UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA

Bases de Gröbner e Aplicações em Aproximações de Padé e Codificação

por

Juliane Golubinski Capaverde

Dissertação submetida como requisito parcial
para a obtenção do grau de

Mestre em Matemática Aplicada

Prof. Dr. Vilmar Trevisan
Orientador

Porto Alegre, julho de 2009.

CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Capaverde, Juliane Golubinski

Bases de Gröbner e Aplicações em Aproximações de Padé e Codificação / Juliane Golubinski Capaverde.—Porto Alegre: PPGMAp da UFRGS, 2009.

81 p.: il.

Dissertação (mestrado) —Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Matemática Aplicada, Porto Alegre, 2009.
Orientador: Trevisan, Vilmar

Dissertação: Matemática Aplicada
Bases de Gröbner, aproximações de Padé, códigos lineares

Bases de Gröbner e Aplicações em Aproximações de Padé e Codificação

por

Juliane Golubinski Capaverde

Dissertação submetida ao Programa de Pós-Graduação em Matemática Aplicada do Instituto de Matemática da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do grau de

Mestre em Matemática Aplicada

Linha de Pesquisa: Computação Científica

Orientador: Prof. Dr. Vilmar Trevisan

Banca examinadora:

Prof. Dr. Antônio Paques
PPGMAT/UFRGS-RS

Prof. Dr. José Afonso Barrionuevo
PPGMAp/UFGRS-RS

Profa. Dra. María Rosario Robbiano Bustamante
Universidad Católica del Norte - Chile

Profa. Dra. Renata Raposo Del-Vecchio
Departamento de Análise/UFF-RJ

Dissertação apresentada em
julho de 2009.

Prof. Dr. Waldir Leite Roque
Coordenador

AGRADECIMENTOS

Aos meus incansáveis pais, a quem devo boa parte do que sou, obrigada pelo amor, pela dedicação, e pelo esforço para que nada me faltasse. À minha irmã, agradeço pela amizade, companheirismo e imensurável apoio. Agradeço aos três pela compreensão nos momentos em que estive ausente devido aos estudos.

Agradeço ao meu orientador pelo incentivo constante desde a graduação, e pela confiança no meu potencial.

Aos queridos colegas da pós-graduação, obrigada pela ajuda nas disciplinas e pela companhia nas horas de descontração. Vocês fizeram com que essa caminhada fosse muito mais leve, e até divertida.

Agradeço ainda à banca examinadora pela atenção e pelas correções que ajudaram a aperfeiçoar este trabalho.

Conteúdo

LISTA DE SÍMBOLOS

$K[x_1, \ldots, x_n]$	Anel de polinômios em n variáveis com coeficientes em K
$\mathbf{M}$	Conjunto de todos os monômios em $K[x_1, \ldots, x_n]$
$\mathbf{T}$	Conjunto de todos os termos em $K[x_1, \ldots, x_n]$
$tl(f)$	Termo líder de f
$ml(f)$	Monômio líder de f
$cl(f)$	Coeficiente líder de f
$deg(f)$	grau do polinômio f
$\overline{f}^F$	Resto da divisão de f por F
$\langle S \rangle$	Ideal ou submódulo gerado pelo conjunto S
$\langle tl(S) \rangle$	Ideal ou submódulo gerado pelos termos líderes dos elementos de S
mmc	Mínimo múltiplo comum
$spol(f, g)$	S-polinômio de f e g
$K[x_1, \ldots, x_n]/I$	Anel quociente de $K[x_1, \ldots, x_n]$ pelo ideal I
$N_G(f)$	Forma normal de f módulo G
$\dim_K L$	Dimensão do K-espaço vetorial L
$\mathcal{B}(I)$	Base monomial de I
$V(S)$	Variedade definida pelo conjunto de polinômios S
$I(V)$	Ideal em $K[x_1, \ldots, x_n]$ definido por V
$\sqrt{I}$	Radical do ideal I
M/N	Módulo quociente de M por N
$\mathbb{F}_q$	Corpo finito com q elementos
$d(u, v)$	Distância de Hamming entre u e v
$\omega(u)$	Peso de Hamming de u
$d(C)$	Distância mínima do código C

RESUMO

Nesta dissertação estudamos algumas aplicações da teoria das bases de Gröbner, visando principalmente a utilização dessas técnicas na teoria de códigos. Apresentamos um algoritmo para obter a base de Gröbner reduzida do ideal de um conjunto finito de pontos, e descrevemos um método para encontrar aproximações de Padé de polinômios multivariados. Terminamos apresentando o procedimento desenvolvido por J. Farr e S. Gao para a construção e decodificação de códigos lineares via bases de Gröbner.

ABSTRACT

In this master thesis we study some applications of Gröbner bases theory, aiming using these techniques in coding theory. We present an algorithm for computing the reduced Gröbner basis of the vanishing ideal of a finite set of points, and describe a method for finding Padé approximations of multivariate polynomials. We finish presenting the procedure developed by J. Farr and S. Gao for construction and decoding of linear codes via Gröbner bases.

1 INTRODUÇÃO

O conceito de bases de Gröbner foi introduzido em 1965 por B. Buchberger [3], que desenvolveu também um algoritmo para obter tais bases. Inicialmente, a importância de seu trabalho não foi devidamente reconhecida; apenas nos anos 80 pesquisadores começaram uma investigação mais profunda da nova teoria. Desde então muitas generalizações e uma ampla variedade de aplicações foram desenvolvidas, e o crescente interesse nesta teoria deve-se ao fato de que ela fornece ferramentas computacionais aplicáveis a um grande número de problemas em matemática, ciência, engenharia e ciência da computação.

A teoria de códigos é uma das áreas em que as bases de Gröbner são empregadas, na construção e decodificação de códigos. O objetivo da teoria de códigos é transmitir informações eficientemente e de forma segura através de um canal "ruidoso", ou seja, um canal que pode alterar parte da informação transmitida. Exemplos de tais situações são a comunicação via satélite e o armazenamento de dados em CDs. Para isso, adiciona-se uma certa quantidade de informação à mensagem que será transmitida, a fim de tornar possível a correção dos erros que possivelmente ocorrerão na transmissão. Assim, busca-se um código capaz de corrigir uma grande quantidade de erros adicionando apenas uma pequena quantidade de informação à mensagem original.

Nosso objetivo principal é estudar o método desenvolvido por J. Farr e S. Gao para construção e decodificação de uma classe de códigos lineares, que tem como casos particulares vários códigos conhecidos e largamente utilizados na prática. A construção destes códigos é baseada em um conjunto finito de pontos, e é feita através de uma base de Gröbner do ideal relacionado a este conjunto. Faz-se necessário, desta forma, um algoritmo eficiente para obter uma tal base. A base de Gröbner

deste ideal é também importante no processo de decodificação, que tem como passo fundamental o cálculo da aproximação de Padé de um certo polinômio.

A primeira parte da dissertação, que compreende o segundo e o terceiro capítulos, constitui uma introdução à teoria das bases de Gröbner. No segundo capítulo, definimos as ordens monomiais e apresentamos o algoritmo da divisão para polinômios multivariados. Isto é feito numa tentativa de generalizar a teoria conhecida para polinômios em uma variável. A seguir, no terceiro capítulo, introduzimos o conceito de base de Gröbner e descrevemos o algoritmo de Buchberger, através do qual pode-se obter uma base de Gröbner a partir de um conjunto finito qualquer de geradores do ideal.

Na segunda parte da dissertação são apresentadas algumas aplicações das bases de Gröbner, baseadas no trabalho de J. Farr e S. Gao [7, 8, 9]. Primeiramente, apresentamos um algoritmo que calcula a base de Gröbner do ideal de um conjunto finito de pontos. Visamos especialmente a aplicação deste algoritmo na teoria de códigos, no entanto ele é útil também em outros ramos da matemática, como interpolação e estatística, nos quais algoritmos eficientes, que permitem cálculos com um grande número de pontos, são necessários.

No capítulo subsequente, apresentamos uma técnica para obter aproximações de Padé utilizando o algoritmo do capítulo anterior e a teoria de bases de Gröbner de submódulos de módulos livres sobre o anel de polinômios. Para isso, generelizamos a teoria desenvolvida nos primeiros capítulos para o caso de submódulos. Novamente, estudamos as aproximações de Padé com o intuito de empregá-las na decodificação de códigos, mas essas aproximações constituem ferramentas importantes e que são empregadas também em diversas áreas da matemática e outras ciências.

No último capítulo, primeiramente é feita uma breve introdução à teoria de códigos, e então apresentamos o método proposto em [7] para a construção e de-

codificação de uma classe de códigos lineares. Para isso, empregamos a teoria e os algoritmos desenvolvidos nos capítulos anteriores.

2 DIVISÃO DE POLINÔMIOS MULTIVARIADOS

No anel de polinômios em uma variável $K[x]$, para decidir se um polinômio f pertence ou não ao ideal gerado por um conjunto de polinômios $f_1, f_2, \ldots, f_m$, primeiro calculamos o máximo divisor comum de $f_1, \ldots, f_m$ através do Algoritmo de Euclides. Então o polinômio f pertence ao ideal gerado por $f_1, \ldots, f_m$ se, e somente se, o resto da divisão de f por $mdc(f_1, \ldots, f_m)$ é zero.

A teoria das bases de Gröbner pode ser vista como uma generalização deste procedimento para polinômios a várias variáveis. Dado um conjunto finito de polinômios multivariados sobre um corpo, existe um algoritmo, conhecido como Algoritmo de Buchberger, que encontra um novo conjunto de polinômios, chamado de base de Gröbner, que gera o mesmo ideal que o conjunto original. A base de Gröbner tem a seguinte propriedade: um dado polinômio pertence ao ideal gerado pela base de Gröbner se, e somente se, a forma normal do polinômio com respeito à base de Gröbner é igual a zero. Esta forma normal a que nos referimos é calculada utilizando um procedimento análogo ao algoritmo da divisão do caso univariado, com a diferença que agora dividiremos um polinômio por um conjunto finito de polinômios. Assim, a base de Gröbner tem o mesmo papel do mdc do caso univariado.

Nosso primeiro passo em direção à generalização mencionada acima buscará estender o algoritmo da divisão para polinômios a várias variáveis. Nossas principais referências neste capítulo serão os livros de W. Adams e P. Loustaunau [1], T. Becker e V. Weispfenning [2], e D. Cox, J. Little e D. O'Shea [5].

2.1 Ordens Monomiais

Lembremos como funciona o algoritmo da divisão no caso de uma variável. Dado um polinômio não-nulo $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$, com $a_i \in K$ e $a_n \neq 0$, o grau de f, denotado por $deg(f)$, é o maior expoente de x que aparece em f, o termo líder de f, denotado por $tl(f)$, é o termo de maior grau de f, e o coeficiente líder de f, denotado por $cl(f)$ é o coeficiente do termo líder de f. Assim, temos $deg(f) = n$, $tl(f) = a_n x^n$ e $cl(f) = a_n$. Dados dois polinômios f e g em $K[x]$, no primeiro passo do algoritmo da divisão calcula-se $h = f - \frac{tl(f)}{tl(g)}g$. A ideia é subtrair de f um múltiplo apropriado de g de forma que o termo líder de f seja cancelado. Esse múltiplo é $\frac{tl(f)}{tl(g)}g$. Depois repetimos esse processo para o polinômio h obtido, até que o grau do polinômio seja menor do que o grau de g.

Para generalizar esse procedimento para o caso de mais variáveis, precisaremos estabelecer uma ordenação para os monômios em $K[x_1, \ldots, x_n]$, com a finalidade de decidir quem é o termo líder de um polinômio multivariado.

Seguiremos aqui a convenção de que um *monômio* em $K[x_1, \ldots, x_n]$ é um produto de potências da forma $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, com $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{N}$, e um *termo* é um monômio com um coeficiente, ou seja, um termo t tem a forma $t = cx_1^{\alpha_1} \cdots x_n^{\alpha_n}$, com $c \in K$. Para abreviar a notação, escreveremos o monômio $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ na forma

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} = x^{\alpha},$$

onde $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. O conjunto de todos os monômios e o conjunto de todos os termos serão denotados por $\mathbf{M}$ e $\mathbf{T}$, respectivamente, ou seja,

$$\mathbf{M} = \{x^{\alpha} : \ \alpha \in \mathbb{N}^n\}$$

$$\mathbf{T} = \{cx^{\alpha} : \ c \in K, \alpha \in \mathbb{N}^n\}$$

Definição 2.1.1. Dizemos que uma relação de ordem $\succ$ em $\mathbf{M}$ é uma *ordem monomial* se satisfaz as seguintes condições:

(i) $\succ$ é uma ordem total, isto é, dados monômios $x^\alpha \neq x^\beta$, ou $x^\alpha \succ x^\beta$ ou $x^\beta \succ x^\alpha$;

(ii) Se $x^\alpha \succ x^\beta$, então $x^\alpha x^\gamma \succ x^\beta x^\gamma$, para todo $x^\gamma \in \mathbf{M}$;

(iii) $\succ$ é uma boa ordem, isto é, todo subconjunto não vazio de $\mathbf{M}$ admite um menor elemento.

A condição (i) garante que não há ambigüidade na escolha do menor monômio de uma coleção finita e, portanto, os monômios que aparecem em um polinômio f podem ser listados de forma única em ordem crescente ou decrescente. A condição (ii), por sua vez, assegura que a ordenação não é alterada quando multiplicamos f por um monômio x^γ. O resultado abaixo ([12, pág. 56]), que apresentaremos sem demonstração, nos diz que podemos substituir a condição (iii) por outra e obter uma definição equivalente de ordem monomial.

Proposição 2.1.1. *Seja $\succ$ uma ordem em $\mathbf{M}$ que satisfaz as condições (i) e (ii) da Definição 2.1.1. Então $\succ$ é uma boa ordem se e somente se $x_i \succ 1$ para $i = 1, \ldots, n$.*

Corolário 2.1.1. *Se $\succ$ é uma ordem monomial e x^α divide x^β, então $x^\alpha \preceq x^\beta$.*

Demonstração. Se $x^\alpha | x^\beta$, então existe x^γ tal que $x^\beta = x^\alpha x^\gamma$. Como $\succ$ é ordem monomial, temos $x_i \succ 1$ para $i = 1, \ldots, n$, o que implica $x^\gamma \succeq 1$. Logo, $x^\beta = x^\alpha x^\gamma \succeq x^\alpha$. □

Eis alguns exemplos de ordens monomiais.

Exemplo 2.1.1. *Ordem lexicográfica.* Dados $x^\alpha, x^\beta \in \mathbf{M}$, definimos

$$x^\alpha \succ_{lex} x^\beta \iff \exists j \in \{1, \ldots, n\} \text{ tal que } \alpha_j > \beta_j \text{ e } \alpha_i = \beta_i, \ \forall\, i < j.$$

Ou seja: o primeiro expoente em x^α que não coincidir com o correspondente em x^β tem que ser maior.

Exemplo 2.1.2. *Ordem lexicográfica graduada.* Dados $x^\alpha, x^\beta \in \mathbf{M}$, definimos

$$x^\alpha \succ_{lg} x^\beta \iff \begin{cases} \sum_{i=1}^n \alpha_i > \sum_{i=1}^n \beta_i \text{ ou} \\ \sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i \text{ e } x^\alpha \succ_{lex} x^\beta. \end{cases}$$

Exemplo 2.1.3. *Ordem lexicográfica reversa graduada.* Dados $x^\alpha, x^\beta \in \mathbf{M}$, definimos

$$x^\alpha \succ_{lrg} x^\beta \iff \begin{cases} \sum_{i=1}^n \alpha_i > \sum_{i=1}^n \beta_i \text{ ou} \\ \sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i \text{ e } \exists j \in \{1, \ldots, n\} \text{ tal que} \\ \alpha_j < \beta_j \text{ e } \alpha_i = \beta_i, \ \forall\ i > j. \end{cases}$$

Proposição 2.1.2. *As ordens lexicográfica, lexicográfica graduada e lexicográfica reversa graduada são ordens monomiais.*

Demonstração. Faremos a verificação apenas para a ordem lexicográfica, uma vez que para as demais a prova é análoga. Temos que mostrar que $\succ_{lex}$ satisfaz às três condições da Definição 2.1.1.

Sejam $x^\alpha,\ x^\beta \in \mathbf{M}$ tais que $x^\alpha \neq x^\beta$. Então x^α e x^β diferem em ao menos um expoente. Seja j o menor tal que $\alpha_j \neq \beta_j$. Então $\alpha_k = \beta_k$ para cada $k < j$, e $\alpha_j > \beta_j$ ou $\beta_j > \alpha_j$. Assim, temos $x^\alpha \succ_{lex} x^\beta$ ou $x^\beta \succ_{lex} x^\alpha$.

Sejam agora $x^\alpha,\ x^\beta,\ x^\gamma \in \mathbf{M}$ tais que $x^\alpha \succ_{lex} x^\beta$. Então existe j tal que $\alpha_j > \beta_j$ e $\alpha_k = \beta_k$ para cada $k < j$. Segue que $\alpha_j + \gamma_j > \beta_j + \gamma_j$ e $\alpha_k + \gamma_k = \beta_k + \gamma_k$ para cada $k < j$. Logo,

$$x^\alpha x^\gamma = x^{\alpha+\gamma} \succ_{lex} x^{\beta+\gamma} = x^\beta x^\gamma$$

Resta mostrar agora que $\succ_{lex}$ é uma boa ordem. Pela Proposição 2.1.1, basta mostrar que $x_i \succ_{lex} 1$ para toda variável x_i. Como $x_i = x_1^0 \cdots x_i^1 \cdots x_n^0$ e $1 = x_1^0 \cdots x_n^0$, segue da definição da ordem lexicográfica que $x_i \succ_{lex} 1$. □

O resultado a seguir nos fornece uma caracterização de uma boa ordem, e será empregado futuramente para mostrar que o Algoritmo da Divisão envolve somente um número finito de passos.

Proposição 2.1.3. *Seja $\succ$ uma ordem em $\mathbf{M}$. Então $\succ$ é uma boa ordem se e somente se toda sequência decrescente $m_1 \succeq m_2 \succeq \cdots$ em $\mathbf{M}$ estabiliza, isto é, existe N tal que $m_k = m_N$ para todo $k \geq N$.*

Demonstração. Suponhamos que $\succ$ não é uma boa ordem. Então existe um subconjunto não-vazio $S \subset \mathbf{M}$ que não admite um menor elemento. Seja $m_1 \in S$. Como S não admite menor elemento, existe $m_2 \in S$ tal que $m_1 \succ m_2$. Repetindo o mesmo raciocínio para m_2, e assim sucessivamente, obtemos uma sequência estritamente decrescente infinita $m_1 \succ m_2 \succ m_3 \succ \cdots$.

Reciprocamente, suponhamos que $\succ$ é uma boa ordem. Seja $m_1 \succeq m_2 \succeq \cdots$ uma sequência decrescente em $\mathbf{M}$. Tomando $S = \{m_1, m_2, \ldots\} \subset \mathbf{M}$, como $\succ$ é uma boa ordem, temos que S admite um menor elemento, digamos m_N. Então $m_k = m_N$ para todo $k \geq N$. □

Definição 2.1.2. Dada uma ordem monomial $\succ$ em $\mathbf{M}$ e um polinômio $f \in K[x_1, \ldots, x_n]$, $f \neq 0$, podemos escrever f na forma

$$f = \sum_{i=1}^{k} c_i x^{\alpha_i}$$

onde $c_i \in K$ são constantes e $x^{\alpha_i} \in \mathbf{M}$ são monômios para $1 \leq i \leq k$, com $x^{\alpha_1} \succ x^{\alpha_2} \succ \ldots \succ x^{\alpha_k}$. Para f escrito desta forma, definimos:

(i) o *termo líder* de f, denotado por $tl(f)$, por $tl(f) = c_1 x^{\alpha_1}$;

(ii) o *monômio líder* de f, denotado por $ml(f)$, por $ml(f) = x^{\alpha_1}$;

(iii) o *coeficiente líder* de f, denotado por $cl(f)$, por $cl(f) = c_1$;

(iv) o *grau* de f, denotado por $deg(f)$, por $deg(f) = \alpha_1$.

Estenderemos a relação $\succ$ ao conjunto dos termos $\mathbf{T}$, definindo, para termos $ax^\alpha,\ bx^\beta \in \mathbf{T}$,

$$ax^\alpha \succ bx^\beta \iff x^\alpha \succ x^\beta.$$

Note que o termo, o monômio e o coeficiente líderes, bem como o grau de um polinômio, podem variar de acordo com a ordem monomial utilizada. Vejamos um exemplo.

Exemplo 2.1.4. Consideremos os seguintes termos em $\mathbb{Q}[x_1, x_2, x_3]$: $4x_1x_2x_3^2$, $4x_1^3$, $-5x_2^4$, e $7x_1x_2^2x_3$. Vejamos como estes termos são ordenados de acordo com cada uma das ordens monomiais dos exemplos acima.

(a) Ordem lexicográfica: $4x_1^3 \succ_{lex} 7x_1x_2^2x_3 \succ_{lex} 4x_1x_2x_3^2 \succ_{lex} -5x_2^4$

(b) Ordem lexicográfica graduada: $7x_1x_2^2x_3 \succ_{lg} 4x_1x_2x_3^2 \succ_{lg} -5x_2^4 \succ_{lg} 4x_1^3$

(c) Ordem lexicográfica reversa graduada: $-5x_2^4 \succ_{lrg} 7x_1x_2^2x_3 \succ_{lrg} 4x_1x_2x_3^2 \succ_{lrg} 4x_1^3$

Desta forma, o polinômio $f = 4x_1x_2x_3^2+4x_1^3-5x_2^4+7x_1x_2^2x_3 \in \mathbb{Q}[x_1, x_2, x_3]$ tem um termo líder diferente com respeito a cada uma das ordens monomiais, conforme podemos ver na tabela abaixo.

	$tl(f)$	$ml(f)$	$cl(f)$	$deg(f)$
$\succ_{lex}$	$4x_1^3$	x_1^3	4	$(3,0,0)$
$\succ_{lg}$	$7x_1x_2^2x_3$	$x_1x_2^2x_3$	7	$(1,2,1)$
$\succ_{lrg}$	$-5x_2^4$	x_2^4	-5	$(0,4,0)$

Vejamos agora algumas propriedades dos monômios líderes.

Proposição 2.1.4. *Sejam $f, g \in K[x_1, \ldots, x_n]$ polinômios não nulos e $\prec$ uma ordem monomial. Então*

(i) $ml(f \cdot g) = ml(f) \cdot ml(g)$;

(ii) $ml(f + g) \preceq \max\{ml(f), ml(g)\}$.

Demonstração. Sejam $f = \sum_{i=1}^{r} a_i x^{\alpha_i}$ e $g = \sum_{j=1}^{s} b_j x^{\beta_j}$, onde $a_i, b_j \in K$ e $x^{\alpha_1} \succeq x^{\alpha_2} \succeq \cdots \succeq x^{\alpha_r}$ e $x^{\beta_1} \succeq x^{\beta_2} \succeq \cdots \succeq x^{\beta_s}$. Segue que $ml(f) = x^{\alpha_1}$ e $ml(g) = x^{\beta_1}$.

O produto $f \cdot g$ é dado por

$$f \cdot g = \left(\sum_{i=1}^{r} a_i x^{\alpha_i} \right) \left(\sum_{j=1}^{s} b_j x^{\beta_j} \right) = \sum_{i=1}^{r} \sum_{j=1}^{s} a_i b_j x^{\alpha_i} x^{\beta_j}.$$

Como $x^{\alpha_1} \succ x^{\alpha_i}$ para $i = 2, \ldots, r$, temos $x^{\alpha_1} x^{\beta_j} \succ x^{\alpha_i} x^{\beta_j}$ para $i = 2, \ldots, r$ e $j = 1, \ldots, s$. Analogamente, $x^{\alpha_i} x^{\beta_1} \succ x^{\alpha_i} x^{\beta_j}$ para $j = 2, \ldots, s$ e $i = 1, \ldots, r$. Assim, temos que

$$x^{\alpha_1} x^{\beta_1} \succ x^{\alpha_1} x^{\beta_j} \succ x^{\alpha_i} x^{\beta_j}$$

para $i = 2, \ldots, r$ e $j = 2, \ldots, s$, e portanto, $ml(f \cdot g) = x^{\alpha_1} x^{\beta_1} = ml(f) ml(g)$.

Para a soma $f + g$ temos duas possibilidades: ou $tl(f) = -tl(g)$ ou $tl(f) \neq -tl(g)$. Se $tl(f) = -tl(g)$, então os termos líderes de f e g se cancelam na soma, restando apenas monômios menores, logo

$$ml(f + g) \prec \max\{ml(f), ml(g)\} = \max\{x^{\alpha_1}, x^{\beta_1}\} = x^{\alpha_1} = x^{\beta_1}.$$

Se $tl(f) \neq -tl(g)$, então não há cancelamento dos termos líderes e $ml(f + g) = \max\{ml(f), ml(g)\}$. □

Da Proposição 2.1.4 decorrem outras propriedades, como as que seguem:

(a) $tl(f \cdot g) = tl(f) \cdot tl(g)$

(b) $cl(f \cdot g) = cl(f) \cdot cl(g)$

(c) $deg(f \cdot g) = deg(f) + deg(g)$

(d) $tl(f + g) \preceq \max\{tl(f), tl(g)\}$

Veremos agora uma forma de definir ordens monomiais através de matrizes. Seja M uma matriz $m \times n$ com entradas reais, cujos vetores-linha são $w_1, \ldots, w_m$. Definimos $\succ$ da seguinte forma:

$$x^{\alpha} \succ x^{\beta} \iff \exists j \in \{1, \ldots, m\} \text{ tal que } \alpha \cdot w_j > \beta \cdot w_j \text{ e } \alpha \cdot w_i = \beta \cdot w_i, \forall\ i < j,$$

onde $\gamma \cdot v$ denota o produto escalar de γ por v, ou seja, se $\gamma = (\gamma_1, \ldots, \gamma_n)$ e $v = (v_1, \ldots, v_n)$ então $\gamma \cdot v = \gamma_1 v_1 + \cdots + \gamma_n v_n$.

Proposição 2.1.5. *A matriz M define uma ordem monomial se e somente se as seguintes condições são satisfeitas:*

(i) Não existe $\alpha \in \mathbb{Z}^n \backslash \{0\}$ tal que $w_i \cdot \alpha = 0$ para todo $i = 1, \ldots, m$;

(ii) A primeira entrada não-nula de cada coluna de M é positiva.

Demonstração. Seja $\succ$ a ordem definida por M. Suponhamos que M satisfaz (i) e (ii). Vejamos que $\succ$ é uma ordem total. Sejam $x^\alpha, x^\beta \in \mathbf{M}$, $x^\alpha \neq x^\beta$. Então $\alpha \neq \beta$, ou ainda, $\alpha - \beta \neq 0$. Pela condição (i), os produtos $w_i \cdot (\alpha - \beta)$ não são todos nulos. Seja $j \in \{1, \ldots, m\}$ o menor índice tal que $w_j \cdot (\alpha - \beta) \neq 0$. Segue que $w_j \cdot (\alpha - \beta) > 0$ ou $w_j \cdot (\alpha - \beta) < 0$. Logo $x^\alpha \succ x^\beta$ ou $x^\beta \succ x^\alpha$.

Sejam agora $x^\alpha, x^\beta, x^\gamma \in \mathbf{M}$ tais que $x^\alpha \succ x^\beta$. Então existe j tal que $w_j \cdot \alpha > w_j \cdot \beta$ e $w_i \cdot \alpha = w_i \cdot \beta$ para todo $i < j$. Segue que

$$w_j \cdot (\alpha + \gamma) = w_j \cdot \alpha + w_j \cdot \gamma > w_j \cdot \beta + w_j \cdot \gamma = w_j \cdot (\beta + \gamma)$$

e

$$w_i \cdot (\alpha + \gamma) = w_i \cdot \alpha + w_i \cdot \gamma = w_i \cdot \beta + w_i \cdot \gamma = w_i \cdot (\beta + \gamma)$$

para todo $i < j$. Logo, $x^\alpha x^\gamma = x^{\alpha+\gamma} \succ x^{\beta+\gamma} = x^\beta x^\gamma$.

Resta mostrar que $\succ$ é uma boa ordem. Pela Proposição 2.1.1 basta mostrar que $x_i \succ 1$ para $i = 1, \ldots, n$. Temos que $x_i = x^{e_i}$, onde e_i é o vetor que tem 1 na i-ésima coordenada e 0 nas demais. O produto $e_i \cdot w_k$ é igual à i-esima coordenada do vetor w_k, e a i-ésima coluna da matriz M é formada pelas i-ésimas coordenadas dos vetores $w_1, \ldots, w_m$. Suponhamos que a primeira entrada não-nula da coluna i está na linha j. Então temos que $e_i \cdot w_j > 0$ e $e_i \cdot w_k = 0$ para $k < j$. Logo, $x_i = x^{e_i} \succ x^0 = 1$.

Reciprocamente, suponhamos que $\succ$ é uma ordem monomial. Suponhamos ainda que existe $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n \backslash \{0\}$ tal que $\alpha \cdot w_i = 0$ para $i = 1, \ldots, m$. Definimos $\beta = (\beta_1, \ldots, \beta_n)$ e $\gamma = (\gamma_1, \ldots, \gamma_n)$, onde

$$\beta_i = \begin{cases} \alpha_i, & \text{se } \alpha_i > 0 \\ 0, & \text{caso contrário} \end{cases}$$

$$\gamma_i = \begin{cases} -\alpha_i, & \text{se } \alpha_i < 0 \\ 0, & \text{caso contrário} \end{cases}$$

Assim, $\beta, \gamma \in \mathbb{N}^n$ e $\alpha = \beta - \gamma$. Ademais,

$$0 = \alpha \cdot w_i = (\beta - \gamma) \cdot w_i = \beta \cdot w_i - \gamma \cdot w_i$$

ou seja, $\beta \cdot w_i = \gamma \cdot w_i$ para $i = 1, \ldots, m$. Isso implica que os monômios x^β, x^γ são tais que $x^\beta \neq x^\gamma$, mas não temos $x^\beta \succ x^\gamma$ nem $x^\gamma \succ x^\beta$, o que contradiz o fato de que $\succ$ é ordem total.

Agora, temos que $x_i \succ 1$ para $i = 1, \ldots, n$. Logo, existe j tal que $e_i \cdot w_j > 0$ e $e_i \cdot w_k = 0$ para $k < j$. Como $e_i \cdot w_j$ é a entrada da coluna i e linha j da matriz M, temos que a primeira entrada não-nula da coluna i é a da linha j, que é positiva. □

Exemplo 2.1.5. Vamos comparar os monômios xyz^2 e x^3z^2 usando a ordem monomial definida pela matriz

$$M = \begin{pmatrix} 2 & 4 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Como

$$(2,4,3) \cdot (1,1,2) = 12$$

$$(2,4,3) \cdot (3,0,2) = 12$$

$$(1,1,1) \cdot (1,1,2) = 4$$

$$(1,1,1) \cdot (3,0,2) = 5$$

temos que $xyz^2 \prec x^3z^2$.

Exemplo 2.1.6. A ordem monomial definida pela matriz identidade de ordem n

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{2.1}$$

é a ordem lexicográfica do Exemplo 2.1.1. Na verdade, de acordo com um resultado de L. Robbiano [16], toda ordem monomial em $K[x_1, \ldots, x_n]$ é definida por alguma matriz.

2.2 Algoritmo da Divisão em $K[x_1, \ldots, x_n]$

Dados polinômios $f, f_1, \ldots, f_m$ em $K[x_1, \ldots, x_n]$, em analogia ao caso em uma variável, queremos encontrar $q_1, \ldots, q_m, r \in K[x_1, \ldots, x_n]$ tais que

$$f = q_1 f_1 + \cdots + q_m f_m + r \tag{2.2}$$

Buscamos uma generalização do algoritmo da divisão do caso univariado que nos permita encontrar uma tal representação. A principal diferença é que agora queremos dividir um polinômio por um conjunto de polinômios. Contudo, isto não será um obstáculo. Usaremos a mesma idéia do caso univariado, que é cancelar os termos do dividendo usando os termos líderes dos divisores, de forma que os termos introduzidos sejam sempre menores do que os cancelados, e repetir este processo até que ele não possa mais ser realizado. Aqui aparece outra diferença do caso de uma variável: não nos limitaremos ao cancelamento apenas do termo líder do dividendo.

Definição 2.2.1. Dados polinômios $f, f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$, dizemos que f é *reduzido módulo* $f_1, \ldots, f_m$ se f é combinação linear de monômios que não são divisíveis por nenhum dos termos líderes de $f_1, \ldots, f_m$. Mais geralmente, dizemos que f é reduzido módulo $G \subseteq K[x_1, \ldots, x_n]$ se nenhum dos termos de f é divisível por $tl(g)$, para todo $g \in G$.

Algoritmo 2.2.1 Algoritmo da Divisão

Entrada: $f, f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$
Saída: $q_1, \ldots, q_m, r \in K[x_1, \ldots, x_n]$
$q_1 \leftarrow 0;\ q_2 \leftarrow 0;\ \ldots;\ q_m \leftarrow 0;\ r \leftarrow 0$
$p \leftarrow f$
while $p \neq 0$ **do**
  $i \leftarrow 1$
  $d \leftarrow$ false
  while $i \leq m$ **and** $d =$ false **do**
    if $tl(f_i)$ divide $tl(p)$ **then**
      $q_i \leftarrow q_i + \dfrac{tl(p)}{tl(f_i)}$
      $p \leftarrow p - \dfrac{tl(p)}{tl(f_i)} f_i$
      $d \leftarrow$ true
    else
      $i \leftarrow i + 1$
    end if
  end while
  if $d =$ false **then**
    $r \leftarrow r + tl(p)$
    $p \leftarrow p - tl(p)$
  end if
end while

Proposição 2.2.1. *Sejam $f, f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$. Fixe uma ordem monomial $\succ$ em $\mathbf{M}$. Então existem polinômios $q_1, \ldots, q_m, r \in K[x_1, \ldots, x_n]$ tais que*

$$f = q_1 f_1 + \cdots + q_m f_m + r$$

com $tl(q_i f_i) \preceq tl(f)$, para todo $i = 1, \ldots, m$, e r reduzido módulo $f_1, \ldots, f_m$.

Demonstração. Provaremos a existência de $q_1, \ldots, q_m$ e r satisfazendo as condições do enunciado mostrando que o Algoritmo da Divisão (Algoritmo 2.2.1) funciona. Para isso, mostraremos primeiro que a igualdade

$$f = q_1 f_1 + \cdots + q_m f_m + p + r, \tag{2.3}$$

com $tl(q_i f_i) \preceq tl(f)$, para $i = 1, \ldots, m$, e $tl(p) \preceq tl(f)$, vale em cada iteração do algoritmo.

A igualdade (2.3) obviamente é válida para os valores iniciais $q_1 = q_2 = \cdots = q_m = 0$, $p = f$ e $r = 0$.

Suponhamos agora que (2.3) vale em uma determinada iteração do algoritmo. Temos duas possibilidades para a próxima iteração:

Caso 1: Algum $tl(f_i)$ divide $tl(p)$: Neste caso, atualizamos o valor de q_i e de p, e as demais variáveis não serão modificadas. Para ver que a igualdade (2.3) continua valendo, basta mostrar então que o valor de $q_i f_i + p$ permanece o mesmo após as atualizações. Denotaremos por q_i^* e p^* os novos valores de q_i e p. Assim,

$$q_i^* = q_i + \frac{tl(p)}{tl(f_i)}$$

$$p^* = p - \frac{tl(p)}{tl(f_i)} f_i$$

e portanto

$$\begin{aligned} q_i^* f_i + p^* &= \left(q_i + \frac{tl(p)}{tl(f_i)}\right) f_i + p - \frac{tl(p)}{tl(f_i)} f_i \\ &= q_i f_i + \frac{tl(p)}{tl(f_i)} f_i + p - \frac{tl(p)}{tl(f_i)} f_i \\ &= q_i f_i + p \end{aligned}$$

Analisemos agora os termos líderes de $q_i^* f_i$ e p^*. Temos que

$$tl(q_i^* f_i) \preceq max\left(tl(q_i f_i), tl\left(\frac{tl(p)}{tl(f_i)} f_i\right)\right)$$

e como $tl(q_i f_i) \preceq tl(f)$ e $tl\left(\frac{tl(p)}{tl(f_i)} f_i\right) = tl(p) \preceq tl(f)$, segue que $tl(q_i^* f_i) \preceq tl(f)$. Como $tl\left(\frac{tl(p)}{tl(f_i)} f_i\right) = tl(p)$ e $p^* = p - \frac{tl(p)}{tl(f_i)} f_i$, temos que $tl(p^*) \prec tl(p) \preceq tl(f)$. Logo, a igualdade (2.3) permanece válida, e as desigualdades entre os termos líderes são preservadas.

Caso 2: Nenhum $tl(f_i)$ divide $tl(p)$: Neste caso, atualizaremos os valores de r e p, e as demais variáveis não serão alteradas. Como no caso anterior, basta mostrar que a soma $r + p$ não é alterada. Denotando por r^* e p^* os novos valores de r e p, respectivamente, temos

$$r^* = r + tl(p)$$

$$p^* = p - tl(p)$$

logo

$$\begin{aligned} r^* + p^* &= r + tl(p) + p - tl(p) \\ &= r + p \end{aligned}$$

Além disso, como $p^* = p - tl(p)$, temos que $tl(p^*) \prec tl(p) \preceq tl(f)$. Portanto, a igualdade (2.3) e as desigualdades entre os termos líderes também são preservadas neste caso.

O algoritmo termina quando $p = 0$ e, neste caso, por (2.3) temos

$$f = q_1 f_1 + \cdots + q_m f_m r,$$

com $tl(q_i f_i) \preceq tl(f)$. Como só adicionamos termos a r quando tais termos não são divisíveis por nenhum dos $tl(f_i)$, segue que r é reduzido módulo $f_1, \ldots, f_m$.

Resta mostrar que o algoritmo de fato termina. A cada iteração o valor de p é atualizado, e nos dois casos vistos acima temos $tl(p^*) \prec tl(p)$. Portanto, se o algoritmo não terminasse obteríamos uma sequência estritamente decrescente infinita em $\mathbf{T}$, o que é uma contradição. □

Definição 2.2.2. Sejam $f, r \in K[x_1, \ldots, x_n]$ e $F = \{f_1, \ldots, f_m\} \subset K[x_1, \ldots, x_n]$ tais que $f = q_1 f_1 + \ldots + q_m f_m + r$, com $q_1, \ldots, q_m \in K[x_1, \ldots, x_n]$ e r reduzido módulo F. Dizemos que r é um *resto* da divisão de f por F, escrevendo $r = \overline{f}^F$.

Exemplo 2.2.1. Consideremos os polinômios $f = x^2 + xy + y^3$, $f_1 = x + y^2$ e $f_2 = xy + y \in \mathbb{Q}[x, y]$. Vamos dividir f por $F = \{f_1, f_2\}$, utilizando a ordem lexicográfica.

$x^2 + xy + y^3$	$x + y^2$	$xy + y$	resto
$-(x^2 + xy^2)$	x		
$-xy^2 + xy + y^3$	$-y^2$		
$-(-xy^2 - y^4)$			
$xy + y^4 + y^3$	y		
$-(xy + y^3)$			
y^4			y^4

Aplicando o Algoritmo da Divisão, obtemos $f = (x - y^2 + y) \cdot f_1 + 0 \cdot f_2 + y^4$. Agora dividiremos f por F trocando a ordem dos polinômios f_1 e f_2.

$x^2 + xy + y^3$	$xy + y$	$x + y^2$	resto
$-(x^2 + xy^2)$		x	
$-xy^2 + xy + y^3$	$-y$		
$-(-xy^2 - y^2)$			
$xy + y^3 + y^2$	1		
$-(xy + y)$			
$y^3 + y^2 - y$			$y^3 + y^2 - y$

Trocando a ordem dos polinômios no algoritmo, obtemos $f = x \cdot f_1 + (-y + 1) \cdot f_2 + y^3 + y^2 - y$. Podemos ver neste exemplo que os "quocientes"e o resto não são únicos. Vejamos o que acontece se trocamos a ordem lexicográfica pela ordem lexicográfica graduada.

y^3+x^2+xy	y^2+x	$xy+y$	resto
$-(y^3+xy)$	y		
x^2			x^2

Neste caso, obtemos $f = y \cdot f_1 + 0 \cdot f_2 + x^2$.

Exemplo 2.2.2. Consideremos $f = xy + x$, $f_1 = xy^2 + 1$, $f_2 = x^2y - x \in \mathbb{Q}[x, y]$. Dividindo f por $F = \{f_1, f_2\}$, empregando a ordem lexicográfica, obtemos $f = 0 \cdot f_1 + 0 \cdot f_2 + xy + x$. Por outro lado, é fácil ver que $f = x \cdot f_1 - y \cdot f_2 \in \langle F \rangle$.

3 BASES DE GRÖBNER

Neste capítulo faremos uma introdução ao conceito de base de Gröbner, bem como descreveremos um algoritmo para obter uma tal base a partir de um conjunto qualquer de geradores de um ideal, conhecido como Algoritmo de Buchberger. Temos ainda como referência os livros [1, 2, 5].

3.1 Ideais Monomiais e Bases de Gröbner

Definição 3.1.1. Um ideal I de $K[x_1, \ldots, x_n]$ é dito um *ideal monomial* se I é gerado por algum conjunto de monômios, isto é, se existir $S \subseteq \mathbf{M}$, com $1 \notin S$, tal que $I = \langle S \rangle$.

Podemos caracterizar os monômios pertencentes a um ideal monomial da seguinte forma:

Proposição 3.1.1. *Sejam $S \subseteq \mathbf{M}$ e $I = \langle S \rangle$. Então $x^\beta \in I$ se e somente se existe $x^\alpha \in S$ tal que x^β é divisível por x^α.*

Demonstração. Se $x^\alpha | x^\beta$ para algum $x^\alpha \in S$, então obviamente $x^\beta \in I$.

Reciprocamente, suponhamos que $x^\beta \in I$. Então existem polinômios $f_1, \ldots, f_k \in K[x_1, \ldots, x_n]$ e monômios $x^{\alpha_1}, \ldots, x^{\alpha_k} \in S$ tais que

$$x^\beta = \sum_{i=1}^{k} f_i x^{\alpha_i}$$

Agora, cada f_i é combinação linear de monômios. Assim, podemos escrever a igualdade acima na forma

$$x^\beta = \sum_{j=1}^{\ell} c_j x^{\beta_j} \tag{3.1}$$

onde $c_1, \ldots, c_\ell \in K$ são constantes e cada monômio x^{β_j} é múltiplo de algum $x^{\alpha_i} \in S$. Mas a igualdade polinomial (3.1) nos dá que $\ell = 1$, $c_1 = 1$ e $\beta_1 = \beta$. Logo x^β é múltiplo de algum $x^{\alpha_i} \in S$. □

Definição 3.1.2. Dado um subconjunto $S \subseteq K[x_1, \ldots, x_n]$, denotamos por $tl(S)$ o conjunto dos termos líderes dos elementos de S, isto é,

$$tl(S) = \{tl(f) : f \in S\}.$$

O *ideal dos termos líderes de* S, denotado por $\langle tl(S)\rangle$, é o ideal monomial gerado pelos termos líderes dos elementos de S, ou seja,

$$\langle tl(S)\rangle = \langle \{tl(f) : f \in S\}\rangle.$$

Se S é um subconjunto de $K[x_1, \ldots, x_n]$ e I é o ideal gerado por S, qual a relação entre os ideais monomiais $\langle tl(S)\rangle$ e $\langle tl(I)\rangle$? Vejamos alguns exemplos.

Exemplo 3.1.1. Consideremos $I = \langle f\rangle \subseteq K[x_1, \ldots, x_n]$. Como $tl(fg) = tl(f)tl(g)$, temos que $\langle tl(I)\rangle = \langle tl(f)\rangle$.

Exemplo 3.1.2. Consideremos o ideal $I = \langle x^2 - y, x - y\rangle \subset \mathbb{Q}[x, y]$, e fixemos a ordem lexicográfica. Então

$$\langle tl(x^2 - y), tl(x - y)\rangle = \langle x^2, x\rangle = \langle x\rangle.$$

No entanto,

$$y^2 - y = (x^2 - y) - (x + y)(x - y) \in I$$

e

$$tl(y^2 - y) = y^2 \notin \langle x\rangle.$$

O exemplo 3.1.2 mostra que, em geral, se $I = \langle S\rangle$, $\langle tl(I)\rangle \neq \langle tl(S)\rangle$. Claramente vale sempre a inclusão $\langle tl(S)\rangle \subseteq \langle tl(I)\rangle$.

Definição 3.1.3. Dado um ideal I de $K[x_1, \ldots, x_n]$, dizemos que um subconjunto finito $G \subseteq I$ é uma *base de Gröbner* de I se $\langle tl(G)\rangle = \langle tl(I)\rangle$. Dizemos simplesmente que G é uma base de Gröbner se G é uma base de Gröbner do ideal gerado por G.

Exemplo 3.1.3. Consideremos $f_1 = y - z^2$, $f_2 = x - z^3 \in \mathbb{Q}[x, y, z]$. Sejam $F = \{f_1, f_2\}$ e $I = \langle F\rangle$. Escolhendo a ordem lexicográfica, com $x \succ y \succ z$, temos $tl(f_1) = y$ e $tl(f_2) = x$. Suponhamos que existe $f \in \langle F\rangle$ tal que $tl(f) \notin \langle tl(F)\rangle = \langle x, y\rangle$. Então $tl(f) = z^m$, para algum $m \in \mathbb{N}$, o que implica que $f \in \mathbb{Q}[z]$. Por outro lado, como $f \in \langle F\rangle$, existem $h_1, h_2 \in \mathbb{Q}[x, y, z]$ tais que

$$f = h_1 \cdot (y - z^2) + h_2 \cdot (x - z^3).$$

Como x não aparece em f, podemos fazer $x = z^3$, o que nos dá

$$f(z) = h_1(z^3, y, z) \cdot (y - z^2).$$

Mas isso implica que $y - z^2$ divide f, contradizendo o fato de que a única variável que aparece em f é z. Concluímos que F é uma base de Gröbner com respeito à ordem lexicográfica.

No entanto, se utilizamos a ordem lexicográfica graduada, temos $tl(f_1) = -z^2$ e $tl(f_2) = -z^3$, e $\langle tl(F)\rangle = \langle z^2\rangle$. Tomando $f = z \cdot f_1 - f_2 = yz - x \in \langle F\rangle$, temos que $tl(f) = yz \notin \langle tl(F)\rangle$. Logo, F não é base de Gröbner com respeito à ordem lexicográfica graduada.

A seguir veremos algumas propriedades das bases de Gröbner, que nos permitem começar a vislumbrar a importância destas bases.

Proposição 3.1.2. *Sejam $I \subseteq K[x_1, \ldots, x_n]$ um ideal e $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner de I, e seja $f \in K[x_1, \ldots, x_n]$. Então existe um único $r \in K[x_1, \ldots, x_n]$ reduzido módulo G tal que $f = h + r$, para algum $h \in I$.*

Demonstração. Pelo Algoritmo da Divisão, temos que

$$f = q_1 g_1 + \ldots + q_\ell g_\ell + r$$

onde r é reduzido módulo G e $h = q_1 g_1 + \ldots + q_\ell g_\ell \in I$. Assim, está provada a existência de r.

Para mostrar a unicidade, suponhamos que

$$f = h_1 + r_1 = h_2 + r_2$$

onde $h_1, h_2 \in I$ e r_1, r_2 são reduzidos módulo G. Então temos

$$\begin{aligned} r_1 - r_2 &= (f - h_1) - (f - h_2) \\ &= f - h_1 - f + h_2 \\ &= h_2 - h_1 \in I \end{aligned}$$

Se $r_1 - r_2 \neq 0$, então $tl(r_1 - r_2) \in \langle tl(I) \rangle = \langle tl(g_1), \ldots, tl(g_\ell) \rangle$. Pela Proposição 3.1.1, segue que $tl(g_i) | tl(r_1 - r_2)$ para algum $i \in \{1, \ldots, \ell\}$. Mas isso é impossível, pois nenhum dos termos de r_1 e r_2 é divisível pelos termos líderes de $g_1, \ldots, g_\ell$, visto que r_1, r_2 são reduzidos módulo G. Logo $r_1 - r_2 = 0$ e, portanto, $r_1 = r_2$. □

A Proposição 3.1.2 implica que o resto da divisão de f por G é único se G é uma base de Gröbner, não importando a ordem em que o elementos de G são listados no Algoritmo da Divisão.

Proposição 3.1.3. *Seja I um ideal de $K[x_1, \ldots, x_n]$. Se G é uma base de Gröbner de I, então $f \in I$ se e somente se $\overline{f}^G = 0$.*

Demonstração. Seja f um elemento não nulo de $K[x_1, \ldots, x_n]$, e seja $r = \overline{f}^G$. Se $r = 0$, como $G \subseteq I$, temos que $f \in I$.

Suponhamos agora que $f \in I$, e que $r \neq 0$. Podemos escrever

$$f = \sum_{g \in G} h_g g + r.$$

Como $f \in I$ e cada $g \in I$, segue que $r \in I$. Logo $tl(r) \in \langle tl(I) \rangle = \langle tl(G) \rangle$. Portanto, $tl(r)$ é divisível pelo termo líder de algum elemento de G, o que é uma contradição, pois r é reduzido módulo G. Logo $r = 0$. □

Esta propriedade das bases de Gröbner resolve o problema de pertinência a um ideal: para decidir se um polinômio f pertence ou não ao ideal I, basta dividir f por G.

Nosso objetivo agora é mostrar que todo ideal possui uma base de Gröbner, e que uma base de Gröbner é de fato uma base do ideal. Para isso, provaremos primeiro que todo ideal de $K[x_1, \ldots, x_n]$ admite um conjunto finito de geradores. Começaremos pelo caso monomial.

Proposição 3.1.4. *Seja $I \subseteq K[x_1, \ldots, x_n]$ um ideal monomial. Então I admite uma base finita, isto é, existe um subconjunto finito de monômios $F \subseteq I$ tal que $I = \langle F \rangle$. Além disso, tal subconjunto finito pode ser extraído de qualquer conjunto de monômios que gere I.*

Demonstração. Procederemos por indução no número de variáveis n.

Se I é um ideal monomial em $K[x]$, então existe $S \subseteq \mathbb{N}$ tal que $I = \langle \{x^m | m \in S\} \rangle$. Seja m_0 o menor elemento de S; segue que $I = \langle x^{m_0} \rangle$.

Seja agora $I \subseteq K[x_1, \ldots, x_n, y] = K[x_1, \ldots, x_n][y]$ um ideal monomial. Podemos supor $I \neq \langle 0 \rangle$. Então existe $f_1 = f_1^*(x_1, \ldots, x_n)y^{d_1} \in I \backslash \langle 0 \rangle$, onde $f_1^* \in K[x_1, \ldots, x_n]$ denota um monômio e $d_1 = \min\{\text{grau de } f \text{ em } y | f \in I \backslash \langle 0 \rangle\}$.

Se $I = \langle f_1 \rangle$ acabou. Se não, existe $f_2 = f_2^*(x_1, \ldots, x_n)y^{d_2} \in I \backslash \langle f_1 \rangle$, onde f_2^* é um monômio e $d_2 = \min\{\text{grau de } f \text{ em } y | f \in I \backslash \langle f_1 \rangle\}$.

Note que $d_2 \geq d_1$, pois $f_2 \in I \backslash \langle 0 \rangle$, logo o grau de f_2 em y, d_2, não pode ser menor do que o grau de f_1 em y, d_1, que é mínimo.

Se não existisse um conjunto finito de monômios que gera I, poderíamos repetir o procedimento acima e obter uma seqüência infinita $f_1, f_2, \ldots$, onde cada f_m é da forma $f_m = f_m^*(x_1, \ldots, x_n)y^{d_m} \in I \backslash \langle f_1, \ldots, f_{m-1} \rangle$, com $f_m^* \in K[x_1, \ldots, x_n]$ um monômio e d_m mínimo.

Seja $I^* \subseteq K[x_1, \ldots, x_n]$ o ideal gerado pelos monômios $f_1^*, f_2^*, \ldots$. Pela hipótese de indução, existe $N \in \mathbb{N}$ tal que $I^* = \langle f_1^*, \ldots, f_N^* \rangle$. Em particular, $f_{N+1}^* \in I^*$, e portanto temos que f_{N+1}^* é divisível por f_i^* para algum $1 \leq i \leq N$. Então podemos escrever $f_{N+1}^* = g f_i^*$, com $g \in K[x_1, \ldots, x_n]$. Logo

$$f_{N+1} = f_{N+1}^* y^{d_{N+1}} = g f_i^* y^{d_{N+1}}$$

e como $d_{N+1} \geq d_i$, podemos escrever

$$\begin{aligned} f_{N+1} &= g f_i^* y^{d_{N+1}-d_i} y^{d_i} \\ &= g y^{d_{N+1}-d_i} (f_i^* y^{d_i}) \\ &= g y^{d_{N+1}-d_i} f_i \end{aligned}$$

o que implica $f_{N+1} \in \langle f_i \rangle \subseteq \langle f_1, \ldots, f_N \rangle$, o que é uma contradição, pois f_{N+1} foi escolhido de forma que $f_{N+1} \in I \backslash \langle f_1, \ldots, f_N \rangle$.

Portanto, temos que todo ideal monomial de $K[x_1, \ldots, x_n]$ admite uma base finita.

Sejam agora $I \subseteq K[x_1, \ldots, x_n]$ um ideal monomial e $M \subseteq I$ um conjunto de monômios tal que $I = \langle M \rangle$. Suponhamos que $F = \{f_1, \ldots, f_N\}$ é um conjunto finito de geradores de I. Descartando alguns elementos se necessário, podemos supor que não ocorrem relações de divisibilidade em F, isto é, se $f_i = g f_j$, com $g \in K[x_1, \ldots, x_n]$, então $f_i = f_j$. Vamos mostrar que $F \subseteq M$.

Seja $f_i \in F$. Como $\langle F \rangle = \langle M \rangle$, existem $g \in M$ e $h \in K[x_1, \ldots, x_n]$ tais que $f_i = gh$.

Por outro lado, $g \in \langle F \rangle$, logo existem $f_j \in F$ e $q \in K[x_1, \ldots, x_n]$ tais que $g = f_j q$. Portanto temos

$$f_i = gh = f_j(qh)$$

ou seja, f_j divide f_i, o que implica $f_i = f_j$. Segue que

$$g = f_j q = f_i q = g(hq)$$

e como f_i e g são monômios, temos $q = h = 1$ e $f_i = g \in M$. □

Uma conseqüência deste resultado é a existência das bases de Gröbner.

Corolário 3.1.1. *Seja $I \subseteq K[x_1, \ldots, x_n]$ um ideal. Então I admite uma base de Gröbner.*

Demonstração. Consideremos o ideal monomial gerado pelos termos líderes de elementos de I, $\langle tl(I)\rangle$. Pela Proposição 3.1.4, existe um subconjunto finito $F \subseteq tl(I) = \{tl(f)|f \in I\}$ tal que $\langle tl(I)\rangle = \langle F\rangle$. Seja $G = \{f \in I|tl(f) \in F\}$. Segue que $F = tl(G)$, e portanto $\langle tl(G)\rangle = \langle F\rangle = \langle tl(I)\rangle$, ou seja, G é uma base de Gröbner de I. □

Podemos agora demonstrar o famoso Teorema da Base de Hilbert. A prova deste teorema apresentada aqui mostra que uma base de Gröbner é, de fato, uma base do ideal.

Teorema 3.1.1. *Todo ideal de $K[x_1, \ldots, x_n]$ é finitamente gerado.*

Demonstração. Seja $I \subseteq K[x_1, \ldots, x_n]$ um ideal. Pela Proposição 3.1.4 existe $F = \{f_1, \ldots, f_N\} \subseteq I$ tal que $\langle tl(F)\rangle = \langle tl(I)\rangle$. Mostraremos que $I = \langle f_1, \ldots, f_N\rangle$.

Suponhamos, por absurdo, que $I \neq \langle f_1, \ldots, f_N\rangle$. Então $I\backslash\langle f_1, \ldots, f_N\rangle \neq \emptyset$. Seja $f \in I\backslash\langle f_1, \ldots, f_N\rangle$ de termo líder mínimo, e seja $f^* = tl(f)$. Como $f^* \in \langle tl(F)\rangle$, existe $1 \leq i \leq N$ tal que $tl(f_i)$ divide f^*, ou seja, $f^* = tl(f_i)g$ para algum monômio g. Seja $h = f - f_i g$. Então $h \in I$ e $h \notin \langle f_1, \ldots, f_N\rangle$, pois $h \in \langle f_1, \ldots f_N\rangle$ implicaria $f \in \langle f_1, \ldots, f_N\rangle$. Agora,

$$tl(f) = f^* = tl(f_i)g = tl(f_i g)$$

logo $h = 0$ ou $tl(h) \prec tl(f)$, o que contradiz a minimalidade do termo líder de f. Portanto $h = 0$ e $f = f_i g \in \langle f_1, \ldots, f_N\rangle$. □

Um anel A para o qual vale a afirmação do Teorema 3.1.1, ou seja, todo ideal de A é finitamente gerado, é dito um anel *Noetheriano*. Assim, $K[x_1, \ldots, x_n]$ é um anel Noetheriano.

Corolário 3.1.2. *Seja $\{I_m\}_{m\geq 1}$ uma cadeia ascendente de ideais em $K[x_1, \ldots, x_n]$, isto é,*

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots$$

Então existe N tal que $I_n = I_N$ para todo $n \geq N$.

Demonstração. Seja $I = I_1 \cup I_2 \cup I_3 \cup \ldots$. É facil ver que I é um ideal. Pelo Teorema da Base de Hilbert, I tem uma base finita, digamos $I = \langle f_1, \ldots, f_r \rangle$. Assim, cada $f_i \in I = I_1 \cup I_2 \cup \ldots$, ou seja, $f_i \in I_{m_i}$ para algum $m_i \in \mathbb{N}$. Tomando $N = \max\{m_1, \ldots, m_r\}$, temos que $f_i \in I_N$ para todo i. Logo $I \subseteq I_N \subseteq I$, e portanto $I = I_N = I_n$ para todo $n \geq N$. □

Claramente, um ideal não tem uma única base de Gröbner. Além disso, algumas bases podem conter polinômios desnecessários. Em geral, se G é uma base de Gröbner e $g \in G$ é tal que $tl(g) \in \langle tl(G\backslash\{g\})\rangle$, então $G\backslash\{g\}$ é ainda uma base de Gröbner do ideal $\langle G\rangle$. De fato, se $tl(g) \in \langle tl(G\backslash\{g\})\rangle$, então $\langle tl(G\backslash\{g\})\rangle = \langle tl(G)\rangle$. Isso motiva a seguinte definição.

Definição 3.1.4. Seja G uma base de Gröbner. Dizemos que G é *mínima* se $cl(g) = 1$ e $tl(g) \notin \langle tl(G\backslash\{g\})\rangle$, para todo $g \in G$.

Dada uma base de Gröbner $G = \{g_1, \ldots, g_\ell\}$ de um ideal I, para obter uma base de Gröbner mínima de I basta eliminar todos os polinômios $g_i \in G$ para os quais existe $j \neq i$ tal que $tl(g_j)|tl(g_i)$, e dividir os polinômios g_i que restarem por $cl(g_i)$ para torná-los mônicos. Portanto todo ideal possui uma base de Gröbner mínima.

Proposição 3.1.5. *Se G e H são bases de Gröbner mínimas de um ideal $I \subseteq K[x_1, \ldots, x_n]$ com respeito à mesma ordem monomial, então $tl(G) = tl(H)$.*

Demonstração. Sejam G e H bases de Gröbner mínimas do ideal I. Seja $h \in H$. Como $h \in I$ e G é uma base de Gröbner de I, existe $g \in G$ tal que $tl(g)|tl(h)$. Como $g \in I$ e H é uma base de Gröbner de I, existe $h' \in H$ tal que $tl(h')|tl(g)$. Logo $tl(h')|tl(h)$, o que implica que $h' = h$, pois H é uma base de Gröbner mínima. Segue que $tl(g) = tl(h)$.

Portanto, para cada elemento $g \in G$ existe um único elemento $h \in H$ tal que $tl(g) = tl(h)$. Analogamente, para cada $h \in H$ existe um único $g \in G$ tal que $tl(h) = tl(g)$. Logo, $tl(G) = tl(H)$. □

A Proposição 3.1.5 implica que duas bases de Gröbner mínimas de um ideal com respeito à mesma ordem monomial têm o mesmo número de elementos. Contudo, as bases de Gröbner mínimas não são únicas; um ideal pode conter muitos polinômios com o mesmo termo líder. Para obter unicidade, é necessário impor uma condição mais forte sobre os polinômios da base.

Definição 3.1.5. Seja G uma base de Gröbner. Dizemos que G é *reduzida* se $cl(g) = 1$ e g é reduzido módulo $G\backslash\{g\}$, para todo $g \in G$.

Proposição 3.1.6. *Seja $I \neq \langle 0\rangle$ um ideal de $K[x_1, \ldots, x_n]$. Então, para uma dada ordem monomial, I tem uma única base de Gröbner reduzida.*

Demonstração. Seja $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner mínima de I. Definimos

$$\begin{aligned}
h_1 &= \overline{g_1}^{H_1}, && \text{onde } H_1 = \{g_2, \ldots, g_\ell\}\\
h_2 &= \overline{g_2}^{H_2}, && \text{onde } H_2 = \{h_1, g_3, \ldots, g_\ell\}\\
h_3 &= \overline{g_3}^{H_3}, && \text{onde } H_3 = \{h_1, h_2, g_4, \ldots, g_\ell\}\\
&\vdots\\
h_\ell &= \overline{g_\ell}^{H_\ell}, && \text{onde } H_\ell = \{h_1, h_2, \ldots, h_{\ell-1}\}
\end{aligned}$$

Mostraremos que o conjunto $H = \{h_1, \ldots, h_\ell\}$ obtido é a base de Gröbner reduzida que procuramos.

Como G é uma base de Gröbner mínima, temos que $tl(g_1)$ não é divisível por nenhum dos termos líderes de $g_2, \ldots, g_\ell$. Logo, quando dividimos g_1 por $g_2, \ldots, g_\ell$, o termo líder de g_1 será um termo do resto. Portanto, $tl(h_1) = tl(g_1)$. Assim, o conjunto $\{h_1, g_2, \ldots, g_\ell\}$ obtido trocando g_1 por h_1 é também uma base de Gröbner mínima de I.

Analogamente, $tl(g_2)$ não é divisível por nenhum dos termos líderes de h_1, g_3, $\ldots$, g_ℓ, logo $tl(h_2) = tl(g_2)$, e $\{h_1, h_2, g_3, \ldots, g_\ell\}$ é base de Gröbner mínima de I.

Repetindo este raciocínio sucessivamente, obtemos que $\{h_1, \ldots, h_\ell\}$ é base de Gröbner mínima de I. Além disso, como h_i é o resto da divisão de g_i por $H_i = \{h_1, \ldots, h_{i-1}, g_{i+1}, \ldots, g_\ell\}$, temos que h_i é reduzido módulo H_i, e portanto é também reduzido módulo $H \backslash \{h_i\}$, pois $tl(h_j) = tl(g_j)$ para todo $j \in \{1, \ldots, \ell\}$. Assim, concluímos que H é uma base de Gröbner reduzida.

Provaremos agora a unicidade. Suponhamos que $G = \{g_1, \ldots, g_\ell\}$ e $H = \{h_1, \ldots, h_\ell\}$ são bases de Gröbner reduzidas de I. Como bases reduzidas são também mínimas, a Proposição 3.1.5 nos garante que G e H têm o mesmo número de elementos, e podemos assumir ainda, reenumerando os polinômios se necessário, que $tl(h_i) = tl(g_i)$, para cada $i \in \{1, \ldots, \ell\}$.

Seja $i \in \{1, \ldots, \ell\}$, e suponhamos que $g_i \neq h_i$. Como $g_i - h_i \in I$, existe j tal que $tl(h_j) | tl(g_i - h_i)$, e como $tl(g_i - h_i) \prec tl(h_i)$, temos que $j \neq i$. Mas isso implica que $tl(h_j) = tl(g_j)$ divide um termo de g_i ou h_i, o que contradiz o fato de que G e H são bases de Gröbner reduzidas. Logo, $g_i = h_i$. □

3.2 Algoritmo de Buchberger

Uma vez que está garantida a existência das bases de Gröbner, passamos agora ao cálculo efetivo de tais bases. Nesta seção, introduziremos o conceito de

S-polinômios e apresentaremos uma caracterização das bases de Gröbner que nos leverará a um algoritmo para obtê-las partindo de uma base qualquer de um ideal.

Definição 3.2.1. Dados dois polinômios não nulos $f, g \in K[x_1, \ldots, x_n]$, o *S-polinômio* de f e g é definido por

$$spol(f,g) = mmc\,(ml(f), ml(g)) \left(\frac{f}{tl(f)} - \frac{g}{tl(g)} \right)$$

Provaremos que uma base de Gröbner para um ideal gerado por um conjunto finito de polinômios é construída pela adjunção de S-polinômios.

Lema 3.2.1. *Sejam $p_1, \ldots, p_N \in K[x_1, \ldots, x_n]$ polinômios com monômio líder x^δ. Se existem constantes $c_1, \ldots, c_N$ tais que*

$$ml\left(\sum_{i=1}^{N} c_i p_i \right) \prec x^\delta,$$

então existem constantes d_{jk} tais que

$$\sum_{i=1}^{N} c_i p_i = \sum_{j=1}^{N} \sum_{k=1}^{N} d_{jk}\, spol(p_j, p_k). \tag{3.2}$$

Além disso, temos

$$ml(spol(p_j, p_k)) \prec x^\delta, \ \forall j, k.$$

Demonstração. Seja $l_i = cl(p_i)$. Como todos os polinômios p_i têm o mesmo monômio líder x^δ e $ml\left(\sum_{i=1}^{N} c_i p_i\right) \prec x^\delta$, segue que

$$\sum_{i=1}^{N} c_i l_i = 0.$$

Como $ml(p_j) = ml(p_k) = x^\delta$, temos que $mmc(ml(p_j), ml(p_k)) = x^\delta$, logo

$$\begin{aligned} spol(p_j, p_k) &= x^\delta \left(\frac{p_j}{l_j x^\delta} - \frac{p_k}{l_k x^\delta} \right) \\ &= p'_j - p'_k \end{aligned}$$

onde $p'_j = \frac{p_j}{l_j}$ e $p'_k = \frac{p_k}{l_k}$.

Podemos escrever:

$$\begin{aligned}
\sum_{i=1}^{N} c_i p_i &= c_1 l_1 p'_1 + c_2 l_2 p'_2 + \cdots + c_N l_N p'_N \\
&= c_1 l_1 p'_1 + c_1 l_1 p'_2 - c_1 l_1 p'_2 + \cdots + c_1 l_1 p'_N - c_1 l_1 p'_N + \\
&\quad + c_2 l_2 p'_2 + c_2 l_2 p'_3 - c_2 l_2 p'_3 + \cdots + c_2 l_2 p'_N - c_2 l_2 p'_N + \\
&\quad + \cdots + \\
&\quad + c_{N-1} l_{N-1} p'_{N-1} + c_{N-1} l_{N-1} p'_N - c_{N-1} l_{N-1} p'_N + \\
&\quad + c_N l_N p'_N \\
&= c_1 l_1 (p'_1 - p'_2) + (c_1 l_1 + c_2 l_2)(p'_2 - p'_3) + \cdots + \\
&\quad + \left(\sum_{i=1}^{N-1} c_i l_i \right) (p'_{N-1} - p'_N) + \left(\sum_{i=1}^{N} c_i l_i \right) p'_N \\
&= \sum_{i=1}^{N-1} c'_i spol(p_i, p_{i+1}),
\end{aligned}$$

onde $c'_i = \sum_{k=1}^{i} c_k l_k$. Fazendo

$$d_{jk} = \begin{cases} c'_j, & \text{se } k = j+1 \\ 0, & \text{caso contrário} \end{cases}$$

obtemos a equação (3.2).

Agora, como $tl(p'_j) = tl(p'_k) = x^\delta$, temos que os termos líderes se cancelam na diferença $p'_j - p'_k$, e restam apenas os termos menores. Portanto $ml(spol(p_j, p_k)) = ml(p'_j - p'_k) \prec x^\delta$. □

Lema 3.2.2. *Sejam $p, q \in K[x_1, \ldots, x_n]$ polinômios e x^α, x^β monômios tais que*

$$ml(x^\alpha p) = ml(x^\beta q).$$

Então existe $x^\gamma \in \mathbf{M}$ tal que

$$spol(x^\alpha p, x^\beta q) = x^\gamma spol(p, q).$$

Demonstração. Sejam $x^\rho = ml(p)$, $x^\sigma = ml(q)$ e $x^\delta = ml(x^\alpha p) = ml(x^\beta q)$. Seja ainda $\eta = (\eta_1, \ldots, \eta_n)$, onde $\eta_i = \max\{\rho_i, \sigma_i\}$. Assim,

$$x^\eta = mmc(tl(p), tl(q)).$$

Temos ainda que

$$\begin{aligned} spol(x^\alpha p, x^\beta q) &= x^\delta \left(\frac{x^\alpha p}{cl(p)x^\delta} - \frac{x^\beta q}{cl(q)x^\delta} \right) \\ &= x^\delta \left(\frac{p}{cl(p)x^{\delta-\alpha}} - \frac{q}{cl(q)x^{\delta-\beta}} \right) \\ &= x^\delta \left(\frac{p}{cl(p)x^\rho} - \frac{q}{cl(q)x^\sigma} \right) \\ &= x^\delta \left(\frac{p}{tl(p)} - \frac{q}{tl(q)} \right) \end{aligned}$$

Mas $\eta_i = \max\{\rho_i, \sigma_i\} \leq \delta_i$, logo $\delta_i - \eta_i \geq 0$. Tomando $\gamma = \delta - \eta$ temos

$$\begin{aligned} spol(x^\alpha p, x^\beta q) &= x^\gamma \cdot x^\eta \left(\frac{p}{tl(p)} - \frac{q}{tl(q)} \right) \\ &= x^\gamma spol(p, q). \end{aligned}$$

□

Teorema 3.2.1. *Seja $I \subseteq K[x_1, \ldots, x_n]$ um ideal. Então uma base finita G de I é uma base de Gröbner se e somente se para todo par $p, q \in G$ temos $\overline{spol(p,q)}^G = 0$.*

Demonstração. Suponhamos que G é uma base de Gröbner. Sejam $p, q \in G \subseteq I$. Então $spol(p, q) \in I$, e como G é uma base de Gröbner de I, segue que $\overline{spol(p,q)}^G = 0$.

Reciprocamente, suponhamos que $G = \{g_1, \ldots, g_N\}$ é tal que para todo par $g_i, g_j \in G$, $\overline{spol(g_i, g_j)}^G = 0$. Para provar que G é uma base de Gröbner, temos que mostrar que $\langle tl(G) \rangle = \langle tl(I) \rangle$. Sabemos que $\langle tl(G) \rangle \subseteq \langle tl(I) \rangle$. Seja então $f \in I$. Como $I = \langle G \rangle$, existem $h_1, \ldots, h_N \in K[x_1, \ldots, x_n]$ tais que

$$f = h_1 g_1 + \ldots + h_N g_N.$$

Como cada h_i é uma soma de termos, podemos escrever

$$f = \sum_{\alpha} \sum_{g \in G} c_{\alpha,g} x^{\alpha} g,$$

onde $c_{\alpha,g} \in K$ e $\alpha \in \mathbb{N}^n$.

Seja $x^{\delta} = \max\{x^{\alpha} ml(g) | c_{\alpha,g} \neq 0\}$, e

$$f^* := \sum_{x^{\alpha} ml(g) = x^{\delta}, g \in G} c_{\alpha,g} x^{\alpha} g.$$

Assim,

$$f = f^* + \text{ termos menores.}$$

Por boa ordenação, podemos tomar x^{δ} mínimo com esta propriedade, ou seja, para qualquer expressão de f na forma

$$f = \sum_{\beta} \sum_{g \in G} b_{\beta,g} x^{\beta} g,$$

com $b_{\beta,g} \in K$ e $\beta \in \mathbb{N}^n$, temos

$$x^{\delta} \preceq \max\{x^{\beta} ml(g) | b_{\beta,g} \neq 0\}.$$

Queremos ver agora que $ml(f^*) = x^{\delta}$. Suponhamos, por absurdo, que $ml(f^*) \prec x^{\delta}$. Pelo Lema 3.2.1, existem constantes $b_{jk} \in K$ tais que

$$f^* = \sum_{x^{\alpha} ml(g) = x^{\delta}, g \in G} c_{\alpha,g} x^{\alpha} g = \sum_{j,k} b_{jk} spol(x^{\alpha_j} g_j, x^{\alpha_k} g_k)$$

com $ml(spol(x^{\alpha_j} g_j, x^{\alpha_k} g_k)) \prec x^{delta}, \forall j, k$.

Pelo Lema 3.2.2, para cada j, k existe $\gamma_{jk} \in \mathbb{N}^n$ tal que

$$spol(x^{\alpha_j} g_j, x^{\alpha_k} g_k) = x^{\gamma_{jk}} spol(g_j, g_k).$$

Logo

$$f^* = \sum_{j,k} b_{jk} x^{\gamma_{jk}} spol(g_j, g_k)$$

e como $ml(spol(x^{\alpha_j} g_j, x^{\alpha_k} g_k)) \prec x^\delta$, segue que

$$ml(x^{\gamma_{jk}} spol(g_j, g_k)) = x^{\gamma_{jk}} ml(spol(g_j, g_k)) \prec x^\delta.$$

Por hipótese, $\overline{spol(g_j, g_k)}^G = 0$; segue daí que existem polinômios $q_1, \ldots, q_N$ tais que

$$spol(g_j, g_k) = q_1 g_1 + \cdots + q_N g_N$$

com $ml(q_i g_i) \preceq ml(spol(g_j, g_k))$. Como cada q_i é uma soma de monômios, podemos escrever

$$spol(g_j, g_k) = \sum_{\eta} \sum_{g \in G} a_{\eta,g} x^\eta g$$

com $x^\eta ml(g) \preceq ml(spol(g_j, g_k))$. Logo

$$x^{\gamma_{jk}} spol(g_j, g_k) = \sum_{\eta} \sum_{g \in G} a_{\eta,g} x^{\eta+\gamma_{jk}} g$$

com $x^{\eta+\gamma_{jk}} ml(g) \preceq ml(x^{\gamma_{jk}} spol(g_j, g_k)) = x^{\gamma_{jk}} ml(spol(g_j, g_k)) \prec x^\delta$. Segue que f^*, e portanto f, pode ser escrito na forma

$$\sum_{\mu} \sum_{g \in G} c'_{\mu,g} x^\mu g,$$

com cada termo $c'_{\mu,g} x^\mu g$ menor do que x^δ, o que contradiz a minimalidade de x^δ. □

O Teorema 3.2.1 fornece um critério para determinar se um conjunto $G = \{f_1, \ldots, f_m\}$ é uma base de Gröbner, e é a base para o algoritmo: para cada par de

polinômios f_i, f_j em G, calculamos $spol(f_i, f_j)$ e verificamos se $\overline{spol(f_i, f_j)}^G = 0$. Se $\overline{spol(f_i, f_j)}^G \neq 0$, acrescentamos o polinômio $f_{m+1} = \overline{spol(f_i, f_j)}^G$ ao conjunto G, e repetimos o processo para este novo G.

Algoritmo 3.2.1 Algoritmo de Buchberger

Entrada: $F = \{f_1, \ldots, f_m\}$.
Saída: $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner do ideal $\langle F \rangle$.
$G \leftarrow F$
$B \leftarrow \{\{p, q\} | p, q \in G, p \neq q\}$
while $B \neq \emptyset$ **do**
  selecione $\{p, q\} \in B$
  $B \leftarrow B \backslash \{\{p, q\}\}$
  $h \leftarrow spol(p, q)$
  $h \leftarrow$ resto da divisão de h por G
  if $h \neq 0$ **then**
    $B \leftarrow B \cup \{\{g, h\} | g \in G\}$
    $G \leftarrow G \cup \{h\}$
  end if
end while

Teorema 3.2.2. *Seja $F = \{f_1, \ldots, f_m\} \subset K[x_1, \ldots, x_n]$. Então o Algoritmo 3.2.1 constrói uma base de Gröbner G do ideal $\langle F \rangle$.*

Demonstração. Primeiramente mostraremos que em qualquer iteração do algoritmo temos que (i) $G \subset \langle F \rangle$ e (ii) $\overline{spol(p, q)}^G = 0$ para todo par $p, q \in G$ tal que $\{p, q\} \notin B$. Procederemos por indução no número de iterações.

Inicialmente as condições acima se verificam, pois todos os pares estão em B e $G = F$.

Na primeira iteração temos um único par $\{p, q\}$ que não está em B. Se $h = \overline{spol(p, q)}^G = 0$, então $G = F$. Se $h \neq 0$, então o conjunto G é atualizado e temos $G = F \cup \{h\}$, e assim $\overline{spol(p, q)}^G = 0$. Como h é a redução de $spol(p, q)$ módulo $f_1, \ldots, f_m$, temos que existem $q_1, \ldots, q_m$ tais que $spol(p, q) = q_1 f_1 + \cdots + q_m f_m + h$. Como $spol(p, q) \in \langle F \rangle$, segue que $h \in \langle F \rangle$, e portanto $G \subset \langle F \rangle$.

Suponhamos que as condições são satisfeitas na i-ésima iteração do algoritmo, ou seja,

$$G_i \subset \langle F \rangle \text{ e } \overline{spol(p,q)}^{G_i} = 0 \ \forall p, q \in G_i \text{ tais que } \{p,q\} \notin B_i$$

onde G_i e B_i denotam os conjuntos G e B na i-ésima iteração. Então temos

$$B_{i+1} = B_i \backslash \{\{p,q\}\}$$

$$h = \overline{spol(p,q)}^{G_i}$$

Se $h = 0$, então $G_{i+1} = G_i \subset \langle F \rangle$ e $\{\{p,q\} \mid p,q \in G_{i+1},\ \{p,q\} \notin B_{i+1}\} \subset \{\{p,q\} \mid p,q \in G_i e \{p,q\} \notin B_i\}$, logo as condições (i) e (ii) são satisfeitas.

Se $h \neq 0$, então os conjuntos B e G são atualizados, e temos

$$B_{i+1} = B_{i+1} \cup \{\{g,h\} | g \in G_i\}$$

$$G_{i+1} = G_i \cup \{h\}$$

logo

$$\{\{u,v\} | \{u,v\} \notin B_{i+1}\} = \{\{u,v\} | \{u,v\} \notin B_i\} \cup \{\{p,q\}\}.$$

Como $G_i \subset G_{i+1}$ e $\overline{spol(u,v)}^{G_i} = 0$ para todo par $\{u,v\} \notin B_i$, segue que $\overline{spol(u,v)}^{G_{i+1}} = 0$ para todo par $\{u,v\} \notin B_i$. E ainda, $\overline{spol(p,q)}^{G_{i+1}} = 0$. Logo a condição (ii) é satisfeita. Além disso, $G_i \subset \langle F \rangle$ e $h \in \langle F \rangle$, e portanto $G_{i+1} \subset \langle F \rangle$.

O algoritmo termina quando $B = \emptyset$, e então temos $\langle G \rangle = \langle F \rangle$ e $\overline{spol(p,q)}^{G} = 0$ para todo par $p, q \in G$. Pelo Teorema 3.2.1, o conjunto G obtido é uma base de Gröbner do ideal $\langle F \rangle$.

Para ver que o algoritmo realmente termina, observamos que o polinômio $\overline{spol(p,q)}^{G}$ é reduzido módulo G. Logo, se $\overline{spol(p,q)}^{G} \neq 0$, nenhum de seus termos é divisível por algum dos termos líderes dos elementos de G; em particular, $tl(\overline{spol(p,q)}^{G}) \notin \langle tl(G) \rangle$. Assim, cada vez que um novo elemento é acrescentado a G, o ideal $\langle tl(G) \rangle$ aumenta. Mas pelo Corolário 3.1.2 esta cadeia de ideais não pode crescer indefinidamente. □

Exemplo 3.2.1. Sejam $f_1 = x^2 + y^2 + 1$ e $f_2 = x^2y + 2xy + x$ polinômios em $\mathbb{Z}_5[x, y]$. Vamos calcular uma base de Gröbner do ideal $I = \langle f_1, f_2 \rangle$, usando a ordem lexicográfica com $x \succ y$. A tabela abaixo mostra os resultados do Algoritmo 3.2.1 neste caso.

	h	G	B
início		$\{f_1, f_2\}$	$\{\{f_1, f_2\}\}$
1ª iter.	$f_3 = 3xy + 4x + y^3 + y$	$\{f_1, f_2, f_3\}$	$\{\{f_1, f_3\}, \{f_2, f_3\}\}$
2ª iter.	$f_4 = 4y^5 + 3y^4 + y^2 + y + 3$	$\{f_1, f_2, f_3, f_4\}$	$\{\{f_2, f_3\}, \{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}\}$
3ª iter.	0	$\{f_1, f_2, f_3, f_4\}$	$\{\{f_1, f_4\}, \{f_2, f_4\}, \{f_3, f_4\}\}$
4ª iter.	0	$\{f_1, f_2, f_3, f_4\}$	$\{\{f_2, f_4\}, \{f_3, f_4\}\}$
5ª iter.	0	$\{f_1, f_2, f_3, f_4\}$	$\{\{f_3, f_4\}\}$
6ª iter.	0	$\{f_1, f_2, f_3, f_4\}$	$\emptyset$

O conjunto $G = \{f_1, f_2, f_3, f_4\}$ é uma base de Gröbner do ideal I. Note que f_2 pode ser removido de G, pois o conjunto $\{f_1, f_3, f_4\}$ ainda é uma base de Gröbner do ideal I, uma vez que

$$\langle tl(I) \rangle = \langle x^2, x^2y, 3xy, 4y^5 \rangle = \langle x^2, 3xy, 4y^5 \rangle.$$

De fato, para obter a base de Gröbner reduzida de I basta dividir f_1, f_3, f_4 pelos seus respectivos termos líderes para torná-los mônicos. Assim, a base reduzida de I é dada por

$$\{x^2 + y^2 + 1, xy + 3x + 2y^3 + 2y, y^5 + 2y^4 + 4y^2 + 4y + 2\}$$

O algoritmo de Buchberger baseia-se no conceito de S-polinômio e no algoritmo da divisão de polinômios multivariados. A divisão é um processo computacionalmente custoso, e por isso é vantajoso evitá-la sempre que possível. A versão do algoritmo de Buchberger apresentada acima pode ser melhorada utlizando-se critérios

conhecidos que permitem desconsiderar certos S-polinômios, evitando assim algumas divisões.

No entanto, mesmo com as melhores versões do algoritmo atualmente conhecidas, existem exemplos de ideais para os quais o cômputo de uma base de Gröbner leva muito tempo ou consome uma quantidade enorme de espaço de armazenamento. De fato, sabe-se que a complexidade do algoritmo de Buchberger é duplamente exponencial.

4 BASES MONOMIAIS E IDEAIS ZERO DIMENSIONAIS

Neste capítulo apresentaremos algumas aplicações das bases de Gröbner. Primeiramente, veremos como utilizar a teoria das bases de Gröbner para encontrar uma base do quociente $K[x_1, \ldots, x_n]/I$ como espaço vetorial sobre K. A seguir, apresentaremos um algoritmo para calcular a base de Gröbner do ideal de um conjunto finito de pontos.

4.1 Anéis Quociente e Bases Monomiais

Seja I um ideal de $K[x_1, \ldots, x_n]$. Dados dois polinômios $f, g \in K[x_1, \ldots, x_n]$, dizemos que f e g são congruentes módulo I, e escrevemos $f \equiv g \pmod{I}$, se $f - g \in I$. A relação de congruência assim definida é uma relação de equivalência, e portanto o conjunto das classes de equivalência forma uma partição de $K[x_1, \ldots, x_n]$. Cada $f \in K[x_1, \ldots, x_n]$ está em uma única classe de equivalência, dada por $f + I = \{f + h : h \in I\}$, e geralmente denotada por $\overline{f}$. O conjunto de todas as classes de equivalência é denotado por $K[x_1, \ldots, x_n]/I$. É facil ver que a relação satisfaz as seguintes propriedades:

Se $f_1 \equiv g_1 \pmod{I}$ e $f_2 \equiv g_2 \pmod{I}$, então

(i) $f_1 + f_2 \equiv g_1 + g_2 \pmod{I}$

(ii) $f_1 \cdot f_2 \equiv g_1 \cdot g_2 \pmod{I}$

(iii) $h \cdot f_1 \equiv h \cdot g_1 \pmod{I}$, para qualquer $h \in K[x_1, \ldots, x_n]$.

Definindo as operações

$$\overline{f} + \overline{g} = \overline{f+g},\ \overline{f} \cdot \overline{g} = \overline{f \cdot g},$$

as propriedades (i) e (ii) implicam que as duas operações estão bem definidas e que $K[x_1, \ldots, x_n]/I$ com as operações acima é um anel comutativo. Além disso, $K[x_1, \ldots, x_n]/I$ é um espaço vetorial sobre K. Veremos agora como resolver os seguintes problemas relacionados a esta construção:

(a) Determinar um conjunto de representantes das classes de $K[x_1, \ldots, x_n]/I$;

(b) Determinar uma base de $K[x_1, \ldots, x_n]/I$ como espaço vetorial sobre K.

Seja $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner. Sabemos pela Proposição 3.1.2 que para cada $f \in K[x_1, \ldots, x_n]$ existe um único elemento $r \in K[x_1, \ldots, x_n]$, reduzido módulo G, tal que $r = \overline{f}^G$.

Definição 4.1.1. Dados um polinômio $f \in K[x_1, \ldots, x_n]$ e uma base de Gröbner G, a *forma normal* de f módulo G, denotada por $N_G(f)$, é o resto da divisão de f por G.

Proposição 4.1.1. *Sejam $f, g \in K[x_1, \ldots, x_n]$, I um ideal de $K[x_1, \ldots, x_n]$ e G uma base de Gröbner de I. Então $f \equiv g \pmod{I}$ se e somente se $N_G(f) = N_G(g)$. Portanto, $\{N_G(f) : f \in K[x_1, \ldots, x_n]\}$ é um conjunto de representantes das classes de $K[x_1, \ldots, x_n]/I$.*

Demonstração. Suponhamos que $f \equiv g \pmod{I}$. Pelo Algoritmo da Divisão, existem $q_1, q_2 \in I$ tais que $f = q_1 + N_G(f)$ e $g = q_2 + N_G(g)$. Então temos que

$$f - g = (q_1 - q_2) + (N_G(f) - N_G(g)).$$

Como nenhum dos termos de $N_G(f)$ e $N_G(g)$ é divisível pelos termos líderes de $g_1, \ldots, g_\ell$, os termos de $N_G(f) - N_G(g)$ também não são. Logo, $N_G(f) - N_G(g)$ é

reduzido módulo G, e pela Proposição 3.1.2, segue que $N_G(f) - N_G(g) = N_G(f-g)$. Como $f \equiv g \pmod{I}$, temos que $f - g \in I$, logo $N_G(f-g) = 0$. Daí

$$N_G(f) - N_G(g) = N_G(f-g) = 0$$

e, portanto, $N_G(f) = N_G(g)$.

Reciprocamente, se $N_G(f) = N_G(g)$, então

$$f - g = (f - N_G(f)) - (g - N_G(g))$$

e como $f - N_G(f) \in I$ e $g - N_G(g) \in I$, segue que $f - g \in I$. Logo, $f \equiv g \pmod{I}$. □

Para resolver o segundo problema proposto, faremos a seguinte definição.

Definição 4.1.2. Fixada uma ordem monomial $\succ$ em $K[x_1, \ldots, x_n]$, para qualquer conjunto $S \subseteq K[x_1, \ldots, x_n]$ definimos

$$\mathcal{B}(S) = \{x^\alpha : \alpha \in \mathbb{N}^n \text{ e } x^\alpha \text{ não é divisível por nenhum } tl(g), g \in S\}.$$

O conjunto $\mathcal{B}(S)$ assim definido tem a seguinte propriedade: se $S_1 \subseteq S_2$ então $\mathcal{B}(S_2) \subseteq \mathcal{B}(S_1)$. De fato, se $x^\alpha \in \mathcal{B}(S_2)$, então x^α não é divisível por nenhum $tl(g), g \in S_2$; como $S_1 \subseteq S_2$, em particular x^α não é divisível por nenhum $tl(g), g \in S_1$. Logo $x^\alpha \in \mathcal{B}(S_1)$.

Proposição 4.1.2. *Sejam I um ideal de $K[x_1, \ldots, x_n]$ e $G \subset I$. Então $\mathcal{B}(I) = \mathcal{B}(G)$ se e somente se G é uma base de Gröbner de I.*

Demonstração. Pela Proposição 3.1.1, temos que

$$\mathcal{B}(I) = \{x^\alpha : x^\alpha \notin \langle tl(I) \rangle\}$$

Se G é uma base de Gröbner de I, então $\langle tl(I) \rangle = \langle tl(G) \rangle$, logo

$$\mathcal{B}(I) = \{x^\alpha : x^\alpha \notin \langle tl(I) \rangle\} = \{x^\alpha : x^\alpha \notin \langle tl(G) \rangle\} = \mathcal{B}(G).$$

Reciprocamente, suponhamos que $\mathcal{B}(I) = \mathcal{B}(G)$. Para mostrar que G é uma base de Gröbner de I, mostraremos que $\langle tl(I)\rangle = \langle tl(G)\rangle$. Seja $cx^\alpha = tl(f), f \in K[x_1, \ldots, x_n]$, onde $c \in K$ e x^α é um monômio. Suponhamos que $tl(f) \notin \langle tl(G)\rangle$. Então $x^\alpha \notin \langle tl(G)\rangle$, o que implica que $x^\alpha \in \mathcal{B}(G) = \mathcal{B}(I)$. Logo, $x^\alpha \notin \langle tl(I)\rangle$. □

O próximo resultado nos diz que $\mathcal{B}(I)$ é a base de $K[x_1, \ldots, x_n]/I$ que procurávamos.

Proposição 4.1.3. *Seja I um ideal de $K[x_1, \ldots, x_n]$. Então uma base do K-espaço vetorial $K[x_1, \ldots, x_n]/I$ consiste das classes de todos os monômios em $\mathcal{B}(I)$.*

Demonstração. Seja $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner de I. Pela Proposição 4.1.1, sabemos que para cada polinômio $f \in K[x_1, \ldots, x_n]$, $f + I = N_G(f) + I$ em $K[x_1, \ldots, x_n]/I$. Como $N_G(f)$ é reduzido módulo G, por definição, $N_G(f)$ é uma combinação K-linear de monômios que não são divisíveis pelos termos líderes de $g_1, \ldots, g_\ell$, ou seja, $N_G(f)$ é uma combinação K-linear de monômios de $\mathcal{B}(G) = \mathcal{B}(I)$. Logo, o conjunto das classes dos monômios de $\mathcal{B}(I)$ gera o espaço vetorial $K[x_1, \ldots, x_n]/I$.

Resta mostrar que as classes dos monômios de $\mathcal{B}(I)$ são linearmente independentes. Sejam $x^{\alpha_1}, \ldots, x^{\alpha_t} \in \mathcal{B}(I)$, e sejam $c_1, \ldots, c_t \in K$ tais que

$$c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t} \equiv 0 \pmod{I}. \tag{4.1}$$

Então, pela Proposição 4.1.1, $N_G(c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t}) = N_G(0) = 0$. Mas como os monômios de $c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t}$ não são divisíveis pelos termos líderes de $g_1, \ldots, g_\ell$, temos que $c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t}$ é reduzido módulo G. Logo, $N_G(c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t}) = c_1 x^{\alpha_1} + \cdots + c_t x^{\alpha_t} = 0$, o que implica que as constantes c_i em (4.1) são todas nulas. □

O conjunto $\mathcal{B}(I)$ é chamado de *base monomial* de $K[x_1, \ldots, x_n]/I$ com respeito à ordem monomial dada. Por simplicidade, diremos que $\mathcal{B}(I)$ é uma base monomial de I.

Exemplo 4.1.1. Vimos no Exemplo 3.2.1 que a base de Gröbner reduzida do ideal $I = \langle f_1, f_2 \rangle \subset \mathbb{Z}_5[x, y]$, onde $f_1 = x^2 + y^2 + 1$ e $f_2 = x^2y + 2xy + x$, em relação à ordem lexicográfica é dada por

$$G = \{g_1, g_2, g_3\}$$

onde

$$g_1 = x^2 + y^2 + 1,\ g_2 = xy + 3x + 2y^3 + 2y,\ g_3 = y^5 + 2y^4 + 4y^2 + 4y + 2.$$

Então

$$\mathcal{B}(I) = \mathcal{B}(G) = \{1, x, y, y^2, y^3, y^4\},$$

e $\dim_{\mathbb{Z}_5}(\mathbb{Z}_5[x, y]/I) = 6$.

4.2 Teorema dos Zeros de Hilbert

Seja $\overline{K}$ o fecho algébrico do corpo K. Dado um subconjunto $S \subseteq K[x_1, \ldots, x_n]$, definimos a *variedade*, denotada por $V(S)$, em $\overline{K}^n$ por

$$V(S) = \{P \in \overline{K}^n : f(P) = 0 \text{ para todo } f \in S\}.$$

Dado um subconjunto $V \subseteq \overline{K}^n$, definimos o ideal $I(V)$ em $K[x_1, \ldots, x_n]$ por

$$I(V) = \{f \in K[x_1, \ldots, x_n] : f(P) = 0 \text{ para todo } P \in V\}.$$

A seguir enunciaremos o famoso Teorema dos Zeros de Hilbert, que estabelece a relação entre um ideal I e o ideal $I(V(I))$. Para demonstrá-lo, faremos uso do seguinte resultado, conhecido como a versão fraca do Teorema dos Zeros, que apresentaremos sem demonstracão [11, pág. 411].

Teorema 4.2.1. *Seja I um ideal em $K[x_1, \ldots, x_n]$. Então $V(I) = \emptyset$ se e somente se $I = K[x_1, \ldots, x_n]$.*

Definição 4.2.1. Dado um ideal $I \subseteq K[x_1, \ldots, x_n]$, definimos o *radical* de I, denotado $\sqrt{I}$, por

$$\sqrt{I} = \{f \in K[x_1, \ldots, x_n] : \text{ existe } e \in \mathbb{N} \text{ tal que } f^e \in I\}.$$

Teorema 4.2.2. (Teorema dos Zeros de Hilbert) *Seja I um ideal em $K[x_1, \ldots, x_n]$. Então $I(V(I)) = \sqrt{I}$.*

Demonstração. Se $f \in \sqrt{I}$, então $f^e \in I$ para algum $e \geq 1$. Se $(a_1, \ldots, a_n) \in \overline{K}^n$ é um zero de I, então $0 = f^e(a_1, \ldots, a_n) = (f(a_1, \ldots, a_n))^e$, logo $f(a_1, \ldots, a_n) = 0$. Portanto, $\sqrt{I} \subseteq I(V(I))$.

Seja agora $f \in I(V(I))$. Como $0 \in \sqrt{I}$, podemos supor que $f \neq 0$. Consideraremos o ideal J do anel de polinômios em $n+1$ variáveis $K[x_1, \ldots, x_n, y]$ gerado por I e $yf - 1$; assim, se $(a_1, \ldots, a_n, b) \in \overline{K}^{n+1}$ é um zero de J, então $(a_1, \ldots, a_n) \in \overline{K}^n$ deve ser um zero de I. Mas $(yf - 1)(a_1, \ldots, a_n, b) = bf(a_1, \ldots, a_n) - 1 = -1$, para todos os zeros $(a_1, \ldots, a_n)$ de I em $\overline{K}^n$. Logo, $V(J) = \emptyset$, e pelo Teorema 4.2.1, $J = K[x_1, \ldots, x_n, y]$. Segue que $1 \in J$, e então podemos escrever

$$1 = \sum_{i=1}^{t-1} g_i f_i + g_t(yf - 1), \tag{4.2}$$

com $f_i \in I$ e $g_i \in K[x_1, \ldots, x_n, y]$. Fazendo $y = 1/f$ em 4.2 obtemos a seguinte igualdade em $K(x_1, \ldots, x_n)$

$$1 = \sum_{i=1}^{t-1} g_i \left(x_1, \ldots, x_n, \frac{1}{f}\right) f_i(x_1, \ldots, x_n). \tag{4.3}$$

Se e é um inteiro positivo maior do que o grau de cada g_i em y, então multiplicando a equação (4.3) por f^e obtemos

$$f^e = \sum_{i=1}^{t-1} f^e(x_1, \ldots, x_n) g_i \left(x_1, \ldots, x_n, \frac{1}{f}\right) f_i(x_1, \ldots, x_n), \tag{4.4}$$

onde cada $f^e(x_1, \ldots, x_n)g_i(x_1, \ldots, x_n, \frac{1}{f}) \in K[x_1, \ldots, x_n]$. Logo $f^e \in I$, ou seja, $f \in \sqrt{I}$ e, portanto, $I(V(I)) \subseteq \sqrt{I}$. □

Proposição 4.2.1. *Sejam $I \subseteq K[x_1, \ldots, x_n]$ um ideal e $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner de I. Então as seguintes afirmações são equivalentes:*

(i) A variedade $V(I)$ é finita;

(ii) Para cada $i = 1, \ldots, n$, exite $j \in \{1, \ldots, \ell\}$ tal que $ml(g_j) = x_i^\nu$, para algum $\nu \in \mathbb{N}$;

(iii) A dimensão do K-espaço vetorial $K[x_1, \ldots, x_n]/I$ é finita.

Demonstração. (i)⇒(ii) Suponhamos que $V(I)$ é finita. Se $V(I) = \emptyset$, então, pelo Teorema 4.2.1, $I = K[x_1, \ldots, x_n]$, logo $G = \{1\}$ e (ii) é satisfeita com $j = 1$ e $\nu = 0$ para todo $i = 1, \ldots, n$.

Suponhamos então que $V(I) = \{P_1, \ldots, P_m\} \neq \emptyset$. Seja $i \in \{1, \ldots, n\}$. Para cada $j \in \{1, \ldots, m\}$, seja $a_{ij} \in \overline{K}$ a i-ésima coordenada do ponto P_j, e seja $f_j \in K[x_i]$ um polinômio não nulo tal que $f_j(a_{ij}) = 0$. Um tal polinômio f_j sempre existe, pois $\overline{K}$ é extensão algébrica de K. Seja $f = f_1 f_2 \cdots f_m \in K[x_i] \subset K[x_1, \ldots, x_n]$. Como $f(P_j) = f_1(a_{ij}) \cdots f_m(a_{ij})$ e $f_j(a_{ij}) = 0$, temos que $f \in I(V(I))$, e pelo Teorema 4.2.2, existe e tal que $f^e \in I$. Ademais, como $f \in K[x_i]$, temos que $ml(f^e) = x_i^{te}$ para algum $t \in \mathbb{N}$. Como o monômio líder de todo elemento de I é divisível pelo monômio líder de algum elemento de G, existe um elemento de G cujo monômio líder é uma potência de x_i.

(ii)⇒(iii) Suponhamos que para cada $i = 1, \ldots, n$ existe j tal que $ml(g_j) = x_i^{\nu_i}$. Pela Proposição 4.1.3, uma base de $K[x_1, \ldots, x_n]/I$ é o conjunto das classes dos monômios reduzidos módulo G. Por definição, um monômio $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ é reduzido módulo G se x^α não é divisível pelos termos líderes de $g_1, \ldots, g_\ell$. Se $\alpha_i \geq \nu_i$, então x^α é divisível pelo termo líder do elemento g_j correspondente a i. Logo, existe apenas um número finito de monômios reduzidos módulo G e, portanto, $\dim_K K[x_1, \ldots, x_n]/I$ é finita.

(iii)⇒(i) Suponhamos agora que $\dim_K K[x_1,\ldots,x_n]/I$ é finita. Seja $P = (a_1,\ldots,a_n) \in V(I)$. Mostraremos que para cada $i = 1,\ldots,n$ há apenas um número finito de possíveis valores para a_i.

Seja $i \in \{1,\ldots,n\}$. Como, por hipótese, $K[x_1,\ldots,x_n]/I$ é um K-espaço vetorial de dimensão finita, as potências $1, x_i, x_i^2, x_i^3, \ldots$ de x_i são linearmente dependentes módulo I. Então existe $m \in \mathbb{N}$ e constantes $c_j \in K,\ j = 0,1,\ldots,m$, não todas nulas, tais que $f = c_0 + c_1x_i + c_2x_i^2 + \cdots + c_mx_i^m \in I$. Como este polinômio está em I, temos que $f(P) = c_0 + c_1a_i + c_2a_i^2 + \cdots + c_ma_i^m = 0$, ou seja, a_i é raiz do polinômio f. Como f tem apenas um número finito de raízes em $\overline{K}$, há um número finito de possíveis valores de a_i. □

Definição 4.2.2. Um ideal $I \subset K[x_1,\ldots,x_n]$ é dito um ideal *zero dimensional* se I satisfaz qualquer uma das condições equivalentes da Proposição 4.2.1. Neste caso, definimos o *grau* de I como sendo a dimensão do K-espaço vetorial $K[x_1,\ldots,x_n]/I$.

4.3 Ideais de Conjuntos Finitos de Pontos

Nesta seção apresentaremos o algoritmo desenvolvido por Farr e Gao em [7] para encontrar uma base de Gröbner do ideal de um conjunto finito de pontos. Dado um conjunto de pontos distintos $V = \{P_1,\ldots,P_m\}$, denotaremos o ideal $I(V)$ por $I(P_1,\ldots,P_m)$.

Proposição 4.3.1. *Sejam $P_1,\ldots,P_m \in K^n$ pontos distintos, $I = I(P_1,\ldots,P_m)$ e $G \subset I$. Então G é uma base de Gröbner de I se e somente se $|\mathcal{B}(G)| = m$.*

Demonstração. Sabemos que as classes dos monômios em $\mathcal{B}(I)$ formam uma K-base de $K[x_1,\ldots,x_n]/I$; logo, $|\mathcal{B}(I)| = dim_K K[x_1,\ldots,x_n]/I$.

Considere $\phi : K[x_1,\ldots,x_n] \longrightarrow K^m$ definida por

$$\phi(f) = (f(P_1),\ldots,f(P_m))$$

A aplicação ϕ assim definida é um homomorfismo de K-espaços vetoriais. De fato,

$$\begin{aligned}
\phi(f+g) &= (f(P_1)+g(P_1),\ldots,f(P_m)+g(P_m)) \\
&= (f(P_1),\ldots,f(P_m))+(g(P_1),\ldots,g(P_m)) \\
&= \phi(f)+\phi(g) \\
\phi(af) &= (af(P_1),\ldots,af(P_m)) \\
&= a(f(P_1),\ldots,f(P_m)) \\
&= a\phi(f) \qquad (4.5)
\end{aligned}$$

para quaisquer $f, g \in K[x_1,\ldots,x_n]$ e $a \in K$.

O núcleo do homomorfismo ϕ é dado por $\ker(\phi) = \{f \in K[x_1,\ldots,x_n] : f(P_i) = 0,\ i = 1,\ldots,m\}$, ou seja, $\ker(\phi) = I(P_1,\ldots,P_m) = I$.

Seja $(b_1,\ldots,b_m) \in K^m$, e suponhamos que $P_i = (a_{i1},\ldots,a_{in})$. Como os pontos $P_1,\ldots,P_m$ são dois a dois distintos, dados $i \neq j$, existe $k \in \{1,\ldots,n\}$ tal que $a_{ik} \neq a_{jk}$. Definimos então

$$f_{i,j} = \frac{x_k - a_{jk}}{a_{ik} - a_{jk}}$$

Assim, temos que $f_{i,j}(P_i) = 1$ e $f_{i,j}(P_j) = 0$. Consideramos agora, para cada $i = 1,\ldots,m$, o polinômio

$$f_i = b_i \prod_{\substack{j=1 \\ j\neq i}}^{m} f_{i,j}$$

Então $f_i(P_i) = b_i$ e $f_i(P_j) = 0$, para todo $j \neq i$. Finalmente, consideramos o polinômio

$$f = f_1 + \cdots + f_m$$

Este polinômio é tal que $f(P_i) = f_i(P_i) = b_i$. Logo $\phi(f) = (b_1,\ldots,b_m)$ e, portanto, ϕ é sobrejetora.

Pelo Teorema dos Isomorfismos para espaços vetoriais, $K[x_1,\ldots,x_n]/I$ e K^m são isomorfos como K-espaços vetoriais, o que implica

$$|\mathcal{B}(I)| = dim_K K[x_1,\ldots,x_n]/I = dim_K K^m = m.$$

Se G é uma base de Gröbner de I, então $\mathcal{B}(G) = \mathcal{B}(I)$ e, portanto, $|\mathcal{B}(G)| = m$. Reciprocamente, suponhamos que $\mathcal{B}(G) = m$; então $|\mathcal{B}(G)| = |\mathcal{B}(I)|$. Como $G \subset I$, temos que $\mathcal{B}(I) \subseteq \mathcal{B}(G)$, o que implica $\mathcal{B}(G) = \mathcal{B}(I)$, e pela Proposição 4.1.2 temos que G é base de Gröbner de I. □

Proposição 4.3.2. *Sejam $V \subset K^n$ um conjunto finito, $G = \{g_1, \ldots, g_\ell\}$ uma base de Gröbner do ideal $I(V)$ e $P = (a_1, \ldots, a_n) \notin V$. Se g_i é o polinômio em G de menor termo líder tal que $g_i(P) \neq 0$, então o conjunto*

$$\tilde{G} = \{\tilde{g}_1, \ldots, \tilde{g}_{i-1}, \tilde{g}_{i+1}, \ldots, \tilde{g}_\ell, g_{i1}, \ldots, g_{in}\},$$

onde

$$\tilde{g}_j = g_j - \frac{g_j(P)}{g_i(P)} \cdot g_i, \quad j \neq i, \ \ e$$
$$g_{ik} = (x_k - a_k) \cdot g_i, \quad 1 \leq k \leq n,$$

é uma base de Gröbner do ideal $I(V \cup \{P\})$.

Demonstração. Como $P \notin V$, pelo menos um polinômio em $I(V)$ deve não se anula em P; logo, g_i existe.

Se $Q = (b_1, \ldots, b_n) \in V$, então

$$\tilde{g}_j(Q) = g_j(Q) - \frac{g_j(P)}{g_i(P)} \cdot g_i(Q),$$

e como $g_j, g_i \in I(V)$, temos $g_j(Q) = g_i(Q) = 0$, logo $\tilde{g}_j(Q) = 0$. Temos ainda que

$$\tilde{g}_j(P) = g_j(P) - \frac{g_j(P)}{g_i(P)} \cdot g_i(P) = g_j(P) - g_j(P) = 0,$$

logo, $\tilde{g}_j \in I(V \cup \{P\})$.

Agora,

$$g_{ik}(Q) = (b_k - a_k) \cdot g_i(Q) = 0,$$

pois $g_i(Q) = 0$, e

$$g_{ik}(P) = (a_k - a_k) \cdot g_i(P) = 0;$$

Logo, temos também que $g_{ik} \in I(V \cup \{P\})$. Portanto $\tilde{G} \subset I(V \cup \{P\})$.

Seja $x^\alpha = ml(g_i)$. Afirmamos que $\mathcal{B}(\tilde{G}) = \mathcal{B}(G) \cup \{x^\alpha\}$.

Se g_i e g_j têm monômios líderes iguais, para $j \neq i$, então $G\backslash\{g_j\}$ também é base de Gröbner de $I(V)$. Logo, podemos assumir que $ml(g_j) \neq ml(g_i)$ para $j \neq i$. Pela forma como g_i foi escolhido, temos que, para $j \neq i$, ou $g_j(P) = 0$ ou $ml(g_j) \succ ml(g_i) = x^\alpha$. Se $g_j(P) = 0$, então $\tilde{g}_j = g_j$ e, consequentemente, $ml(\tilde{g}_j) = ml(g_j)$. Se $ml(g_j) \succ ml(g_i)$, então $ml(\tilde{g}_j) = \max\{ml(g_j), ml(g_i)\} = ml(g_j)$. Portanto, temos $ml(\tilde{g}_j) = ml(g_j)$ para todo $j \neq i$. Para os polinômios g_{ik} temos $ml(g_{ik}) = ml(x_k - a_k) \cdot ml(g_i) = x_k \cdot x^\alpha$. Segue que $\mathcal{B}(G) \subset \mathcal{B}(\tilde{G})$.

Suponhamos agora que x^β é tal que $x^\beta \in \mathcal{B}(\tilde{G})\backslash\mathcal{B}(G)$. Então temos que $x_k x^\alpha \nmid x^\beta$ para cada $k = 1, \ldots, n$, mas $x^\alpha | x^\beta$, o que implica que $x^\beta = x^\alpha$.

Logo, $\mathcal{B}(\tilde{G}) = \mathcal{B}(G) \cup \{x^\alpha\}$, e então temos $|\mathcal{B}(\tilde{G})| = |\mathcal{B}(G)| + 1$. Como G é base de Gröbner de $I(V)$, pela Proposição 4.3.1 temos que $|\mathcal{B}(G)| = |V|$. Assim, $|\mathcal{B}(\tilde{G})| = |V| + 1$, e aplicando a Proposição 4.3.1 novamente obtemos que $\tilde{G}$ é base de Gröbner de $I(V \cup \{P\})$. □

Nosso objetivo é obter um algoritmo que, dado um conjunto finito de pontos distintos $\{P_1, \ldots, P_m\} \subset K^n$, retorna a base de Gröbner reduzida de $I(P_1, \ldots, P_m)$. Para isso, utilizaremos a Proposição 4.3.2. Suponhamos que na Proposição 4.3.2 a base de Gröbner G do ideal $I(P_1, \ldots, P_m)$ é reduzida. Vimos que

$$tl(\tilde{G}) = \{tl(g_1), \ldots, tl(g_{i-1}), tl(g_{i+1}), \ldots, tl(g_\ell), x_1 tl(g_1), \ldots, x_n tl(g_i)\}.$$

Para $j \neq i$, os termos de $\tilde{g}_j$ são combinações dos termos de g_j e g_i. Como g_j e g_i são reduzidos módulo $G\backslash\{g_j\}$ e $G\backslash\{g_i\}$, respectivamente, seus termos não são divisíveis pelos elementos de $tl(\tilde{G})\backslash\{tl(g_j)\}$. Logo, $\tilde{g}_j$ é reduzido módulo $\tilde{G}\backslash\{\tilde{g}_j\}$. Quanto aos polinômios g_{ik}, temos que eles não são necessariamente reduzidos módulo $\tilde{G}\backslash\tilde{g}_{ik}$. O Algoritmo 4.3.1 calcula a base de Gröbner reduzida de $I(P_1, \ldots, P_m)$, utilizando uma classe específica de ordens monomiais para simplificar o processo.

Algoritmo 4.3.1 Base de Gröbner do Ideal de um Conjunto Finito de Pontos

Entrada: $P_1, \ldots, P_m \in K^n$ e uma ordem mononial tal que $x_1 \prec x_2 \prec \cdots \prec x_n$.
Saída: $G = \{g_1, \ldots, g_\ell\}$ base de Gröbner reduzida do ideal $I(P_1, \ldots, P_m)$.
$G \leftarrow \{1\}$ {o i-ésimo polinômio em G é denotado por g_i}
for $k = 1, \ldots, m$ **do**
  encontre o menor i tal que $g_i(P_k) \neq 0$
  for $j = i+1, \ldots, |G|$ **do**
    $g_j \leftarrow g_j - \frac{g_j(P_k)}{g_i(P_k)} \cdot g_i$
  end for
  $G \leftarrow G \backslash \{g_i\}$
  for $j = 1, \ldots, n$ **do**
    if $x_j \cdot tl(g_i)$ não é divisível por nenhum termo líder de G **then**
      $h \leftarrow N_G((x_j - a_j) \cdot g_i)$
      insira h em ordem em G
    end if
  end for
end for

Proposição 4.3.3. *Sejam $P_1, \ldots, P_m$ pontos distintos em K^n. Então o Algoritmo 4.3.1 calcula a base de Gröbner reduzida do ideal $I(P_1, \ldots, P_m)$. Além disso, os elementos da base obtida estão em ordem crescente em relação aos termos líderes.*

Demonstração. Mostraremos que para $k = 1, \ldots, m$ o conjunto G obtido ao fim da k-ésima iteração é a base de Gröbner reduzida do ideal $I(P_1, \ldots, P_k)$.

Inicialmente, $G = \{1\}$. Na primeira iteração, temos $i = 1$ e $g_i = 1$. Como $|G| = 1$ e $i + 1 = 2$, o comando do primeiro loop interno não é executado. Em seguida $G = \emptyset$. No segundo loop interno, para $j = 1, \ldots, n$ os polinômios $x_j - a_j$ são inseridos em G, nesta ordem, de forma que ao fim da primeira iteração do algoritmo temos $G = \{x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n\}$, onde $a_1, \ldots, a_n$ são as coordenadas do ponto P_1. Queremos ver que o conjunto G é realmente a base de Gröbner reduzida de $I(P_1)$. Claramente, $G \subset I(P_1)$. Se $f \in I(P_1)\backslash\{0\}$, então $f(P_1) = 0$, de onde segue que f é um polinômio não constante, o que implica que $x_j | tl(f)$ para algum $j \in \{1, \ldots, n\}$. Logo, $tl(f) \in \langle tl(G) \rangle = \langle x_1, \ldots, x_n \rangle$. Portanto, G é base de Gröbner de $I(P_1)$. O fato de que a base é reduzida é evidente.

Suponhamos agora que o conjunto G obtido ao fim da $(r-1)$-ésima iteração, que denotaremos por G_{r-1}, é a base de Gröbner reduzida de $I(P_1, \ldots, P_{r-1})$, com seus elementos em ordem crescente de acordo com os termos líderes. Denotaremos por G_r o conjunto G construído na r-ésima iteração. Seja i o menor tal que $g_i(P_r) \neq 0$. Para $j < i$ os elementos $g_j \in G_{r-1}$ não são alterados. Os elementos g_j com $j > i$ são substituídos por

$$\tilde{g}_j = g_j - \frac{g_j(P_r)}{g_i(P_r)} \cdot g_i$$

e g_i é retirado. Até aqui temos

$$G_r = \{g_1, \ldots, g_{i-1}, \tilde{g}_{i+1}, \ldots, \tilde{g}_\ell\}$$

onde $\ell = |G_{r-1}|$. Como os elementos de G_{r-1} estão em ordem crescente de acordo com o termo líder, o elemento de menor termo líder tal que $g_j(P_r) \neq 0$ é o elemento g_i. Assim, para $j < i$ temos $g_j(P_r) = 0$, logo

$$g_j = g_j - \frac{g_j(P_r)}{g_i(P_r)} \cdot g_i = \tilde{g}_j.$$

Sabemos que $tl(\tilde{g}_j) = tl(g_j)$, logo os elementos de G_r estão na ordem desejada.

A seguir são inseridos em G_r as formas normais dos polinômios $g_{ij} = (x_j - a_j) \cdot g_i$. Note que somente são colocados em G_r aqueles g_{ij} tais que $tl(g_{ij}) = x_j \cdot tl(g_i)$ não é divisível pelo termo líder de nenhum elemento de G_r, pois caso contrário $\langle tl(G_r)\rangle = \langle tl(G_r \cup \{g_{ij}\}\rangle$, e g_{ij} pode ser omitido. Uma vez que os polinômios são inseridos em G_r na posição correta, os elementos de G_r estão ordenados da forma que pretendíamos. Além disso, pelas observações que seguem a Proposição 4.3.2, temos que os elementos $\tilde{g}_j$ são reduzidos módulo $G_r \backslash \{\tilde{g}_j\}$. Como a ordem monomial utilizada é tal que $x_1 \prec x_2 \prec \cdots \prec x_n$, temos $tl(g_{is}) = x_s \cdot tl(g_i) \prec x_t tl(g_i) = tl(G_{it})$, para $s < t$. Logo, a forma normal de g_{is} continua reduzida módulo G_r quando os demais polinômios são inseridos. Assim, pela Proposição 4.3.2, o conjunto G_r obtido é a base de Gröbner reduzida de $I(P_1, \ldots, P_r)$. □

Exemplo 4.3.1. Consideremos o corpo $K = \mathbb{Z}_5 = \mathbb{Z}/5\mathbb{Z}$, e os pontos $P_1 = (2, 3), P_2 = (1, 2)$ e $P_3 = (0, 4)$ em $\mathbb{Z}_5^2$. A tabela abaixo mostra os resultados do

Algoritmo 4.3.1 para estes pontos, utilizando a ordem lexicográfica com $x_1 \prec x_2$. Começamos com $G = \{1\}$, e ao final da k-ésima iteração, para $k = 1, 2, 3$, o conjunto G é a base de Gröbner reduzida de $I(P_1), I(P_1, P_2)$ e $I(P_1, P_2, P_3)$, respectivamente.

k	i	$\tilde{g}_j$ e g_{ij}	G
1	1	$g_{11} = x_1 + 3$ $g_{12} = x_2 + 2$	$\{x_1 + 3, x_2 + 2\}$
2	1	$\tilde{g}_2 = x_2 + 4x_1 + 4$ $g_{11} = x_1^2 + 2x_1 + 2$	$\{x_1^2 + 2x_1 + 2, x_2 + 4x_1 + 4\}$
3	1	$\tilde{g}_2 = x_2 + x_1^2 + x_1 + 1$ $g_{11} = x_1^3 + 2x_1^2 + 2x_1$	$\{x_1^3 + 2x_1^2 + 2x_1, x_2 + x_1^2 + x_1 + 1\}$

Veremos agora como o Algoritmo 4.3.1 pode ser aplicado para resolver um problema de interpolação. Dados pontos $P_1, \ldots, P_m \in K^n$ e $r_1, \ldots, r_m \in K$, procuramos o "menor"polinômio $f \in K[x_1, \ldots, x_n]$ tal que

$$f(P_i) = r_i,\ 1 \leq i \leq m. \tag{4.6}$$

O próximo resultado mostra como o polinômio interpolador pode ser encontrado simplesmente computando a base de Gröbner reduzida do ideal dos pontos aumentados $(P_1, r_1), \ldots, (P_m, r_m)$.

Proposição 4.3.4. *Dada uma ordem monomial em $K[x_1, \ldots, x_n]$, seja G a base de Gröbner reduzida de $I = I(P_1, \ldots, P_m)$ e $\mathcal{B}(I) = \{x^{\alpha_1}, \ldots, x^{\alpha_m}\}$ a base monomial correspondente.*

(i) Para quaisquer $r_1, \ldots, r_m \in K$, existe um único polinômio $f = \sum_{i=1}^{m} a_i x^{\alpha_i}$, onde $a_i \in K$, que satisfaz (4.6);

(ii) Defina uma nova variável z e considere a ordem monomial em $K[x_1, \ldots, x_n, z]$ que estende a ordem monomial dada em $K[x_1, \ldots, x_n]$ e é tal que z é maior que todos os monômios em $x_1, \ldots, x_n$. Então a base de Gröbner reduzida do ideal $I((P_1, r_1), \ldots, (P_m, r_m))$ é $G \cup \{z - f\}$, onde f é o polinômio em (i).

Demonstração. (i) Na demonstração da Proposição 4.3.1 vimos como encontar um polinômio $g \in K[x_1, \ldots, x_n]$ tal que $g(P_i) = r_i$ para $i = 1, \ldots, m$. Para obter um polinômio f da forma desejada, basta tomar $f = N_G(g)$.

(ii) Note que $z - f \in I((P_1, r_1), \ldots, (P_m, r_m))$, e pela ordem monomial considerada temos $tl(z-f) = z$. Como os anéis $K[x_1, \ldots, x_n, z]/I((P_1, r_1), \ldots, (P_m, r_m))$ e $K[x_1, \ldots, x_n]/I$ são isomorfos, temos que $\mathcal{B}(I((P_1, r_1), \ldots, (P_m, r_m))) = \mathcal{B}(I)$, e portanto $G \cup \{z - f\}$ é a base de Gröbner reduzida de $I((P_1, r_1), \ldots, (P_m, r_m))$. □

Em [8] é feita uma comparação entre o Algoritmo 4.3.1 e dois outros algoritmos existentes para a obtenção de bases de Gröbner deste tipo de ideais. Esta comparação, baseada em experimentos computacionais, indica que o Algoritmo 4.3.1 tem melhor desempenho que os demais quando o número de variáveis é pequeno em relação ao número de pontos. Este é o caso das aplicações em teoria de códigos, em que normalmente tem-se $n \leq 3$.

5 APROXIMAÇÕES DE PADÉ GENERALIZADAS

Neste capítulo, uma generalização da teoria das bases de Gröbner é feita para o caso de módulos livres de posto finito sobre o anel de polinômios $K[x_1, \ldots, x_n]$. Em seguida apresentamos uma aplicação da técnica de bases de Gröbner para encontrar aproximações de Padé no caso multivariado, utilizando os resultados do capítulo anterior e a teoria de bases de Gröbner desenvolvida para submódulos. Os resultados deste capítulo serão aplicados adiante na construção e decodificação de códigos lineares.

5.1 Módulos

Primeiramente, revisaremos alguns conceitos e resultados da teoria de módulos que serão necessários.

Sejam A um anel comutativo e $(M, +)$ um grupo abeliano. Dizemos que M é um A-*módulo* se existe uma operação binária (multiplicação por escalar)$A \times M \longrightarrow M$, $(a, \mathbf{m}) \mapsto a\mathbf{m}$, que satisfaz

(i) $a(\mathbf{m} + \mathbf{n}) = a\mathbf{m} + a\mathbf{n}$, para todo $a \in A$ e $\mathbf{m}, \mathbf{n} \in M$

(ii) $(a + b)\mathbf{m} = a\mathbf{m} + b\mathbf{m}$, para todo $a, b \in A$ e $\mathbf{m} \in M$

(iii) $a(b\mathbf{m}) = (ab)\mathbf{m}$, para todo $a, b \in A$ e $\mathbf{m} \in M$

(iv) $1\mathbf{m} = \mathbf{m}$, para todo $\mathbf{m} \in M$

Um *submódulo* de um A-módulo M é um subconjunto de M que é um A-módulo. Para $\mathbf{m}_1, \ldots, \mathbf{m}_s \in M$,

$$N = \{a_1\mathbf{m}_1 + \cdots + a_s\mathbf{m}_s : a_1, \ldots, a_s \in A\} \subseteq M$$

é um submódulo de M, chamado *submódulo gerado* por $\mathbf{m}_1, \ldots, \mathbf{m}_s$, denotado por $\langle \mathbf{m}_1, \ldots, \mathbf{m}_s \rangle$.

Dizemos que M é um A-módulo *livre* se M possui uma base, ou seja, um conjunto de geradores linearmente independentes. M é dito um A-módulo livre de *posto* r se r é o número de elementos da base. Assim, existem $\mathbf{m}_1, \ldots, \mathbf{m}_r \in M$ tais que

$$M = A\mathbf{m}_1 + \cdots + A\mathbf{m}_r$$

e todo elemento $\mathbf{m} \in M$ pode ser escrito de forma única como

$$\mathbf{m} = a_1\mathbf{m}_1 + \cdots + a_r\mathbf{m}_r$$

com $a_1, \ldots, a_r \in A$. Dizemos simplesmente que M é um A-módulo livre de *posto finito* se M tem base finita.

Para dois A-módulos M e M', uma função $\phi : M \longrightarrow M'$ é um *homomorfismo de A-módulos* se

$$\begin{aligned} \phi(\mathbf{m}_1 + \mathbf{m}_2) &= \phi(\mathbf{m}_1) + \phi(\mathbf{m}_2) \\ \phi(a\mathbf{m}_1) &= a\phi(\mathbf{m}_1) \end{aligned}$$

para todo $a \in A$ e $\mathbf{m}_1, \mathbf{m}_2 \in M$. Se ϕ é uma bijeção, então é dito um *isomorfismo*, e nesse caso escrevemos $M \cong M'$.

O produto cartesiano $A^r = \{(a_1, \ldots, a_r) | a_i \in A, i = 1, \ldots, r\}$ é um A-módulo livre de posto r. O conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$, onde $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)$ é o vetor que tem 1 na i-ésima coordenada e 0 nas demais, é chamado de *base canônica* de A^r. Se M é um A-módulo livre de posto r, então $M \cong A^r$.

Proposição 5.1.1. *Se A é um anel Noetheriano, então todo submódulo de A^r é finitamente gerado.*

Demonstração. Seja M um submódulo de A^r. Procederemos por indução em r. Se $r = 1$, então M é um ideal de A, e pelo Teorema da Base de Hilbert (Teorema 3.1.1) é finitamente gerado.

Para $r > 1$, definimos o conjunto

$$I = \{a \in A : a \text{ é a primeira coordenada de um elemento de } M\}.$$

É fácil ver que I é um ideal de A, e pelo Teorema da Base de Hilbert, I é finitamente gerado, digamos $I = \langle a_1, \ldots, a_t \rangle$. Sejam $\mathbf{m}_1, \ldots, \mathbf{m}_t \in M$ tais que a primeira coordenada de $\mathbf{m}_i$ é a_i para $i = 1, \ldots, t$. Considere $M' = \{(b_2, \ldots, b_r) : (0, b_2, \ldots, b_r) \in M\}$. M' é um submódulo de A^{r-1} e, por indução, é finitamente gerado, digamos $M' = \langle \mathbf{n}'_1, \ldots, \mathbf{n}'_\ell \rangle$. Para $i = 1, \ldots, \ell$, seja $\mathbf{n}_i$ o elemento de A^r com 0 na primeira coordenada e as coordenadas de $\mathbf{n}'_i$ nas demais $r-1$ coordenadas. Note que $\mathbf{n}_i \in M$. Afirmamos que $M = \langle \mathbf{m}_1, \ldots, \mathbf{m}_t, \mathbf{n}_1, \ldots, \mathbf{n}_\ell \rangle$.

Seja $\mathbf{m} \in M$. A primeira coordenada de $\mathbf{m}$, pode ser escrita como

$$d_1 a_1 + \cdots + d_t a_t.$$

Consideramos

$$\mathbf{m}' = \mathbf{m} - (d_1 \mathbf{m}_1 + \cdots + d_t \mathbf{m}_t);$$

então $\mathbf{m}' \in M$ e sua primeira coordenada é zero. Logo podemos escrever

$$\mathbf{m}' = c_1 \mathbf{n}_1 + \cdots + c_\ell \mathbf{n}_\ell$$

e, portanto,

$$\mathbf{m} = \mathbf{m}' + \sum_{i=1}^{t} d_i \mathbf{m}_i = \sum_{i=1}^{\ell} c_i \mathbf{n}_i + \sum_{i=1}^{t} d_i \mathbf{m}_i$$

como queríamos. □

Definição 5.1.1. Um A-módulo é dito *Noetheriano* se todo submódulo de M é finitamente gerado.

Proposição 5.1.2. *Seja M um A-módulo. Então M é Noetheriano se, e somente se, para toda cadeia ascendente $M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n \subseteq \cdots$ de submódulos de M, existe n_0 tal que $M_n = M_{n_0}$ para todo $n \geq n_0$.*

Demonstração. Sejam M um A-módulo Noetheriano e

$$M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n \subseteq \cdots$$

uma cadeia ascendente de submódulos de M. Consideremos o conjunto $N = \cup_{n=1}^{\infty} M_n$; N é um submódulo de M e, portanto, é finitamente gerado, digamos $N = \langle \mathbf{m}_1, \ldots, \mathbf{m}_s \rangle$. Como $\mathbf{m}_1, \ldots, \mathbf{m}_s \in N$, para cada $i = 1, \ldots, s$ existe n_i tal que $\mathbf{m}_i \in M_{n_i}$. Seja $n_0 = \max\{n_1, \ldots, n_s\}$; então $\mathbf{m}_i \in M_{n_0}$ para todo $i = 1, \ldots, s$, e $N = M_{n_0}$. Logo, $M_n = M_{n_0}$ para todo $n \geq n_0$.

Para a outra implicação, suponhamos que existe um submódulo N de M que não é gerado por um conjunto finito de elementos. Seja $\mathbf{m}_1 \in N$. Então existe $\mathbf{m}_2 \in N$ tal que $\mathbf{m}_2 \notin \langle \mathbf{m}_1 \rangle$, logo $\langle \mathbf{m}_1 \rangle \subsetneq \langle \mathbf{m}_1, \mathbf{m}_2 \rangle$. Continuando desta maneira, obtemos uma cadeia estritamente ascendente de submódulos de M. □

5.2 Bases de Gröbner de Submódulos

Seja $A = K[x_1, \ldots, x_n]$. Generalizaremos a teoria de bases de Gröbner para submódulos de A-módulos livres de posto finito, através de construções semelhantes àquelas que fizemos para ideais. Começaremos com a definição de monômios e ordens monomiais.

Para um A-módulo livre M de posto r, fixamos uma base $\{\mathbf{z}_1, \ldots, \mathbf{z}_r\}$ de M. Um *monômio* em M é um elemento da forma $x^\alpha \mathbf{z}_i$, onde $x^\alpha \in A$ é um monômio em A e $\mathbf{z}_i$ é um elemento da base.

Se $\mathbf{x} = x^\alpha \mathbf{z}_i$ e $\mathbf{y} = x^\beta \mathbf{z}_j$ são monômios em M, dizemos que $\mathbf{x}$ *divide* $\mathbf{y}$ se $i = j$ e $x^\alpha | x^\beta$. Neste caso, existe um monômio $x^\gamma \in A$ tal que $x^\beta \mathbf{z}_j = x^\gamma (x^\alpha \mathbf{z}_i)$, e então

definimos

$$\frac{\mathbf{y}}{\mathbf{x}} = \frac{x^\beta \mathbf{z}_i}{x^\alpha \mathbf{z}_i} = \frac{x^\beta}{x^\alpha} = x^\gamma.$$

Semelhantemente, um *termo* em M tem a forma $c\mathbf{x}$, onde $c \in K$ e $\mathbf{x} \in M$ é um monômio. Se $\mathbf{x} = cx^\alpha \mathbf{z}_i$ e $\mathbf{y} = dx^\beta \mathbf{z}_j$ são termos em M, dizemos que $\mathbf{x}$ *divide* $\mathbf{y}$ se $i = j$ e $x^\alpha | x^\beta$, e escrevemos

$$\frac{\mathbf{y}}{\mathbf{x}} = \frac{dx^\beta}{cx^\alpha}.$$

Definimos ordens monomiais em M de forma análoga ao caso do anel de polinômios.

Definição 5.2.1. Uma ordem $\succ$ no conjunto dos monômios de M é uma *ordem monomial* se satisfaz as seguintes condições:

(i) $\succ$ é uma ordem total, isto é, dados monômios $\mathbf{x} \neq \mathbf{y} \in M$, ou $\mathbf{x} \succ \mathbf{y}$ ou $\mathbf{y} \succ \mathbf{x}$;

(ii) Se $\mathbf{x} \succ \mathbf{y}$, então $x^\alpha \mathbf{x} \succ x^\alpha \mathbf{y}$, para quaisquer monômios $\mathbf{x}, \mathbf{y} \in M$ e $x^\alpha \in A$;

(iii) $\mathbf{x} \prec x^\alpha \mathbf{x}$, para todo monômio $\mathbf{x} \in M$ e todo $x^\alpha \in A$ com $x^\alpha \neq 1$.

Pode-se mostrar que toda ordem monomial em M é uma boa ordem.

Fixada uma ordem monomial $\succ$ em M, para todo $\mathbf{f} \in M$, $\mathbf{f} \neq 0$, podemos escrever

$$\mathbf{f} = a_1 \mathbf{x}_1 + \cdots + a_m \mathbf{x}_m$$

onde $a_i \in K \backslash \{0\}$ e $\mathbf{x}_i$ são monômios em M tais que $\mathbf{x}_1 \succ \mathbf{x}_2 \succ \cdots \succ \mathbf{x}_m$, e definimos:

(i) o *monômio líder* de $\mathbf{f}$, $ml(\mathbf{f}) = \mathbf{x}_1$;

(ii) o *termo líder* de $\mathbf{f}$, $tl(\mathbf{f}) = a_1 \mathbf{x}_1$;

(iii) o *coeficiente líder* de $\mathbf{f}$, $cl(\mathbf{f}) = a_1$.

Dada uma ordem monomial $\succ_A$ em $A = K[x_1, \ldots, x_n]$, há duas formas naturais de obter ordens monomiais em M, que são frequentemente utilizadas. A primeira delas é comparar os dois monômios utilizando a ordem monomial em A, e em caso de empate, comparar os elementos da base que aparecem nos monômios. Assim, temos, por exemplo

$$x^\alpha \mathbf{z}_i \prec x^\beta \mathbf{z}_j \iff \begin{cases} x^\alpha \prec_A x^\beta \\ \text{ ou} \\ x^\alpha = x^\beta \text{ e } i < j \end{cases}$$

A segunda forma de obter uma ordem monomial em M é comparar primeiro os elementos da base, e desempatar utilizando a ordem monomial em A, como por exemplo

$$x^\alpha \mathbf{z}_i \prec x^\beta \mathbf{z}_j \iff \begin{cases} i < j \\ \text{ ou} \\ i = j \text{ e } x^\alpha \prec_A x^\beta \end{cases}$$

As duas formas acima não são as únicas maneiras de definir ordens monomiais em módulos. Tal como no caso de ideais, podemos definir uma ordem monomial sobre os monômios de um módulo através de uma matriz. Dada uma matriz $T = (W|U)$ de ordem $\ell \times (n + r)$ com entradas reais, definimos uma relação de ordem $\succ$ sobre os monômios de M da seguinte forma: seja (w_i, u_i) a i-ésima linha de T, onde w_i e u_i são vetores-linha com n e r componentes, respectivamente; então $x^\alpha \mathbf{z}_i \succ x^\beta \mathbf{z}_j$ se e somente se existe $k \in \{1, \ldots, \ell\}$ tal que $(w_k, u_k) \cdot (\alpha, \mathbf{e}_i) > (w_k, u_k) \cdot (\beta, \mathbf{e}_j)$ e $(w_m, u_m) \cdot (\alpha, \mathbf{e}_i) = (w_m, u_m) \cdot (\beta, \mathbf{e}_j)$, para todo $m < k$, onde $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ é a base canônica de A^r.

Dando continuidade à construção de bases de Gröbner para módulos, passaremos agora ao algoritmo da divisão.

Definição 5.2.2. Sejam $\mathbf{g} \in M$ e $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_s\}$ um conjunto de elementos não-nulos de M. Dizemos que $\mathbf{g}$ é *reduzido módulo* F se $\mathbf{g} = \mathbf{0}$ ou nenhum monômio que aparece em $\mathbf{g}$ é divisível por $ml(\mathbf{f}_i)$, para $i = 1, \ldots, s$.

O algoritmo da divisão para módulos é muito semelhante ao Algoritmo 2.2.1. Dados $\mathbf{f}, \mathbf{f}_1, \ldots, \mathbf{f}_s \in M$, o algoritmo retorna quocientes $a_1, \ldots, a_s \in A$ e um resto $\mathbf{g} \in M$ reduzido módulo $\mathbf{f}_1, \ldots, \mathbf{f}_s$.

Algoritmo 5.2.1 Algoritmo da Divisão para Módulos

Entrada: $\mathbf{f}, \mathbf{f}_1, \ldots, \mathbf{f}_s \in M$ com $\mathbf{f}_i \neq \mathbf{0}$
Saída: $a_1, \ldots, a_s \in A, \mathbf{g} \in M$
$a_1 \leftarrow 0;\ a_2 \leftarrow 0;\ \ldots;\ a_s \leftarrow 0;\ \mathbf{g} \leftarrow \mathbf{0}$
$\mathbf{p} \leftarrow \mathbf{f}$
while $\mathbf{p} \neq \mathbf{0}$ **do**
  $i \leftarrow 1$
  $d \leftarrow$ false
  while $i \leq s$ **and** $d =$ false **do**
    if $tl(\mathbf{f}_i)$ divide $tl(\mathbf{p})$ **then**
      $a_i \leftarrow a_i + \dfrac{tl(\mathbf{p})}{tl(\mathbf{f}_i)}$
      $\mathbf{p} \leftarrow \mathbf{p} - \dfrac{tl(\mathbf{p})}{tl(\mathbf{f}_i)}\mathbf{f}_i$
      $d \leftarrow$ true
    else
      $i \leftarrow i + 1$
    end if
  end while
  if $d =$ false **then**
    $\mathbf{g} \leftarrow \mathbf{g} + tl(\mathbf{p})$
    $\mathbf{p} \leftarrow \mathbf{p} - tl(\mathbf{p})$
  end if
end while

Proposição 5.2.1. *Dados* $\mathbf{f}, \mathbf{f}_1, \ldots, \mathbf{f}_s \in M$, *com* $\mathbf{f}_i \neq \mathbf{0}$ *para* $i = 1, \ldots, s$, *o Algoritmo 5.2.1 produz* $a_1, \ldots, a_s \in A$ *e* $\mathbf{g} \in M$ *tais que*

$$\mathbf{f} = a_1\mathbf{f}_1 + \cdots + a_s\mathbf{f}_s + \mathbf{g}, \tag{5.1}$$

com $\mathbf{g}$ *reduzido módulo* $\mathbf{f}_1, \ldots, \mathbf{f}_s$ *e* $ml(a_i)ml(\mathbf{f}_i) \preceq ml(\mathbf{f})$, $1 \leq i \leq s$.

A prova é exatamente a mesma da Proposição 2.2.1. O resto $\mathbf{g}$ neste caso também será denotado por $\overline{\mathbf{f}}^F$.

Para um conjunto $S \subset M$, definimos o *submódulo dos termos líderes* de S como sendo o submódulo de M

$$\langle tl(S)\rangle = \langle tl(\mathbf{f}) : \mathbf{f} \in S\rangle.$$

Podemos então definir bases de Gröbner em módulos.

Definição 5.2.3. Dado um submódulo N de M, dizemos que um conjunto finito G é uma *base de Gröbner* de N se $\langle tl(G)\rangle = \langle tl(N)\rangle$. Dizemos simplesmente que G é uma base de Gröbner se G é uma base de Gröbner do submódulo gerado por G.

Seguem resultados análogos aos que temos para ideais:

(i) Se G é base de Gröbner de um submódulo $N \subseteq M$, então $N = \langle G\rangle$.

(ii) Todo submódulo não-nulo de M tem uma base de Gröbner.

(iii) Se G é base de Gröbner, então o resto da divisão de um elemento $\mathbf{f}$ por G é único.

(iv) Dados um elemento $\mathbf{f} \in M$ e uma base de Gröbner G, $\mathbf{f} \in \langle G\rangle$ se, e somente se, o resto da divisão de $\mathbf{f}$ por G é zero.

Queremos agora generalizar a noção de S-polinômio; para isso, precisamos definir o mínimo múltiplo comum de dois monômios em M. Dados $x^\alpha \mathbf{z}_i, x^\beta \mathbf{z}_j$ monômios em M, definimos

$$mmc(x^\alpha \mathbf{z}_i, x^\beta \mathbf{z}_j) = \begin{cases} \mathbf{0}, \text{ se } i \neq j \\ mmc(x^\alpha, x^\beta)\mathbf{z}_i, \text{ se } i = j. \end{cases}$$

Dados dois elementos não-nulos $\mathbf{f}, \mathbf{g} \in M$, definimos o *S-polinômio* de $\mathbf{f}$ e $\mathbf{g}$ por

$$spol(\mathbf{f}, \mathbf{g}) = \frac{mmc(ml(\mathbf{f}), ml(\mathbf{g}))}{tl(\mathbf{f})}\mathbf{f} - \frac{mmc(ml(\mathbf{f}), ml(\mathbf{g}))}{tl(\mathbf{g})}\mathbf{g}$$

Continuaremos usando o termo "S-polinômio", embora neste caso $spol(\mathbf{f}, \mathbf{g})$ seja um elemento do módulo M, e não do anel de polinômios. Obtemos um resultado semelhante ao caso polinomial.

Proposição 5.2.2. *Seja G um conjunto finito de elementos não-nulos de M. Então G é uma base de Gröbner se, e somente se, $\overline{spol(\mathbf{g}, \mathbf{h})}^{G} = \mathbf{0}$ para todo $\mathbf{g}, \mathbf{h} \in G$.*

A prova é análoga àquela do Teorema 3.2.1. Através desta Proposição, obtemos um algoritmo semelhante ao Algoritmo de Buchberger (Algoritmo 3.2.1) para calcular bases de Gröbner de submódulos a partir de uma base qualquer.

Algoritmo 5.2.2 Bases de Gröbner de Submódulos

Entrada: $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_s\} \subset M \backslash \{\mathbf{0}\}$.
Saída: $G = \{\mathbf{g}_1, \ldots, \mathbf{g}_\ell\}$ uma base de Gröbner do submódulo $\langle F \rangle$.
$G \leftarrow F$
$B \leftarrow \{\{\mathbf{p}, \mathbf{q}\} | \mathbf{p}, \mathbf{q} \in G, \mathbf{p} \neq \mathbf{q}\}$
while $B \neq \emptyset$ **do**
  selecione $\{\mathbf{p}, \mathbf{q}\} \in B$
  $B \leftarrow B \backslash \{\{\mathbf{p}, \mathbf{q}\}\}$
  $\mathbf{h} \leftarrow spol(\mathbf{p}, \mathbf{q})$
  $\mathbf{h} \leftarrow$ resto da divisão de $\mathbf{h}$ por G
  if $\mathbf{h} \neq \mathbf{0}$ **then**
    $B \leftarrow B \cup \{\{\mathbf{g}, \mathbf{h}\} | \mathbf{g} \in G\}$
    $G \leftarrow G \cup \{\mathbf{h}\}$
  end if
end while

Podemos também definir bases de Gröbner reduzidas.

Definição 5.2.4. Uma base de Gröbner G é dita *reduzida* se, para todo $\mathbf{g} \in G$, $\mathbf{g}$ é reduzido módulo $G \backslash \{\mathbf{g}\}$ e $cl(\mathbf{g}) = 1$.

E neste caso continua valendo o seguinte resultado.

Proposição 5.2.3. *Dada uma ordem monomial em M, todo submódulo de M possui uma única base de Gröbner reduzida com respeito a esta ordem monomial.*

A seguir, faremos uso das bases de Gröbner de submódulos para solucionar problemas análogos aos do capítulo anterior.

Dado um submódulo N do A-módulo M, definimos

$$M/N = \{\mathbf{f} + N : \mathbf{f} \in M\}.$$

M/N é o grupo quociente com a soma usual de classes. Definindo

$$a \cdot (\mathbf{f} + N) = a\mathbf{f} + N$$

para todo $a \in A$ e $\mathbf{f} \in M$, M/N tem a estrutura de A-módulo, e é chamado de módulo quociente de M por N. Além da estrutura de A-módulo, M/N também tem a estrutura de K-espaço vetorial. Vamos determinar:

(a) um conjunto de representantes das classes de M/N;

(b) uma base do K-espaço vetorial M/N.

Se G é uma base de Gröbner do submódulo N, então para cada $\mathbf{f} \in M$, o resto da divisão de $\mathbf{f}$ por G é único. Assim como no caso de ideais, chamaremos o resto de $\mathbf{f}$ na divisão por G de *forma normal* de $\mathbf{f}$ módulo G e denotaremos por $N_G(\mathbf{f})$.

Proposição 5.2.4. *Sejam $\mathbf{f}, \mathbf{g} \in M$, N um submódulo de M e G uma base de gröbner de N. Então $\mathbf{f} + N = \mathbf{g} + N$ em M/N se, e somente se, $N_G(\mathbf{f}) = N_G(\mathbf{g})$. Portanto, $\{N_G(\mathbf{f}) : \mathbf{f} \in M\}$ é um conjunto de representantes das classes de M/N.*

Definindo, para cada $S \subseteq M$, o conjunto

$$\mathcal{B}(S) = \{\text{monômios } \mathbf{x} \in M : ml(\mathbf{g}) \text{ não divide } \mathbf{x}, \text{ para qualquer } \mathbf{g} \in S\}$$

temos resultados análogos aos que tínhamos no caso de ideais.

Proposição 5.2.5. *Seja N um submódulo de M. Então uma base do K-espaço vetorial M/N consiste de todas as classes dos monômios em $\mathcal{B}(N)$.*

Proposição 5.2.6. *Sejam N um submódulo de M e $G \subset N$. Então $\mathcal{B}(N) \subseteq \mathcal{B}(G)$, e $\mathcal{B}(G) = \mathcal{B}(N)$ se, e somente se, G é uma base de Gröbner de N.*

5.3 Aproximações de Padé Generalizadas

Nesta seção apresentamos o método descrito por Farr e Gao em [9] para encontrar aproximações de Padé empregando a teoria de bases de Gröbner, que constitui uma parte importante do algoritmo de decodificação dos códigos lineares abordados no próximo capítulo.

Continuaremos denotando $A = K[x_1, \ldots, x_n]$. Dados um polinômio $f \in A$ e um ideal $I \subset A$, queremos encontrar $a, b \in A$ tais que

$$b \cdot f \equiv a \pmod{I}. \tag{5.2}$$

O *grau da aproximação* é dado pelo número total de coeficientes em a e b. Estamos interessados em encontrar soluções de (5.2) no caso em que I é um ideal zero dimensional, e o grau da aproximação é o grau de I. Uma abordagem possível é considerar (5.2) como um sistema linear homogêneo, onde os coeficientes de a e b são incógnitas, e resolver o sistema através de técnicas da álgebra linear. Abordaremos o problema sob o ponto de vista das bases de Gröbner.

Considere o conjunto de todas as soluções

$$M_f = \{(a, b) \in A^2 : a \text{ e } b \text{ satisfazem } (5.2)\}.$$

É fácil ver que M_f é um A-módulo. Na verdade, M_f é um submódulo de A^2. Mostraremos que uma solução (a, b) de (5.2) está contida na base de Gröbner reduzida do submódulo M_f com respeito a uma determinada ordem monomial.

A função

$$\begin{array}{rcl} \phi : A^2 & \rightarrow & A + \mathbf{z} \cdot A \\ (a, b) & \mapsto & \mathbf{z} \cdot b - a \end{array} \tag{5.3}$$

é um isomorfismo de A-módulos. Portanto, $M = A + \mathbf{z} \cdot A$ é um A-módulo livre de posto 2, com base $\{\mathbf{1}, \mathbf{z}\}$, onde $\mathbf{1} = 1 + \mathbf{z} \cdot 0$. A imagem de M_f, que continuaremos a denotar por M_f, é o submódulo de M

$$M_f = \{\mathbf{z} \cdot b - a : f \cdot b - a \in I\}. \tag{5.4}$$

Os elementos de A podem ser vistos como elementos do módulo M através da inclusão

$$\begin{array}{rcl} A & \hookrightarrow & A + \mathbf{z} \cdot A \\ g & \mapsto & g + \mathbf{z} \cdot 0 \end{array} \tag{5.5}$$

Assim, temos $A \subset M$.

Lema 5.3.1. *Seja $\succ_W$ uma ordem monomial em A definida por uma matriz $W \in \mathbb{R}^{\ell \times m}$, e sejam $x^{\alpha_1}, x^{\alpha_2}$ monômios em A. Então podemos definir uma ordem monomial $\succ$ em $M = A + \mathbf{z} \cdot A$ de forma que $x^{\alpha_1} \prec x^{\alpha_2} \cdot \mathbf{z}$ e $x^{\alpha_1}, x^{\alpha_2} \cdot \mathbf{z}$ são consecutivos, ou seja, não existe nenhum monômio em M entre eles.*

Demonstração. Seja w_i a i-ésima linha de W, e $c_i = w_i \cdot \alpha_1 - w_i \cdot \alpha_2$, para $1 \leq i \leq \ell$. Definimos a matriz

$$T = \begin{pmatrix} & & & 0 & c_1 \\ & W & & \vdots & \vdots \\ & & & 0 & c_\ell \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

e consideramos a ordem monoimial $\succ$ definida por T. Para $1 \leq i \leq \ell$ temos

$$(w_i, 0, c_i) \cdot (\alpha_1, 1, 0) = w_i \cdot \alpha_1 + 0 + 0 = w_i \cdot \alpha_1 \tag{5.6}$$

e

$$\begin{array}{rcll} (w_i, 0, c_i) \cdot (\alpha_2, 0, 1) & = & w_1 \cdot \alpha_2 + 0 + c_i & \\ & = & w_i \cdot \alpha_2 + w_i \cdot \alpha_1 - w_i \cdot \alpha_2 & \\ & = & w_i \cdot \alpha_1 & (5.7) \end{array}$$

e para a última linha de T temos

$$(0, \dots, 0, 1) \cdot (\alpha_1, 1, 0) = 0$$
$$(0, \dots, 0, 1) \cdot (\alpha_2, 0, 1) = 1$$

logo $x^{\alpha_1} \prec x^{\alpha_2} \cdot \mathbf{z}$.

Resta mostrar que não há nenhum monômio em M entre x^{α_1} e $x^{\alpha_2} \cdot \mathbf{z}$. Os monômios em M são da forma x^β ou $x^\beta \cdot \mathbf{z}$. Mostraremos primeiramente, que não há monômios da forma x^β entre x^{α_1} e $x^{\alpha_2} \cdot \mathbf{z}$. Suponhamos que $x^\beta \succ x^{\alpha_1}$; queremos provar que nesse caso temos também $x^\beta \succ x^{\alpha_2} \cdot \mathbf{z}$. Se $x^\beta \succ x^{\alpha_1}$, então como

$$(0, \ldots, 0, 1) \cdot (\beta, 1, 0) = 0 \quad \text{e} \quad (0, \ldots, 0, 1) \cdot (\alpha_1, 1, 0) = 0$$

temos que existe $j \in \{1, \ldots, \ell\}$ tal que $(w_j, 0, c_j) \cdot (\beta, 1, 0) > (w_j, 0, c_j) \cdot (\alpha_1, 1, 0)$ e $(w_i, 0, c_i) \cdot (\beta, 1, 0) = (w_i, 0, c_i) \cdot (\alpha_1, 1, 0)$ para $i < j$. Das equações (5.6) e (5.7) segue que

$$(w_i, 0, c_i) \cdot (\alpha_1, 1, 0) = (w_i, 0, c_i) \cdot (\alpha_2, 0, 1) \tag{5.8}$$

para $1 \leq i \leq \ell$. Logo,

$$(w_j, 0, c_j) \cdot (\beta, 1, 0) > (w_j, 0, c_j) \cdot (\alpha_1, 1, 0) = (w_j, 0, c_j) \cdot (\alpha_2, 0, 1)$$

e

$$(w_i, 0, c_i) \cdot (\beta, 1, 0) = (w_i, 0, c_i) \cdot (\alpha_1, 1, 0) = (w_i, 0, c_i) \cdot (\alpha_2, 0, 1)$$

para $i < j$ e, portanto, $x^\beta \succ x^{\alpha_2} \cdot \mathbf{z}$.

Mostraremos agora que não há monômios da forma $x^\beta \cdot \mathbf{z}$ entre x^{α_1} e $x^{\alpha_2} \cdot \mathbf{z}$. Procedemos de forma análoga ao caso anterior. Se $x^\beta \cdot \mathbf{z} \prec x^{\alpha_2} \cdot \mathbf{z}$, como

$$(0, \ldots, 0, 1) \cdot (\beta, 0, 1) = 1 \quad \text{e} \quad (0, \ldots, 0, 1) \cdot (\alpha_2, 0, 1) = 1$$

existe j tal que $(w_j, 0, c_j) \cdot (\beta, 0, 1) < (w_j, 0, c_j) \cdot (\alpha_2, 0, 1)$ e $(w_i, 0, c_i) \cdot (\beta, 0, 1) = (w_i, 0, c_i) \cdot (\alpha_2, 0, 1)$ para $i < j$. Da identidade (5.8), segue que $x^\beta \cdot \mathbf{z} \prec x^{\alpha_1}$. □

Lema 5.3.2. *Seja $I \subset A$ um ideal zero dimensional de grau t, e fixemos uma ordem monomial em $M = A + \mathbf{z} \cdot A$. Então, para qualquer $f \in A$ e M_f o submódulo de M definido acima,*

(i) o módulo quociente M/M_f tem dimensão t como K-espaço vetorial;

(ii) $\{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$ *é uma base de Gröbner de* M_f *se, e somente se,* $|\mathcal{B}(\mathbf{g}_1, \ldots, \mathbf{g}_s)| = t$.

Demonstração. (i) Através da inclusão (5.5), temos $I \subset M_f$, pois para cada $g \in I$ temos

$$g = g + \mathbf{z} \cdot 0 = \mathbf{z} \cdot b - a$$

onde $a = -g$ e $b = 0$; logo $f \cdot b - a = f \cdot 0 - (-g) = g \in I$ e, portanto, $g \in M_f$. Além disso, $\mathbf{z} - f \in M_f$, pois

$$\mathbf{z} - f = \mathbf{z} \cdot b - a$$

onde $a = f$ e $b = 1$, logo $f \cdot b - a = f \cdot 1 - f = 0 \in I$.

Como a dimensão de M/M_f não depende da ordem monomial, faremos uso da ordem $\succ$ definida da seguinte forma: se $\mathbf{x} = x^\alpha \mathbf{z}_1$ e $\mathbf{y} = x^\beta \mathbf{z}_2$ são monômios em M, onde $\mathbf{z}_1, \mathbf{z}_2 \in \{\mathbf{1}, \mathbf{z}\}$, então

$$\mathbf{x} \succ \mathbf{y} \Longleftrightarrow \begin{cases} \mathbf{z}_1 = \mathbf{z} \text{ e } \mathbf{z}_2 = \mathbf{1} \\ \text{ou} \\ \mathbf{z}_1 = \mathbf{z}_2 \text{ e } x^\alpha \succ_A x^\beta \end{cases} \tag{5.9}$$

onde $\succ_A$ é uma ordem monomial em A.

Afirmamos que se $\{g_1, \ldots, g_s\}$ é uma base de Gröbner de I com respeito a $\succ_A$, então $G = \{g_1, \ldots, g_s, \mathbf{z} - f\}$ é uma base de Gröbner de M_f com respeito a $\succ$. Pelas observações feitas acima, temos $G \subset M_f$, logo $\langle tl(G) \rangle \subseteq \langle tl(M_f) \rangle$. Vejamos que $\langle tl(G) \rangle \supseteq \langle tl(M_f) \rangle$.

Dado $\mathbf{g} \in M_f$ temos duas possibilidades: ou $tl(\mathbf{g}) = x^\alpha \cdot \mathbf{z}$, ou $tl(\mathbf{g}) = x^\alpha \cdot \mathbf{1} = x^\alpha$. Se $tl(\mathbf{g}) = x^\alpha \cdot \mathbf{z}$, então $tl(\mathbf{g})$ é divisível por $tl(\mathbf{z} - f) = \mathbf{z}$ (de acordo com a ordem monomial $\succ$ definida por (5.9)). Se $tl(\mathbf{g}) = x^\alpha$, isso significa que em $\mathbf{g}$ não aparecem termos da forma $cx^\beta \cdot \mathbf{z}$, ou seja, $\mathbf{g} = 0 \cdot \mathbf{z} + a = a$. Como $\mathbf{g} = \mathbf{z} \cdot 0 - (-a) \in M_f$, segue que $f \cdot 0 - a = -a \in I$ e, portanto, $\mathbf{g} = a \in I$. Neste caso, $tl(\mathbf{g}) = tl(a)$ é

divisível por $tl(g_i)$ para algum $i \in \{1, \ldots, s\}$. Concluímos que G é base de Gröbner de M_f.

Mostraremos agora que $\mathcal{B}(G) = \mathcal{B}(g_1, \ldots, g_s)$. Por definição, temos

$$\mathcal{B}(g_1, \ldots, g_s, \mathbf{z} - f) = \{\mathbf{x} \text{ monômio em } M : tl(g_i) \nmid \mathbf{x}, \text{ para } i = 1, \ldots, s, \text{ e } \mathbf{z} \nmid \mathbf{x}\}.$$

Como $\mathbf{z}$ não divide $\mathbf{x}$, para todo $\mathbf{x} \in \mathcal{B}(G)$, então $\mathcal{B}(G) \subset A$, isto é, os monômios $\mathbf{x}$ de $\mathcal{B}(G)$ são da forma $\mathbf{x} = x^\alpha \cdot \mathbf{1} = x^\alpha$. Agora, como $tl(g_i)$ não divide x^α para todo $i = 1, \ldots, s$ e $x^\alpha \in \mathcal{B}(G)$, segue que $\mathcal{B}(G) = \mathcal{B}(g_1, \ldots, g_s)$.

Finalmente, como I tem grau t, temos $|\mathcal{B}(G)| = |\mathcal{B}(g_1, \ldots, g_s)| = t$, e uma vez que as classes dos elementos de $\mathcal{B}(G)$ formam uma K-base de M/M_f, temos $\dim_K M/M_f = t$.

(ii) Se $\{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$ é uma base de Gröbner de M_f, então $|\mathcal{B}(\mathbf{g}_1, \ldots, \mathbf{g}_s)| = |\mathcal{B}(M_f)| = \dim_K M/M_f = t$.

Reciprocamente, se $|\mathcal{B}(\mathbf{g}_1, \ldots, \mathbf{g}_s)| = t$, então como $\mathcal{B}(M_f) \subseteq \mathcal{B}(\mathbf{g}_1, \ldots, \mathbf{g}_s)$ e $|\mathcal{B}(M_f)| = t$, temos que $\mathcal{B}(\mathbf{g}_1, \ldots, \mathbf{g}_s) = \mathcal{B}(M_f)$, logo $\{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$ é base de Gröbner de M_f. □

Teorema 5.3.1. *Seja $I \subset A$ um ideal zero dimensional de grau t, e fixemos uma ordem monomial em A. Seja $\mathcal{B}(I) = \{1 = x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_t}\}$ a base monomial de I com respeito à ordem monomial fixada, listados em ordem crescente. Então para qualquer $f \in A$ e quaisquer inteiros positivos t_1 e t_2 tais que $t_1 + t_2 = t + 1$, existe um par de polinômios da forma*

$$a = \sum_{i=1}^{t_1} a_i x^{\alpha_i}, \quad b = \sum_{i=1}^{t_2} b_i x^{\alpha_i}, \tag{5.10}$$

não ambos nulos, que satisfazem (5.2). Além disso, existe um par (a, b) da forma acima contido na base de Gröbner reduzida de M_f com respeito a uma certa ordem monomial.

Demonstração. Como as classes dos elementos de $\mathcal{B}(I)$ formam uma base de A/I, temos que

$$\begin{aligned} f \cdot b - a &= f \cdot \left(\sum_{i=1}^{t_2} b_i x^{\alpha_i} \right) - \left(\sum_{i=1}^{t_1} a_i x^{\alpha_i} \right) \\ &\equiv \sum_{j=1}^{t} c_j x^{\alpha_j} \pmod{I} \end{aligned}$$

onde os coeficientes c_j são combinações lineares dos a_i e b_i. A congruência $f \cdot b - a \equiv 0 \pmod{I}$ implica que $c_j = 0$ para $j = 1, \ldots, t$, e portanto temos t equações lineares homogêneas para determinar $t_1 + t_2 = t + 1$ coeficientes a_i, b_i, ou seja, temos um sistema linear homogêneo com mais incógnitas do que equações. Isso garante a existência de uma solução não-nula.

Se o par (a, b) é uma solução de (5.2) da forma (5.10), então pela correspondência (5.3), temos pelo menos um elemento $b \cdot \mathbf{z} - a$ no módulo M_f, com a e b da forma (5.10).

Suponhamos que a ordem monomial $\succ_W$ em A é definida por uma matriz W. Então pelo Lema 5.2.1, podemos definir uma ordem monomial $\succ_T$ em M de forma que $x^{\alpha_{t_1}} \prec_T x^{\alpha_{t_2}} \cdot \mathbf{z}$ e estes monômios são consecutivos. Como $x^{\alpha_1} \prec_W x^{\alpha_2} \prec_W \cdots \prec_W x^{\alpha_t}$, pela forma como é definida a ordem $\succ_T$, temos

$$x^{\alpha_1} \prec_T x^{\alpha_2} \prec_T \cdots \prec_T x^{\alpha_t}$$

e

$$x^{\alpha_1} \cdot \mathbf{z} \prec_T x^{\alpha_2} \cdot \mathbf{z} \prec_T \cdots \prec_T x^{\alpha_t} \cdot \mathbf{z}.$$

Além disso, como $x^{\alpha_{t_1}} \prec_T x^{\alpha_{t_2}} \cdot \mathbf{z}$, temos $x^{\alpha_i} \prec_T x^{\alpha_{t_2}} \cdot \mathbf{z}$ para $i = 1, \ldots, t_1$.

Seja $b \cdot \mathbf{z} - a \in M_f$. Se $b \cdot \mathbf{z} - a$ não é da forma (5.10), então $tl(a) \succ_T x^{\alpha_{t_1}}$ ou $tl(b) \succ_T x^{\alpha_{t_2}}$. Se $tl(a) \succ_T x^{\alpha_{t_1}}$, como $x^{\alpha_{t_1}}$ e $x^{\alpha_{t_2}} \cdot \mathbf{z}$ são consecutivos, temos que $tl(a) \succ_T x^{\alpha_{t_2}} \cdot \mathbf{z}$. Se $tl(b) \succ_T x^{\alpha_{t_2}}$, então $tl(b) \cdot \mathbf{z} \succ_T x^{\alpha_{t_2}} \cdot \mathbf{z}$. Agora, isso implica que

$$tl(b \cdot \mathbf{z} - a) = \max\{tl(b \cdot \mathbf{z}), tl(a)\} = \max\{tl(b) \cdot \mathbf{z}, tl(a)\} \succ_T x^{\alpha_{t_2}} \cdot \mathbf{z}.$$

Portanto, se não há solução da forma (5.10) em M_f, temos $tl(\mathbf{g}) \succ_T x^{\alpha_{t_2}} \cdot \mathbf{z}$ para todo $\mathbf{g} \in M_f$. Então se G é uma base de Gröbner de M_f com respeito a $\succ_T$, os monômios $x^{\alpha_1}, \ldots, x^{\alpha_{t_1}}, x^{\alpha_1} \cdot \mathbf{z}, \ldots, x^{\alpha_{t_2}} \cdot \mathbf{z}$ não são divisíveis por nenhum $tl(\mathbf{g})$, para $\mathbf{g} \in M_f$, e portanto

$$\{x^{\alpha_1}, \ldots, x^{\alpha_{t_1}}, x^{\alpha_1} \cdot \mathbf{z}, \ldots, x^{\alpha_{t_2}} \cdot \mathbf{z}\} \subseteq \mathcal{B}(G)$$

o que nos dá $|\mathcal{B}(G)| \geq t_1 + t_2 = t + 1$, contradizendo o Lema 5.3.2. □

O Teorema 5.3.1 garante que pelo menos uma solução está contida na base de Gröbner de M_f com respeito a uma certa ordem monomial. Uma forma de obter esta base é sugerida em [9]: dada uma base de Gröbner G do ideal I, conforme vimos na demonstração do Lema 5.3.2 o conjunto $G \cup \{\mathbf{z} - f\}$ é uma base de Gröbner de M_f com respeito a uma determinada ordem monomial. A seguir uma base de Gröbner de M_f com respeito à ordem monomial a que se refere o Teorema 5.3.1 pode ser obtida através de um algoritmo de conversão de bases de Gröbner. Tais algoritmos, como o "FGLM"[10] por exemplo, transformam uma base de Gröbner com respeito a qualquer ordem monomial em uma base de Gröbner com respeito a qualquer outra ordem. Em geral, é mais vantajoso computacionalmente aplicar algoritmos de conversão ao invés de empregar o Algoritmo 5.2.2. De acordo com [9], este método para o cálculo de aproximações de Padé é mais eficiente do que a abordagem baseada na álgebra linear quando o número de variáveis é pequeno em relação ao grau da aproximação.

6 CONSTRUÇÃO E DECODIFICAÇÃO DE CÓDIGOS LINEARES VIA BASES DE GRÖBNER

Neste capítulo apresentamos o método proposto por Farr e Gao em [7] para a construção e decodificação de códigos lineares. Uma breve introdução aos códigos lineares é feita no início deste capítulo, baseada nos livros de Pless [14], Lint [13] e Pretzel [15].

6.1 Códigos Lineares

A Teoria de Códigos estuda técnicas que visam detectar e corrigir erros que ocorrem na transmissão de informação através de instrumentos eletrônicos. Assim, temos a seguinte situação: uma mensagem é transmitida através de um canal com ruído. A mensagem distorcida é recebida e então processada, de forma a recuperar a mensagem original com a maior precisão possível.

A mensagem transmitida é composta de *símbolos* ou *caracteres* de um certo conjunto finito que chamaremos de *alfabeto*. A mensagem é *codificada* a fim de tornar possível detecter, e talvez também corrigir, quaisquer erros que possam ser introduzidos pelo canal ruidoso. Apresentaremos agora alguns conceitos básicos da Teoria de Códigos.

Definição 6.1.1. Seja A um alfabeto com q símbolos. Uma *A-palavra* de comprimento m é uma sequência de m símbolos de A. As palavras são denotadas na forma vetorial como $(a_1, \ldots, a_m)$, ou ainda como um bloco de símbolos $a_1 \cdots a_m$. O conjunto das A-palavras de comprimento m é denotado por A^m.

Um (m, k)-*código* C sobre A é um conjunto de q^k *palavras-código* em A^m. Um *esquema de codificação* é uma função bijetora $f : A^k \longrightarrow C$ que leva qualquer

A-palavra x de comprimento k em uma palavra-código $u = f(x)$. O número m é chamado de *comprimento* do código, e k é a *dimensão* do código.

O alfabeto é em geral um corpo finito. Assim, se A é o corpo finito $(\mathbb{F}, +, \cdot)$, duas operações podem ser definidas sobre as palavras em A^m. Se $u = (u_1, \ldots, u_m)$ e $v = (v_1, \ldots, v_m)$ são palavras e $a \in A$, definimos

$$\begin{aligned} u + v &= (u_1 + v_1, \ldots, u_m + v_m) \\ au &= (au_1, \ldots, au_m) \end{aligned}$$

Munido destas operações, A^m é um espaço vetorial sobre A. Em princípio, o conjunto das palavras-código pode ser um subconjunto qualquer de A^m. Contudo, restringiremos nossa atenção a uma classe especial de códigos que possuem uma estrutura adicional, conveniente para o processo de codificação e decodificação.

Definição 6.1.2. Seja $A = \mathbb{F}_q$ um alfabeto, onde $\mathbb{F}_q$ é um corpo finito com q elementos. Um código C de comprimento m é dito um *código linear* se C é um subespaço de A^m.

Para detectar ou corrigir erros de códigos, precisamos de uma noção de distância entre as palavras.

Definição 6.1.3. Sejam u, v palavras em A^m. A *distância de Hamming* entre u e v, denotada por $d(u, v)$, é o número de coordenadas nas quais u e v diferem. O *peso de Hamming* de u, denotado por $\omega(u)$, é o número de coordenadas não-nulas de u.

A distância de Hamming é uma métrica, isto é, para quaisquer $u, v, w \in A^m$, d satisfaz as seguintes propriedades:

1. $d(u, v) \geq 0$
2. $d(u, v) = 0$ se e somente se $u = v$

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

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

A distância e o peso de Hamming estão relacionados pela identidade

$$d(u,v) = \omega(u-v).$$

Quando ocorrem erros na transmissão de uma mensagem, o receptor lê uma palavra v, embora tenha sido enviada uma palavra u.

Definição 6.1.4. Se u é uma palavra-código e v é uma palavra recebida, o *erro* é a diferença $e = v - u$.

Se o erro e tem peso de Hamming t, ou seja, se a palavra recebida difere da palavra transmitida em t coordenadas, então dizemos que um erro de peso t ocorreu, ou ainda que t erros ocorreram.

Definição 6.1.5. Seja t um inteiro positivo. Dizemos que um código $C \subset \mathbb{F}_q^m$ *corrige t erros* se, para cada palavra recebida $v \in \mathbb{F}_q^m$, há no máximo um vetor $u \in C$ tal que $d(u,v) \leq t$.

A correção do erro é feita por verossimilhança, que consiste em procurar a palavra-código mais próxima da palavra recebida.

Definição 6.1.6. A *distância mínima* de um código C é definida por

$$d(C) = \min_{\substack{u,v \in C \\ u \neq v}} d(u,v)$$

Proposição 6.1.1. *Um código C pode detectar todos os erros de peso menor do que ou igual a s se e somente se $d(C) \geq s+1$.*

Demonstração. Se duas palavras-código estão a uma distância de no máximo s, então uma pode ser distorcida na outra por um erro de peso no máximo s. Neste

caso, não é possível detectar todos os erros de peso até s. Por outro lado, se quaisquer duas palavras-código estão a uma distância de no mínimo $s+1$, então qualquer erro de peso s distorcerá uma palavra-código em uma palavra que não está em C. Assim, todos os erros de peso até s podem ser detectados, por exemplo, comparando a palavra recebida com todas as palavras-código. □

Proposição 6.1.2. *Um código C pode corrigir todos os erros de peso até t se e somente se $d(C) \geq 2t+1$.*

Demonstração. Se C não pode corrigir t erros, então existe uma palavra $w \in \mathbb{F}_q^n$ e palavras-código $u, v \in C$ tais que $d(u,w) \leq t$ e $d(v,w) \leq t$. Isso nos dá

$$d(u,v) \leq d(u,w) + d(v,w) \leq 2t,$$

logo $d(C) < 2t+1$.

Suponhamos agora que existem $u = (u_1, \ldots, u_m), v = (v_1, \ldots, v_m) \in C$ tais que $d(u,v) \leq 2t$. Se $d(u,v) < t$, então C não corrige t erros. Se $d(u,v) \geq t$, seja $w = (w_1, \ldots, w_m)$ tal que $w_i = u_i$ para todo i tal que $u_i = v_i$, $w_i = u_i$ para as t primeiras coordenadas em que $u_i \neq v_i$, e $w_i = v_i$ para os índices i restantes. Então $d(u,w) \leq t$ e $d(v,w) \leq t$. Logo C não corrige t erros. □

Se C é um código linear, C é um subespaço vetorial de $\mathbb{F}_q^m$ e, portanto, tem uma base.

Definição 6.1.7. Seja C um código linear. Uma matriz cujas linhas formam uma base de C é dita uma *matriz geradora* de C.

Pela definição acima, se G é uma matriz geradora de C, qualquer $u = vG$, com $v \in \mathbb{F}_q^k$ é uma palavra-código.

Definição 6.1.8. Seja C um (m,k)-código linear. Uma matriz H de ordem $(m-k) \times m$ é dita uma *matriz de paridade* de C se um vetor $v \in \mathbb{F}_q^m$ está em C se e somente se $Hv^{\mathrm{T}} = 0$.

6.2 Bases de Gröbner e Códigos Lineares

Nesta seção, apresentaremos os métodos para a construção e decodificação de códigos lineares formulados em [7], que são baseados na teoria de bases de Gröbner.

6.2.1 Construção de Códigos Lineares

Construiremos um (m, k)-código linear sobre um corpo finito $\mathbb{F}_q$ da seguinte forma: Escolhemos um conjunto de m pontos distintos $V = \{P_1, \ldots, P_m\} \subset \mathbb{F}_q^n$, onde n é um inteiro tal que $q^n \geq m$. Seja $I = I(V) \subset \mathbb{F}_q[x_1, \ldots, x_n]$ o ideal de V. Fixamos uma ordem monomial $\succ_A$ em $A = \mathbb{F}_q[x_1, \ldots, x_n]$ e consideramos a base monomial $\mathcal{B}(I) = \{x^{\alpha_1}, \ldots, x^{\alpha_m}\}$ com respeito à ordem $\succ_A$, listados em ordem crescente. Definimos

$$L_k = \left\{ f(x_1, \ldots, x_n) = \sum_{i=1}^{k} a_i x^{\alpha_i} : a_i \in \mathbb{F}_q \right\}, \tag{6.1}$$

ou seja, L_k é o conjuntos das combinações lineares de $x^{\alpha_1}, \ldots, x^{\alpha_k}$ sobre $\mathbb{F}_q$. A seguir definimos o código C por

$$C = \{(f(P_1), \ldots, f(P_m)) : f \in L_k\}. \tag{6.2}$$

Considere a matriz

$$\mathcal{M} = \begin{pmatrix} x^{\alpha_1}(P_1) & x^{\alpha_2}(P_1) & \cdots & x^{\alpha_m}(P_1) \\ x^{\alpha_1}(P_2) & x^{\alpha_2}(P_2) & \cdots & x^{\alpha_m}(P_2) \\ \vdots & \vdots & \ddots & \vdots \\ x^{\alpha_1}(P_m) & x^{\alpha_2}(P_m) & \cdots & x^{\alpha_m}(P_m) \end{pmatrix} \tag{6.3}$$

onde $x^{\alpha}(P)$ denota o valor do monômio x^{α} no ponto P. Se $c = (c_1, \ldots, c_m) \in \mathbb{F}_q^m$ é tal que $\mathcal{M}c = 0$, então

$$\begin{cases} c_1 x^{\alpha_1}(P_1) + \cdots + c_m x^{\alpha_m}(P_1) = 0 \\ \vdots \\ c_1 x^{\alpha_1}(P_m) + \cdots + c_m x^{\alpha_m}(P_m) = 0 \end{cases} \tag{6.4}$$

ou seja, o polinômio $f = c_1 x^{\alpha_1} + \cdots + c_m x^{\alpha_m}$ é tal que $f(P_i) = 0$ para $i = 1, \ldots, m$, logo $f \in I$. Mas isso implica que

$$c_1 x^{\alpha_1} + \cdots + c_m x^{\alpha_m} \equiv 0 \pmod{I} \tag{6.5}$$

e como $x^{\alpha_1}, \ldots, x^{\alpha_m}$ são linearmente independentes módulo I, temos que $c_1 = c_2 = \cdots = c_m = 0$. Portanto, $\mathcal{M}$ tem posto m. Agora, C é o espaço gerado pelas k primeiras colunas de $\mathcal{M}$, logo C é um (m, k)-código linear sobre $\mathbb{F}_q$.

Os códigos Reed-Solomon, Reed-Muller e hermitianos são casos particulares do código descrito acima [7].

6.2.2 Decodificação

Sejam $r = (r_1, \ldots, r_m)$ uma palavra recebida e $V = \{P_1, \ldots, P_m\}$ o conjunto de pontos usados para definir o código C. Definimos o conjunto de pontos de $\mathbb{F}_q^{n+1}$

$$V_r = \{(P_1, r_1), \ldots, (P_m, r_m)\}$$

e o conjunto de polinômios em $\mathbb{F}_q[x_1, \ldots, x_n, z]$

$$M(V_r) = \{h = u \cdot z + v : u, v \in \mathbb{F}_q[x_1, \ldots, x_n],\ h(P_i, r_i) = 0 \text{ para todo } (P_i, r_i) \in V_r\}.$$

$M(V_r)$ é um $\mathbb{F}_q[x_1, \ldots, x_n]$-módulo. Se nenhum erro ocorreu, então

$$r = (f(P_1), \ldots, f(P_m))$$

para algum polinômio $f \in L_k$, e $M(V_r) = M_f$, onde M_f é o módulo definido no capítulo anterior.

Pelo Lema 5.3.1, podemos definir uma ordem monomial $\succ$ em $M(V_r)$ tal que z e x^{α_k} são monômios consecutivos e $z \succ x^{\alpha_k}$. O primeiro passo do procedimento de decodificação é encontrar a base de Gröbner reduzida G de $M(V_r)$ com respeito a esta ordem monomial.

Para encontrar esta base, primeiro obtemos a base de Gröbner reduzida do ideal $I(V_r)$, através do Algoritmo 4.3.1. Pela Proposição 4.3.4, esta base tem a forma $G \cup \{z - g\}$, onde G é a base de Gröbner reduzida de $I(P_1, \ldots, P_m)$ e $g \in \mathbb{F}_q[x_1, \ldots, x_n]$ é tal que $g(P_i) = r_i$ para $i = 1, \ldots, m$. Agora, pela demonstração da parte (i) do Lema 5.3.2, $G \cup \{z - g\}$ é a base de Gröbner reduzida do módulo $M_g = \{bz - a : a, b \in \mathbb{F}_q[x_1, \ldots, x_n],\ bg \equiv a \pmod{I}\}$. Como $bz - a \in M_g$ se e somente se $bg - a \in I$, ou seja, $(bg-a)(P_i) = b(P_i)g(P_i) - a(P_i) = b(P_i)r_i - a(P_i) = 0$, temos que $M_g = M(V_r)$. Portanto, $G \cup \{z - g\}$ é a base de Gröbner reduzida de $M(V_r)$ com respeito à ordem monomial da parte (i) do Lema 5.3.2. Para obter a base reduzida com respeito à ordem do Lema 5.3.1, basta utilizar um algoritmo de conversão, como já foi mencionado no final do capítulo anterior.

Procuramos um polinômio localizador de erros, ou seja, um polinômio w tal que $w(P_i) = 0$ se ocorreu um erro no símbolo r_i. Se w é um polinômio localizador de erros e $(f(P_1), \ldots, f(P_m))$ é a palavra enviada, então $w \cdot (z - f) \in M(V_r)$, pois se ocorreu um erro em r_i, temos $w(P_i) = 0$, e se não ocorreu, temos $(z - f)(P_i, r_i) = r_i - f(P_i) = 0$. Esperamos que um tal polinômio apareça em um dos elementos da base G. Assim, consideramos $g = w \cdot z + v \in G$ tal que $tl(g)$ é divisível por z e é mínimo, e tomamos w como polinômio localizador de erros.

Consideramos agora o conjunto $V_r' = V_r \backslash \{(P_i, r_i) : w(P_i) = 0\}$. Se w é um polinômio localizador de erros e $f \in L_k$ é tal que $(f(P_1), \ldots, f(P_m))$ é a palavra enviada, então para cada r_i tal que $(P_i, r_i) \in V_r'$, temos $r_i = f(P_i)$. A seguir encontramos a base de Gröbner reduzida G' de

$$M(V_r') = \{h = u \cdot z + v : h(P_i, r_i) = 0 \text{ para todo } (P_i, r_i) \in V_r'\}.$$

Temos que $z - f \in M(V_r')$, pois $(z-f)(P_i, r_i) = r_i - f(P_i) = 0$ para todo $(P_i, r_i) \in V_r'$. Além disso, $tl(z - f) = z$, pois como $f \in L_k$, $tl(f) \preceq x^{\alpha_k} \prec z$.

Se $g \in A$ é um polinômio tal que $g(P_i) = r_i$ para todo $(P_i, r_i) \in M(V_r')$, então $M_g = M(V_r')$, e o Teorema 5.3.1 garante que o elemento $z - f$ está na base de Gröbner reduzida de M_g.

O procedimento de decodificação descrito acima depende da existência de um polinômio localizador de erros na base G, e pode falhar se todos os elementos da forma $w \cdot (z - f)$, onde w é um polinômio localizador de erros, têm termos líderes divisíveis por $tl(h)$ para algum $h \in M(V_r)$. Segundo Farr e Gao [7], isso pode acontecer em duas situações: quando ocorrem mais erros do que o código é capaz de corrigir, e quando os pontos $P_1, \ldots, P_m$ possuem uma certa estrutura geométrica.

De acordo com [7], experimentos computacionais indicam que o método de decodificação funciona bem em geral. No entanto, pouco se sabe ainda sobre as distâncias mínimas dos códigos e de que forma elas dependem da estrutura geométrica dos pontos usados para definir os códigos.

Bibliografia

[1] Adams, W. W., and Loustaunau, P. *An Introduction to Gröbner Bases*, vol. 3 of *Graduate Studies in Mathematics*. American Mathematical Society, 1994.

[2] Becker, T., Weispfenning, V., and Kredel, H. *Gröbner Bases: A Computational Approach to Commutative Algebra*. Springer-Verlag, New York, 1993.

[3] Buchberger, B. *An algorithm for finding the basis elements in the residue class ring modulo a zero dimensional polynomial ideal*. PhD thesis, Mathematical Institute, University of Innsbruck, Austria, 1965. English translation published in the Journal of Symbolic Computation 41 (2006) 475–511.

[4] Costa, A. V., and Vainsencher, I. Bases de Gröbner: resolvendo equações polinomiais. In *Atas da XIII Escola de Álgebra* (1994), vol. 1, IMECC-UNICAMP, pp. 111–184.

[5] Cox, D., Little, J., and O'Shea, D. *Ideals, Varieties, and Algorithms*, 2nd ed. Springer-Verlag, New York, 1997.

[6] Cox, D., Little, J., and O'Shea, D. *Using Algebraic Geometry*, 2nd ed. Springer-Verlag, New York, 1998.

[7] Farr, J. B., and Gao, S. Gröbner bases, Padé approximation, and decoding of linear codes. In *Coding Theory and Quantum Computing*, vol. 381 of *Contemporary Mathematics*. American Mathematical Society, 2005.

[8] Farr, J. B., and Gao, S. Computing Gröbner bases for vanishing ideals of finite sets of points. In *Applied Algebra, Algebraic Algorithms and Error-*

Correcting Codes (Berlin / Heidelberg, 2006), vol. 3857 of *Lecture Notes in Computer Science*, Springer, pp. 118–127.

[9] Farr, J. B., and Gao, S. Gröbner bases and generalized Padé approximation. *Mathematics of Computation 75* (2006), 461–473.

[10] Faugère, J. C., Gianni, P., Lazard, D., and Mora, T. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *Journal of Symbolic Computation 16* (1993), 329–344.

[11] Hungerford, T. W. *Algebra*. Springer Verlag, New York, 1974.

[12] Kreuzer, M., and Robbiano, L. *Computational Commutative Algebra 1*. Springer-Verlag, Heidelberg, 2000.

[13] Lint, J. H. v. *Introduction to Coding Theory*, 2nd ed. Springer-Verlag, Berlin, 1992.

[14] Pless, V. *Introduction to the Theory of Error-Correcting Codes*, 2nd ed. John Wiley and Sons, New York, 1990.

[15] Pretzel, O. *Error-Correcting Codes and Finite Fields*. Oxford Applied Mathematics and Computer Science Series. Clarendon Press, New York, 1992.

[16] Robbiano, L. Term orderings on the polynomial ring. In *Proceedings of EUROCAL '85* (Berlin / Heidelberg, 1985), vol. 204 of *Lecture Notes in Computer Science*, Springer, pp. 513–517.

Índice