cdots) \\
&\Rightarrow (a_0s_0, a_0s_1 + a_1s_0, \cdots) = (a_0, a_1, \cdots).
\end{aligned}$$

Logo,

$$\begin{aligned}
a_0s_0 &= a_0 \\
a_0s_1 + a_1s_0 &= a_1 \\
&\vdots
\end{aligned}$$

Como todos os elementos das sequências obtidas pertencem ao anel de integridade A, logo s_0 é elemento neutro multiplicativo de A, ou seja, $s_0 = 1_A$. Por igualdade de sequências, obtemos que

$$\begin{aligned} a_0s_1 + a_1s_0 = a_1 &\Rightarrow a_0s_1 + a_1 1_A = a_1 \\ &\Rightarrow a_0s_1 + a_1 = a_1 \\ &\Rightarrow a_0s_1 = 0_A. \end{aligned}$$

Como a_0 é um elemento qualquer da sequência p, temos que $s_1 = 0_A$. De maneira análoga, se substituirmos s_0 e s_1 na próxima igualdade $a_0s_2 + a_1s_1 + a_2s_0 = a_2$, obtemos que $s_2 = 0_A$, logo de maneira recursiva obtemos que

$$s_1 = s_2 = \cdots = 0_A.$$

Portanto, existe um $h \in A[X]$ tal que $h = (1_A, 0_A, 0_A, \cdots)$.

ii) **Lei do anulamento do produto:** $pq = 0_{A[X]} \Rightarrow p = 0_{A[X]}$ ou $q = 0_{A[X]}$.

Suponha que a igualdade a seguir é válida

$$\begin{aligned} pq = 0_{A[X]} &\Rightarrow (a_0, a_1, \cdots)(b_0, b_1, \cdots) = (0_A, 0_A, \cdots) \\ &\Rightarrow (a_0b_0, a_0b_1 + a_1b_0, \cdots) = (0_A, 0_A, \cdots). \end{aligned}$$

Logo,

$$a_0b_0 = 0_A \tag{A.3}$$

$$a_0b_1 + a_1b_0 = 0_A \tag{A.4}$$

$$\vdots$$

Como A é anel de integridade, da igualdade A.3 obtemos que $a_0 = 0_A$ ou $b_0 = 0_A$. Sem perda de generalidade, considere $a_0 = 0_A$ e $b_0 \neq 0_A$. Desse modo, de A.4 temos

$$\begin{aligned} a_0b_1 + a_1b_0 = 0_A &\Rightarrow 0_Ab_1 + a_1b_0 = 0_A \\ &\Rightarrow a_1b_0 = 0_A. \end{aligned} \tag{A.5}$$

Assim, da igualdade A.5 obtemos que $a_1 = 0_A$, pois $b_0 \neq 0_A$. Portanto, de maneira recursiva, temos que $a_0 = a_1 = a_2 = \cdots = a_n = \cdots = 0_A$, logo $p = (0_A, 0_A, \cdots) = 0_{A[X]}$.

Por outro lado, se considerarmos $a_0 \neq 0_A$ e $b_0 = 0_A$, obtemos que $q = (0_A, 0_A, \cdots) = 0_{A[X]}$. Logo, vale a lei do anulamento do produto em $A[X]$.

Portanto, $A[X]$ é um anel comutativo com unidade e não possui divisores de zero, sendo assim, $A[X]$ é um domínio de integridade.

■

Seja a sequência $(a_0, a_1, \cdots) \in A[X]$, então o símbolo $(a_0, a_1, \cdots)^n$ denomina o elemento

$$\underbrace{(a_0, a_1, \cdots)(a_0, a_1, \cdots)\cdots(a_0, a_1, \cdots)}_{n \text{ vezes}}.$$

Utilizando as operações de adição e multiplicação definidas em $A[X]$, temos

$$\begin{aligned}
(a_0, 0_A, 0_A, \cdots) &= (a_0, 0_A, 0_A, \cdots)(1_A, 0_A, 0_A, \cdots) \\
(0_A, a_1, 0_A, 0_A, \cdots) &= (a_1, 0_A, 0_A, \cdots)(0_A, 1_A, 0_A, 0_A, \cdots) \\
&\vdots \\
\underbrace{(0_A, \cdots, 0_A, a_n, 0_A, 0_A, \cdots)}_{a_n \text{ na posição } n+1} &= (a_n, 0_A, 0_A, \cdots)\underbrace{(0_A, \cdots, 0_A, 1_A, 0_A, 0_A, \cdots)}_{1_A \text{ na posição } n+1},
\end{aligned}$$

note que, a_n encontra-se na posição $n+1$, pois nossa sequência começa na coordenada a_0, consequentemente 1_A encontra-se na mesma posição $n+1$.

Observe que

$$\begin{aligned}
(1_A, 0_A, 0_A, \cdots) &= (0_A, 1_A, 0_A, 0_A, \cdots)^0 \\
(0_A, 1_A, 0_A, 0_A, \cdots) &= (0_A, 1_A, 0_A, 0_A, \cdots)^1 \\
(0_A, 0_A, 1_A, 0_A, 0_A \cdots) &= (0_A, 1_A, 0_A, 0_A, \cdots)^2 \\
&\vdots \\
\underbrace{(0_A, \cdots, 0_A, 1_A, 0_A, 0_A, \cdots)}_{1_A \text{ na posição } n+1} &= (0_A, 1_A, 0_A, 0_A, \cdots)^n.
\end{aligned}$$

Portanto, podemos escrever uma sequência como

$$\begin{aligned}
(a_0, a_1, \cdots, a_n, 0_A, 0_A, \cdots) =& [(a_0, 0_A, 0_A, \cdots)(0_A, 1_A, 0_A, 0_A, \cdots)^0] \\
&+ [(a_1, 0_A, 0_A, \cdots)(0_A, 1_A, 0_A, 0_A, \cdots)^1] \\
&+ [(a_2, 0_A, 0_A, \cdots)(0_A, 1_A, 0_A, 0_A, \cdots)^2] \\
&+ \cdots \\
&+ [(a_n, 0_A, 0_A, \cdots)(0_A, 1_A, 0_A, 0_A, \cdots)^n].
\end{aligned}$$

Para simplificar essa expressão, vamos usar o símbolo x para representar o elemento $(0_A, 1_A, 0_A, 0_A, \cdots)$ e a_i para respresentar $(a_i, 0_A, 0_A, \cdots)$, tal que $i \in \mathbb{N}$.

Logo, podemos representar o elemento $p = (a_0, a_1, \cdots, a_n, 0, \cdots) \in A[x]$ por $p(x) = a_0 + a_1 x + \cdots + a_n x^n$.

Definição A.1.1. Seja $(A, +, \cdot)$ um domínio de integridade. O conjunto $A[X]$ dos polinômios numa variável sobre A é um **anel de polinômios** e pode ser representado por

$$A[X] = \left\{ \sum_{i=0}^{n} a_i x^i \mid n \in \mathbb{N} \text{ e } a_i \in A \right\}.$$

Definição A.1.2. Se $p(x) = \sum_{i=0}^{n} a_i x^i$ é um polinômio não nulo em $A[X]$ com $a_n \neq 0$, define-se o **grau** de $p(x)$ como sendo o inteiro n. O coeficiente a_n será chamado de **coeficiente líder** de $p(x)$. Quando o grau do polinômio $p(x)$ é igual a n e o coeficiente líder de $p(x)$ for igual a 1, o polinômio é dito **mônico**.

Se $p(x) = 0$, não se define grau de $p(x)$. Para $p(x) \neq 0$, denotaremos o grau por $gr(p(x))$. Observe que, $gr(p(x)) = 0$ se, e somente se, $p(x) = a_0 \in A \setminus \{0\}$.

Proposição A.1.3. *Seja A domínio de integridade. Temos que*

i) Se $p(x), q(x) \in A[X] \setminus \{0\}$, então $gr(p(x)q(x)) = gr(p(x)) + gr(q(x))$.

ii) Os elementos invertíveis de $A[X]$ são os elementos invertíveis de A.

Demonstração. i) Sejam $p(x) = a_0 + a_1 x + \cdots + a_n x^n, q(x) = b_0 + b_1 x + \cdots + b_m x^m \in A[X] \setminus \{0\}$, com a_n e b_m diferentes de zero, temos que

$$\begin{aligned} p(x)q(x) &= (a_0 + a_1 x + \cdots + a_n x^n)(b_0 + b_1 x + \cdots + b_m x^m) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \cdots + a_n b_m x^{n+m}. \end{aligned}$$

Como a_n e b_m são elementos não nulos de A, e A é um domínio de integridade, então $a_n b_n \neq 0$. Logo, pela Definição A.1.2, temos

$$gr(p(x)q(x)) = n + m = gr(p(x)) + gr(q(x)).$$

ii) Seja $p(x)$ um elemento invertível de $A[X]$, logo existe um elemento $q(x) \in A[X]$ tal que $p(x)q(x) = 1_{A[X]}$. Pela Proposição A.1.3 ii) temos que $A[X]$ é domínio de integridade, logo $p(x) \neq 0_{A[X]}$ e $q(x) \neq 0_{A[X]}$, e pela Proposição A.1.3 i), temos que

$$gr(p(x)q(x)) = gr(p(x)) + gr(q(x)) = gr(1_{A[X]}) = 0.$$

Desse modo, $gr(p(x)) = 0$ e $gr(q(x)) = 0$, então $p(x), q(x) \in A \setminus \{0\}$, logo, $p(x)$ é um elemento invertível em A. Por outro lado, a recíproca é imediata pois, seja $a_0 \in A$ um elemento invertíverel em A, então podemos representar a_0 como um polinômio de grau 0, portanto a_0 é um elemento invertível em $A[X]$.

■

A partir de agora vamos supor que A é um corpo qualquer $\mathbb{K}$, logo pela Proposição A.1.3 ii) $\mathbb{K}[X]$ é um domínio de integridade.

Definição A.1.3. Sejam $p(x), q(x) \in \mathbb{K}[X]$, $p(x)$ e $q(x)$ são ditos **polinômios associados** se, e somente se, existe um elemento não nulo $k \in \mathbb{K}$ tal que $p(x) = kq(x)$.

Seja $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{K}[X]$, com $a_n \neq 0$, logo, $p(x) = a_n(\frac{a_0}{a_n} + \frac{a_1}{a_n}x + \cdots + x^n)$. Portanto, todo polinômio não nulo é associado a um único polinômio mônico.

A.2 Divisão de Polinômios

Tendo em vista o Algoritmo de *Euclides*, nessa seção vamos desenvolver a teoria da divisibilidade para poliniômios.

Teorema A.2.1. ***(Divisão Euclidiana para polinômios)*** *Dados dois polinômios $p(x)$ e $g(x)$ pertencentes ao anel de polinômios $\mathbb{K}[X]$, com coeficientes num corpo $\mathbb{K}$, $g(x)$ não nulo e o coeficiente líder de $g(x)$ invertível então é possível determinar únicos* ***polinômio quociente*** *$q(x)$ e* ***polinômio resto*** *$r(x)$ em $\mathbb{K}[X]$ tais que*

$$p(x) = g(x)q(x) + r(x), \text{ com } r(x) = 0 \text{ ou } gr(r(x)) < gr(g(x)).$$

Demonstração. Existência: Suponhamos que $p(x) = a_0 + a_1x + \cdots + a_nx^n$ e $g(x) = b_0 + b_1x + \cdots + b_mx^m$ com $m \geq 0$ e b_m invertível, separando em casos temos:

i) Para $p(x) = 0$. Neste caso, $q(x) = r(x) = 0$, logo $0 = g(x)0 + 0$.

ii) Para $p(x) \neq 0$ e $gr(p(x)) < gr(g(x))$. Neste caso, tomemos $q(x) = 0$ e $r(x) = p(x)$, logo $p(x) = g(x)0 + p(x)$.

iii) Para $p(x) \neq 0$ e $gr(p(x)) \geq gr(g(x))$. Neste caso, utilizaremos indução sobre o grau do polinômio p(x).

Se $n = 0$, como por hipótese, $gr(p(x))$ é maior ou igual ao $gr(g(x))$, temos $gr(g(x)) = 0$, logo $p(x)$ e $g(x)$ são polinômios constantes não nulos: $p(x) = a_0$ e $g(x) = b_0$, onde b_0 é invertível por hipótese. Agora temos uma divisão possível no corpo $\mathbb{K}$, de modo que $q(x) = \frac{a_0}{b_0}$ e $r(x) = 0$, logo $a_0 = b_0\frac{a_0}{b_0} + 0$.

Suponha que $n > 0$, e considere, como hipótese de indução, que a proposição seja válida para todo polinômio de grau menor que j, com $0 \leq j < n$.

Considere o polinômio $p_1(x)$ pertencente a $\mathbb{K}[X]$ definido por

$$p_1(x) = p(x) - \frac{a_n}{b_m}x^{n-m}g(x), \tag{A.6}$$

isolando $p(x)$ obtemos

$$p(x) = \frac{a_n}{b_m}x^{n-m}g(x) + p_1(x).$$

Se $p_1(x) = 0$ ou $gr(p_1(x)) < gr(g(x))$, então $q(x) = \frac{a_n}{b_m}x^{n-m}$ e $r(x) = p_1(x)$.

Caso contrário, tem-se $gr(p_1(x)) \geq gr(g(x))$ e $gr(p_1(x)) < n$, pois o coeficiente líder de $p(x)$ é igual ao coeficiente líder do polinômio $\frac{a_n}{b_m}x^{n-m}g(x)$. Portanto, por hipótese de indução, existem $q_1(x)$ e $r_1(x)$ tais que

$$p_1(x) = g(x)q_1(x) + r_1(x), \text{ com } r_1(x) = 0 \text{ ou } gr(r_1(x)) < gr(g(x)). \tag{A.7}$$

Igualando A.6 com A.7, temos

$$p(x) - \frac{a_n}{b_m}x^{n-m}g(x) = g(x)q_1(x) + r_1(x) \Rightarrow p(x) = \frac{a_n}{b_m}x^{n-m}g(x) + g(x)q_1(x) + r_1(x)$$
$$\Rightarrow p(x) = \left[\frac{a_n}{b_m}x^{n-m} + q_1(x)\right]g(x) + r_1(x),$$

com $r_1(x) = 0$ ou $gr(r_1(x)) < gr(g(x))$.

Portanto, existem polinômios $q(x) = \frac{an}{bn}x^{n-m} + q_1(x)$ e $r(x) = r_1(x)$.

Unicidade: Suponha, por absurdo, que existam $q_1(x), r_1(x), q_2(x), r_2(x) \in \mathbb{K}[X]$ tais que

$$p(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x),$$

com $r_i(x) = 0$ ou $gr(r_i(x)) < gr(g(x))$, $i \in \{1, 2\}$.

Assim,

$$[q_1(x) - q_2(x)]g(x) = r_2(x) - r_1(x).$$

Se $q_1(x) \neq q_2(x)$, e $b_m \neq 0$, pela Proposição A.1.3 temos

$$gr(r_2(x) - r_1(x)) = gr([q_1(x) - q_2(x)]g(x)) = gr(q_1(x) - q_2(x)) + gr(g(x)).$$

Logo, $gr(r_2(x) - r_1(x)) > gr(g(x))$, o que é absurdo, pois

$$gr(g(x)) > max\{gr(r_1(x)), r_2(x))\} \geq gr(r_2(x) - r_1(x)).$$

Portanto, $q_1(x) = q_2(x)$ e consequentemente

$$r_1(x) = p(x) - q_1(x)g(x) = p(x) - q_2(x)g(x) = r_2(x).$$

Conclui-se que os polinômios $q(x)$ e $r(x)$ existem e são únicos.

■

A partir de agora vamos desenvolver um algoritmo capaz de calcular máximos divisores comuns entre polinômios.

Dados $p(x), g(x) \in \mathbb{K}[X]$, com $g(x) \neq 0$, podemos aplicar a divisão Euclidiana para polinômios sucessivamente com o resto, obtendo

$$p(x) = g(x)q_1(x) + r_1(x), \text{ com } r_1(x) = 0 \text{ ou } gr(r_1(x)) < gr(g(x)); \tag{A.8}$$

$$g(x) = r_1(x)q_2(x) + r_2(x), \text{ com } r_2(x) = 0 \text{ ou } gr(r_2(x)) < gr(r_1(x)); \tag{A.9}$$

$$r_1(x) = r_2(x)q_3(x) + r_3(x), \text{ com } r_3(x) = 0 \text{ ou } gr(r_3(x)) < gr(r_2(x)); \tag{A.10}$$

$$\vdots$$

$$r_{n-2}(x) = r_{n-1}(x)q_n(x) + r_n(x), \text{ com } r_n(x) = 0 \text{ ou } gr(r_n(x)) < gr(r_{n-1}(x)); \tag{A.11}$$

$$r_{n-1}(x) = r_n(x)q_{n+1}(x) + r_{n+1}(x), \text{ com } r_{n+1}(x) = 0. \tag{A.12}$$

Como o resto diminui a cada passo, o processo não pode continuar indefinidamente, logo alguma dessas divisões deve ser exata. Suponha que $r_{n+1}(x)$ seja o primeiro resto nulo, como está esquematizado acima.

Teorema A.2.2. *O polinômio $r_n(x)$ acima existe e é o máximo divisor comum de dois polinômios $p(x)$ e $g(x)$ em $\mathbb{K}[X]$, não simultaneamente nulos, podendo ser escrito da forma*

$$r_n(x) = \lambda(x)p(x) + \mu(x)g(x), \text{ tal que existam } \lambda(x), \mu(x) \in \mathbb{K}[X].$$

Demonstração. Inicialmente vamos mostrar que $r_n(x)$ é um divisor comum de $p(x)$ e $g(x)$. De fato, de A.12, temos que $r_n(x) \mid r_{n-1}(x)$, logo dessa relação e de A.11, temos que $r_n(x) \mid r_{n-2}(x)$. Repetindo esse processo até que de A.10 obtemos que $r_n(x) \mid r_1(x)$, juntamente de A.9 temos que $r_n \mid g(x)$, e por fim de A.8 temos que $r_n \mid p(x)$.

Seja $t(x) \in \mathbb{K}[X]$ um polinômio qualquer, suponha que $t(x) \mid p(x)$ e $t(x) \mid g(x)$, de A.8 obtemos que $t(x) \mid r_1(x)$, dessa relação e de A.9, temos que $t(x) \mid r_2(x)$. Repetindo esse processo até A.12 obtemos que $t(x) \mid r_n(x)$.

Logo, para qualquer polinômio $t(x)$, $t(x) \leq r_n(x)$. Portanto $r_n(x)$ é o máximo divisor comum de $p(x)$ e $g(x)$.

Agora vamos provar a existência dos polinômios $\lambda(x)$ e $\mu(x)$. Por indução sobre n, temos:

Para $n = 1$, da primeira divisão A.8 obtemos que

$$r_1(x) = p(x) - q_1(x)g(x), \tag{A.13}$$

ou seja, $r_1(x)$ foi escrito em função de $p(x)$ e $g(x)$. Portanto, existem $\lambda_1(x)$ e $\mu_1(x)$ tais que $\lambda_1(x) = 1$ e $\mu_1(x) = -q_1(x)$.

Para $n = 2$, substituindo A.13 em A.9, temos que

$$g(x) = [p(x) - g(x)q_1(x)]q_2(x) + r_2(x) \Rightarrow r_2(x) = g(x) - p(x)q_2(x) + g(x)q_1(x)q_2(x)$$
$$\Rightarrow r_2(x) = -q_2(x)p(x) + [1 + q_1(x)q_2(x)]g(x).$$

Note que, novamente pudemos escrever $r_2(x)$ em função de $p(x)$ e $g(x)$. Portanto, existem $\lambda_2(x)$ e $\mu_2(x)$ tais que $\lambda_2(x) = -q_2(x)$ e $\mu_2(x) = 1 + q_1(x)q_2(x))$.

Considere, como hipótese de indução, que existam $\lambda_i(x)$ e $\mu_i(x)$ para todo subíndice $i \in \mathbb{N}^*$ tal que $i < n$.

Para n, de A.11 e pela hipótese de indução, obtemos

$$\begin{aligned} r_n(x) &= r_{n-2}(x) - r_{n-1}(x)q_n(x) \\ &= \lambda_{n-2}(x)p(x) + \mu_{n-2}(x)g(x) - \lambda_{n-1}(x)q_n(x)p(x) - \mu_{n-1}(x)q_n(x)g(x) \\ &= [\lambda_{n-2}(x) - \lambda_{n-1}(x)q_n(x)]p(x) + [\mu_{n-2}(x) - \mu_{n-1}(x)q_n(x)]g(x) \\ &= \lambda_n(x)p(x) + \mu_n(x)g(x). \end{aligned}$$

Portanto, o resultado vale para todo valor de n. Em particular, colocando $\lambda(x) = \lambda_n(x)$ e $\mu(x) = \mu_n(x)$, temos

$$r_n(x) = \lambda(x)p(x) + \mu(x)g(x).$$

■

Por razões de ordem prática, de agora em diante, denotaremos o máximo divisor comum ou MDC entre dois poilinômios $p(x)$ e $g(x)$ em $\mathbb{K}[X]$ por $mdc(p(x), g(x))$.

O teorema acima prova a existência de um máximo divisor comum de dois polinômios e nos fornece um algoritmo para calculá-lo, nos garantindo que

$$mdc(p(x), g(x)) = mdc(g(x), r_1(x)) = mdc(r_1(x), r_2(x)) = \cdots = mdc(r_{n-1}(x), r_n(x)).$$

Portanto, como $r_n(x) \mid r_{n-1}(x)$ temos que $mdc(r_{n-1}(x), r_n(x)) = r_n(x)$, consequentemente $mdc(p(x), g(x)) = r_n(x)$, ou seja, nesse processo, o máximo divisor comum de $p(x)$ e $g(x)$ é o último resto diferente de zero.

Para dividir $p(x)$ por $g(x)$ vamos utilizar o seguinte esquema:

	$q(x)$
$p(x)$	$g(x)$
$r(x)$	

Desse modo, tem-se assim o seguinte **algoritmo de *Euclides* para polinômios:**

	$q_1(x)$	$q_2(x)$	$q_3(x)$	$\cdots$	$\cdots$	$q_n(x)$	$q_{n+1}(x)$
$p(x)$	$g(x)$	$r_1(x)$	$r_2(x)$	$\cdots$	$r_{n-2}(x)$	$r_{n-1}(x)$	$r_n(x)$
$r_1(x)$	$r_2(x)$	$r_3(x)$	$\cdots$	$\cdots$	$r_n(x)$	0	

Portanto, como existe $mdc(p(x), g(x))$ não nulo, para $p(x)$ e $g(x)$ não simultaneamente nulos em $\mathbb{K}[X]$, logo MDC é associado a um único polinômio mônico que também é um MDC. Sendo assim, o único MDC mônico de $p(x)$ e $g(x)$.

ANEXO A – RESULTADOS DO ALGORITMO DE EUCLIDES PARA POLINÔMIOS

Proposição A.0.1. *Dois polinômios $p(x)$ e $g(x)$ são primos entre si em $\mathbb{K}[X]$ se, e somente se, existem polinômios $\lambda(x)$ e $\mu(x)$ tais que*

$$\lambda(x)p(x) + \mu(x)g(x) = 1.$$

Proposição A.0.2. *Sejam $p(x), g(x), h(x) \in \mathbb{K}[X]$. Se $p(x) \mid g(x)h(x)$ e $mdc(p(x), g(x)) = 1$, então $p(x) \mid h(x)$.*

Proposição A.0.3. *No anel $\mathbb{K}[X]$, todo elemento não invertível e irredutível é primo.*

Proposição A.0.4. *Todo polinômio não invertível possui pelo menos um divisor primo mônico.*

Proposição A.0.5. *Sejam $p(x), p_1(x), p_2(x) \cdots, p_n(x)$ polinômios primos mônicos tais que $p \mid (p_1(x))(p_2(x)) \cdots (p_n(x))$, então $p(x) = p_i(x)$ para algum $i \in \mathbb{N}^*$,*

Teorema A.0.6. *Todo polinômio mônico em $\mathbb{K}[X]$, não constante e não irredutível, se escreve como produto de polinômios de $\mathbb{K}[X]$ irredutíveis e mônicos. Essa escrita é única a menos da ordem dos fatores.*

Corolário 5. *Para todo poliômio mônico $p(x) \in \mathbb{K}[X] \setminus \mathbb{K}$, existem $n \geq 1$, polinômios mônicos irredutíveis distintos $p_1(x), \cdots, p_n(x)$, inteiros positivos $\alpha_1, \cdots, \alpha_n$ e $a \in \mathbb{K}^*$ tais que*

$$p(x) = ap_1(x)^{\alpha_1} \cdots p_n(x)^{\alpha_2}.$$