Universidade Federal de Sergipe
Centro de Ciências Exatas e Tecnologia
Departamento de Matemática
Pós-Graduação em Matemática

Códigos Cíclicos: uma introdução aos códigos corretores de erros

São Cristóvão – SE
2017

Universidade Federal de Sergipe
Centro de Ciências Exatas e Tecnologia
Departamento de Matemática
Pós-Graduação em Matemática

Códigos Cíclicos: uma introdução aos códigos corretores de erros

por

CANUTO RUAN SANTOS ARAGÃO

sob a orientação do

Prof. Dr. Kalasas Vasconcelos de Araújo

São Cristóvão – SE
2017

UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Dissertação submetida à aprovação pelo Programa de Pós-Graduação em Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Códigos cíclicos: Uma introdução aos códigos corretores de erros

por

Canuto Ruan Santos Aragão

Aprovada pela banca examinadora:

Prof. Dr. Kalasas Vasconcelos De Araujo - UFS
Orientador

Prof. Dr. Andre Vinicius Santos Doria - UFS
Primeiro Examinador

Prof. Dr. Filipe Dantas dos Santos - UFS
Segundo Examinador

São Cristóvão, 13 de Junho de 2017

Dedicado àqueles alunos de disciplinas algébricas que, não satisfeitos com a beleza da matemática pura, sempre acabam questionando: "Onde eu vou aplicar tudo isso na prática, professor(a)?"

Agradecimentos

Agradeço e dedico essa conquista a meus pais, Marta e Rosenaldo, que sempre cuidaram tão bem de mim, a minha irmã Ruana, que desde minha aprovação no vestibular sempre me motivou e foi minha fã, a minha avó Rosalva que sempre investiu na minha educação e acreditou que havia potencial em mim, a meu avô Aragão (Pintinho), a minha avó Dionélia (vó senhora) e a Vera Lúcia (sogrinha querida), por todo o carinho e por sempre deixarem claro que estariam por perto caso eu precisasse, minha esposa Manuela por sempre aguentar meus desabafos durante essa tragetória e meu filho Kadu, que sempre alegra meus dias.

Agradeço e dedico essa vitória também aos professores e colegas de curso, especialmente: Lázaro, Sóstenes, Glauber, Deusdete, Wesley, Edson e meu orientador Kalasas. Mais do que colegas, éramos uma equipe, todos sempre prontos e dispostos a ajudar um ao outro.

E, é claro, àquele que não por acaso está sempre em todos os agradecimentos: Deus. Obrigado por mais essa bênção.

Resumo

Um código cíclico é um tipo específico de código linear. Sua relevância consiste no fato de que todas suas principais informações são intrinsecas à estrutura dos ideais no anel quociente $K[x]/(x^n - 1)$ via um isomorfismo. Neste trabalho, caracterizamos os códigos cíclicos em correspondência biunívoca com os ideais deste anel quociente. Apresentaremos também sua matriz geradora, a matriz de paridade e aboradaremos sua codificação e decodificação.

Palavras-chave: códigos corretores de erros, códigos cíclicos, matriz geradora, matriz de paridade, anel quociente.

Abstract

A cyclic code is a specific type of linear code. Its relevance consists in the fact that all its main information is intrinsic to the structure of the ideals in the quotient ring $K[x]/(x^n - 1)$ via an isomorphism. In this work, we characterize the cyclic codes in biunivocal correspondence with the ideals of this quotient ring. We will also present its generating matrix, the parity matrix and we will discuss its codification and decoding.

Keywords: Error correction codes, cyclic codes, generating matrix, parity matrix, quotient ring.

Sumário

Introdução

Em diversas situações do cotidiano os códigos corretores de erros estão presentes: no uso do celular, internet e wi-fi, por exemplo. Tais códigos tem como principal finalidade corrigir informações que sofreram algum tipo de alteração durante sua transmissão ou armazenamento.

Pensar em tudo o que há por trás do envio de uma simples mensagem por telefone ou envio de fotos por sondas espaciais é um estímulo ao estudo deste tema. Além disso, por ser uma teoria largamente utilizada e com problemas em aberto, desperta também o interesse profissional.

A teoria sobre códigos corretores de erros nasceu em 1948, fundada pelo matemático C. E. Shannon, foi bastante desenvolvida pelos matemáticos da época e, a partir dos anos 70, despertaram o interesse de engenheiros que ajudaram a continuar no desenvolvimento da mesma. Hoje, sempre que se deseja transmitir dados com confiabilidade, os códigos corretores são utilizados.

Neste trabalho, exploramos aplicações para conceitos abstratos bastante conhecidos na matemática, alguns desde o ensino básico, tendo como finalidade principal a caracterização dos códigos corretores lineares cíclicos. Para isto, o texto foi dividido em três capítulos da seguinte forma:

No primeiro capítulo abordaremos conceitos de Álgebra necessários para o entendimento do tema, tais como: Anéis de Polinômios sobre corpos finitos, Ideais de um Anel e Anéis Quocientes.

O segundo capítulo tráz o conceito de código linear e como funciona a sua codificação e decodificação.

Por fim, traremos no capítulo três o conceito de código linear cíclico, cujo diferencial é a velocidade de codificação e decodificação.

O texto requer do leitor certo conhecimento algébrico, como por exemplo: o conceito de Espaço Vetorial, base, dimensão e dependência linear.

Capítulo 1

Polinômios

1.1 Anéis de Polinômios

1.1.1 Polinômios

Nesse texto decidimos abordar a definição de Polinômios de forma pouco convencional, afim de dar ao leitor uma oportunidade de estudar essa teoria de forma alternativa.

Definição 1 *Seja A um anel comutativo. Uma sequência infinita* $(a_0, a_1, a_2, a_3, ...)$ *com* $a_i \in A$*, onde todos os elementos são nulos exceto para uma quantidade finita de índices é chamada* ***Polinômio****. Os* $a_i's$ *são chamados coeficientes do Polinômio.*

Diremos que dois polinômios $(a_0, a_1, a_2, a_3, ...)$ e $(b_0, b_1, b_2, b_3, ...)$ são iguais se $a_i = b_i$, $\forall i \in \{0, 1, 2, ...\}$. Além disso, a sequência nula $(0, 0, 0, ...)$ será chamada de polinômio nulo e representada apenas por 0.
Neste trabalho denotaremos por $A[x]$ o conjunto de todos os polinômios $p(x)$ com coeficientes em A, mais adiante explicaremos o aparecimento desse "x"em nossa notação.

Definição 2 *Chamaremos de* ***grau de*** $p(x)$*, ou simplesmente* $gr(p(x))$*, o índice i tal que*

$$gr(p(x)) = max\{i | a_i \neq 0\}$$

Obs.: *Seguiremos a convenção de que* $gr(0) = \infty$.

1.1.2 Operações em $A[x]$

Teorema 1 *As seguintes operações em* $A[x]$ *estão bem definidas:*

Adição:

$(a_0, a_1, a_2, a_3, ...) + (b_0, b_1, b_2, b_3, ...) = (c_0, c_1, c_2, c_3, ...)$

onde $c_i = a_i + b_i \forall i \in \{0, 1, 2, ...\}$.

Multiplicação:

$(a_0, a_1, a_2, a_3, ...) \cdot (b_0, b_1, b_2, b_3, ...) = (d_0, d_1, d_2, d_3, ...)$

onde $d_i = \sum_{j=0}^{i} a_j b_{i-j} = a_0 b_i + a_1 b_{i-1} + \cdots + a_{i-1} b_1 + a_i b_0 \forall i \in \{0, 1, 2, ...\}$.
Prova: *Sejam* $p(x), q(x) \in A[x]$, *se* $p(x)$ *ou* $q(x)$ *é o polinômio nulo, então* $p(x) + q(x)$ *é* $p(x)$ *ou* $q(x)$ *e* $p(x) \cdot q(x) = 0$. *Suponhamos então* $p(x)$ *e* $q(x)$ *não nulos, com graus* n *e* m *respectivamente. Se tomarmos um valor* $t > max\ \{n, m\}$, *então* $a_t + b_t = 0$, *onde* a_i *e* b_i *são os coeficientes de* $p(x)$ *e* $q(x)$ *respectivamente. Além disso, pela definição 2, teríamos* $gr(p(x) + q(x))$ *igual a* n *ou* m. *Dessa forma* $p(x) + q(x)$ *é uma sequência infinita com um número finito de coeficientes não nulos e portanto está em* $A[x]$.

Para mostrar que a multiplicação $p(x) \cdot q(x)$ *também é um polinômio em* $A[x]$, *tomemos* $t > n + m$, *assim teremos* $d_t = 0$, *onde* d_t *representa o coeficiente de índice* t *do polinômio* $p(x) \cdot q(x)$. *De fato, se tivermos* $j > n$, *então* $a_j = 0$, *donde* $a_j b_{t-j} = 0$. *Caso tenhamos* $j \leqslant n$, *então* $-j \geqslant -n$, *dái* $t > n + m$ *implica* $t - j > n + m - n$, *ou seja,* $t - j > m$ *e portanto* $a_j b_{t-j} = 0$ *pois* $b_{t-j} = 0$. *Sendo assim, todas as parcelas do somatório que define* d_t *seriam nulas e consequentemente* $d_t = 0$. *Dessa forma, mais uma vez temos uma sequência infinita com um número finito de termos não nulos, logo* $p(x) \cdot q(x)$ *está em* $A[x]$.

■

Teorema 2 *O conjunto* $A[x]$ *munido das operações de soma e multiplicação definidas acima é um Anel comutativo.*

Prova: *Basta verificarmos que* $A[x]$ *satisfaz as propriedades de um Anel comutativo:*
Considere os polinômios em $A[x]$: $p(x) = (a_0, a_1, a_2, ...)$, $q(x) = (b_0, b_1, b_2, ...)$ *e* $s(x) = (c_0, c_1, c_2, ...)$.

$\mathbf{A_1}$*) Associatividade:*
$(p(x) + q(x)) + s(x) =$
$= ((a_0, a_1, a_2, ...) + (b_0, b_1, b_2, ...)) + (c_0, c_1, c_2, ...) =$
$= (a_0 + b_0, a_1 + b_1, a_2 + b_2, ...) + (c_0, c_1, c_2, ...) =$
$= (a_0 + b_0 + c_0, a_1 + b_1 + c_c, a_2 + b_2 + c_2, ...) =$
$= (a_0, a_1, a_2, ...) + (b_0 + c_0, b_1 + c_1, b_2 + c_2, ...) =$

$= (a_0, a_1, a_2, ...) + ((b_0, b_1, b_2, ...) + (c_0, c_1, c_2, ...)) =$
$= p(x) + (q(x) + s(x))$

A_2) *Comutatividade:*
$p(x) + q(x) =$
$= (a_0, a_1, a_2, ...) + (b_0, b_1, b_2, ...)$
$= (a_0 + b_0, a_1 + b_1, a_2 + b_2, ...)$
$= (b_0 + a_0, b_1 + a_1, b_2 + a_2, ...)$
$= (b_0, b_1, b_2, ...) + (a_0, a_1, a_2, ...)$
$= q(x) + p(x)$

A_3) *Existência de um elemento neutro para adição:*
O polinômio nulo $0 = (0, 0, 0, ...)$ *é tal que* $p(x) + 0 = 0 + p(x) = p(x)$ *(A verificação deste fato é trivial).*

A_4) *Todo elemento em* $A[x]$ *possui inverso aditivo:*
Sendo $p(x) = (a_0, a_1, a_2, ...)$, *então* $-p(x) = (-a_0, -a_1, -a_2, ...) \in A[x]$ *é tal que* $p(x) + (-p(x)) = 0$

M_1) *A multiplicação é associativa:*
O n-ésimo coeficiente de $(p(x) \cdot q(x)) \cdot s(x)$, *pela definição será:*

$((a_0, a_1, a_2, ...) \cdot (b_0, b_1, b_2, ...)) \cdot (c_0, c_1, c_2, ...) =$

$= (d_0, d_1, d_2, ...) \cdot (c_0, c_1, c_2, ...)$, *com* $d_i = \sum_{j=0}^{i} a_j b_{i-j}$

$= \sum_{i=0}^{n} d_i c_{n-i}$

$= \sum_{i=0}^{n} (\sum_{j=0}^{i} a_j b_{i-j}) c_{n-i}$, *tomando* $u = j, v = i - j, w = n - i$, *temos* $u + v + w = n$, *dessa forma o somatório pode ser reescrito desta forma:*

$= \sum_{u+v+w=n} a_u b_v c_w$, *com* $u, v, w \geq 0$ *(*)*

De forma análoga podemos analisar o n-ésimo coeficiente de $p(x) \cdot (q(x) \cdot s(x))$*:*

$(a_0, a_1, a_2, ...) \cdot ((b_0, b_1, b_2, ...) \cdot (c_0, c_1, c_2, ...)) =$

$= (a_0, a_1, a_2, ...) \cdot (d_0, d_1, d_2, ...)$, *com* $d_i = \sum_{j=0}^{i} b_j c_{i-j}$

$= \sum_{i=0}^{n} a_i d_{n-i}$

$= \sum_{i=0}^{n} a_i (\sum_{j=0}^{n-i} b_j c_{n-i-j})$, *tomando* $u = i, v = n-i, w = n-i-j$, *temos* $u + v + w = n$, *dessa forma o somatório pode ser reescrito desta forma:*

$= \sum_{u+v+w=n} a_u b_v c_w$, *com* $u, v, w \geq 0$ *(**)*

Por (*) e (**), temos $(p(x) \cdot q(x)) \cdot s(x) = p(x) \cdot (q(x) \cdot s(x))$

M_2*) A multiplicação é comutativa:*

Sendo o Anel A comutativo, então

$p(x) \cdot q(x) = \sum_{j=0}^{i} a_j b_{i-j} = \sum_{j=0}^{i} b_{i-j} a_j$, *fazendo a mudança de variável* $k = i - j$, *temos* $\sum_{k=0}^{i} b_k a_{i-k} = q(x) \cdot p(x)$

M_3*) Existência de um elemento neutro para a multiplicação:*

O Elemento $1 = (1, 0, 0, 0, ...) \in A[x]$ *é tal que* $p(x) \cdot 1 = 1 \cdot p(x) = p(x)$. *A veracidade dessa afirmação é clara, pois no somatório* $\sum_{j=0}^{i} a_j b_{i-j} = a_0 b_i + a_1 b_{i-1} + \cdots + a_{i-1} b_1 + a_i b_0$, *que define a multiplicação de dois polinômios, todas as parcelas da soma são zeradas sempre que* $i - j \neq 0$

AM*) A multiplicação é distributiva com relação à adição:*

Tomando $p(x) \cdot (q(x) + s(x)) =$

$= (a_0, a_1, a_2, ...) \cdot ((b_0, b_1, b_2, ...) + (c_0, c_1, c_2, ...))$

$= (a_0, a_1, a_2, ...) \cdot (b_0 + c_0, b_1 + c_1, b_2 + c_2, ...)$

$= \sum_{j=0}^{i} a_j (b_{i-j} + c_{i-j}) = \sum_{j=0}^{i} a_j b_{i-j} + \sum_{j=0}^{i} a_j c_{i-j}$

$= p(x) \cdot q(x) + p(x) \cdot s(x)$

■

Teorema 3 *Seja* $A[x]$ *o anel dos polinômios sobre um anel A. Se* $A^* \subset A[x]$ *é o conjunto de todos os polinômios da forma* $(a, 0, 0, ...)$, $a \in A$, *então* A^* *é subanel de* $A[x]$ *e isomorfo a A.*

Prova: *Defina a aplicação* $\phi : A \to A^*$, $a \mapsto \phi(a) = (a, 0, 0, ...)$. *Dessa forma, temos*

$\phi(a + b) = (a + b, 0, 0, ...) = (a, 0, 0, ...) + (b, 0, 0, ...) = \phi(a) + \phi(b)$
e
$\phi(a \cdot b) = (a \cdot b, 0, 0, ...) = (a, 0, 0, ...) \cdot (b, 0, 0, ...) = \phi(a) \cdot \phi(b)$.

Além disso, $\phi(1_A) = (1, 0, 0, ...) = 1_{A[x]}$. *Isto mostra que* ϕ *é um homomorfismo, resta verificar se é bijetor. Perceba que se* $\phi(a) = (0, 0, 0, ...)$, *então*

$a = 0$, *em outras palavras esse homomorfismo possui núcleo nulo e, portanto, injetivo.*

E quanto à sobrejetividade de ϕ*, como os elementos em* A^* *são da forma* $(a, 0, 0, ...)$*, com* $a \in A$*, segue que, para todo* $(a, 0, 0, ...) \in A^*$*, existe* $a \in A$ *tal que* $\phi(a) = (a, 0, 0, ...)$*.*

■

1.1.3 A notação polinomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$

Definição 3 *Um polinômio do tipo* $(a, 0, 0, ...) \in A[x]$*, será chamado de* ***polinômio constante*** *e o representaremos apenas por:* a*. Além disso, chamaremos de* x *o polinômio* $(0, 1, 0, 0, ...) \in A[x]$*.*

Proposição 1 $x^n = (0, 0, ..., 1, 0, ...)$ *com* 1 *na n-ésima posição.*

Prova por indução em n: *Por definição de potência e multiplicação de polinômios, temos*
$x^0 = 1$
$x^1 = x = (0, 1, 0, 0, ...)$
$x^2 = x \cdot x = (0, 0, 1, 0, ...)$
$x^3 = x^2 \cdot x = (0, 0, 0, 1, ...)$

Mostrando assim que a proposição é válida para alguns valores de n. Suponhamos agora a proposição válida para $n-1$*, ou seja,* $x^{n-1} = (0, 0, ..., 1, 0, ...)$ *com* 1 *na entrada de índice* $n - 1$*, então*

$x^n = x^{n-1} \cdot x = (0, 0, ..., 0, 1, 0, ...)$ *com* 1 *na n-ésima posição.*

■

Tendo em mãos a proposição e a definição acima, é fácil ver que

$$ax^n = (a, 0, 0, ...) \cdot (0, 0, ..., 1, 0, ...) = (0, 0, ..., a, 0, ...)$$

com a na n-ésima posição. Daí segue que, dado um polinômio de grau n $(a_0, a_1, a_2, a_3, ...)$ em $A[x]$ pode ser reescrito da seguinte forma:

$$(a_0, a_1, a_2, a_3, ...) =$$

$$= (a_0, 0, 0, 0, ...) + (0, a_1, 0, 0, ...) + (0, 0, 0, a_3, ...) + ... + (0, 0, 0, ..., a_n, ...)$$

$$= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ... + a_n x^n$$

Cada uma das parcelas $a_i x^i$ da soma será chamada de termo, sendo o termo não nulo de maior índice chamado de **termo líder** de $p(x)$, ou simplesmente $\mathbf{TL}(p(x))$.

Perceba que da definição de igualdade de polinômios, temos que se $b_0 + b_1x + b_2x^2 + b_3x^3 + ... + b_mx^m$ é uma outra representação para o polinômio $(a_0, a_1, a_2, a_3, ...)$, então $n = m$ e $a_i = b_i \forall i$. Em outras palavras a representação dada acima é única.

Proposição 2 *Seja A um domínio de integridade. Temos que*
i) $\forall\ p(x), q(x) \in A[x] - \{0\}$, $gr(p(x) \cdot q(x)) = gr(p(x)) + gr(q(x))$
ii) $A[x]$ é um domínio de integridade.
iii) Os elementos invertíveis de $A[x]$ são os elementos invertíveis de A.

Prova: *i)Sejam $p(x) = a_0 + a_1x + ... + a_nx^n$ e $q(x) = b_0 + b_1x + ... + b_mx^m$, com $a_n \neq 0$ e $b_m \neq 0$. Logo, $p(x) \cdot q(x) = a_0b_0 + ... + a_nb_mx^{n+m}$. Como A é domínio, $a_nb_m \neq 0$, logo*

$$gr(p(x) \cdot q(x)) = n + m = gr(p(x)) + gr(q(x))$$

ii)Decorre de i), afinal se $p(x) \cdot q(x) \neq 0$, então $a_nb_m \neq 0$, logo $a_n \neq 0$ e $b_m \neq 0$, donde $p(x) \neq 0$ e $q(x) \neq 0$.

iii) É claro que todo elemento invertível em A também o será em $A[x]$, basta então mostrarmos que dado um elemento $p(x) \in A[x]$ invertível, então $p(x)$ também será invertível em A.
Sendo $p(x)$ invertível, existe então $q(x) \in A[x]$ tal que $p(x) \cdot q(x) = 1$, logo $p(x) \neq 0$ e $q(x) \neq 0$, donde $gr(p(x)) + gr(q(x)) = gr(p(x) \cdot q(x)) = gr(1) = 0$. Isso implica dizer que $gr(p(x)) = gr(q(x)) = 0$ e portanto $p(x)$ e $q(x) \in A$ com $p(x) \cdot q(x) = 1$. Logo, $p(x)$ é invertível em A.

■

Definição 4 *Se $gr(p(x)) = n$ e $a_n = 1$, diremos que $p(x)$ é um polinômio mônico.*

A partir deste ponto trabalharemos apenas com polinômios sobre um corpo K. A primeira observação que pode ser feita nesse caso é que todo polinômio é mônico em $K[x]$ a menos de uma multiplicação por uma constante. Em outras palavras, temos $p(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n = a_n \cdot (a_n^{-1}a_0 + a_n^{-1}a_1x + a_n^{-1}a_2x^2 + ... + x^n)$

A definição e proposição abaixo somente far-se-ão necessárias quando formos tratar de códigos cíclicos.

Definição 5 *Seja $p(x) \in K[x]$. Chamaremos de **polinômio recíproco** de $p(x)$ o polinômio*

$$p^*(x) = x^n \cdot p\left(\frac{1}{x}\right)$$

Ressaltaremos aqui apenas umas das propriedades dos polinômios recíprocos:
i) $(g(x) \cdot d(x))^* = g^*(x) \cdot d^*(x)$

Prova: Sejam $g(x) = \sum_{i=0}^{n} a_i x^i$ e $d(x) = \sum_{i=0}^{m} b_i x^i$. Dessa forma, temos

$$g(x) \cdot d(x) = \sum_{i}^{n+m} c_i x^i, \text{ sendo } c_i = \sum_{j+k=i} a_i b_k, \text{ donde}$$
$$(g(x) \cdot d(x))^* = x^{n+m} \cdot \sum_{i}^{n+m} c_i \left(\frac{1}{x}\right)^i.$$

Por outro lado, temos $g^*(x) = x^n \cdot \sum_{i=0}^{n} a_i \left(\frac{1}{x}\right)^i$ e $d^*(x) = x^m \cdot \sum_{i=0}^{m} b_i \left(\frac{1}{x}\right)^i$. Daí

$$g^*(x) \cdot d^*(x) = x^n \cdot x^m \sum_{i}^{n+m} c_i \left(\frac{1}{x}\right)^i, \text{ sendo } c_i = \sum_{j+k=i} a_i b_k.$$

Portanto, temos $(g(x) \cdot d(x))^* = g^*(x) \cdot d^*(x)$. ■

A seguinte proposição decorre imediatamente da propriedade i).

Proposição 3 *Se $p(x)$ e $g(x)$ em $K[x]$ são tais que $g(x)$ divide $p(x)$, então $g^*(x)$ divide $p^*(x)$.*

1.2 Divisibilidade em $K[x]$

O algorítmo da divisão entre polinômios é algo simples, lecionado inclusive no ensino básico, o que nos leva a esquecer da força algébrica que há por trás deste procedimento. O teorema logo abaixo e a sua prova estão aí para nos lembrar como e porque esta técnica funciona.

Teorema 4 *(Algoritmo da divisão) Seja K um corpo e $p(x)$, $g(x) \in K[x]$ com $g(x) \neq 0$. Então existem únicos polinômios $q(x)$, $r(x) \in K[x]$ tais que*

$$p(x) = q(x)g(x) + r(x)$$

com $r(x) = 0$ ou $gr(r(x)) < gr(g(x))$.

Prova: *(Existência) Se $p(x) = 0$ ou $gr(p(x)) < gr(g(x))$, basta tomarmos $r(x) = p(x)$ e $q(x) = 0$. Por outro lado, suponhamos $gr(p(x)) \geq gr(g(x))$. Nesse caso a prova será por indução completa em $n = gr(p(x))$. É fácil verificar a veracidade para o caso $gr(p(x)) = gr(g(x)) = 0$, pois seriam da forma $p(x) = a$ e $g(x) = b$, com a, $b \in K$, donde $a = ab^{-1}b + 0$, ou seja,*

$q(x) = ab^{-1}$ *e* $r(x) = 0$. *Vamos agora supor nossa afirmação válida para todo polinômio com grau menor que n.*

O polinômio $h(x) = p(x) - \dfrac{TL(p(x))}{TL(g(x))} \cdot g(x)$ *tem grau menor que o de n, tendo em vista que* $p(x)$ *e* $\dfrac{TL(p(x))}{TL(g(x))} \cdot g(x)$ *possuem o mesmo termo líder. Sendo assim, por hipótese indutiva, existem* $q_1(x)$ *e* $r(x) \in K[x]$ *tais que* $h(x) = q_1(x)g(x) + r(x)$, *com* $r(x) = 0$ *ou* $gr(r(x)) < gr(g(x))$. *Logo,*

$$h(x) = p(x) - \frac{TL(p(x))}{TL(g(x))} \cdot g(x) = q_1(x)g(x) + r(x)$$

$$\Rightarrow p(x) = \left(q_1(x) + \frac{TL(p(x))}{TL(g(x))}\right) \cdot g(x) + r(x),$$ *com* $r(x) = 0$ *ou* $gr(r(x)) < gr(g(x))$.

Logo, pelo princípio de indução, nossa afirmação é verdadeira para todo n inteiro positivo.

(Unicidade) Para provar a unicidade de $q(x)$ *e* $r(x)$, *vamos supor que existem* $q_2(x)$ *e* $r_2(x) \in K[x]$ *tais que* $p(x) = q_2(x)g(x) + r_2(x)$, *com* $r_2(x) = 0$ *ou* $gr(r_2(x)) < gr(g(x))$. *Dessa forma temos*

$$q_2(x)g(x) + r_2(x) = p(x) = q(x)g(x) + r(x), \text{ donde}$$
$$(q(x) - q_2(x))g(x) = r_2(x) - r(x)$$

Perceba que temos acima uma igualdade de polinômios, desta forma o grau do polinômio que está a esquerda da igualdade deve ser igual ao grau do polinômio que está a direita. Entretanto, pela Proposição 2, se $q(x) - q_2(x) \neq 0$ *temos* $gr((q(x) - q_2(x))g(x)) = gr(q(x) - q_2(x)) + gr(g(x))$ *que será um valor maior ou igual a* $gr(g(x))$. *Por outro lado, como* $r_2(x)$ *e* $r(x)$ *possuem grau estritamente menor que* $gr(g(x))$, *então* $gr(r_2(x) - r(x))$ *é estritamente menor que* $gr(g(x))$, *o que gera uma contradição. Desta forma temos* $q(x) - q_2(x) = 0$, *ou equivalentemente* $q(x) = q_2(x)$, *o que nos leva a* $r_2(x) - r(x) = 0$, *ou seja,* $r_2(x) = r(x)$.

■

1.3 Ideais em $K[x]$

Definição 6 *Um subconjunto não vazio I de um dado anel A é um* ***ideal*** *se as seguintes propriedades forem satisfeitas:*

i) Se $a, b \in I$, então $a + b \in I$;

ii) Se $r \in A$ e $a \in I$, então $r \cdot a \in I$.

Note que, de acordo com as propriedades acima, dado um ideal $I \subset A$, $0 \in I$, qualquer que seja este ideal. Além disso, $I = A$ e $I = \{0\}$ são Ideais.

Exemplo 1 *O subconjunto $P \subset \mathbb{Z}$, onde P é o conjunto dos números pares é um Ideal do Anel $\mathbb{Z}$.*

Exemplo 2 *Se $a \in A$, então o conjunto $I(a) = \{c \cdot a : c \in A\}$ é um Ideal. Este em particular, será chamado de* ***Ideal principal*** *gerado por a. Mais geralmente, se $I(a_1, ..., a_n) = \{c_1 \cdot a_1 + ... + c_n \cdot a_n : c_1, ..., c_n \in A\}$ é um Ideal de A. Os elementos $a_1, ..., a_n$ são chamados geradores deste.*

Proposição 4 *Todo ideal em $K[x]$ é da forma $I(p(x))$, com $p(x) \in K[x]$.*

Prova: *Seja I um ideal de $K[x]$. Se $I = \{0\}$ basta tomarmos $p(x) = 0$ e a proposição é válida. Caso $I \neq \{0\}$, tomemos $p(x) \in I$ não nulo com menor grau possível. Vamos provar que I é gerado por $p(x)$, ou seja, $I = I(p(x))$. Como $p(x) \in I$, então $I(p(x)) \subset I$. Para provar que $I \subset I(p(x))$ vamos considerar um polinômio qualquer $g(x) \in I$. Pelo Algorítmo da Divisão, existem polinômios $q(x)$ e $r(x)$, com $gr(r(x)) = 0$ ou $gr(r(x)) < gr(p(x))$ tais que*

$$g(x) = p(x) \cdot q(x) + r(x).$$

Como $-p(x) \cdot q(x) \in I$, temos

$$r(x) = g(x) - p(x) \cdot q(x) \in I$$

Dessa forma, se $r(x) \neq 0$, teríamos um elemento em I de grau menor que $p(x)$, o que não é possível. Sendo assim, tem-se $r(x) = 0$, logo $g(x) = p(x) \cdot q(x)$, donde $g(x) \in I(p(x))$. Contudo, $I \subset I(p(x))$ e, portanto, $I = I(p(x))$.

■

Um fato importante sobre a proposição acima é que o polinômio $p(x)$ que gera o ideal $I(p(x))$ não é único, mas há uma relação entre os polinômios que geram este ideal. Se $g(x) \in K[x]$ também gera $I(p(x))$, então dizemos que ele é **associado** à $p(x)$. Sabendo que dois polinômios associados geram

o mesmo ideal conclui-se que eles são iguais a menos de uma multiplicação por um polinômio constante:

Dado um ideal $I \subset K[x]$, com $I = I(p(x)) = I(g(x))$, temos

$$p(x) = g(x) \cdot a(x) \text{ e } g(x) = p(x) \cdot b(x).$$

Se $p(x) = 0$, então $g(x) = 0$, logo são iguais. Se $p(x) \neq 0$, das relações acima obtemos

$$\begin{gathered} p(x) = p(x) \cdot a(x) \cdot b(x) \\ \Rightarrow p(x) \cdot a(x) \cdot b(x) - p(x) = 0 \\ \Rightarrow p(x) \cdot (a(x) \cdot b(x) - 1) = 0, \text{ sendo } K[x] \text{ um Domínio} \\ \Rightarrow a(x) \cdot b(x) = 1 \end{gathered}$$

Dessa forma, $a(x)$ e $b(x)$ são invertíveis em $K[x]$ e portanto invertíveis em K, de acordo com a Proposição 2. Contudo, $a(x)$ e $b(x)$ são polinômios constantes.

Corolário 1 *Seja $I \neq \{0\}$ um ideal de $K[x]$. Existe apenas um polinômio mônico em $K[x]$ que gera I.*

Prova: *De fato, pela proposição 3, todo ideal em $K[x]$ é gerado por um polinômio $p(x) \in K[x]$, dessa forma consideremos o coeficiente do seu termo líder como sendo a_n. Se $a_n = 1$, nada temos a demonstrar quanto à existência, caso $a_n \neq 1$ consideremos o polinômio $s(x) = a_n^{-1} \cdot p(x)$. Por definição de Ideal, $s(x) \in I(p(x))$. Além disso, $p(x) = a_n \cdot s(x)$, donde $s(x)$ também gera o ideal $I(p(x))$. Resta mostrar que $s(x)$ é único. Seja $t(x) \in K[x]$ um polinômio mônico que gera $I(p(x))$. Dessa forma $s(x)$ e $t(x)$ são associados e portanto são iguais a menos de uma multiplicação por um polinômio constante, porém multiplicar um deles por um polinômio constante diferente de 1 acabaria por torna-lo não mônico, logo $s(x)$ e $t(x)$ são iguais.*

■

Definição 7 *Um Anel onde todo Ideal é principal é chamado de* ***Anel Principal****, caso seja um domínio*[1] *de integridade podemos chamar de* ***Domínio Princial*** *ou* ***Domínio de Ideais Principais, DIP****.*

Pelas proposições 4 e 2-ii), temos que $K[x]$ é um Domínio de Ideais Principais, por exemplo.

[1] O anel A é domínio de integridade quando: dados $a, b \in A$, se $a \neq 0$ e $b \neq 0$, então $a \cdot b \neq 0$

1.4 O Anel K[x]/I

1.4.1 Congruência Módulo I e Anéis Quocientes

Definição 8 *Seja I um ideal do Anel A e sejam $a, b \in A$. Diremos que a é congruente a b módulo I sempre que $a - b \in I$. Notação: $a \equiv b(modI)$.*

Exemplo 3 *Os casos de congruência módulo n no Anel $\mathbb{Z}$, para algum $n \in \mathbb{Z}$ são congruências módulo $I(n)$, mais precisamente, $\mathbb{Z}_n = \mathbb{Z}/(n)$.*

Exemplo 4 *Seja F o Anel de todas as funções contínuas de $\mathbb{R}$ em $\mathbb{R}$ e seja $I \subset F$ o ideal das funções g tais que $g(2) = 0$. Se $f(x) = x^2 + 6$ e $h(x) = 5x$, temos*

$$(f - h)(2) = f(2) - h(2) = 0$$

Logo, $f \equiv h(modI)$.

Proposição 5 *Seja I um ideal do Anel A. Se $a \equiv b(modI)$ e $c \equiv d(modI)$, então*
i) $a + c \equiv b + d(modI)$;
ii) $a \cdot c \equiv b \cdot d(modI)$.

Prova: *i) Temos por hipótese que $(a - b) \in I$ e $(c - d) \in I$, sendo I um ideal, a soma de dois de seus elementos também estará em I. Desta forma, $(a-b+c-d) \in I$, donde $(a+c-(b+d)) \in I$, e portanto $a+c \equiv b+d(modI)$.*
ii)Pela definição de Ideal, podemos afirmar que $c \cdot (a-b) \in I$ e $b \cdot (c-d) \in I$, dessa forma $c \cdot (a-b) + b \cdot (c-d) \in I$, donde $(ac-bd) \in I$, logo $ac \equiv bd(modI)$.

■

Definição 9 *Seja A um anel. Dado $a \in A$, chamaremos de classe de congruência de a módulo I o conjunto de todos elementos de A que são congruentes a ele módulo I. Denotaremos essa classe por $\overline{a}$.*

Denotaremos por A/I o conjunto de todas as classes de equivalência de um dado Anel A módulo I.

Definem-se em A/I duas operações:

Adição: $\overline{a} + \overline{b} = \overline{a + b}$
Multiplicação: $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$

Veja que de fato as leis acima definem efetivamente operações em A/I, pois sendo
$\overline{a} = \{z : a \equiv z(modI)\}$ e $\overline{b} = \{z' : b \equiv z'(modI)\}$, pela Proposição 5 verifica-se que

$$\overline{a} + \overline{b} = \{z + z' : z \in \overline{a}, z' \in \overline{b}\} = \{z + z' : a + b \equiv z + z'(mod I)\} = \overline{a + b}$$

e de forma análoga,

$$\overline{a} \cdot \overline{b} = \{z \cdot z' : z \in \overline{a}, z' \in \overline{b}\} = \{z \cdot z' : a \cdot b \equiv z \cdot z'(mod I)\} = \overline{a \cdot b}$$

Com as operações definidas acima, A/I é um Anel, chamado de Anel Quociente. Onde $\overline{0}$ é o elemento neutro da adição, $\overline{1}$ é o elemento neutro da multiplicação e $\overline{-a}$ é o inverso aditivo de $\overline{a}$.

1.4.2 $K[x]/I$

Seja I um Ideal em $K[x]$. Sendo $K[x]$ um DIP, existe então um $p(x) \in K[x]$ mônico tal que $I = I(p(x))$. Se $p(x)$ é nulo temos $K[x]/(0) = K[x]$, ou seja, se $g(x) \in K[x]$, então $\overline{g(x)} = \{r(x) \in K[x] : g(x) - r(x) \in I(0)\} = \{g(x)\}$. Caso tenhamos $p(x) = a$, onde $a \in K - \{0\}$, então $K[x]/(a) = (0)$, ou seja, dado novamente um $g(x) \in K[x]$, então $\overline{g(x)} = \{r(x) \in K[x] : g(x) - r(x) \in I(a)\} = \{r(x) \in K[x] : r(x) \in I(a) = K[x]\} = \overline{0}$. Vejamos agora o caso não trivial, onde

$$n = gr(p(x)) > 0.$$

Temos, $K[x]/I(p(x)) = \{\overline{g(x)} : g(x) \in K[x]\}$, onde $\overline{g(x)} = \{h(x) \in K[x] : g(x) - h(x) \in I(p(x))\}$. Segundo o algorítmo da divisão em $K[x]$, existem únicos $q(x)$ e $r(x)$, com $gr(r(x)) = 0$ ou $gr(r(x)) < gr(p(x)) = n$ tais que $g(x) = q(x)p(x) + r(x)$, sendo que desta última igualdade obtemos:

$$\overline{g(x)} = \overline{q(x)p(x) + r(x)} = \overline{q(x)p(x)} + \overline{r(x)} = \overline{r(x)}$$

Portanto, toda classe $\overline{g(x)} \in K[x]/I(p(x))$ possui um representante de grau $\leq n-1$, sendo que este representante é único, pois se tomarmos $r_2(x) \in K[x]$, com $gr(r_2(x)) \leq n-1$ e tal que $\overline{g(x)} = \overline{r(x)} = \overline{r_2(x)}$, teríamos $\overline{r(x)} - \overline{r_2(x)} = \overline{0}$, donde $\overline{r(x) - r_2(x)} = \overline{0}$. Se $r(x) \neq r_2(x)$ teremos $r(x) - r_2(x) = p(x)s(x)$, com $s(x) \in K[x]$, o que seria uma contradição, visto que o grau do polinômio que está à esquerda da igualdade é sempre $\leq n - 1$ e o grau do que está à direita é sempre $\geq n$. Logo $r(x) = r_2(x)$.

Contudo, podemos afirmar que $K[x]/I(p(x)) = \{\overline{r(x)} : r(x) \in K[x]$ e $gr(r(x)) \leq n - 1\}$.

Analisando ainda o caso não trivial em que $n = gr(p(x)) > 0$, mostraremos agora que o anel quociente $K[x]/I$ **é um espaço vetorial sobre** K sendo o conjunto $\{1, \overline{x}, \overline{x^2}, ..., \overline{x^{n-1}}\}$ uma base para ele.

Seja $\overline{K} = \{\overline{a} : a \in K\} \subset K[x]/I$ o conjunto das classes dos polinômios constantes em $K[x]$. A aplicação

$$\varphi_{|K} : K \to K[x]/I,\ a \to \overline{a}$$

define um homomorfismo de núcleo nulo, consequentemente injetivo, e possui como imagem o conjunto $\overline{K}$. Dessa forma, podemos concluir que $\overline{K}$ é ismorfo a K e faremos a identificação $\overline{a} := a$, para todo $a \in K$ e obtemos a inclusão $K \subset K[x]/I$.

Pelas conclusões acima, podemos agora escrever:
$K[x]/I(p(x)) = \{\overline{r(x)} : r(x) \in K[x] \text{ e } gr(r(x)) \leq n-1\}$
$= \{\overline{b_0 + ... + b_{n-1}x^{n-1}} : b_i \in K\}$
$= \{\overline{b_0} + \overline{b_1}\overline{x} + ... + \overline{b_{n-1}}\overline{x^{n-1}} : b_i \in K\}$
$= \{b_0 + b_1\overline{x} + ... + b_{n-1}\overline{x^{n-1}} : b_i \in K\}$

Note que dessa forma podemos afirmar que todos os elementos de $K[x]/I(p(x))$ são combinações lineares de $1, \overline{x}, ..., \overline{x^{n-1}} \in K[x]/I(p(x))$ com coeficientes em K. Além disso, as propriedades de Espaço Vetorial com relação à soma entre vetores e multiplicação por escalar são satisfeitas, pois $K[x]/I$ é Anel e com respeito à multiplicação por elementos de K é distributivo, comutativo e $1 \cdot \overline{g(x)} = \overline{g(x)}$, $\forall\ \overline{g(x)} \in K[x]/I$. Por fim, mostraremos que $1, \overline{x}, ..., \overline{x^{n-1}} \in K[x]/I(p(x))$ não só gera esse espaço vetorial como também é um conjunto linearmente independente e por isso uma base. Suponha que

$\overline{c_0 + c_1\overline{x} + ... + c_{n-1}\overline{x^{n-1}}} = \overline{0}$, então
$\overline{c_0 + ... + c_{n-1}x^{n-1}} = \overline{0}$, ou seja,
$c_0 + ... + c_{n-1}x^{n-1} = p(x) \cdot q(x)$, porém, como $gr(p(x)) = n$, temos necessariamente que $q(x) = 0$, donde $c_0 = c_1 = ... = c_{n-1} = 0$.

■

A partir daqui escreveremos $K[x]/I(p(x))$ apenas como $K[x]/(p(x))$.

Proposição 6 *Todo ideal de $K[x]/p(x)$ é da forma $I(\overline{f(x)})$ onde $f(x)$ é um divisor de $p(x)$.*
Prova: *Seja I um ideal de $K[x]/p(x)$. Considere o conjunto dos polinômios em $K[x]$ tais que sua classe pertence ao ideal I:*

$$J = \{g(x) \in K[x]; \overline{g(x)} \in I\}$$

A princípio perceba que J é um ideal de $K[x]$:

i)Se $g_1(x), g_2(x) \in J$, então $\overline{g_1(x)}, \overline{g_2(x)} \in I$, dessa forma temos

$$\overline{g_1(x)} + \overline{g_2(x)} = \overline{g_1(x) + g_2(x)} \in I$$

e, consequentemente, $g_1(x) + g_2(x) \in J$

ii)Se $g(x) \in J$ e $h(x) \in K[x]$, como $\overline{g(x)} \in I$ e $\overline{h(x)} \in K[x]/p(x)$ temos

$$\overline{g(x)} \cdot \overline{h(x)} = \overline{g(x) \cdot h(x)} \in I, \text{ logo } g(x) \cdot h(x) \in J$$

Note que $J \neq \{0\}$, *pois* $\overline{p(x)} = \overline{0} \in I$, *e portanto* $p(x) \in J$. *Além disso, segue da proposição 4 que existe* $f(x) \in K[x]$, *não nulo, tal que* $J = I(f(x))$. *Dessa forma, temos que* $f(x)$ *é divisor de* $p(x)$.

Contudo, veja que podemos escrever o ideal I *da seguinte forma:*

$$I = \{\overline{g(x)}; g(x) \in J\} = \{\overline{h(x)f(x)}; \overline{h(x)} \in K[x]/p(x)\} = I(\overline{f(x)}).$$

■

Capítulo 2

Códigos Lineares

Afinal, o que é um *código*? A melhor maneira de entender esse conceito é imagina-lo como sendo um idioma, composto de palavras. A título de ilustração considere o seguinte exemplo:

Imagine que você quer passar uma ordem ao seu computador, cujo idioma obviamente não é igual ao seu, pois estes só entendem o chamado código binário. Obviamente você terá que traduzir para que ele entenda, essa mudaça de idioma é o que chamamos de *codificar*, essa primeira codificação vai gerar o que chamaremos de *código fonte*. Codificada a ordem você agora precisa transmití-la ao computador, por um caminho que chamaremos de *Canal* (rede wi-fi, por exemplo), que por vezes pode causar interferência no que está sendo transmitido e, por conta disso, sua ordem chega ao computador com *erros*, digamos que um dos dígitos trocados (digo dígitos porque o código binário é composto por números), por conta disso há uma má interpretação e o computador acaba por não cumprir tal ordem. É aí que nasce a necessidade de um "código corretor de erros".

Dando continuidade ao exemplo acima, suponha agora que os comandos, que chamaremos de *palavras* no decorrer do texto, possíveis para enviarmos ao computador, já escritos em código binário, são 11 e 00. Além disso, seja 11 a ordem que nós enviamos. Caso de fato ocorra uma interferência e o código chegue ao computador como 10, ele não saberá sequer qual dos dois dígitos está errado e portanto não poderá corrigir o erro. Mas se implementarmos nosso código reeescrevendo-o como 1101 e 0010? Essa seria uma segunda codificação, afim de aprimorar nosso código fonte, teríamos agora o chamado *código corretor de erros*. Perceba que havendo novamente um erro na transmissão e ao enviar a ordem 1101 ela chegue ao computador como 0101, fica mais claro qual é o dígito errado, já que comparando com as duas únicas ordens conhecidas ela difere apenas um dígito da primeira, enquanto que da segunda ela difere três dígitos. Essa quantidade de dígitos diferentes

entre as palavras 0101, 1101 e 0010 é o que chamamos mais tarde de *distância* entre elas.

Agora que o leitor está mais familiarizado com o objeto de estudo deste trabalho, vejamos algumas definições que vão formalizar as ideias transmitidas no texto acima.

2.1 Definições

Formalmente, dado um conjunto finito A e um natural n, chamamos de código corretor de erros um subconjunto próprio qualquer de A^n, sendo este último o conjunto de todas as n-uplas com elementos em A. O fato desse subconjunto ter que ser próprio vai ser melhor entendido mais a frente, o leitor perceberá que, se o código corretor fosse o A^n, então a distância mínima entre as palavras seria 1 e, consequentemente, o código corretor não seria capaz de corrigir nada, perdendo assim seu sentido.

A seguinte definição formaliza a idéia de distância entre palavras de um código.

Definição 10 *(Métrica de Hamming[1]) Dados dois elementos $u, v \in A^n$, a distância de hamming entre u e v é definida da seguinte maneira:*

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

ou seja, a distância entre u e v é a quantidade caracteres diferentes entre eles, como mostra o seguinte caso:

Sendo $A = \{0,1\}$, em A^4 temos
$d(0101, 1101) = 1$ e $d(0101, 0010) = 3$

A distância definida acima satisfaz as propriedades usuais de métrica na matemática:
(i) Positividade: $d(u,v) \geq 0$;
(ii) Simetria: $d(u,v) = d(v,u)$;
(iii) Desigualdade triangular: $d(u,v) \leq d(u,w) + d(w,v)$.

As duas primeiras propriedades seguem diretamente da definição, provaremos a seguir apenas a desigualdade triangular.

Prova: A princípio, recorde da definição de distância de hamming que a contribuição das i-ésimas coordenadas de u e v para $d(u,v)$ é igual a zero se $u_i = v_i$, e igual a um se $u_i \neq v_i$. Analisando cada um desses dois casos, temos: No caso em que a contribuição é zero, obviamente a contribuição das

[1] Em homenagem a Richard Hamming, que introduzio o conceito fundamental sobre códigos de Hamming: Error detecting and error correcting codes, em 1950.

i-ésimas coordenadas a $d(u,v)$ é menor ou igual a das i-ésimas coordenadas a $d(u,w)+d(w,v)$ (que por sua vez podem ser $=0,1$ ou 2).

No outro caso, temos que $u_i \neq v_i$ e, portanto, não podemos ter $u_i = w_i$ e $w_i = v_i$. Consequentemente, a contribuição das i-ésimas coordenadas de u e v para $d(u,w)+d(w,v)$ é maior ou igual a um, que é a contribuição das i-ésimas coordenadas a $d(u,v)$.

■

Chamaremos de **distância mínima** de um código C o número

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

Definição 11 *(disco e esfera em C) Dados um elemento $a \in A^n$ e um número real $t \geq 0$, definimos o disco e a esfera de centro a e raio t como sendo, respectivamente, os conjuntos*

$$D(a,t) = \{u \in A^n; d(u,a) \leq t\},$$
$$S(a,t) = \{u \in A^n; d(u,a) = t\}.$$

Dado um código C com distância mínima d, defini-se

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

onde $[t]$ representa a parte inteira de t.

Lema 1 *Seja C um código com distância mínima d. Se c e c' são palavras distintas de C, então*

$$D(c,\kappa) \cap D(c',\kappa) = \emptyset.$$

Prova: *Suponha, por absurso, $D(c,\kappa) \cap D(c',\kappa) \neq \emptyset$. Dessa forma, existe x nesta intersecção tal que $d(x,c) \leq \kappa$ e $d(x,c') \leq \kappa$. Pela simetria e desigualdade triangular, temos*

$$d(c,c') \leq d(c,x) + d(x,c') \leq 2\kappa.$$

Como $\kappa = \left[\frac{d-1}{2}\right]$, perceba que temos $2\kappa = d-1$ se d for ímpar e $2\kappa < d-1$ se d for par. Logo,

$$d(c,c') \leq d(c,x) + d(x,c') \leq 2\kappa \leq d-1,$$

absurdo, pois $d(c,c') \geq d$.

■

Teorema 5 *Seja C um código com distância mínima d. Então C pode corrigir até $\kappa = \left[\frac{d-1}{2}\right]$ erros e detectar até $d-1$ erros.*

Prova: *Se durante a transmissão de uma palavra c do código cometemos t erros, com $t \leq \kappa$, recebendo assim a palavra r, então $d(r,c) = t \leq \kappa$; dessa forma temos que c é univocamente determinado por r, pois, de acordo com o lema 1, a distância de r a qualquer outra palavra do código é maior que κ.*

Por outro lado, dada uma palavra c do código, podemos nela introduzir até $d-1$ erros sem que ela se torne qualquer outra palavra do código, e assim, é possível detectar o erro.

■

Vale ressaltar uma pequena sutileza no teorema acima no que se refere à detectar até $d-1$ erros: conseguimos detectar apenas que há erro na palavra recebida, entretanto não sabemos quais ou quantos dígitos são os errados.

Definição 12 *Diremos que uma função $F : A^n \to A^n$ é uma **isometria** de A^n se ela preserva distâncias de Hamming. Ou seja,*

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

Definição 13 *Sejam C e C' dois códigos em A^n, diremos que eles são **equivalentes** se existir uma isometria F de A^n tal que $F(C) = C'$.*

Teorema 6 [2] *Sejam C e C' dois códigos em A^n. Temos que C e C' são equivalentes se, e somente se, existem uma permutação π de $\{1, ..., n\}$ e bijeções $f_1, ..., f_n$ de A tais que*

$$C' = \{(f_{\pi(1)}(x_{\pi(1)}), ..., (f_{\pi(n)}(x_{\pi(n)}); \; (x_1, ..., x_n) \in C\}$$

O teorema acima é muito importante porque caracteriza códigos equivalentes: Sejam C e C' dois códigos em A^n, dizemos que estes são equivalente se, e somente se, um pode ser obtido a partir do outro por meio de uma sequência de operações do tipo:
i)Substituição das letras numa dada posição fixa em todas as palavras do código por meio de uma bijeção de A.
ii) Permutação das posições das letras em todas as palavras do código, mediante uma permutação fixa de $\{1, ..., n\}$.

Exemplo 5 *Considere $A = \{1, 2, 3, 4, 5\}$ e o código $C = \{112, 123, 345, 224\} \subset A^3$. Vamos através do teorema 6, obter um código C' equivalente à C. Tomemos f_1, f_2, f_3 como sendo as seguintes bijeções em A:*

[2] A prova deste teorema pode ser vista no apêndice da referência [1].

$$\begin{array}{ccc} f_1 : A \to A & f_2 : A \to A & f_3 : A \to A \\ 1 \mapsto 1 & 1 \mapsto 1 & 1 \mapsto 2 \\ 2 \mapsto 2 & 2 \mapsto 3 & 2 \mapsto 3 \\ 3 \mapsto 3 & 3 \mapsto 4 & 3 \mapsto 1 \\ 4 \mapsto 4 & 4 \mapsto 2 & 4 \mapsto 5 \\ 5 \mapsto 5 & 5 \mapsto 5 & 5 \mapsto 4 \end{array}$$

Tomemos também a permutação de $\{1,2,3\}$, $\pi = (2,3)$. *Dessa forma, aplicando as operações i) e ii) descritas acima, temos*

$$\begin{array}{ccccc} 112 & \dashrightarrow & f_1(1)f_2(1)f_3(2) & \dashrightarrow & 131 \\ 123 & \dashrightarrow & f_1(1)f_2(2)f_3(3) & \dashrightarrow & 113 \\ 345 & \dashrightarrow & f_1(3)f_2(4)f_3(5) & \dashrightarrow & 342 \\ 224 & \dashrightarrow & f_1(2)f_2(2)f_3(4) & \dashrightarrow & 253 \end{array}$$

Contudo, nosso código equivalente a C *é* $C' = \{131, 113, 342, 253\}$.

Denotaremos por K um corpo finito com q elementos tomado como alfabeto. Dessa forma, dado um n natural, temos que K^n é um espaço vetorial de dimensão n sobre o corpo K.

Definição 14 *Um código* $C \subset K^n$ *será chamado de* ***código linear*** *se for um subespaço vetorial de* K^n.

Perceba que dado um corpo finito K com q elementos, um código linear $C \subset K^n$ de dimensão k possui exatamente q^k elementos, pois, sendo $\{v_1, ..., v_k\}$ uma base de C todos os seus elementos são escritos como combinação linear dessa base

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k,$$

onde os λ_i, $i = 1, ..., k$, são elementos de K. Dessa forma podemos construir a relação

$$|C| = q^k, \text{ donde } \dim C = k = \log_q q^k = \log_q |C|.$$

Definição 15 *Seja* $u \in K^n$, *chamaremos de peso de* u *o inteiro*

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

ou seja, o peso de u é a quantidade de caracteres diferentes de 0 que ele possui. Dessa forma temos que, $\omega(u) = d(u, 0)$.

Definição 16 *O peso de um código linear* C *é o inteiro*

$$\omega(C) := min\{\omega(u) : u \in C - \{0\}\}$$

Proposição 7 *Seja $C \subset K^n$ um código linear com distância mínima d. Temos que*
1) $\forall\ x, y \in K^n$, $d(x,y) = \omega(x-y)$.
2) $d = \omega(C)$.

Prova: *Perceba que dadas as definições de distância de hamming e peso, o item 1) se torna óbvio. O item 2) decorre do fato que tomados $x, y \in C$, com $x \neq y$, temos $z = x - y \in C - \{0\}$ e $d(x,y) = \omega(z)$, dessa forma se x e y forem o par de elementos em C com a menor distância possível, o peso de z será também mínimo e portanto $d = d(x,y) = \omega(z) = \omega(C)$.*

■

Definição 17 *Dados dois códigos lineares em K^n, dizemos que estes são* ***linearmente equivalentes*** *se existir uma isometria linear $T : K^n \to K^n$ tal que $T(C) = C'$.*

Esta definição é muito parecida com a definição 13, e nos trará um resultado semelhante ao teorema 6:

Dois códigos lineares são linearmente equivalentes se, e somente se, cada um deles pode ser obtido a partir do outro mediante uma sequência de operações do tipo:
i) Multiplicação dos elementos numa dada posição fixa por um escalar não nulo em todas as palavras.
ii) Permutação das posições de todas as palavras do código, mediante uma permutação fixa de $\{1, 2, ..., n\}$.

Em outras palavras, dado um código $C \in K^n$, um código equivalente a ele é do tipo:

$$C' = \{c_1 x_{\pi(1)}, ..., c_n x_{\pi(n)};\ (x_1, ..., x_n) \in C\}$$

O caminho até este resultado está mais detalhado em [1].

2.2 Matriz Geradora

Seja $C \subset K^n$ um código linear. Dessa forma C é um subespaço vetorial e portanto é gerado por uma base $\beta = \{v_1, ..., v_k\}$, onde $k \leq n$. Chamaremos de **matriz geradora do código C** associada à base β a matriz:

$$\mathrm{G} = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kn} \end{pmatrix}$$

Com uma matriz geradora como esta podemos fazer uma codificação, veja: Dado um código fonte K^k montaremos uma transformação linear afim de transforma-lo num código corretor de erros $C \in K^n$

$$\begin{array}{c} T : K^k \to K^n \\ x \mapsto xG \end{array}$$

donde $T(K^k) = C$.

A matriz G não é única, pois depende da base escolhida de C, mas assim como podemos obter uma nova base através de operações elementares[3] em vetores de uma base dada, o mesmo podemos fazer com a matriz G, obtendo assim uma matriz G' geradora do mesmo código.

Exemplo 6 *Tomemos* $K = \mathbb{Z}_2$ *e*

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

A transformação linear abaixo possui como imagem um código corretor gerado pela matriz G:

$$\begin{array}{c} T : K^3 \to K^5 \\ x \mapsto xG \end{array}$$

Dada uma palavra $(1, 0, 1)$ *em* K^3*, depois de codificada pela transformação acima ela se tornará* $(0, 1, 0, 1, 0) \in K^5$

O processo de decodificação de uma palavra $y \in K^n$ de um código linear gerado por uma matriz G consiste em encontrar um $x \in K^k$ que seja solução do sistema $xG = y$. Algo que facilita muito esse processo é o escalonamento da matriz G, pois como foi dito acima podemos utilizar as operações elementares para obter uma nova matriz geradora do mesmo código.

Definição 18 *Dizemos que uma matriz geradora G está na forma padrão se*

$$G = (Id_k|A),$$

ou seja, as primeiras k colunas da matriz G formam a matriz identidade de ordem k e A é uma matriz de ordem $k \times (n - k)$.

Note que dada uma matriz G geradora de um código C, as seguintes operações sobre G são equivalentes às operações i) e ii) em C, descritas na página 22 deste trabalho:
C1) Permutação das colunas.

[3] As três operações elementares são descritas em livros de Álgebra Linear ou até mesmo nos capítulos sobre matrizes nos livros de ensino básico.

C2) Multiplicação de uma coluna por um escalar não nulo.
Dessa forma, efetuando operações C1 e C2 na matriz G obteremos uma matriz G', que por sua vez vai gerar um código C' linearmente equivalente ao código C.

Teorema 7 *Dado um código C, existe um código C' com matriz geradora na forma padrão.*

A prova deste teorema resume-se a escalonamento de matrizes, feito utilizando as operações elementares já comentadas. O fato das linhas de uma matriz geradora G serem linearmente independentes garante que nenhuma delas será nula, mesmo depois do escalonamento. Além disso, G tendo k linhas, se uma das k primeiras colunas for nula, basta utilizarmos a operação C1.

2.3 Códigos Duais

Nesta sessão definiremos a matriz *teste de paridade* de um código corretor C, uma matriz talvez até mais relevante que a matriz geradora de C, sendo uma de suas funções a de verificar se uma dada palavra pertence ou não a um determinado código.

Definiremos o seguinte produto interno em K^n:
Dados $u = (u_1, ..., u_n)$ e $v = (v_1, ..., v_n)$ em K^n, temos

$$\langle u, v\rangle = u_1v_1 + ... + u_nv_n \in K$$

É fácil ver que este produto é simétrico, $\langle u, v\rangle = \langle v, u\rangle$, e também bilinear $\langle u + \lambda w, v\rangle = \langle u, v\rangle + \lambda\langle w, v\rangle$, para todo $\lambda \in K$.

Proposição 8 *Dado um código $C \subset K^n$ com matriz geradora G, o conjunto*

$$C^\perp = \{v \in K^n;\ \langle u, v\rangle = 0,\ \forall u \in C\} \subset K^n,$$

conhecido na álgebra linear como complemento ortogonal de C, é um subespaço vetorial de K^n que chamaremos de ***código dual*** [4] *de C. Além disso, $x \in C^\perp \Leftrightarrow Gx^t = 0$.*

Prova: *Considere $u, v \in C^\perp$ e $\lambda \in K$. Temos, para todo $x \in C$, que*

$$\langle u + \lambda v, x\rangle = \langle u, x\rangle + \lambda\langle v, x\rangle = 0,$$

e, portanto, $u + \lambda v \in C^\perp$, provando assim que $C^\perp$ é subespaço vetorial de K^n. Quanto à segunda afirmação temos que $x \in C^\perp$ se, e somente se, x é ortogonal (produto interno nulo) a todos os elementos de C, em particular aos elementos de uma dada base de C, logo $x \in C^\perp$ se, e somente se, $Gx^t = 0$, visto que as linhas de G formam uma base para C.

[4] Em Álgebra Linear, o espaço dual de um dado espaço vetorial V sobre K é o espaço dos funcionais linear $f : V \to K$, definição esta diferente da de código dual.

■

Proposição 9 *Seja $C \subset K^n$ um código linear de dimensão k com matriz geradora $G = (Id_k|A)$, na forma padrão. Então*
i)dim $C^{\perp} = n - k$;
ii)$H = (-A^t|Id_{n-k})$ é uma matriz geradora de $C^{\perp}$.

Prova: *i) Pela proposição 8, $x = (x_1, ..., x_n) \in C^{\perp}$ se, e somente se, $Gx^t = 0$. Donde,*

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & a_{1(k+1)} & \cdots & a_{1n} \\ 0 & 1 & \cdots & 0 & a_{2(k+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & a_{k(k+1)} & \cdots & a_{kn} \end{pmatrix} \cdot (x_1, ..., x_n)^t = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} x_1 + 0 + \cdots + 0 + a_{1(k+1)}x_{k+1} + \cdots + a_{1n}x_n \\ 0 + x_2 + \cdots + 0 + a_{2(k+1)}x_{k+1} + \cdots + a_{2n}x_n \\ \vdots \\ 0 + 0 + \cdots + x_k + a_{k(k+1)}x_{k+1} + \cdots + a_{kn}x_n \end{pmatrix} = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{array}{l} x_1 = -a_{1(k+1)}x_{k+1} - \cdots - a_{1n}x_n \\ x_2 = -a_{2(k+1)}x_{k+1} - \cdots - a_{2n}x_n \\ \vdots \\ x_k = -a_{k(k+1)}x_{k+1} - \cdots - a_{kn}x_n \end{array} \Leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = -A \begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix} \ (*)$$

Portanto, sendo q o número de elementos no corpo finito K, o código $C^{\perp}$ possui q^{n-k} palavras, já que os a_{ij} da matriz A são fixos e existem q escolhas possíveis para cada um dos $x_{k+1}, ..., x_n$. Logo, $C^{\perp}$ tem dimensão $n - k$.[5]

ii)Por conta do bloco Id_{n-k} podemos garantir que as linhas da matriz H são linearmente independentes, portanto elas foram uma base para um espaço vetorial de dimensão $n - k$. Além disso, cada uma dessas linhas são ortogonais a G [6] *e, por isso, o espaço gerado por elas está contido em $C^{\perp}$ cuja dimensão também é $n - k$. Segue daí que o espaço gerado pelas linhas de H é $C^{\perp}$.*

■

Diremos que a matriz geradora do código linear $C^{\perp}$, $H = (-A^t|Id_{n-k})$, escrita dessa forma, **é também uma matriz na forma padrão**. A intenção

[5] Em caso de dúvida quanto à relação entre número de elementos e dimensão, ver página 21 deste trabalho.

[6] Perceba que as linhas de H são justamente os vetores x que satisfazem $(*)$ com $x_{k+m} = 1$ na linha m.

aqui é de complementar a definição dada sobre a forma padrão da matriz geradora.

O Corolário a seguir generaliza o item i) da proposição acima para qualquer código em K^n, independentemente de sua matriz estar na forma padrão. A prova deste resultado pode ser encontrada em [1] e não será posta aqui porque utiliza resultados que fogem aos nossos objetivos.

Corolário 1 *Se D é um código linear em K^n de dimensão k, então $D^\perp$ é um código de dimensão $n-k$.*

Lema 2 *Suponha que C seja um código de dimensão k em K^n com matriz geradora G. Uma matriz H de ordem $(n-k)\times n$, com coeficientes em K e com linhas linearmente independentes, é uma matriz geradora de $C^\perp$ se, e somente se,*

$$G \cdot H^t = 0.$$

Prova: *Perceba inicialmente que as linhas linearmente independentes de H geram um subespaço vetorial de K^n de dimensão $n-k$, que é igual à dimensão de $C^\perp$. Por outro lado, representando por $h_1, ..., h_{n-k}$ e por $g_1, ..., g_k$ as linhas de H e G respectivamente, temos que*

$$(G \cdot H^t)_{i,j} = \langle g_i, h_j \rangle.$$

Desta forma, $G \cdot H^t = 0$ equivale a dizer que todos os vetores do subespaço gerado pelas linhas de H estão em $C^\perp$, mas como este subespaço contido em $C^\perp$ tem a mesma dimensão que ele, conclui-se que são o mesmo.

■

Um leitor com certo conhecimento em Álgebra Linear poderia perguntar-se por que não provar o seguinte resultado usando o fato das dimensões de C e $C^\perp$ se completarem, $\dim C + \dim C^\perp = n = \dim K^n$, como habitualmente é feito em livros de álgebra linear. Porém, em corpos finitos, esses dois subespaços vetoriais não necessariamente são disjuntos e dessa forma não teríamos uma soma direta. A título de ilustração, veja que no exemplo 8 na página 29 deste trabalho, $C^\perp \subset C$.

Corolário 2 $(C^\perp)^\perp = C$.
Prova: *Sejam G e H respectivamente matrizes geradoras de C e $C^\perp$. Logo, $G \cdot H^t = 0$. Tomando transpostas dessa última igualdade, temos que $H \cdot G^t = 0$, logo, G é matriz geradora de $(C^\perp)^\perp$.*

■

Proposição 10 *Seja C um código linear e suponhamos que H seja uma matriz geradora de $C^\perp$. Temos então que*

$$v \in C \text{ se, e somente se, } Hv^t = 0.$$

Prova: *Este resultado decorre imediatamente do corolário 2 e da Proposição 8, pois estes mostram que* $v \in C$ *se, e somente se,* $v \in (C^{\perp})^{\perp}$ *se, e somente se,* $Hv^t = 0$.

■

Perceba que fazendo uso da proposição acima podemos determinar se um vetor v pertence ou não a um dado código apenas utilizando a matriz H de seu código dual. Esta matriz é chamada de **matriz teste de paridade**.

A matriz teste de paridade é de fato muito significativa, pois dado $y \in K^n$ seria necessário analisar se o sistema $Gx = y$ admite solução, afim de descobrir se y pertence ou não ao código gerado por G, algo que requer um custo computacional maior do que verificar se $Hy^t = 0$.

Dado um código $C \subset K^n$ com matriz teste de paridade H e um vetor $v \in K^n$, chamaremos o vetor Hv^t de **síndrome** de v.

Exemplo 7 *Tomemos o corpo finito* $\mathbb{Z}_3$ *e um código* $C \subset \mathbb{Z}_3^5$ *com matriz geradora*

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

Perceba que multiplicando a primeira e a segunda linha da matriz G *por 2, obteremos a matriz* G' *na forma padrão:*

$$G' = \begin{pmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{pmatrix}.$$

Pela proposição 9 (ii), é fácil obter a matriz teste de paridade H

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

Dados $v_1 = (1, 2, 1, 2, 0)$ *e* $v_2 = (1, 0, 2, 0, 1)$ *em* $\mathbb{Z}_3^5$*, vejamos qual destes é uma palavra pertencente ao código* C*:*

$$Hv_1^t = \begin{pmatrix} 0 & 0 \end{pmatrix} \text{ e } Hv_2^t = \begin{pmatrix} 2 & 0 \end{pmatrix}.$$

Logo, $v_1 \in C$ *e* $v_2 \notin C$.

A matriz teste de paridade traz também informações importantes sobre o peso do código, como mostra a seguinte proposição.

Proposição 11 *Seja $H_{(n-k)\times n}$ a matriz teste de paridade de um código $C \subset K^n$, com dim $C = k$. Temos que o peso de C é maior do que ou igual a s se, e somente se, quaisquer $s-1$ colunas de H são linearmente independentes.*

Prova: *($\Leftarrow$) Vamos supor inicialmente que qualquer conjunto de $s-1$ colunas de H é linearmente independente. Seja $c = (c_1, c_2, ..., c_n)$ uma palavra não nula em C e sejam h_{ij} os elementos de H. Como $Hc^t = 0$, temos*

$$\begin{aligned} h_{11}c_1 + h_{12}c_2 + ...h_{1n}c_n &= 0 \\ h_{21}c_1 + h_{22}c_2 + ...h_{2n}c_n &= 0 \\ &\vdots \\ h_{(n-k)1}c_1 + h_{(n-k)2}c_2 + ...h_{(n-k)n}c_n &= 0 \end{aligned}$$

$$\Leftrightarrow \begin{pmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{(n-k)1} \end{pmatrix} c_1 + \begin{pmatrix} h_{12} \\ h_{22} \\ \vdots \\ h_{(n-k)2} \end{pmatrix} c_2 + \cdots + \begin{pmatrix} h_{1n} \\ h_{2n} \\ \vdots \\ h_{(n-k)n} \end{pmatrix} c_n = 0.$$

Chamando de h^i a coluna i da matriz H, temos

$$h^1c_1 + h^2c_2 + ... + h^nc_n = 0$$

Perceba que se c possui $t < s$ componentes não nulas, ou seja, $t = \omega(c) < s$, teríamos uma combinação linear nula de t colunas de H,

$h^{i_1}c_{i_1} + h^{i_2}c_{i_2} + ... + h^{i_t}c_{i_t} = 0$, *onde h^{i_t} representa alguma das colunas de H.*

mas como $t \leq s-1$ então essas t colunas são L.I., e por tanto essas t componentes de c são obrigatoriamente nulas, donde $c = 0$, o que é uma contradição. Logo, $\omega(c) \geq s$, donde $\omega(C) \geq s$.

($\Rightarrow$) Reciprocamente, suponhamos que $\omega(C) \geq s$. Suponhamos também, por absurdo, que H tenha um conjunto de $s-1$ colunas que são linearmente dependentes: $h^{i_1}, h^{i_2}, ..., h^{i_{s-1}}$. Dessa forma, pela definição de L.D., existem $c_{i_1}, c_{i_2}, ..., c_{i_{s-1}}$ em K, não todos nulos, tais que

$$h^{i_1}c_{i_1} + h^{i_2}c_{i_2} + ... + h^{i_{s-1}}c_{i_{s-1}} = 0.$$

Logo, $c = (0, ..., c_{i_1}, ..., c_{i_{s-1}}, 0, ..., 0) \in C$, e consequentemente, $\omega(c) \leq s - 1 < s$, o que seria um absurdo, visto que nossa hipótese inicial é que $\omega(C) \geq s$.

■

Teorema 8 *Seja H a matriz teste de paridade de um código C. Temos que o peso de C é igual a s se, e somente se, quaisquer $s-1$ colunas de H são*

linearmente independentes e existem s colunas de H linearmente dependentes.

Prova:$(\Rightarrow)$ *Supondo $\omega(C) = s$, a proposição 10 acima nos garante que quaiquer $s-1$ colunas de H são linearmente independentes, além disso, se não existir um conjunto com s colunas de H que sejam linearmente dependentes, então teríamos $\omega(C) \geq s+1$.*

$(\Leftarrow)$ *Supondo agora que todo conjunto de $s-1$ colunas de H são linearmente dependentes e que existem s colunas linearmente dependentes, a proposição 9 nos garante que $\omega(C) \geq s$. Entretanto, $\omega(C)$ não pode ser maior do que s, pois dessa forma, também pela proposição 10, todo conjunto com s colunas de H é linearmente independente, contradizendo nossa hipótese.*

■

O seguinte exemplo demonstra a utilização de alguns resultados apresentados nesta sessão.

Exemplo 8 *Construa um código binário C de comprimento 7, dimensão 4 e distância mínima 3 por meio de uma matriz teste de paridade. Determine uma matriz geradora para este código.*

Solução: *Nossa matriz teste de paridade H possui $n-k = 3$ linhas que são linearmente independentes, além disso, Pelo Teorema 8, para $d = s = 3$, é necessário que quaisquer par de colunas em H seja linearmente independente e que exista algum conjunto com três colunas de H que seja linearmente dependentes. A seguinte matriz satisfaz estes requisitos:*

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

A matriz geradora de C possui $k = 4$ linhas linearmente independentes que geram o código, uma das formas de obtermos essas linhas é o usar o fato mostrado na proposição 10: Seja $c = (c_1, ..., c_7) \in C$, temos $Hc^t = 0$, logo

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{pmatrix} = 0 \Leftrightarrow \begin{array}{r} c_1 + c_2 + c_4 + c_7 = 0 \\ c_2 + c_3 + c_4 + c_5 = 0 \\ c_1 + c_4 + c_5 + c_6 = 0 \end{array}$$

Uma coisa interessante sobre código binário é que o inverso aditivo de um elemento é sempre ele mesmo, dessa forma, resolvendo o sistema linear acima temos

$c = (c_1, c_2, c_2 + c_4 + c_5, c_4, c_5, c_1 + c_4 + c_5, c_1 + c_2 + c_4) =$
$= c_1 \cdot (1,0,0,0,0,1,1) + c_2 \cdot (0,1,1,0,0,0,1) + c_4 \cdot (0,0,1,1,0,1,1) + c_5 \cdot (0,0,1,0,1,1,0)$

Ou seja, todo $c \in C$ é uma combinação dos vetores acima, donde nossa matriz geradora será então

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

O código que construímos acima é um exemplo de *Código de Hamming*, um dos vários tipos de códigos lineares. Um código de Hamming de ordem m sobre $K = \mathbb{Z}_2$ é um código com matriz teste de paridade H_m de ordem $m \times n$, cujas colunas são os elementos de $K^n - \{0\}$ numa ordem qualquer.

2.4 Decodificação

Este é um dos pontos altos deste trabalho, afinal a detecção e correção de erros ocorrem durante a decodificação. Definiremos inicialmente o *vetor erro* e, que nada mais é do que a diferença entre a palavra transmitida c e a palavra recebida r, ou seja, $e = r - c$. Note que desta forma, o peso do vertor erro corresponde justamente ao número de erros cometidos durante a transmissão da palavra. Outro detalhe importante é que o vetor erro possui a mesma síndrome que a palavra recebida:

Como $e = r - c$, então $He^t = H(r-c)^t = Hr^t - Hc^t = Hr^t$, já que $Hc^t = 0$.(*)

De forma análoga ao que foi feito na prova da proposição 11, podemos chamar de h^i a i-ésima coluna de H e se $e = (\alpha_1, ..., \alpha_n)$, temos

$$\sum_{i=1}^{n} \alpha_i h^i = He^t = Hr^t.$$

Lema 3 *Seja C um código linear em K^n com capacidade de correção κ. Se $r \in K^n$ e $c \in C$ são tais que $d(c,r) \leq \kappa$, então existe um único vetor e com $\omega(e) \leq \kappa$, cuja síndrome é igual à síndrome de r e tal que $c = r - e$.*

Prova: *A existência do vetor e satisfazendo as condições acima é óbvia, basta tomarmos $e = r - c$, já que $\omega(e) = d(c,r) \leq \kappa$ e (*). Para provar a unicidade, vamos supor dois vetores que satisfazem o lema, a saber: $e_1 = (\alpha_1, ..., \alpha_n)$ e $e_2 = (\beta_1, ..., \beta_n)$. Dessa forma, seja H a matriz teste de paridade de C, temos*

$$He_1^t = He_2^t \Longrightarrow \sum_{i=1}^{n} \alpha_i h^i = \sum_{i=1}^{n} \beta_i h^i$$
$$\Longrightarrow \sum_{i=1}^{n} (\alpha_i - \beta_i) h^i = 0,$$

sendo esta última uma combinação linear das n colunas de H, entretanto, como os vetores e_1 e e_2 possuem no máximo κ termos não nulos, esta combinação linear é na verdade feita com no máximo $2\kappa \leq d-1$ colunas de H e, pelo Teorema 7, concluímos que a mesma é linearmente independente, donde $\alpha_i = \beta_i$ para todo i, logo $e_1 = e_2$.

■

O resultado acima é de suma importância, pois recebida uma palavra r existirá apenas um vetor e com mesma síndrome, levando em conta as hipóteses do Lema. A pergunta agora é: Como determinar este único vetor e a partir da síndrome Hr^t?

A seguir, apresentaremos um algorítmo capaz de detectar e corrigir até um erro ($\omega(e) \leq 1$). Porém, antes disso, note que supondo um código C com distância mínima $d \geq 3$, matriz teste de paridade H e que o vetor e, introduzido entre a palavra transmitida c e a palavra recebida r, seja tal que $\omega(e) \leq 1$, temos que se $He^t = 0$, então $r \in C$ e se toma $c = r$, ou seja, nenhum erro foi introduzido na transmissão, por outro lado, se $He^t \neq 0$, então $\omega(e) = 1$ e, portanto, e tem apenas uma coordenada não nula. Nesse caso, consideremos que $e = (0, ..., \alpha, ..., 0)$ com $\alpha \neq 0$ na i-ésima posição. Logo,

$$He^t = \alpha h^i,$$

onde h^i é a i-ésima coluna da H. Portanto, não conhecendo e, mas conhecendo

$$He^t = Hr^t = \alpha h^i,$$

podemos determinar e como sendo o vetor com todas as componentes nulas exceto a i-ésima componente que é α. Note que i acima é bem determinado, pois $d \geq 3$, ou seja, com uma distância mínima igual a 2, por exemplo, poderiam haver duas palavras c e c' equidistantes de r.

Algoritmo de decodificação em códigos corretores de um erro:

Seja H a matriz teste de paridade do código C e seja r um vetor recebido. (Suponha $d \geq e$).

(i) Calcule Hr^t.

(ii) Se $Hr^t = 0$, aceite r como sendo a palavra transmitida.

(iii) Se $Hr^t = s^t \neq 0$, compare s^t com as colunas de H.

(iv) Se existirem i e α tais que $s^t = \alpha h^i$, para $\alpha \in K$, então e é a n-upla com α na posição i e zeros nas outras posições. Corrija r pondo $c = r - e$.

(v) Se o contrário de (iv) ocorrer, então mais de um erro foi cometido.

Exemplo 9 *Considere o código C construído no Exemplo 8. Esse código tem matriz teste de paridade:*

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

Seja $r = (0,1,1,1,0,1,0)$ uma palavra recebida, logo

$$He^t = Hr^t = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot h^3.$$

Portanto, $e = (0,0,1,0,0,0,0)$ e, consequentemente,

$$c = r - e = (0,1,0,1,0,1,0).$$

Introduziremos agora algumas definições e resultados necessários na construção de um algorítmo que seja capaz de corrijir um vetor r recebido com mais de um erro. Seja $C \subset K^n$ um código corretor de erros com matriz teste de paridade H. Sejam d a distância mínima de C e $\kappa = \left[\frac{d-1}{2}\right]$. Recorde que e e r possuem a mesma síndrome e, se $\omega(e) = d(r,c) \leq \kappa$, então e é univocamente determinado por r.

Definição 19 *Seja $v \in K^n$. Chamaremos de classe lateral de v segundo C o conjunto*

$$v + C = \{v + c; c \in C\}$$

É fácil ver que

$$v + C = C \Leftrightarrow v \in C$$

Lema 4 *Os vetores u e v de K^n têm a mesma síndrome se, e somente se, $u \in v + C$.*

Prova: *$Hu^t = Hv^t \Leftrightarrow H(u-v)^t = 0 \Leftrightarrow u - v = c$ com $c \in C \Leftrightarrow u = v + c$ com $c \in C \Leftrightarrow u \in v + C$.*

■

As classes laterias possuem as seguintes propriedades:
(i) $v + C = v' + C \Leftrightarrow v - v' \in C$;
(ii) $(v + C) \cap (v' + C) \neq \emptyset \Longrightarrow v + C = v' + C$;
(iii) $\cup_{v \in K^n}(v + C) = K^n$;
(iv) $|(v+C)| = |C| = q^k$, sendo k a dimensão de C e q o número de elementos em K.

Demonstraremos apenas a primeira propriedade, afim de ganhar habilidade no manuseio de classes laterais:

Prova (i): $v + C = v' + C \Leftrightarrow v \in v' + C$ e $v' \in v + C \Leftrightarrow v = v' + c'$ e $v' = v + c$, com $c, c' \in C$, $\Leftrightarrow v - v' = c' \in C$ e $v' - v = c \in C$.

■

Segue imediatamente de (ii)-(iv) acima que o número de classe laterais de C é

$$\frac{q^n}{q^k} = q^{n-k},$$

claro, afinal como duas classes ou são iguais ou são disjuntas, quaisquer dois vetores v e v' em C terão a mesma classe segundo C, apenas os que terão classes diferentes são os vetores que não estão em C, justamente $n-k$ vetores.

Definição 20 *Um vetor de peso mínimo numa classe lateral é chamado de* ***elemento líder*** *dessa classe, podendo haver mais de um líder em cada classe.*

Proposição 12 *Seja C um código linear em K^n com distância mínima d. Se $u \in K^n$ é tal que*

$$\omega(u) \leq \left[\frac{d-1}{2}\right] = \kappa,$$

então u é o único elemento líder de sua classe.

Prova: *Suponhamos dois vetores u e v tais que $\omega(u) \leq \left[\frac{d-1}{2}\right]$ e $\omega(v) \leq \left[\frac{d-1}{2}\right]$. Logo, $u - v = c \in C$, além disso*

$$\omega(u - v) \leq \omega(u) + \omega(v) \leq \left[\frac{d-1}{2}\right] + \left[\frac{d-1}{2}\right] \leq d - 1;$$

sendo assim, teríamos $\omega(u - v) = d(u, v) < d$, logo $u - v = 0$, donde $u = v$.

■

Mostraremos agora um algoritmo capaz de corrigir palavras que tenham sido recebidas com um número de erros menor ou igual à capacidade de

correção do código, que é $\kappa = \left[\frac{d-1}{2}\right]$. Começamos determinando todos os elementos $u \in K^n$ tais que $\omega(u) \leq \kappa$. Dessa forma, pela proposição acima, estaremos encontrando líderes de classes laterais, sendo cada um deles o único líder em sua classe. É claro que não há no código C um vetor u tal que $\omega(u) \leq \kappa$, pois $\omega(C) = d > \kappa$, então perceba que esses vetores que estamos buscando são todos os vetores erro e com até κ erros e que estes serão líderes de classe.

O próximo passo é organizar estes vetores numa tabela e calcular suas respectivas síndromes. Feito isso, basta seguir o algoritmo abaixo.

Algoritmo de Decodificação em códigos corretores de até κ erros
Seja r uma palavra recebida. (i) Calcule a síndrome $Hr^t = s^t$
(ii) Se $s = 0$, aceite r como sendo a palavra transmitida. Caso contrário vá para (iii).
(iii) Se s está na tabela, seja l o elemento líder da classe determinada por s; troque r por $r - l$.
(iv) Se s não está na tabela, então na mensagem recebida foram cometidos mais do que κ erros.

Exemplo 10 *Imaginemos uma situação hipotética em que enviamos uma ordem de movimento à um robô, sendo que este só obedece quatro direções: Norte, Sul, Leste e Oeste. Determinemos inicialmente nosso código fonte utilizando o alfabeto $\mathbb{Z}_2$:*

Direção	*C.F.*
Norte	*01*
Sul	*10*
Leste	*00*
Oeste	*11*

Construiremos aqui um código corretor C com capacidade de correção $\kappa = 2$, para isso, basta construir uma matriz teste de paridade para este código satisfazendo o Teorema 7 com $s = 5$:

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

Como H está na forma padrão, é possível chegar rapidamente à matriz geradora de C:

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

Fazendo a codificação do código fonte para o código corretor C, temos

Direção	*C.F.*	*C.C.*
Norte	*01*	*000111101*
Sul	*10*	*111100010*
Leste	*00*	*000000000*
Oeste	*11*	*111011111*

A seguir temos todos os vetores em $\mathbb{Z}_2^9$ com peso ≤ 2 e suas respectivas síndromes.

vetor	*síndrome*	*vetor*	*síndrome*
000000000	*0000000*	*010000010*	*1011000*
100000000	*1000000*	*010000001*	*0101111*
010000000	*0100000*	*001100000*	*0011000*
001000000	*0010000*	*001010000*	*0010100*
000100000	*0001000*	*001001000*	*0010010*
000010000	*0000100*	*001000100*	*0010001*
000001000	*0000010*	*001000010*	*1101000*
000000100	*0000001*	*001000001*	*0011111*
000000010	*1111000*	*000110000*	*0001100*
000000001	*0001111*	*000101000*	*0001010*
110000000	*1100000*	*000100100*	*0001001*
101000000	*1010000*	*000100010*	*1110000*
100100000	*1001000*	*000100001*	*0000111*
100010000	*1000100*	*000011000*	*0000110*
100001000	*1000010*	*000010100*	*0000101*
100000100	*1000001*	*000010010*	*1111100*
100000010	*0111000*	*000010001*	*0001011*
100000001	*1001111*	*000001100*	*0000011*
011000000	*0110000*	*000001010*	*1111010*
010100000	*0101000*	*000001001*	*0001101*
010010000	*0100100*	*000000110*	*1111001*
010001000	*0100010*	*000000101*	*0001110*
010000100	*0100001*	*000000011*	*1110111*

Com esses dados, o algoritmo de decodificação já pode entrar em ação. Suponhamos uma palavra recebida $r = (111011000)$. Logo $Hr^t = (1110110)^t$. Observe que esta síndrome não se encontra na tabela, o que significa que na palavra r foram cometidos mais do que 2 erros de transmissão. Agora consideremos a palavra recebida $r = (111011100)$. Logo $Hr^t = (1110111)^t$, que está na

tabela e é a síndrome referente à (000000011). *Logo, na palavra recebida haviam dois erros, fazendo a correção, temos* $c = (111011100) - (000000011) = (111011111) \in C$, *palavra esta que para o nosso robô significa "Oeste".*

Capítulo 3

Códigos Cíclicos

O leitor deve ter percebido que a codificação e sobretudo a decodificação apresentadas no capítulo anterior apesar de simples requerem muitas contas e, consequentemente, um custo computacional elevado. Parte daí a necessidade de aperfeiçoar a estrutura dos códigos corretores afim de dar-lhes novas propriedades que facilitem a codificação e a decodificação. O primeiro passo a ser dado nessa direção são os códigos cíclicos.

3.1 Caracterização dos códigos cíclicos

Para este capítulo continuaremos tratando K como sendo um corpo finito e representaremos as coordenadas de um vetor em K^n por $(x_0, ..., x_{n-1})$.

Chamaremos de R_n o Anel das classes residuais em $K[x]$ módulo $(x^n - 1)$, isto é

$$R_n = K[x]/(x^n - 1).$$

Como visto na seção 3.4, este Anel quociente, é também um espaço vetorial sobre K de dimensão n com base $\{1, \overline{x}, \overline{x^2}, ..., \overline{x^{n-1}}\}$ e, portanto, isomorfo a K^n segundo a transformação linear

$$\begin{array}{ccc} \nu : K^n & \to & R_n \\ (a_0, ..., a_{n-1}) & \mapsto & \overline{a_0 + a_1 x + ... + a_{n-1} x^{n-1}} \end{array}$$

O isomorfismo acima mostra que dado um código linear $C \subset K^n$ podemos transportá-lo para R_n, isso dará propriedades adicionais ao código, visto que agora além de subespaço vetorial ele é também um subanel.

Definição 21 *Um código linear $C \subset K^n$ será chamado de **código cíclico** se, para todo $c = (c_0, ..., c_{n-1}) \in C$, o vetor $(c_{n-1}, c_0, ..., c_{n-2})$ também pertence a C.*

Equivalentemente, o código linear C será um código cíclico se, dada a permutação π de $\{0, ..., n-1\}$ definida por

$$\pi(i) = \begin{cases} i-1, & \text{se } i \geq 1 \\ n-1, & \text{se } i = 0 \end{cases}$$

e sendo

$$T_\pi(c_0, c_1, ..., c_{n-1}) = (c_{n-1}, c_0, ..., c_{n-2}),$$

temos que $T_\pi(c) \in C, \forall c \in C$, ou seja, $T_\pi(C) \subset C$.

A transformação linear T_π em K^n traduz-se, por meio da bijeção ν, na multiplicação por $\overline{x}$ em R_n. Ou seja, tomando $c = (c_0, c_1, ..., c_{n-1})$, temos $T_\pi(c) = (c_{n-1}, c_0, ..., c_{n-2})$ e

$$\nu(T_\pi(c)) = \overline{c_{n-1} + c_0 x + ... + c_{n-2}x^{n-1}} = \overline{x} \cdot \overline{c_0 + c_1 x + ... + c_{n-1}x^{n-1}} = \overline{x} \cdot \nu(c).$$

Exemplo 11 *Seja $v \in K^n$. O espaço vetorial*

$$\langle v \rangle = \alpha_1 v + \alpha_2 T_\pi(v) + ... + \alpha_{n-1} T_\pi^{n-1}(v), \text{ com } \alpha_i \in K,$$

é claramente um código linear cíclico (note que $T_\pi^n = Id$). Em particular é cíclico o código $\langle 0 \rangle$.

Lema 5 *Seja V um subespaço vetorial de R_n. Então, V é um ideal de R_n se, e somente se, V é fechado pela multiplicação por $\overline{x}$.*

Prova: *($\Longrightarrow$) Supondo inicialmente V ideal de R_n, por definição de ideal e já que $\overline{x} \in R_n$, temos $\overline{x} \cdot \overline{p(x)} \in V$, $\forall\ \overline{p(x)} \in V$.*

($\Longleftarrow$) Suponhamos agora que V seja fechado com relação à multiplicação por $\overline{x}$. Como V é subespaço vetorial, é fechado com relação à soma. Basta então mostrar que $\overline{g(x)} \cdot \overline{p(x)} \in V$, para todo $\overline{g(x)} \in R_n$ e todo $\overline{p(x)} \in V$. Como V é subespaço de R_n, é claro que $a \cdot \overline{p(x)} \in V$, $\forall\ a \in K$. Como por hipótese,

$$\overline{x \cdot p(x)} = \overline{x} \cdot \overline{p(x)} \in V,$$

então

$$\overline{x^2 \cdot p(x)} = \overline{x} \cdot \overline{x \cdot p(x)} \in V.$$

Indutivamente, obtemos, $\forall\ m \in \mathbb{N}$, que

$$\overline{x^m \cdot p(x)} = \overline{x^m} \cdot \overline{p(x)} \in V.$$

Agora, escrevendo $\overline{g(x)} = \overline{a_0 + a_1 x + ... + a_{n-1}x^{n-1}}$, temos que

$$\overline{g(x)} \cdot \overline{p(x)} = \overline{g(x)p(x)} = \overline{(a_0 + a_1 x + ... + a_{n-1}x^{n-1}) \cdot p(x)} = a_0\overline{p(x)} + a_1\overline{x} \cdot \overline{p(x)} + ... + a_{n-1}\overline{x^{n-1}} \cdot \overline{p(x)} \in V,$$

pois cada parcela da última expressão pertence à V.

■

Teorema 9 *Um subespaço C de K^n é um código cíclico se, somente se, $\nu(C)$ é um ideal de R_n.*

Prova: *Pelo Lema 5, temos que $\nu(C)$ é um ideal de R_n se, e somente se, este subespaço é fechado pela multiplicação por $\overline{x}$. Por outro lado, vimos que a multiplicação por $\overline{x}$ traduz-se por meio de ν^{-1} na ação T_π, descrita no início desta sessão. Portando, dizer que $\nu(C)$ é fechado com relação a multiplicação por $\overline{x}$ é o mesmo que dizer que C é cíclico.*

■

Portanto, pela proposição 6 temos que um código C em K^n é cíclico se, e somente se, $\nu(C) = I(\overline{g(x)})$, onde $g(x) \in K[x]$ é um divisor de $x^n - 1$.

No que se segue, $g(x)$ denotará sempre um divisor de $x^n - 1$, e poremos

$$h(x) = \frac{x^n - 1}{g(x)}.$$

O leitor pode perceber uma certa semelhança entre a base descrita no teorema abaixo e a base de $\langle v \rangle$ do exemplo 10, em suma, temos no teorema a seguir um código cíclico visto em R_n, equanto que no exemplo 10 temos um código cíclico visto em K^n. O corolário 3 mais adiante e sua prova elucidam bem essa relação.

Teorema 10 *Seja $I = I(\overline{g(x)})$, onde $g(x)$ é um divisor de $x^n - 1$ de grau s. Temos que $\overline{g(x)}, \overline{xg(x)}, \overline{x^2g(x)}, ..., \overline{x^{n-s-1}g(x)}$ é uma base de I como espaço vetorial sobre K.*

Prova: *Vamos verificar inicialmente que os elementos acima são linearmente independentes. De fato, suponhamos que*

$$a_0\overline{g(x)} + a_1\overline{xg(x)} + ... + a_{n-s-1}\overline{x^{n-s-1}g(x)} = \overline{0}.$$

Logo,

$$\overline{g(x)} \cdot (\overline{a_0 + a_1x + ... + a_{n-s-1}x^{n-s-1}}) = \overline{0}.$$

Como a classe do polinômio do lado esquerdo da igualdade igual à classe do zero, então para algum $d(x) \in K[x]$, temos que

$$g(x) \cdot (a_0 + a_1x + ... + a_{n-s-1}x^{n-s-1}) = d(x) \cdot (x^n - 1)$$

Daí, segue que

$$a_0 + a_1x + ... + a_{n-s-1}x^{n-s-1} = d(x) \cdot h(x).$$

Como o grau de $h(x)$ é $n - s$, a única possibilidade para a igualdade acima ser verdadeira é se $d(x) = 0$. Logo,

$$a_0 + a_1x + ... + a_{n-s-1}x^{n-s-1} = 0 \text{ e, consequentemente,}$$
$$a_0 = a_1 = ... = a_{n-s-1} = 0$$

Resta mostrar agora que os elementos sitados no teorema geram o espaço vetorial I sobre K. Seja $\overline{p(x)} \in I$, como I é um ideal gerado por $\overline{g(x)}$, temos que

$$\overline{p(x)} = \overline{d(x)} \cdot \overline{g(x)}, \text{ para algum } \overline{d(x)} \in R_n, \text{ logo,}$$
$$p(x) \equiv d(x) \cdot g(x) \ mod(x^m - 1).$$

Pelo algoritmo da divisão, temos que $d(x) = c(x) \cdot h(x) + r(x)$, com $r(x) = a_0 + a_1 x + ... + a_{n-s-1}x^{n-s-1}$, pois $r(x) = 0$ ou $gr(r(x)) < gr(h(x)) = n - s$. Logo,

$$p(x) \equiv d(x) \cdot g(x) \equiv c(x) \cdot h(x) \cdot g(x) + r(x) \cdot g(x) \equiv$$
$$c(x) \cdot (x^n - 1) + r(x) \cdot g(x) \equiv r(x) \cdot g(x) \ mod(x^n - 1)$$

e portanto,

$$p(x) \equiv (a_0 + a_1 x + ... + a_{n-s-1}x^{n-s-1}) \cdot g(x) \ mod(x^n - 1), \text{ donde}$$
$$\overline{p(x)} = a_0\overline{g(x)} + a_1\overline{xg(x)} + ... + a_{n-s-1}\overline{x^{n-s-1}g(x)}.$$

■

Corolário 3 *Dado um código cíclico C, existe $v \in C$ tal que $C = \langle v \rangle$.*

Prova: *Seja $I = \nu(C)$. Logo, I é gerado como espaço vetorial sobre K por $\overline{g(x)} + \overline{xg(x)} + ... + \overline{x^{n-s-1}g(x)}$, onde $g(x)$ é um divisor de $x^n - 1$ de grau s. Portanto, colocando $v = \nu^{-1}(\overline{g(x)})$, temos que C é gerado por $v, T_\pi(v), ..., T_\pi^{n-s-1}(v)$ e, portanto, $C = \langle v \rangle$.*

■

3.2 Matriz Geradora e Matriz Teste de Paridade

O seguinte corolário é decorrente do teorema 10 e diz respeito à matriz geradora de um código cíclico.

Corolário 4 *Seja $g(x) = g_0 + g_1 x + ... + g_s x^s$ um divisor de $x^n - 1$ de grau s. Se $I = I(\overline{g(x)})$, então a dimensão de I sobre K é $n - s$ e o código $C = \nu^{-1}(I)$ tem matriz geradora*

$$G = \begin{pmatrix} \nu^{-1}(\overline{g(x)}) \\ \nu^{-1}(\overline{xg(x)}) \\ \vdots \\ \nu^{-1}(\overline{x^{n-s-1}g(x)}) \end{pmatrix} = \begin{pmatrix} g_0 & g_1 & \cdots & g_s & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_s & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & g_0 & \cdots & \cdots & g_s \end{pmatrix}.$$

Perceba que da forma como definimos nosso $h(x)$, temos que ele é também um divisor de $x^n - 1$, com grau $n - s$. Além disso, o polinômio recíproco de $h(x) = h_0 + h_1 x + ... + h_{n-s}x^{n-s}$ (Definição 5),

$$h^*(x) = x^{n-s} \cdot h\left(\frac{1}{x}\right) = h_{n-s} + h_{n-s-1}x + ... + h_0 x^{n-s},$$

é também um divisor de $x^n - 1$ (Proposição 3, capítulo 1), e portanto, é o polinômio gerador de algum código cíclico que identificaremos adiante.

Teorema 11 *Seja $C = \nu^{-1}(I)$ um código cíclico, onde $I = I(\overline{g(x)})$, com $g(x)$ um divisor de $x^n - 1$ de grau s. Então $C^\perp$ é cíclico e $C^\perp = \nu^{-1}(J)$, onde $J = I(\overline{h^*(x)})$.*

Prova: *Sejam*

$$g(x) = g_0 + g_1 x + ... + g_s x^s \text{ e } h(x) = h_0 + h_1 x + ... + h_{n-s}x^{n-s}.$$

Note que $gr(h(x)) = n - s$, logo $h_{n-s} \neq 0$.
Sejam

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_s & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_s & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & g_0 & \cdots & \cdots & g_s \end{pmatrix}, \text{ matriz geradora do código } C = \nu^{-1}(I), \text{ com } I = I(\overline{g(x)}), \text{ e}$$

$$H = \begin{pmatrix} \nu^{-1}(\overline{h^*(x)}) \\ \nu^{-1}(\overline{xh^*(x)}) \\ \vdots \\ \nu^{-1}(\overline{x^{n-s-1}h^*(x)}) \end{pmatrix} = \begin{pmatrix} h_{n-s} & h_{n-s-1} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_{n-s} & h_{n-s-1} & \cdots & h_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & h_{n-s} & \cdots & \cdots & h_0 \end{pmatrix}.$$

Mostraremos aqui que da forma como foi definida a matriz H, ela é tal que $G \cdot H^t = 0$, e portanto, pelo Lema 2, segue que H é uma matriz geradora de $C^\perp$ e, além disso, pelo Teorema 11, $C^\perp = \nu^{-1}(J)$, onde $J = I(\overline{h^*(x)})$.

É fácil ver que as linhas de H são linearmente independentes.
Além disso, seja $\{e_1, e_2, ..., e_n\}$ a base canônica de K^n. A i-ésima linha de G é

$$G_i = g_0 e_i + g_1 e_{i+1} + ... + g_s e_{i+s},\ 1 \leq i \leq n - s,$$

e a j-ésima coluna de H^t, ou sejam a j-ésima linha de H, é

$$H_i = h_{n-s}e_j + h_{n-s-1}e_{j+1} + ... + h_0 e_{j+n-s},\ 1 \leq j \leq s.$$

Suponhamos que $i \leq j$. O produto interno de G_i por H_j é dado por

$$g_{j-i}h_{n-s} + g_{j-i+1}h_{n-s-1} + ... + g_{n-s}h_{j-i},$$

onde $j - i = 0, ..., s - 1$.

Mas a soma acima é precisamente igual ao coeficiente $x^{n-s+j-i}$ no produto $g(x) \cdot h(x) = x^n - 1$. Como $1 \leq n - s + j - i \leq n - 1$, temos que esse coeficiente é igual a zero. O caso $j \leq i$ é análogo. Logo, $G \cdot H^t = 0$.

■

Note que $C^{\perp}$ é cíclico, pois J é ideal de R_n.

Em suma, o que obtemos do teorema acima é que a matriz H que construímos é geradora de $C^{\perp}$ e portanto a **matriz teste de paridade** do código $C = \nu^{-1}(I)$, em que $I = I(\overline{g(x)})$.

3.3 Codificação e Decodificação de um Código Cíclico

Seja

$$\begin{array}{rcl} \mu : K^s & \to & K[x]_{s-1} \subset K[x] \\ (a_0, ..., a_{s-1}) & \mapsto & \sum_{i=0}^{s-1} a_i x^i \end{array}$$

o isomorfismo de K-espaços vetoriais, onde $K[x]_{s-1}$ é o espaço vetorial dos polinômios de grau menor ou igual a $s-1$. Esse isomorfismo será de grande utilidade no que se segue.

Teorema 12 *Seja $C \subset K^n$ um código cíclico. Suponhamos que $C = \nu^{-1}(I)$, onde $I = I(\overline{g(x)})$, com $g(x)$ um divisor de $x^n - 1$ de grau s. Seja R a matriz $(n-s) \times s$ cuja i-ésima linha é*

$$R_i = -\mu^{-1}(r_i(x)),\ 1 \leq i \leq n-s,$$

onde $r_i(x)$ é o resto da divisão de x^{s-1+i} por $g(x)$. Então, $(R|Id_{n-s})$ é uma matriz geradora de C.

Prova: *Sejam $q_i(x)$ e $r_i(x)$ o quociente e o resto da divisão de x^{s-1+i} por $g(x)$. Logo,*

$$x^{s-1+i} = g(x)q_i(x) + r_i(x),\ \textit{com}\ r_i(x) = 0\ \textit{ou}\ gr(r_i(x)) \leq s-1.$$

Portanto, $\overline{x^{s-1+i} - r_i(x)}$ pertence a I, e se pensarmos nos graus desses polinômios para $i = 1, ..., n-s$ veremos que são linearmente independentes sobre K e portanto uma base para $I(\overline{g(x)})$ como espaço vetorial. Como $\nu^{-1}(\overline{x^{s-1+i} - r_i(x)}) = e_{s-1+i} - \mu^{-1}(r_i(x))$, temos que a matriz

$$\begin{pmatrix} -\mu^{-1}(r_1(x)) & 1 & 0 & \cdots & 0 \\ -\mu^{-1}(r_2(x)) & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\mu^{-1}(r_{n-s}(x)) & 0 & 0 & \cdots & 1 \end{pmatrix}$$

é uma matriz geradora de C.

■

Vimos na sessão 2.2 deste trabalho que, dado um código fonte K^k, podemos fazer uma codificação afim de transformá-lo num código corretor de erros $C \subset K^n$, e que essa codificação consiste em multiplicar cada palavra em K^k pela matriz geradora de C. Com códigos cíclicos isso também funciona, entretanto é possível fazer essa codificação utilizando os polinômios da matriz geradora na forma reduzida como se segue.

Seja K^{n-s} um código fonte e C um código corretor cíclico com matriz na forma padrão $G = (R|Id_{n-s})$. Dado $(a_1, ..., a_{n-s}) \in K^{n-s}$, esse vetor pode ser codificado como elemento de C como se segue:

$$(a_1, ..., a_{n-s})(R|Id_{n-s}) = (b_0, ..., b_{s-1}, a_1, ..., a_{n-s}),$$

onde

$$\begin{gathered}(b_0, ..., b_{s-1}) = -a_1\mu^{-1}(r_1(x)) - \cdots - a_{n-s}\mu^{-1}(r_{n-s}(x)) = \\ -\mu^{-1}(a_1 r_1(x) + \cdots + a_{n-s}r_{n-s}(x)) = \\ -\mu^{-1}(\textstyle\sum_{i=1}^{n-s} a_i r_i(x)).\end{gathered}$$

Teorema 13 *Seja $C \subset K^n$ um código cíclico gerado por um polinômio $g(x)$ de grau s com matriz geradora na forma padrão $(R|Id_{n-s})$ e matriz teste de paridade $H = (Id_s| - R^t)$. Se $v = (v_0, ..., v_{n-1}) \in K^n$, então a síndrome de v com relação à matriz H é dada por*

$$\mu^{-1}(r(x)),$$

onde $r(x)$ é o resto da divisão de $v_0 + v_1x + \cdots + v_{n-1}x^{n-1}$ por $g(x)$.

Prova: *Vimos na sessão 2.4 deste trabalho que a síndrome de v é o vetor $Hv^t = (Id_s| - R^t)v^t$, além disso perceba que podemos escrever*

$$\begin{gathered}Id_s = (\mu^{-1}(1), \mu^{-1}(x), ..., \mu^{-1}(x^{s-1})), \text{ sendo assim,} \\ (Id_s| - R^t)v^t = \\ (\mu^{-1}(1), \mu^{-1}(x), ..., \mu^{-1}(x^{s-1}), \mu^{-1}(r_1(x)), ..., \mu^{-1}(r_{n-s}(x)))v^t = \\ \mu^{-1}(v_0 + v_1x + \cdots + v_{s-1}x^{s-1} + v_s r_1(x) + \cdots + v_{n-1}r_{n-s}(x)),\end{gathered}$$

entretanto,

$$r(x) = v_0 + v_1x + \cdots + v_{s-1}x^{s-1} + v_s r_1(x) + \cdots + v_{n-1}r_{n-s}(x)$$

é o resto da divisão de $v_0 + v_1x + \cdots + v_{n-1}x^{n-1}$ por $g(x)$. Daí segue o resultado.

■

No seguinte exemplo mostraremos na prática a codificação de um código fonte para um código cíclico e o cálculo da síndrome de uma palavra recebida.

Exemplo 12 *Considere o polinômio $x^7 - 1$ sobre $\mathbb{Z}_2$. A fatoração de $x^7 - 1$ é dada por*

$$x^7 - 1 = (1 + x)(1 + x + x^3)(1 + x^2 + x^3).$$

Fatorar polinômios sobre corpos finitos não é uma tarefa simples, seria inclusive um bom tema para uma outra dissertação de mestrado, a referência [4] mostra que já existem algoritmos que facilitam essa tarefa.

Vamos considerar o código $C \subset \mathbb{Z}_2^7$ gerado pelo polinômio $g(x) = x^3 + x^2 + 1$. Analisaremos aqui a codificação e a decodificação em C. É fácil ver que a dimensão de C é 4. Agora, pelo Teorema 12, determinaremos uma matriz geradora para C na forma padrão:

$$\begin{aligned} x^3 &= (x^3 + x^2 + 1) + (x^2 + 1) \\ x^4 &= (x^3 + x^2 + 1) \cdot (x + 1) + (x^2) \\ x^5 &= (x^3 + x^2 + 1) \cdot (x^2 + x + 1) + (x + 1) \\ x^6 &= (x^3 + x^2 + 1) \cdot (x^3 + x^2 + x) + (x^2 + x). \end{aligned}$$

Dessa forma, temos

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

Vimos que dado um vetor $(a_1, a_2, a_3, a_4) \in \mathbb{Z}_2^4$, do código fonte, então a codificação desse vetor é dada por $(b_0, b_1, b_2, a_1, a_2, a_3, a_4)$, onde b_0, b_1 e b_2 são os coeficientes do polinômio

$$\begin{gathered} a_1(x^2 + 1) + a_2(x^2) + a_3(x + 1) + a_4(x^2 + x) = \\ a_1 + a_3 + (a_3 + a_4)x + (a_1 + a_2 + a_4)x^2. \end{gathered}$$

Portanto, a codificação de (a_1, a_2, a_3, a_4) é

$$(a_1 + a_3, a_3 + a_4, a_1 + a_2 + a_4, a_1, a_2, a_3, a_4).$$

Por outro lado, temos a seguinte matriz teste de paridade para C

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

O vetor $(1, 1, 1, 1, 1, 1, 1) \in \mathbb{Z}_2^7$, pelo Teorema 13, possui síndrome relativa a H igual a $\mu^{-1}(r(x))$, onde $r(x)$ é o resto da divisão de $1 + x + x^2 + x^3 + x^4 + x^5 + x^6$ por $g(x) = x^3 + x^2 + 1$. Portanto, $r(x) = 0$ e, consequentemente, a síndrome é 0, donde podemos afirmar que este vetor pertence ao código C.

Vimos como é possível trabalhar com códigos cíclicos por meio de operações sobre polinômios, isso torna os algoritmos de codificação e decodificação mais rápidos e eficientes. Entretanto, a determinação da distância mínima em um

código cíclico é um assunto delicado e é, em parte, uma questão em aberto. Parte daí a necessidade de implementar esses códigos cíclicos, surgindo dessa forma os códigos BCH[1], onde podemos construir famílias de códigos cujas distâncias mínimas possuem cotas inferiores.

[1] B, H e C são as iniciais para Bose, Chaudhuri e Hocquenghem, invetores desse código em 1959-1960

Referências Bibliográficas

[1] HEFEZ, A.; VILLELA, M.L.T. Códigos Corretores de Erros. 2.ed. Rio de Janeiro: IMPA, 2008.

[2] Hungerford, T. W.: Abstract algebra: an introduction, 2nd ed, Saunders College Publ. (1997).

[3] ARAUJO, K. V.: Estruturas Algébricas II, São Cristóvão: UFS, 2009.

[4] http://hdl.handle.net/10183/117745. Último acesso: 17/04/2017.

[5] www.ufsj.edu.br/portal2-repositorio/File/i-ermac/anais/minicursos/mc8.pdf. Último acesso: 17/04/2017.

[6] HEFEZ, A.; VILLELA: Curso de Álgebra. Volume 1. Rio de Janeiro: IMPA, 2014.