Fernando Daniel Moreira Coelho

O Algoritmo LLL e Aplicações

Orientador

Prof. Doutor João Luís Cardoso Soares

Dissertação apresentada para a obtenção do grau de Mestre em Matemática, área de especialização em Matemática para o Ensino.

- À memória de

António Augusto Costa

José Manuel Cunha Torres

Agradecimentos

Ao Professor Doutor João Soares, pela colaboração, sugestões e excelente orientação dadas na elaboração desta tese.

À minha mulher Paula Santos, agora Paula Coelho, pelo incessante apoio, carinho e generosidade. Por ter dito presente em todos os momentos.

Ao meus colegas, Fernando Bernardino e Mateus Mendes, pela disponibilidade e auxílio que sempre me deram.

Aos meus pais e a todos os meus familiares por tudo o que me proporcionaram.

A todos os meus colegas e amigos que me incentivaram a fazer este trabalho. Todos nós temos necessidade de alguém que nos obrigue a realizar aquilo de que somos capazes. É este o papel da amizade.

Conteúdo

Notação

- $\mathcal{S}(C)$ - Expansão linear do conjunto, finito, C.
- $\left(\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array} \right)$ - Matriz com vectores coluna $v_1, \ldots, v_n$.
- $\mathcal{S}(V_i)$ - Expansão linear do conjunto de vectores $\{v_1, \ldots, v_i\}$.
- $proj_F(v)$ ou v_F - Projecção ortogonal de v sobre o espaço F.
- $proj_x\, y$ - Projecção ortogonal do vector y sobre o vector x.
- OGS - Ortogonalização de Gram-Schmidt.
- $\det(A)$ - Determinante de A.
- $\text{adj}(A)$ - Matriz adjunta de A.
- $\text{tr}(A)$ - Traço da matriz A.
- $E_{ij}(\lambda)$ - Matriz elementar que coincide com a identidade excepto na posição (i, j) em que é igual a λ.
- $P_{i+1,i}$ - Matriz de permutação que coincide com a identidade excepto nas colunas i, $i+1$ que estão trocadas.
- $D_{i,i}(-1)$ - Matriz elementar que coincide com a identidade excepto na posição (i, i) que é igual a -1.

- $\mathcal{V}(n)$ - volume da bola unitária de dimensão n.
- $\text{car}(A)$ - característica da matriz A.
- $\lfloor x \rfloor$ - maior inteiro menor ou igual a x.
- $\lceil x \rceil$ - menor inteiro maior ou igual a x.
- $\lfloor x \rceil = \begin{cases} \lceil x \rceil & \text{se} \quad x \geq \lfloor x \rfloor + 1/2 \\ \lfloor x \rfloor & \text{se} \quad x < \lfloor x \rfloor + 1/2 \end{cases}$, isto é, $\lfloor x \rceil$ é o inteiro mais próximo de x.
- $B^0(x, r)$ - Bola aberta de centro x e raio r.
- $B(x, r)$ - Bola fechada de centro x e raio r.
- $a \equiv b (\text{mod m})$ - a e b são congruentes módulo m ou m divide $(a - b)$.
- $\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$, para algum natural m.
- $m.d.c(n, m)$ - máximo divisor comum entre n e m.
- $\|.\|$ - é uma norma, normalmente do tipo $\|x\| = \sqrt{x^T D x}$, para alguma matriz D definida positiva.
- $\|.\|_2$ - é a norma euclidiana, isto é, $\|x\|_2 = \sqrt{x^T x}$.

Introdução

O presente trabalho constitui uma Tese de Mestrado em Matemática para o Ensino da FCTUC. Na sua elaboração seguimos de perto [14]. Inspirámo-nos ainda em [9], [16] e [4].

O objectivo desta dissertação consiste em estudar um método que permita encontrar uma base "reduzida", constituída por vectores "próximos"da ortogonalidade, para um reticulado. Denomina-se por **reticulado** gerado pelos vectores linearmente independentes $v_1, \ldots, v_k$ ao conjunto

$$\mathcal{L} = B\mathbb{Z}^k = \{x \in \mathbb{R}^n : x = a_1 v_1 + \ldots + a_k v_k \, , a_i \in \mathbb{Z}, i = 1, \ldots, k\} \, ,$$

em que B é a matriz com vectores coluna $v_1, \ldots, v_k$. O conjunto $\{v_1, \ldots, v_k\}$ diz-se uma **base de** $\mathcal{L}$. Em particular, se $n = k$ diz-se que $\mathcal{L}$ tem dimensão completa e a sua **norma** ou **determinante** define-se por

$$\det \mathcal{L} = |\det B| \, ,$$

e pode ser interpretado como o volume de um paralelipípedo n-dimensional. Este valor, como veremos no Capítulo 2, é independente da escolha da base. No entanto, apesar de todas as bases de um reticulado de dimensão completa terem o mesmo determinante, nem todas são "equivalentes"do ponto de vista prático (por exemplo, o vector de norma mínima de um reticulado

nem sempre aparece numa sua base). Veja-se o caso de $\mathbb{Z}^2$, $\{(1,0),(0,1)\}$ é uma base com a qual é mais fácil trabalhar do que $\{(3,2),(2,1)\}$, uma vez que a primeira é formada por vectores linearmente independentes de $\mathbb{Z}^2$ de menor comprimento. Lovász (cf. [14] pág. 68) definiu um algoritmo polinomial que permite encontrar uma base reduzida. Esse algoritmo é usualmente descrito como **Algoritmo LLL** (Lenstra, Lenstra e Lovász) ou **Método de redução de base** e foi originalmente proposto em [8].

No Capítulo 1, apresentam-se algumas definições e propriedades elementares sobre espaços vectoriais que serão fundamentais para uma boa compreensão do Algoritmo LLL. Para tal, apresentamos os conceitos de produto interno e a norma por si definida. Estudamos também a unicidade da projecção ortogonal v_F de um vector v sobre um espaço vectorial F. Expomos o processo de **Ortogonalização de Gram-Schmidt** (OGS), processo este que visa obter uma base ortogonal em qualquer espaço vectorial de dimensão finita. No Algoritmo LLL interessar-nos-á actualizar a decomposição QR de uma matriz A', que resulta de A por troca de duas colunas sucessivas; assim como, mostrar que os vectores $v_1^*, \ldots v_n^*$, que resultam do processo de ortogonalização de Gram-Schmidt, permanecem invariáveis se a um vector da sequência $v_1, \ldots, v_n$, somarmos um outro (anterior) previamente multiplicado por um escalar. Ainda neste capítulo introduzimos o conceito de **base reduzida** e apresentamos algumas das suas propriedades.

No Capítulo 2, expomos a definição de reticulado e mostramos que a sua norma não depende da escolha da base. Estudamos a relação presente entre a norma de qualquer vector $x \in \mathcal{L}$ e o comprimento do primeiro vector de uma base reduzida. Como motivação para o Algoritmo LLL, explica-se o

método de **Redução de Gauss em** $\mathbb{R}^2$; este consiste em encontrar uma base (ordenada) reduzida de um reticulado em $\mathbb{R}^2$ cujo primeiro vector é o vector não nulo de norma mínima. Seguidamente, expomos o Algoritmo LLL para o reticulado $\mathbb{Z}^n$ e a sua aplicação para qualquer outro reticulado. Mostramos que este algoritmo tem no máximo $n^2(log_2 n + log_2 T)log_{\frac{4}{3}}2$ iterações e que o tamanho dos números intermediários são limitados polinomialmente. No final, exemplificamos a execução do Algoritmo LLL, em $\mathbb{R}^2$, efectuando todos os cálculos necessários e de uma outra forma recorrendo a uma implementação do Algoritmo LLL, que podemos encontrar na biblioteca de estruturas e algoritmos **NTL - Number Theory Library**.

No Capítulo 3, estudamos algumas aplicações do Algoritmo LLL. Nomeadamente: o *Problema do Vector mais Curto* (*Shortest Vector Problem - SVP*); o *Problema do Vector mais Próximo* (*Closest Vector Problem - SVP*) e o problema de *Aproximação Diofantina simultânea*. Outra aplicação focada é a de determinar a **Forma Normal de Hermite** com auxílio do Algoritmo LLL. Por fim, propomos um método muito simples para verificar a solução de alguns problemas de **Programação Inteira**, substituindo o problema

$$PI = \{x : b' \leq Ax \leq b\} \cap \mathbb{Z}^n \,,$$

por

$$\tilde{PI} = \{y : b' \leq (AU)y \leq b\} \cap \mathbb{Z}^n \,,$$

em que U é uma matriz de inteiros unimodular que resulta por aplicação do Método de redução de base.

No Capítulo 4, descrevemos o sistema criptográfico **RSA** (Rivest, Shamir e Adleman), originalmente proposto em [12]. Seguidamente, veremos de que

forma o Algoritmo LLL pode ser usado para desencriptar mensagens codificadas por este sistema. De modo a ilustrar esta situação, recorremos ao software **Cryptool**, que podemos encontrar em http://www.cryptool.org/. Esta ferramenta permite trabalhar com alguns conceitos de criptografia, entre os quais o código RSA.

Omitem-se as demonstrações de alguns resultados mais conhecidos, já que estas podem ser consultadas em inúmeros livros da respectiva área.

Capítulo 1

Resultados elementares

Neste capítulo, não faremos uma análise exaustiva dos espaços vectoriais, uma vez que não é esse o objectivo desta dissertação, mas preocupar-nos-emos em enunciar algumas definições e demonstrar propriedades essenciais. Estas servem de base para desenvolver o principal objecto desta investigação que é o Método de redução de base, tal como é descrito em [14] pág. 68, também conhecido por Algoritmo LLL (Lenstra, Lenstra and Lovasz) e originalmente proposto em [8].

Na primeira secção apresentamos o conceito axiomático de produto interno num espaço vectorial bem como a norma por si definida. São, também, enunciados dois exemplos de produto interno usados ao longo desta dissertação: o produto interno euclidiano em $\mathbb{R}^n$ e o produto interno associado a uma matriz simétrica definida positiva.

Na segunda secção definimos projecção ortogonal v_F de um vector v sobre um espaço vectorial F e mostramos que é único. Este vector v_F possui uma definição alternativa, concretamente, é o vector que torna mínima a diferença entre v e um outro qualquer vector $u \in F$.

Na terceira secção estudamos também o processo de Ortogonalização de

Gram-Schmidt, que permite obter uma base ortogonal em qualquer espaço vectorial de dimensão finita. Uma consequência deste processo é a de que toda a matriz A se pode decompor na forma $A = QR$, onde Q é uma matriz com colunas ortonormadas e R é uma matriz triangular superior, com elementos diagonais positivos. Como consequência deste processo demonstramos a Desigualdade de Hadamard.

Na quarta secção veremos como actualizar a decomposição QR de uma matriz A', que resulta de A por troca de duas colunas sucessivas. Para além disso, mostramos que os vectores $v_1^*, \ldots v_n^*$ que resultam do processo de Ortogonalização de Gram-Schmidt, quando aplicado a uma sequência de vectores $v_1, \ldots, v_n$ linearmente independentes, permanecem invariáveis se a um vector da sequência, $v_1, \ldots, v_n$, somarmos um outro (anterior) previamente multiplicado por um escalar.

Na última secção introduzimos o conceito de base reduzida e apresentamos algumas das suas propriedades. Como veremos, o conceito de base reduzida é fundamental no Algoritmo LLL.

1.1 Produto interno num espaço vectorial

A definição axiomática de espaço vectorial pode ser encontrada em [9], bem como o desenvolvimento de toda a teoria fundamental dos espaços vectoriais. Supomos que o leitor já tem conhecimento de grande parte desse estudo. Assim sendo, nesta secção recordamos apenas alguns desses conceitos tomando por base [10] e [13], nomeadamente os que são estritamente necessários para uma boa compreensão do Algoritmo LLL.

Um desses conceitos é o de produto interno num espaço vectorial real, noção essa que será primordial no desenvolvimento do Algoritmo LLL.

Definição 1.1.1. *Seja V um espaço vectorial real. Um* ***produto interno*** *em V é uma operação que a cada par de vectores x, y de V faz corresponder um número real, denotado $\langle x, y\rangle$ e chamado* ***produto interno de*** *x* ***por*** *y, que verifica as seguintes propriedades:*

i) $\langle x, y\rangle = \langle y, x\rangle$.

ii) $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$.

iii) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$.

iv) $\langle x, x\rangle \geq 0$, *e* $\langle x, x\rangle = 0$ *se e só se* $x = 0$.

onde x, x' e y designam vectores quaisquer de V e α um número real arbitrário.

De acordo com esta definição, um produto interno em V é uma aplicação de $V \times V$ em $\mathbb{R}$. A segunda e a terceira propriedades podem resumir-se dizendo que esta aplicação é linear "em relação ao primeiro vector"(i.e., mantendo o segundo vector fixo arbitrariamente). Claro que, pela primeira propriedade, o mesmo se verifica em relação ao segundo vector (mantendo fixo o primeiro). Aplicações deste tipo costumam chamar-se **bilineares**.

Definição 1.1.2. *Um espaço vectorial real V em que está definido um produto interno diz-se um* ***espaço com produto interno*** *ou um* ***espaço euclidiano****.*

Dois exemplos de produto interno são os seguintes:

Exemplo 1.1.1. $\langle x, y\rangle = x_1y_1 + \ldots + x_ny_n = x^Ty$ é um produto interno no espaço vectorial $\mathbb{R}^n$. Este produto interno é normalmente denominado por **produto interno usual** ou **produto interno euclidiano**.

Exemplo 1.1.2. $\langle x, y\rangle = x^T D\, y$, com D uma matriz quadrada de ordem n simétrica definida positiva, é um produto interno no espaço vectorial $\mathbb{R}^n$.

Outro conceito, com papel fundamental nesta dissertação, é o de comprimento de um vector que pode ser definido pelo produto interno.

Definição 1.1.3. *Seja V um espaço vectorial real com produto interno e sejam x e y vectores de V. Então:*

i) A ***norma*** *ou* ***comprimento*** *de x é $\|x\| = \sqrt{\langle x, x\rangle}$.*

ii) A ***distância*** *entre x e y é $\|x - y\|$.*

iii) x e y dizem-se ***ortogonais*** *se $\langle x, y\rangle = 0$. Denota-se $x \perp y$.*

Nesta dissertação, denotamos a **norma euclidiana**, definida pelo produto interno do Exemplo 1.1.1 por $\|.\|_2$, isto é, $\|x\|_2 = \sqrt{x^T x}$.
A D**-norma** definida pelo produto interno do Exemplo 1.1.2, sempre que especificada, será representada por $\|.\|$, isto é, $\|x\| = \sqrt{x^T D x}$, para alguma matriz D definida positiva. Quando nos referirmos à norma definida por um produto interno não especificado usaremos também a notação $\|.\|$, isto é $\|x\| = \sqrt{\langle x, x\rangle}$.
Seguidamente, apresentamos algumas propriedades da norma, cujas demonstrações podem ser vistas em [9].

Teorema 1.1.1. *Seja V um espaço vectorial com produto interno e sejam x e y vectores de V. Então:*

i) $|\langle x, y\rangle| \leq \|x\|\,\|y\|$ *(desigualdade de Cauchy-Schwarz).*

ii) $\|x \pm y\| \leq \|x\| + \|y\|$ *(desigualdade triangular).*

iii) Se x e y forem ortogonais tem-se $\|x \pm y\|^2 = \|x\|^2 + \|y\|^2$ *(Teorema de Pitágoras) .*

Sendo $x, y \in V$ não nulos, a desigualdade de Cauchy-Schwarz garante que o quociente $\frac{\langle x,y\rangle}{\|x\|\,\|y\|}$ está entre -1 e 1, o que permite enunciar a primeira das definições seguintes:

Definição 1.1.4. *Seja V um espaço vectorial real com produto interno. Então:*

i) Sendo $x, y \in V$ não nulos, o ***ângulo*** *entre x e y é o número real θ (entre 0 e π) tal que*

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}.$$

ii) Sendo $x \neq 0$, a ***projecção ortogonal*** *de y sobre x é o vector*

$$proj_x\, y = \frac{\langle x, y\rangle}{\|x\|^2}\, x\,.$$

Da primeira destas definições decorre que, sendo θ o ângulo entre x e y, se tem

$$\langle x, y\rangle = \|x\|\,\|y\|\, cos\theta\,.$$

A noção de complemento ortogonal é também referida numa das próximas secções, pelo que relembramos o seu conceito.

Definição 1.1.5. *Seja V um espaço vectorial com produto interno e seja F um subespaço de V. Ao conjunto dos vectores de V que são ortogonais a todos os vectores de F chama-se* ***complemento ortogonal*** *de F. A notação habitual é $F^{\perp}$. Simbolicamente*

$$F^{\perp} = \{x \in V : \langle x, u\rangle = 0\,,\ \textit{para todo } u \in F\}\,. \tag{1.1}$$

Recordamos, ainda, que

Definição 1.1.6. *Sejam F e G subespaços vectoriais. Chama-se* ***soma*** *dos subespaços F e G ao conjunto*

$$F + G = \{v + w : v \in F,\, w \in G\}.$$

Se se tiver $F \cap G = \{0\}$, diz-se que a soma de F com G é ***directa*** *e escreve-se $F \oplus G$.*

Todo o espaço vectorial é igual à soma directa de um subespaço com o seu complemento ortogonal, sendo que este é ainda um subespaço vectorial, como podemos ver no teorema seguinte cuja demonstração pode ser vista em [9] pág. 403.

Teorema 1.1.2. *Seja F um subespaço de um espaço vectorial real V. Então, $F^{\perp}$ é ainda um subespaço de V e $V = F \oplus F^{\perp}$*

Prova-se também (cf. [13] pág. 170) que qualquer conjunto de vectores ortogonais dois a dois são linearmente independentes.

Teorema 1.1.3. *Seja V um espaço vectorial com produto interno. Se $v_1, \ldots, v_k \in V$ são não nulos e dois a dois ortogonais, então são linearmente independentes.*

De acordo com este teorema torna-se importante distinguir uma base, onde todos os vectores são ortogonais dois a dois, de uma base onde tal não acontece.

Definição 1.1.7. *Seja F um subespaço de um espaço vectorial real V. Uma base de F constituída por vectores ortogonais dois a dois diz-se uma* ***base ortogonal*** *de F. Uma base ortogonal totalmente constituída por vectores de norma unitária diz-se uma* ***base ortonormada*** *de F.*

De seguida, mostramos que qualquer vector v de um subespaço vectorial F pode ser escrito como a soma das projecções ortogonais desse vector sobre cada um dos vectores de uma base ortogonal de F.

Teorema 1.1.4. *Seja F um subespaço de um espaço vectorial real V e seja $\{v_1, \ldots, v_k\}$ uma base ortogonal de F. Então, para qualquer vector v de F, tem-se*

$$v = \sum_{i=1}^{k} proj_{v_i} v = \sum_{i=1}^{k} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i \, .$$

Em particular, se a base for ortonormada, tem-se

$$v = \sum_{i=1}^{k} \langle v, v_i \rangle v_i = \sum_{i=1}^{k} (\|v\| \cos \theta_i) \, v_i \, ,$$

onde $\theta_1, \ldots, \theta_k$ são os ângulos de v com $v_1, \ldots, v_k$.

Demonstração: Seja $v = \sum_{i=1}^{k} \alpha_i v_i$, com $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Então, para $j = 1, \ldots, k$, tem-se

$$\langle v, v_j \rangle = \left\langle \sum_{i=1}^{k} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{k} \alpha_i \langle v_i, v_j \rangle = \alpha_j \|v_j\|^2 \, ,$$

uma vez que os vectores $v_1, \ldots, v_k$ são ortogonais dois a dois. Como $v_j \neq 0$, vem

$$\alpha_j = \frac{\langle v, v_j \rangle}{\|v_j\|^2} \, .$$

■

Atendendo ao exposto, podemos colocar a seguinte questão: "*Todo o subespaço possui uma base ortogonal?*" Sobre este assunto pronunciar-nos-emos apenas na Secção 1.3.

1.2 Projecção ortogonal sobre um subespaço

Nesta secção recordamos a definição de projecção ortogonal sobre um subespaço vectorial. Veremos que a diferença entre um vector e a respectiva projecção ortogonal num subespaço é um vector que pertence ao seu complemento ortogonal, que definimos em (1.1).

Definição 1.2.1. *Seja v um vector de um espaço vectorial V com produto interno e seja F um subespaço de V. Um vector $v_F \in F$ diz-se a* ***projecção ortogonal*** *de v sobre F se $v - v_F$ for ortogonal a todos os vectores de F. Também se usa a notação $proj_F v$ em vez de v_F.*

De seguida, mostramos que o vector v_F é único e que a diferença entre v e qualquer outro vector $u \in F$ é mínima quando $u = v_F$. Mostramos, também, que a soma das projecções de v sobre os elementos de uma base ortogonal de F é a projecção ortogonal de v sobre F.

Teorema 1.2.1. *Seja v um vector de um espaço vectorial V com produto interno e seja F um subespaço de V. Então:*

1. *Se existir um vector v_F que seja projecção ortogonal de v sobre F, ele é único e satisfaz*

$$\|v - v_F\| = \min \{\|v - u\| : u \in F\} ,$$

   em particular $\|v - v_F\| \leq \|v\|$.

2. *Se $\{v_1, \ldots, v_k\}$ é uma base ortogonal de F, então, a soma das projecções ortogonais de v sobre os v_i é a projecção ortogonal de v sobre F, isto é*

$$v_F = \sum_{i=1}^{k} proj_{v_i} v = \sum_{i=1}^{k} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i .$$

Demonstração: *1.* Seja v_F projecção ortogonal de v sobre F e seja $u \in F$ arbitrário. Como $v - v_F$ é ortogonal a todos os vectores de F então também é ortogonal a $u - v_F$. Assim, pelo Teorema de Pitágoras (cf. pág. 4) tem-se que

$$\begin{aligned} \|v-u\|^2 &= \|(v-v_F)-(u-v_F)\|^2 \\ &= \|v-v_F\|^2 + \|u-v_F\|^2 . \end{aligned}$$

Segue-se que, para qualquer $u \in F$, tem-se

$$\|v-v_F\| \leq \|v-u\| ,$$

com igualdade se e só se $u = v_F$ (o que prova a unicidade).

2. Consideremos o vector w tal que

$$w = \frac{\langle v, v_1\rangle}{\|v_1\|^2}\, v_1 + \ldots + \frac{\langle v, v_k\rangle}{\|v_k\|^2}\, v_k .$$

É evidente que $w \in F$. Vejamos agora que $v - w$ é ortogonal a todos os vectores de F. Seja $u \in F$ arbitrário, digamos $u = \alpha_1 v_1 + \ldots + \alpha_k v_k$. Tem-se

$$\begin{aligned} \langle v-w, u\rangle &= \langle v, u\rangle - \langle w, u\rangle \\ &= \left\langle v, \sum_{j=1}^{k} \alpha_j v_j \right\rangle - \left\langle \sum_{i=1}^{k} \frac{\langle v, v_i\rangle}{\|v_i\|^2}\, v_i, \sum_{j=1}^{k} \alpha_j v_j \right\rangle \\ &= \sum_{j=1}^{k} \alpha_j \langle v, v_j\rangle - \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{\langle v, v_i\rangle}{\|v_i\|^2}\, \alpha_j \langle v_i, v_j\rangle . \end{aligned}$$

Como $v_1, \ldots, v_k$ são ortogonais dois a dois, tem-se $\langle v_i, v_j\rangle = 0$, se $i \neq j$, pelo que são nulas todas as parcelas em que $i \neq j$ no somatório duplo. Por isso,

$$\begin{aligned} \langle v-w, u\rangle &= \sum_{j=1}^{k} \alpha_j \langle v, v_j\rangle - \sum_{j=1}^{k} \frac{\langle v, v_j\rangle}{\|v_j\|^2}\, \alpha_j \langle v_j, v_j\rangle \\ &= \sum_{j=1}^{k} \alpha_j \langle v, v_j\rangle - \sum_{j=1}^{k} \alpha_j \langle v, v_j\rangle \\ &= 0 . \end{aligned}$$

■

Se F tiver dimensão 1, qualquer vector não nulo w de F constitui uma base ortogonal de F. A expressão para a projecção ortogonal de um vector qualquer v sobre F é então

$$v_F = \frac{\langle v, w\rangle}{\|w\|^2} w\,,$$

o que coincide com a projecção ortogonal de v sobre w. Isto é, a projecção ortogonal de um vector v sobre um subespaço unidimensional F é igual à projecção ortogonal de v sobre qualquer vector não nulo de F.

1.3 Ortogonalização de Gram-Schmidt

Tal como vimos na Secção 1.1, uma base ortogonal possui significativas vantagens sobre as outras bases. O espaço euclidiano real $\mathbb{R}^n$ possui claramente uma base ortogonal (e.g. a base canónica), mas o que é que acontece com outros espaços e com outros produtos internos? Será que todo o espaço vectorial de dimensão finita possui uma base ortogonal? E se possui, como determinar essa base? Nesta secção vamos desenvolver um processo de ortogonalização que permite, a partir de uma base qualquer de um espaço vectorial, determinar uma base ortogonal para esse mesmo espaço.
A partir de uma base qualquer de V pode sempre obter-se uma base ortogonal usando o processo de Ortogonalização de Gram-Schmidt.

Teorema 1.3.1 (Processo de ortogonalização de Gram-Schmidt - OGS)**.** *Seja V um espaço vectorial real. Seja F um subespaço de V e $\{v_1, \ldots, v_k\}$ uma base de F. Define-se*

$$v_1^* \;=\; v_1\,,$$

e sucessivamente, para $j = 2, 3, \ldots, k$,

$$v_j^* = v_j - \sum_{i=1}^{j-1} proj_{v_i^*} v_j \, .$$

Então, $\{v_1^*, \ldots, v_k^*\}$ *é uma base ortogonal de* F*.*

Demonstração: Note-se que cada v_j^* se obtém de v_j subtraindo-lhe uma combinação linear de $v_1, \ldots, v_{j-1}$. Como $v_1, \ldots, v_k$ são linearmente independentes, os vectores $v_1^*, \ldots, v_k^*$ são todos não nulos.
Vamos agora mostrar, por indução, que $v_1^*, \ldots, v_k^*$ são ortogonais dois a dois começando por v_1^* e v_2^*:

$$\begin{aligned} \langle v_2^*, v_1^* \rangle &= \left\langle v_2 - \frac{\langle v_2, v_1^* \rangle}{\|v_1^*\|^2} \, v_1^*, v_1^* \right\rangle \\ &= \langle v_2, v_1^* \rangle - \frac{\langle v_2, v_1^* \rangle}{\|v_1^*\|^2} \langle v_1^*, v_1^* \rangle \\ &= 0. \end{aligned}$$

Seja agora $j > 2$ e suponhamos que $v_1^*, \ldots, v_{j-1}^*$ são dois a dois ortogonais. Vamos provar que v_j^* é ortogonal a cada um desses vectores. Seja $p < j$, calculando o produto interno $\langle v_j^*, v_p^* \rangle$ vem

$$\begin{aligned} \left\langle v_j - \sum_{i=1}^{j-1} \frac{\langle v_j, v_i^* \rangle}{\|v_i^*\|^2} \, v_i^*, v_p^* \right\rangle &= \langle v_j, v_p^* \rangle - \sum_{i=1}^{j-1} \frac{\langle v_j, v_i^* \rangle}{\|v_i^*\|^2} \langle v_i^*, v_p^* \rangle \\ &= \langle v_j, v_p^* \rangle - \frac{\langle v_j, v_p^* \rangle}{\|v_p^*\|^2} \langle v_p^*, v_p^* \rangle \\ &= 0. \end{aligned}$$

Como $v_1^*, \ldots, v_k^*$ são não nulos e ortogonais dois a dois, são linearmente independentes. Como são k e pertencem a F (que tem dimensão k) constituem uma base de F.

■

De acordo com este teorema, sabemos que, para $j = 2, \ldots, k$,

$$v_j^* = v_j - \alpha_{1j} v_1^* - \alpha_{2j} v_2^* - \ldots - \alpha_{j-1,j} v_{j-1}^* \, ,$$

onde, para cada $i = 1, \ldots, j-1$, temos

$$\alpha_{ij} = \frac{\langle v_i^*, v_j \rangle}{\|v_i^*\|^2} \,. \tag{1.2}$$

Para $i \neq j$, tem-se $\langle v_j^*, v_i^* \rangle = 0$ e

$$v_j - v_j^* = \frac{\langle v_1^*, v_j \rangle}{\|v_1^*\|^2} \, v_1^* + \ldots + \frac{\langle v_j^*, v_{j-1} \rangle}{\|v_{j-1}^*\|^2} \, v_{j-1}^* \,, \tag{1.3}$$

ou seja, $v_j - v_j^* \in \mathcal{S}(V_{j-1})$ e desta forma $\langle v_j - v_j^*, u \rangle = 0$, para qualquer $u \in \mathcal{S}(V_{j-1})^\perp$. Podemos então concluir que v_j^* é a projecção de v_j no complemento ortogonal de $\mathcal{S}(V_{j-1})$, caracterizado por

$$\mathcal{S}(V_{j-1})^\perp = \{v \in V : \langle v, u \rangle = 0 \,, \text{ para todo } u \in \mathcal{S}(V_{j-1})\} \,,$$

e em particular $\|v_j^*\| \leq \|v_j\|$.

Observação 1.3.1. No caso do produto interno ser definido por uma matriz simétrica definida positiva, tal como foi referido no Exemplo 1.1.2, o procedimento descrito no Teorema 1.3.1 reduz-se à fórmula

$$v_j^* = v_j - V_{j-1}(V_{j-1}^T D V_{j-1})^{-1} V_{j-1}^T D v_j \,,$$

para $j = 2, 3, \ldots, k$, onde V_{j-1} representa a matriz cujas colunas são os vectores $v_1, \ldots, v_{j-1}$.

Conhecida uma base ortogonal $v_1^*, \ldots, v_k^*$ de F, é imediato obter uma base ortonormada para o mesmo subespaço. Para tal basta dividir cada vector pela sua norma.

Corolário 1.3.1. *Todo o subespaço de V possui pelo menos uma base ortogonal.*

Definição 1.3.1. *Nas condições enunciadas no Teorema 1.3.1,*

$$q_j^* \;=\; \frac{v_j^*}{\|v_j^*\|} \,, \quad j = 1, 2, \ldots, k \,, \tag{1.4}$$

é denominada ***Sequência de Gram-Schmidt****.*

O processo de Ortogonalização de Gram-Schmidt pode ser apresentado na forma matricial, como podemos ver a seguir.

Corolário 1.3.2 (Factorização "rectangular" $\hat{Q}\hat{R}$)**.** *Se $A \in \mathbb{R}^{n\times k}$ tem as colunas linearmente independentes, então A pode decompor-se na forma $A = \hat{Q}\hat{R}$, onde $\hat{Q}$ é $n \times k$ e tem colunas ortonormadas e $\hat{R}$ é $k \times k$ triangular superior não-singular, com elementos diagonais positivos.*

Demonstração: Isto é apenas uma outra forma de descrever o processo de OGS. Designemos as colunas de A por $v_1, \ldots, v_k$ e por $v_1^*, \ldots, v_k^*$ os respectivos vectores coluna, ortogonais dois a dois, que se obtêm das colunas de A por aplicação da OGS.

$$v_j \;=\; \alpha_{1j}v_1^* + \alpha_{2j}v_2^* + \ldots + \alpha_{j-1,j}v_{j-1}^* + v_j^*\,,$$

para $j = 1, 2, \ldots, k$, onde os α_{ij} são números reais definidos por (1.2) quando $i < j$, $\alpha_{ij} = 1$ quando $i = j$ e $\alpha_{ij} = 0$ quando $i > j$. Designando por U a matriz cujas colunas são $v_1^*, \ldots, v_k^*$, estas igualdades podem resumir-se pela igualdade matricial $A = UT$, onde

$$A = UT = \left(\begin{array}{cccc} v_1^* \mid & v_2^* \mid & \ldots \mid & v_k^* \end{array} \right) \left(\begin{array}{ccccccc} 1 & \alpha_{12} & \alpha_{13} & . & . & . & \alpha_{1k} \\ 0 & 1 & \alpha_{23} & . & . & . & \alpha_{2k} \\ 0 & 0 & 1 & & & & . \\ . & . & . & . & & & . \\ . & . & . & & . & & . \\ . & . & . & & & . & . \\ 0 & 0 & 0 & . & . & . & 1 \end{array} \right) . \quad (1.5)$$

Temos aqui A factorizada como o produto de uma matriz $n \times k$ com colunas ortogonais U por uma matriz $k \times k$ triangular superior T com elementos diagonais iguais a um. Para concluir a demonstração resta modificar as

colunas de U. Designemos por D a matriz diagonal $k \times k$ cujos elementos diagonais, respectivamente, são $\|v_1^*\|, \|v_2^*\|, \ldots, \|v_k^*\|$. Então. D é invertível porque os vectores v_j^* são não nulos. Desta forma, as matrizes $\hat{Q} = UD^{-1}$ e $\hat{R} = DT$ satisfazem as condições enunciadas no teorema e tem-se $A = \hat{Q}\hat{R}$.

■

De acordo com [9], muitos autores referem-se à factorização do Corolário 1.3.2 como factorização *"rectangular"* por se poder considerar $n > k$. Se a matriz A é quadrada, caso que nos vai interessar no próximo capítulo, demonstramos no seguinte teorema a unicidade da Factorização $\hat{Q}\hat{R}$.

Teorema 1.3.2 (Factorização "quadrada" QR)**.** *Para cada matriz não singular $A \in \mathbb{R}^{n\times n}$ existe uma única matriz ortogonal Q e uma única matriz triangular superior R, com elementos diagonais positivos, tal que*

$$A = QR\,.$$

Demonstração: Pelo Corolário 1.3.2, apenas necessitamos de provar a unicidade. Consideremos duas factorizações de uma matriz não singular A, ou seja

$$A = Q_1R_1 = Q_2R_2.$$

Então, como as matrizes Q_1 e Q_2 são ortogonais e R_1 e R_2 são triangulares superiores, consideremos

$$U = Q_2^TQ_1 = R_2R_1^{-1}.$$

A matriz $R_2R_1^{-1}$ é ainda uma matriz triangular superior com elementos diagonais positivos, uma vez que a inversa de uma matriz triangular superior, com elementos diagonais positivos $d_1, d_2, \ldots, d_n$, é ainda uma matriz triangular superior com elementos diagonais $d_1^{-1}, d_2^{-1}, \ldots, d_n^{-1}$ (cf. [9] pág. 122) e

o produto de matrizes triangulares superiores é ainda uma matriz triangular superior. Por outro lado, a matriz $Q_2^T Q_1$ é ainda ortogonal (cf. [9] pág. 336) uma vez que resulta do produto de matrizes ortogonais.

Então, a matriz U é uma matriz triangular com colunas ortonormadas e elementos diagonais positivos. Por isso, U terá de ser uma matriz com a seguinte forma

$$U = \begin{pmatrix} u_{11} & 0 & \dots & 0 \\ 0 & u_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{pmatrix} = \left(\begin{array}{c|c|c|c} u_{*1} & u_{*2} & \dots & u_{*n} \end{array} \right),$$

onde u_{*j} representa a *j-ésima* coluna de U, com $1 \leq j \leq n$. Como

$$1 = \|u_{*j}\| = |u_{jj}| = u_{jj}\,,$$

concluímos que U é a matriz identidade I. Então

$$Q_1 = Q_2 \;\text{ e }\; R_1 = R_2\,.$$

■

Uma importante consequência do Teorema 1.3.2 é a de que a decomposição $A = QR$ define completamente a Sequência de Gram-Schmidt.

No próximo corolário mostramos que uma matriz A, não singular, pode ser factorizada de uma outra forma.

Corolário 1.3.3 (Factorização UT)**.** *Para cada matriz não singular $A \in \mathbb{R}^{n\times n}$ existe uma única matriz U de vectores coluna ortogonais e uma única matriz triangular superior T, com diagonal unitária, tal que*

$$A = UT\,.$$

Demonstração: De acordo com (1.5) precisamos apenas de provar a unicidade. Considerem-se, então, duas factorizações de uma matriz A não singular

$$A = UT \quad \text{e} \quad A = U'T' \,,$$

onde U e U' são matrizes com vectores coluna ortogonais e T e T' são matrizes triangulares superiores com diagonal unitária. Designemos por D e E as matrizes diagonais cujos elementos diagonais correspondem, respectivamente, à norma das colunas de U e U'. Então D e E são invertíveis e tem-se

$$\begin{aligned} A = UD^{-1}DT &= U'T' \\ &= U'E^{-1}ET' \,, \end{aligned}$$

logo, pelo Teorema 1.3.2, tem-se $UD^{-1} = U'E^{-1}$ e também

$$DT = ET' \,.$$

Como T e T' têm diagonal unitária e D e E são matrizes diagonais, então os elementos da diagonal principal de DT e ET' coincidem com os de D e E, então

$$D = E \,,$$

e consequentemente

$$U = U' \quad \text{e} \quad T = T'.$$

■

Então, de acordo com o Teorema 1.3.2 e Corolário 1.3.3, podemos dizer que para qualquer matriz não singular A existem matrizes U, D, T, Q e R únicas, com

$$\begin{aligned} A &= UT \\ &= \underbrace{UD^{-1}}_{Q}\underbrace{DT}_{R} \\ &= QR \,, \end{aligned} \tag{1.6}$$

em que U é uma matriz com vectores ortogonais, D é uma matriz diagonal cujos elementos diagonais são, respectivamente, a norma dos vectores coluna de U, T é uma matriz triangular superior com diagonal unitária, Q é uma matriz ortogonal e R é uma matriz triangular superior com os elementos diagonais positivos.

A finalizar esta secção, demonstramos mais um dos resultados que iremos utilizar no desenvolvimento do Algoritmo LLL.

Corolário 1.3.4. *Seja A uma matriz quadrada de ordem n com vectores coluna $v_1, \ldots, v_n$ e $B \in \mathbb{R}^+$ tal que todos os elementos de A são no máximo iguais a B em valor absoluto. Então:*

i) $|\det A| \leq \|v_1\|_2 \ldots \|v_n\|_2$ *(Desigualdade de Hadamard).*

ii) $|\det A| \leq n^{1/2} B^n$.

Demonstração: Podemos assumir que A é uma matriz não singular e que $v_1, \ldots, v_n$ são linearmente independentes, pois caso não o sejam a desigualdade é imediata. Seja $\{v_1^*, \ldots, v_n^*\}$ a base que resulta de $\{v_1, \ldots, v_n\}$ pela OGS relativamente ao produto interno euclidiano, no caso em $k = n$ por (1.5) tem-se

$$\left(\, v_1 \, | \quad \ldots \, | \quad v_n \, \right) = \left(\, v_1^* \, | \quad \ldots \, | \quad v_n^* \, \right) T$$

onde T é $n \times n$ triangular superior com elementos iguais a um na diagonal principal. Então

$$\begin{aligned} \det \left(\, v_1 \, | \quad \ldots \, | \quad v_n \, \right) &= \det \left[\left(\, v_1^* \, | \quad \ldots \, | \quad v_n^* \, \right) T \right] \\ &= \det \left(\, v_1^* \, | \quad \ldots \, | \quad v_n^* \, \right) \det T \\ &= \det \left(\, v_1^* \, | \quad \ldots \, | \quad v_n^* \, \right) . \end{aligned} \tag{1.7}$$

Então, de acordo com (1.7), tem-se

$$\begin{aligned}\left|\det\left(\begin{array}{c|c|c|c} v_1 & \dots & v_n \end{array}\right)\right| &= \left|\det\left(\begin{array}{c|c|c|c} v_1^* & \dots & v_n^* \end{array}\right)\right| \\ &= \|v_1^*\|_2 \dots \|v_n^*\|_2 \\ &\leq \|v_1\|_2 \dots \|v_n\|_2 .\end{aligned}$$

A segunda desigualdade resulta do facto de $\|v_i\|_2 \leq n^{1/2}B$, para todo $i = 1, \dots, n$.

■

De notar que, na Desigualdade de Hadamard, a igualdade ocorre quando $v_1, \dots, v_n$ são ortogonais. Geometricamente significa que o volume de um paralelipípedo é inferior ao produto dos comprimentos das arestas que o definem.

1.4 Actualização da decomposição QR

Dando continuidade à última secção, vejamos agora de que forma é possível actualizar a decomposição $A = QR$ (ou $A = UT$), de uma matriz A' que resulta de uma matriz A não singular, quando efectuamos uma das seguintes operações elementares por colunas:

a) Permuta de duas colunas sucessivas;

b) Adição a uma coluna i de uma coluna j previamente multiplicada por um escalar, com $j < i$.

Consideremos o primeiro caso

a) Permuta de duas colunas sucessivas

Seja $A = \left(\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array} \right)$ uma matriz não singular e A' a matriz que resulta de A por troca de duas colunas v_i, v_{i+1} de A. Se $A = QR$ e $P_{i,i+1} \in \mathbb{R}^{n\times n}$ é uma matriz elementar de permutação, então

$$A' = AP_{i,i+1} = QRP_{i,i+1}\,,$$

ou seja, a partir da decomposição

$$A = Q \underbrace{\begin{pmatrix} r_{11} & \ldots & r_{1i} & r_{1,i+1} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & r_{ii} & r_{i,i+1} & \ldots & r_{in} \\ 0 & \ldots & 0 & r_{i+1,i+1} & \ldots & r_{i+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & r_{nn} \end{pmatrix}}_{R},$$

obtemos

$$A' = Q \underbrace{\begin{pmatrix} r_{11} & \ldots & r_{1,i+1} & r_{1i} & \ldots & r_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & r_{i,i+1} & r_{ii} & \ldots & r_{in} \\ 0 & \ldots & r_{i+1,i+1} & 0 & \ldots & r_{i+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & r_{nn} \end{pmatrix}}_{W}, \quad \begin{array}{l} \\ \\ \\ \\ \\ \rightarrow \textit{linha } i \end{array} \qquad \begin{array}{c} \textit{coluna } i \\ \downarrow \\ \\ \\ \\ \\ \\ \end{array} \tag{1.8}$$

em que $r_{ii}, r_{i+1,i+1} > 0$. A matriz W não é triangular superior e a igualdade (1.8) não é a decomposição QR de A'. No entanto, podemos actualizar esta decomposição fazendo

$$A' = QG^{-1}GW\,,$$

em que G é uma Matriz de Givens ortogonal (cf. [9] pág. 333) da forma

$$G = \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \overset{\textit{coluna } i}{\downarrow} \\ \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & c & s & \dots & 0 \\ 0 & \dots & -s & c & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array}, \quad \rightarrow \textit{linha } i \tag{1.9}$$

com

$$c = \frac{r_{i,i+1}}{\sqrt{r_{i,i+1}^2 + r_{i+1,i+1}^2}} \quad \text{e} \quad s = \frac{r_{i+1,i+1}}{\sqrt{r_{i,i+1}^2 + r_{i+1,i+1}^2}}\,.$$

Consequentemente, obtemos a matriz

$$GW = \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \overset{\textit{coluna } i}{\downarrow} \\ \begin{pmatrix} r_{11} & \dots & r_{1,i+1} & r_{1i} & \dots & r_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & \sqrt{r_{i,i+1}^2 + r_{i+1,i+1}^2} & \frac{r_{i,i+1}r_{ii}}{\sqrt{r_{i,i+1}^2+r_{i+1,i+1}^2}} & \dots & r'_{in} \\ 0 & \dots & 0 & -\frac{r_{i+1,i+1}r_{ii}}{\sqrt{r_{i,i+1}^2+r_{i+1,i+1}^2}} & \dots & r'_{i+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & r_{nn} \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array}, \quad \rightarrow \textit{linha } i$$

que não tem diagonal positiva. No entanto, podemos multiplicar GW por uma matriz elementar $E = D_{i+1,i+1}(-1)$, que coincide com a matriz identidade excepto na componente $(i+1, i+1)$ que é igual a -1, de maneira a que a matriz IGW seja triangular superior com diagonal positiva. A matriz $QG^{-1}I^{-1}$ é ortogonal, dado que a inversa e o produto de matrizes ortogonais

é ainda uma matriz ortogonal (cf. [9] pág. 336). Portanto a factorização QR da matriz A' é

$$A' = \underbrace{QG^{-1}E^{-1}}_{Q'}\underbrace{EGW}_{R'}.$$

Se quisermos actualizar a Factorização UT de A' basta ter em conta em (1.6) que $U = QD$ e $T = D^{-1}R$.

Consideremos agora o segundo caso:

b) Adição a uma coluna i de uma coluna j previamente multiplicada por um escalar, com $j < i$

Seja A' a matriz que resulta de A por adição a uma coluna i de uma outra coluna j previamente multiplicada por um escalar λ, com $j < i$. Sejam $A = QR$ e $A' = Q'R'$ as respectivas factorizações QR. Seja $E = E_{ij}(\lambda) \in \mathbb{R}^{n\times n}$ uma matriz elementar.
A igualdade $A' = AE$ é uma outra forma de dizer que a coluna v_i é substituída pela coluna $v_i + \lambda v_j$. Logo, $A' = QRE$ e como $j < i$, então a componente r_{ii} da matriz R é substituída por $r_{ii} + \lambda r_{ij} = r_{ii}$, uma vez que a matriz R é triangular superior e $r_{ki} = 0$ para todo $i \leq k \leq n$. Portanto, a matriz RE é ainda uma matriz triangular superior com elementos diagonais positivos. Assim, pela unicidade da factorização QR do Teorema 1.3.2, tem-se que

$$R' = RE, \quad e \quad Q = Q'$$

Portanto, é possível somar a um vector um múltiplo inteiro de um outro sem que a Sequência de Gram-Schmidt referida na página 12 se altere.
Como consequência, facilmente se verifica que a sequência de vectores ortogonais $\{v_1^*, \ldots, v_n^*\}$, que resultam da base $\{v_1, \ldots, v_n\}$ pela OGS referida no Teorema 1.3.1, também não se altera se a um deles somarmos um segundo

vector previamente multiplicado por um inteiro e cuja posição na referida sequência é inferior à do primeiro.

1.5 Base reduzida

Nesta secção apresentamos a definição de base reduzida do espaço $\mathbb{R}^n$, que estabelece a ligação entre a norma de um vector que resulta do processo de Ortogonalização de Gram-Schmidt e a norma de um vector de um reticulado. Para uma melhor compreensão desta ligação veja-se a Secção 2.1. Para já limitamo-nos à seguinte definição e respectiva consequência.

Definição 1.5.1 (Base Reduzida)**.** *Seja $\{v_1, \ldots, v_n\}$ uma base (ordenada) de $\mathbb{R}^n$ e $\{v_1^*, \ldots, v_n^*\}$ a correspondente base ortogonal que resulta da OGS. Dizemos que $\{v_1, \ldots, v_n\}$ é uma* ***base reduzida*** *se*

$$\|v_i^*\|^2 \leq 2\|v_{i+1}^*\|^2 \text{ para } 1 \leq i < n\,.$$

Se $\{v_1, v_2, \ldots, v_n\}$ for uma base reduzida e $\{v_1^*, v_2^*, \ldots, v_n^*\}$ a correspondente base ortogonal (ordenada), que resulta da OGS, então, para $j = 2, 3, \ldots, n$, obtemos

$$\|v_1^*\| \leq 2^{(j-1)/2}\|v_j^*\|. \tag{1.10}$$

De modo análogo, fixando o vector v_k^*, para $j = 2, 3, \ldots, n$ e $k \geq j$, obtemos

$$\|v_j^*\|^2 \leq 2^{(k-j)}\|v_k^*\|^2\,. \tag{1.11}$$

Como consequência apresentamos a seguinte proposição.

Proposição 1.5.1. *Seja $\{v_1, \ldots, v_n\}$ uma base reduzida de $\mathbb{R}^n$, e seja $\{v_1^*, v_2^*, \ldots, v_n^*\}$ a correspondente base ortogonal que resulta da OGS. Então,*

para todo $x \in \mathbb{R}^n$ *tal que* $x = v_n^* + a_{n-1}v_{n-1}^* + \ldots + a_1 v_1^*$ *com* $a_1, a_2, \ldots, a_{n-1} \in \mathbb{R}$*, tem-se*

$$\|x\|^2 \leq \|v_n^*\|^2 \left[1 + b^2(2^n - 2)\right] ,$$

sendo $b = \max\{|a_i|,\, i = 1, \ldots, n\}$.

Demonstração: De acordo com o enunciado tem-se

$$\begin{aligned} \|x\|^2 = \left\| v_n^* + \sum_{i=1}^{n-1} a_i v_i^* \right\|^2 &= \|v_n^*\|^2 + |a_i|^2 \sum_{i=1}^{n-1} \|v_i^*\|^2 \\ &\leq \|v_n^*\|^2 + b^2 \sum_{i=1}^{n-1} \|v_i^*\|^2 . \end{aligned}$$

Então, de acordo com 1.11, verifica-se

$$\begin{aligned} \|x\|^2 &\leq \|v_n^*\|^2 + b^2 \sum_{i=1}^{n-1} 2^{n-i} \|v_n^*\|^2 \\ &= \|v_n^*\|^2 \left(1 + b^2 \sum_{i=1}^{n-1} 2^{n-i}\right) \\ &= \|v_n^*\|^2 \left[1 + b^2(2^{n-1} + 2^{n-2} + \ldots + 2)\right] \\ &= \|v_n^*\|^2 \left[1 + b^2 2(2^{n-1} - 1)\right] \\ &= \|v_n^*\|^2 \left[1 + b^2(2^n - 2)\right] . \end{aligned}$$

■

Em particular, se $|a_1, |a_2|, \ldots, |a_n| \leq 1/2$ na proposição anterior, então

$$\|x\| \leq 2^{(n-1)/2} \|v_n^*\| . \tag{1.12}$$

que será usada mais adiante.

Capítulo 2

Algoritmo LLL

Como foi referido no início da última secção, o conceito de base reduzida serve como elo de ligação entre uma base resultante da OGS e um vector de um reticulado $\mathcal{L}$. Assim, na primeira secção, definimos o conceito de reticulado e provamos que a sua norma não depende da escolha da base. Mostramos também a relação existente entre a norma de qualquer vector $x \in \mathcal{L}$, a norma do primeiro vector de uma base reduzida e a norma dos vectores $v_1^*, \ldots, v_n^*$ que resultam da OGS. Note-se que esta norma é definida por um produto interno não especificado.

Na segunda secção, e como motivação para o Algoritmo LLL, explica-se o Método de redução de Gauss em $\mathbb{R}^2$ e a necessidade de recorrer a uma base, de um reticulado, "mais reduzida" que a original.

Na terceira secção, expomos o Algoritmo LLL ou Método de redução de base para o reticulado $\mathbb{Z}^n$ e a sua aplicação para qualquer outro reticulado. Este método permite encontrar em tempo polinomial, uma base reduzida para qualquer reticulado. Com efeito provamos o seu término, ao fim de no máximo $n^2(log_2 n + log_2 T) log_{\frac{4}{3}} 2$ iterações e que o tamanho dos números intermediários são limitados polinomialmente.

Na última secção, apresentamos dois exemplos de execução do Algoritmo LLL, em $\mathbb{R}^2$. No primeiro apresentamos todos os cálculos efectuados. No segundo recurremos a uma implementação do Algoritmo LLL, que podemos encontrar na biblioteca de estruturas e algoritmos NTL - Number Theory Library.

2.1 Definições

No que segue (cf. [5] e [16]) apresentam-se as definições de reticulado e respectiva "medida". Os exemplos foram retirados de [1].

Definição 2.1.1. *Seja* $B = \begin{pmatrix} v_1 & | \ldots | & v_k \end{pmatrix}$ *a matriz com vectores coluna* $v_1, \ldots, v_k \in \mathbb{R}^n$*, linearmente independentes. O* ***reticulado*** *gerado por* $v_1, \ldots, v_k$ *é o conjunto*

$$\mathcal{L} = B\mathbb{Z}^k = \{x \in \mathbb{R}^n : x = a_1 v_1 + \ldots + a_k v_k \, , a_i \in \mathbb{Z}, i = 1, \ldots, k\} \, .$$

O conjunto $\{v_1, \ldots, v_k\}$ *diz-se uma* ***base de*** $\mathcal{L}$*.*

Exemplo 2.1.1. Seja $\mathbb{Z}^n \subset \mathbb{R}^n$ o conjunto de pontos com coordenadas inteiras:

$$\mathbb{Z}^n = \{(a_1, \ldots, a_n) : a_i \in \mathbb{Z}, i = 1, \ldots, n\} \, .$$

O reticulado $\mathbb{Z}^n$ denomina-se **reticulado padrão** - ver Figura 2.1 a).

Exemplo 2.1.2. O conjunto

$$A_2 = \left\{x \in \mathbb{R}^2 : x = a_1(1, 1/2) + a_2(1, -1/2), \, a_1, a_2 \in \mathbb{Z}\right\} ,$$

como é fácil de verificar é um reticulado - ver Figura 2.1 b)

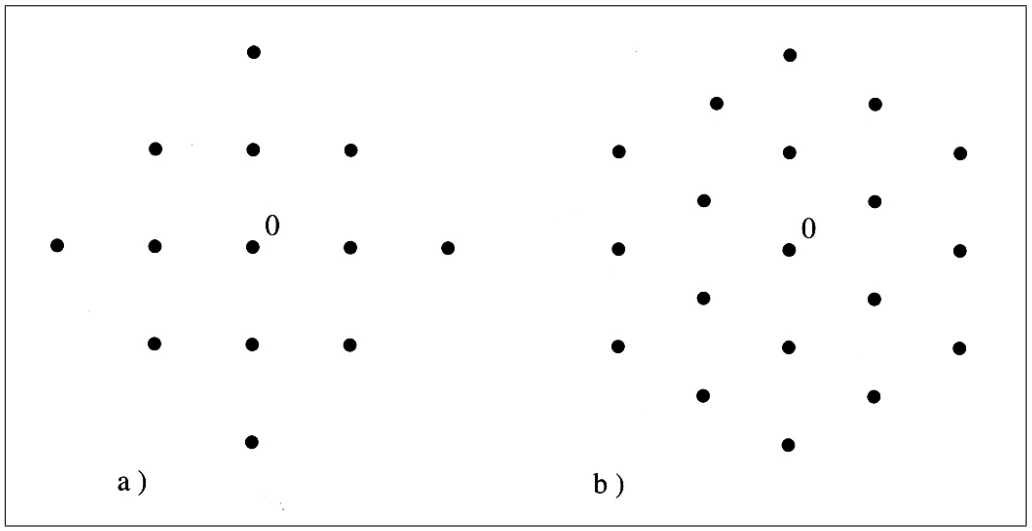

Figura 2.1: a) Reticulado $\mathbb{Z}^2$ b) Reticulado A_2

Definição 2.1.2. *Seja $\mathcal{L} = B\mathbb{Z}^k$ um reticulado de $\mathbb{R}^n$ com base $v_1, \ldots, v_k$. Então, o* ***determinante (ou norma) de*** *$\mathcal{L}$,* $\det \mathcal{L}$*, é definido por*

$$\det \mathcal{L} \equiv \sqrt{\det(B^T B)}.$$

Em particular, se $\mathcal{L}$ tem dimensão completa (i.e. $n = k$), então

$$\det \mathcal{L} \equiv |\det B|,$$

e pode ser interpretado como o volume do paralelipípedo n-dimensional definido por

$$P(B) = P(\mathcal{L}) = \{\lambda_1 v_1 + \ldots + \lambda_n v_n : 0 \leq \lambda_i < 1, i = 1, \ldots, n\},$$

que é, por vezes, referido como o ***paralelipípedo fundamental*** *de $\mathcal{L}$ na base $\{v_1, \ldots, v_n\}$.*

O mesmo reticulado pode ser gerado por diferentes bases, por isso, coloca-se naturalmente a questão da norma estar bem definida. Veja-se a Figura 2.2.

Lema 2.1.1 ([15]). *Se A' é uma matriz, de característica completa por colunas, que resulta de uma outra matriz A por sucessivas operações elementares*[1] *por colunas, então A e A' geram o mesmo reticulado se e só se existe U, uma matriz quadrada de inteiros e unimodular* [2]*, tal que*

$$A' = AU.$$

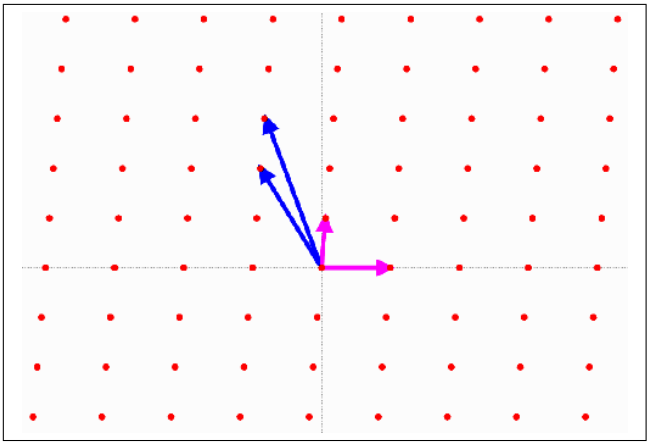

Figura 2.2: O mesmo reticulado gerado por bases diferentes

Portanto, se A e B são matrizes $n \times k$, cujos vectores coluna geram o mesmo reticulado $\mathcal{L}$, de acordo com o Lema 2.1.1, existe uma matriz $k \times k$ de inteiros unimodular U tal que $A = BU$. Então

$$\begin{aligned} \det(A^T A) &= \det(U^T B^T B U) \\ &= \det(U^T)\det(B^T B)\det(U) \\ &= \det(B^T B), \end{aligned}$$

o que significa que o determinante de um reticulado $\mathcal{L} = B\mathbb{Z}^k$, em $\mathbb{R}^n$, não depende da escolha da base B. Veja-se a Figura 2.3.

[1] Permuta de duas colunas; Multiplicação dos elementos de uma coluna por -1; Adição de um múltiplo inteiro de uma coluna a uma outra coluna.

[2] Uma matriz quadrada U diz-se unimodular se $|\det U| = 1$, cf. [14] pág. 48

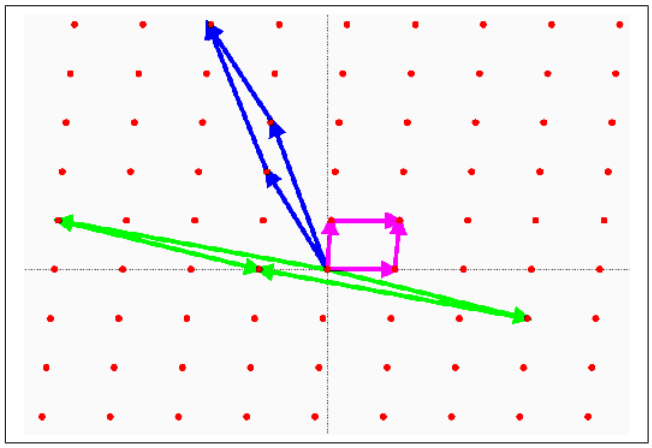

Figura 2.3: Diferentes bases definem paralelipípedos com o mesmo volume

Outra importante propriedade de um reticulado (cf. [1] pág. 286) decorre do próximo lema.

Lema 2.1.2. *Seja $\mathcal{L} \subset \mathbb{R}^n$ um reticulado e π um paralelipípedo fundamental de $\mathcal{L}$. Então, para cada $x \in \mathbb{R}^n$, existem vectores únicos $v \in \mathcal{L}$ e $y \in \pi$ tal que*

$$x = v + y.$$

Demonstração: Suponhamos que o paralelipípedo π tem uma base $\{v_1, \ldots, v_n\}$ de $\mathcal{L}$. Então $\{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$ e $x \in \mathbb{R}^n$ pode ser escrito como

$$x = \alpha_1 v_1 + \ldots + \alpha_n v_n,$$

para $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Seja

$$v = \sum_{i=1}^{n} \lfloor \alpha_i \rfloor v_i \quad \text{e} \quad y = \sum_{i=1}^{n} \{\alpha_i\} v_i,$$

onde $\lfloor . \rfloor$ e $\{.\}$ representam, respectivamente a parte inteira e decimal de um número. É evidente que $v \in \mathcal{L}$, $y \in \pi$ e, portanto,

$$x = v + y.$$

Suponhamos que existem duas decomposições $x = u_1 + y_1$ e $x = u_2 + y_2$, com $u_1, u_2 \in \mathcal{L}$ e $y_1, y_2 \in \pi$. Então

$$y_1 = \sum_{i=1}^{n} \alpha_i v_i \quad \text{e} \quad y_2 = \sum_{i=1}^{n} \beta_i v_i,$$

com $0 \leq \alpha_i, \beta_i < 1$, para $i = 1, \ldots, n$. Então

$$u_1 - u_2 = y_2 - y_1 = \sum_{i=1}^{n} \gamma_i v_i,$$

para $\gamma_i = \alpha_i - \beta_i$. Note-se que $|\gamma_i| < 1$ para $i = 1, \ldots, n$ e que $u_1 - u_2 \in \mathcal{L}$. Como $\{v_1, \ldots, v_n\}$ é uma base de $\mathcal{L}$, os números γ_i são inteiros. Uma vez que $|\gamma_i| < 1$, concluímos que $\gamma_i = 0$ e que $\alpha_i = \beta_i$ para $i = 1, \ldots, n$. Por isso,

$$y_2 = y_1 \quad \text{e} \quad u_1 = u_2$$

. ■

Corolário 2.1.1 ([1])**.** *Seja $\mathcal{L} \in \mathbb{R}^n$ um reticulado e π um paralelipípedo fundamental de $\mathcal{L}$. Então, $\mathbb{R}^n$ é a união disjunta de $\{\pi + v : v \in \mathcal{L}\}$.*

No próximo teorema, cuja demonstração pode ser vista em [1] pág. 287-288, consideramos $B(0, r) = \{x \in \mathbb{R}^n : \|x\| \leq r\}$ - bola fechada centrada na origem e raio r. O número de pontos de um reticulado $\mathcal{L} \in \mathbb{R}^n$ em $B(0, r)$ é representado $|B(0, r) \cap \mathcal{L}|$.

Teorema 2.1.1. *Seja $\mathcal{L} \in \mathbb{R}^n$ um reticulado. Então*

$$\lim_{r \to +\infty} \frac{\operatorname{vol} B(0, r)}{|B(0, r) \cap \mathcal{L}|} = \det \mathcal{L}.$$

De acordo com este teorema, o $\det \mathcal{L}$ pode ser interpretado como o "volume por ponto do reticulado".

O próximo resultado estabelece um limite inferior para a norma de qualquer vector não nulo de um reticulado, limite esse que pode ser calculado em $\mathcal{O}(k^2)$ operações.

Teorema 2.1.2. *Seja $\mathcal{L} \subseteq \mathbb{R}^n$ um reticulado com base $\{v_1, \ldots, v_k\}$ e seja $\{v_1^*, \ldots, v_k^*\}$ a base que resulta de $\{v_1, \ldots, v_k\}$ pelo processo de OGS. Então para todo $x \in \mathcal{L} \setminus \{0\}$ tem-se*

$$\|x\| \geq \min \{\|v_1^*\|, \ldots, \|v_k^*\|\} \ . \tag{2.1}$$

Demonstração: Seja $x = \sum_{j=1}^{k} \lambda_j v_j \in \mathcal{L} \setminus \{0\}$, com $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$, e seja m o maior índice tal que $\lambda_m \neq 0$. Suponhamos que as bases $\{v_1, \ldots, v_k\}$ e $\{v_1^*, \ldots, v_k^*\}$ estão relacionadas através de

$$v_j = \sum_{i=1}^{j} u_{ij} v_i^* \ , \quad j = 1, \ldots, k.$$

Então

$$\begin{aligned} x &= \sum_{j=1}^{m} \lambda_j \left(\sum_{i=1}^{j} u_{ij} v_i^* \right) \\ &= \lambda_m v_m^* + \sum_{i=1}^{m-1} a_i v_i^* \ , \end{aligned}$$

para algum $a_i \in \mathbb{R}$. Então

$$\begin{aligned} \|x\|^2 &= \left\langle \lambda_m v_m^* + \sum_{i=1}^{m-1} a_i v_i^*, \lambda_m v_m^* + \sum_{i=1}^{m-1} a_i v_i^* \right\rangle \\ &= \lambda_m^2 \langle v_m^*, v_m^* \rangle + \sum_{i=1}^{m-1} a_i^2 \langle v_i^*, v_i^* \rangle \\ &\geq \lambda_m^2 \|v_m^*\|^2 \geq \|v_m^*\|^2 \\ &\geq \min \left\{ \|v_1^*\|^2, \ldots, \|v_k^*\|^2 \right\} . \end{aligned}$$

■

Note-se que (2.1) pode também ser interpretado como um limite inferior para a norma do **vector mais curto** de um reticulado - vector não nulo de norma mínima de um reticulado.
No caso da base $\{v_1, \ldots, v_k\}$ ser reduzida, podemos concretizar o seguinte resultado:

Teorema 2.1.3. *Se $\{v_1, \ldots, v_n\}$ é uma base reduzida para o reticulado $\mathcal{L} \subseteq \mathbb{R}^n$ então, para todo $x \in \mathcal{L} \setminus \{0\}$, tem-se*

$$\|v_1\| \leq 2^{(n-1)/2}\|x\|\,. \tag{2.2}$$

Demonstração: Pelo Teorema 2.1.2, tem-se

$$\|x\| \geq \min\{\|v_1^*\|, \|v_2^*\|, \ldots, \|v_n^*\|\}.$$

Por (1.10), uma vez que $\{v_1, \ldots, v_n\}$ é uma base reduzida (ver página 22), tem-se

$$\begin{aligned} \|x\| &\geq \min\{\|v_1^*\|, 2^{-\frac{1}{2}}\|v_1^*\|, \ldots, 2^{-(n-1)/2}\|v_1^*\|\} \\ &\geq 2^{-(n-1)/2}\|v_1^*\| \\ &= 2^{-(n-1)/2}\|v_1\|. \end{aligned}$$

■

Note-se que a norma usada em (2.1) e (2.2) é a norma definida por um produto interno não especificado. As bases ortogonais referidas são ortogonais de acordo com esse produto interno não especificado.

2.2 Redução de Gauss (dimensão 2)

Gauss [3] propôs um procedimento que transforma qualquer base de um reticulado numa outra, onde o primeiro vector é o mais curto desse reticulado.

Este procedimento é conhecido por **Redução de Gauss em** $\mathbb{R}^2$ e está sumariado na Figura 2.4.

Embora todas as bases de um reticulado de dimensão completa tenham o mesmo determinante, nem todas são "equivalentes" do ponto de vista prático. Veja-se o caso de $\mathbb{Z}^2$, gerado pela base canónica de $\mathbb{R}^2$ ou por $\{(3,2),(2,1)\}$. No entanto, $\{(1,0),(0,1)\}$ é uma base com a qual é mais fácil trabalhar do que $\{(3,2),(2,1)\}$, uma vez que a primeira é formada por vectores linearmente independentes de $\mathbb{Z}^2$ de menor comprimento.

No que se segue, mostramos propriedades do método de Redução de Gauss em $\mathbb{R}^2$, conforme [2] pág. 6.

Lema 2.2.1. *Sejam $b_1, b_2 \in \mathbb{R}^2$ vectores linearmente independentes, então*

$$\min\{\|b_2 - tb_1\|^2 : t \in \mathbb{Z}\} = \|b_2 - \lfloor \mu \rceil b_1\|^2 ,$$

em que

$$\mu = \frac{\langle b_1, b_2 \rangle}{\|b_1\|^2}. \tag{2.3}$$

Demonstração: Seja $b_2^* = b_2 - \mu b_1$. Então, $b_2^* \perp b_1$ e, por isso, para todo o $t \in \mathbb{Z}$ tem-se

$$\begin{aligned} \|b_2 - tb_1\|^2 &= \|b_2 - b_2^* - tb_1 + b_2^*\|^2 \\ &= \|\mu b_1 - tb_1 + b_2^*\|^2 \\ &= \|b_2^*\|^2 + (\mu - t)^2 \|b_1\|^2 , \end{aligned}$$

valor este que é mínimo quando $t = \lfloor \mu \rceil$.

Proposição 2.2.1. *O algoritmo da Figura 2.4 termina ao fim de um número finito de iterações. No final, b_1 é o vector não nulo do reticulado que tem norma mínima e $\{b_1, b_2\}$ é uma base (ordenada) reduzida.*

Input: $\{b_1, b_2\}$, base (ordenada) de um reticulado em $\mathbb{R}^2$, tal que $\|b_2\| \geq \|b_1\|$.

Output: $\{b_1, b_2\}$, base (ordenada) reduzida do mesmo reticulado, tal que b_1 é o vector não nulo de norma mínima.

Iteração:

repetir

$t \leftarrow \left\lfloor \frac{\langle b_1, b_2 \rangle}{\|b_1\|^2} \right\rceil$;

$b_1 \leftarrow b_2 - t b_1$;

$b_2 \leftarrow b_1$;

até $\|b_1\| \geq \|b_2\|$.

troca b_1, b_2.

Figura 2.4: Algoritmo de Redução de Gauss em $\mathbb{R}^2$

Demonstração: Seja $\{b_1, b_2\}$ uma base (ordenada) de um reticulado $\mathcal{L}$ em $\mathbb{R}^2$. Sem perda de generalidade consideremos que $\|b_1\| \leq \|b_2\|$. Se

$$v = b_2 - t b_1 \,,$$

para algum $t \in \mathbb{Z}$, então $\{b_1, v\}$ é ainda uma base para o reticulado $\mathcal{L}$. Seja, em particular, v definido por $t = \lfloor \mu \rceil$ com μ definido por (2.3) e seja b_2^* o vector perpendicular a b_1 e que resulta da OGS, então

$$b_2^* = b_2 - \mu b_1 \,. \tag{2.4}$$

Portanto,

$$v = b_2^* - (\lfloor \mu \rceil - \mu) b_1 . \tag{2.5}$$

Note-se que, por (2.4) e (2.5),

$$\begin{aligned} \|b_2\|^2 &= \|b_2^*\|^2 + \mu^2 \|b_1\|^2 \,, \\ \|v\|^2 &= \|b_2^*\|^2 + (\mu - \lfloor \mu \rceil)^2 \|b_1\|^2 \,, \end{aligned}$$

e portanto $\|v\| \leq \|b_2\|$. Se $\|v\| < \|b_1\|$, então $b_1 \leftarrow v$ e $b_2 \leftarrow b_1$ e, tendo ocorrido uma redução em $\|b_1\|$, ocorre uma reiteração. Se $\|v\| \geq \|b_1\|$, então o algoritmo termina com b_1 inalterado e $b_2 = v$. No final tem-se

$$\begin{aligned}\left|\frac{\langle b_1, b_2\rangle}{\|b_1\|^2}\right| &= \left|\frac{\langle b_1, v\rangle}{\|b_1\|^2}\right| \\ &= \left|\frac{\langle b_1, b_2^* - (\lfloor\mu\rceil - \mu)b_1\rangle}{\|b_1\|^2}\right| \\ &= |\mu - \lfloor\mu\rceil| \frac{\|b_1\|^2}{\|b_1\|^2} \\ &= |\mu - \lfloor\mu\rceil| \leq \frac{1}{2}.\end{aligned}$$

Por isso,

$$\begin{aligned}\|b_1\|^2 \leq \|b_2\|^2 &= \left\|b_2^* + \frac{\langle b_1, b_2\rangle}{\|b_1\|^2} b_1\right\|^2 \\ &= \|b_2^*\|^2 + \left|\frac{\langle b_1, b_2\rangle}{\|b_1\|^2}\right|^2 \|b_1\|^2 \\ &\leq \|b_2^*\|^2 + \frac{1}{4}\|b_1\|^2,\end{aligned}$$

então,

$$\|b_1\|^2 = \|b_1^*\|^2 \leq \frac{4}{3}\|b_2^*\|^2 \leq 2\|b_2^*\|^2,$$

donde se conclui que $\{b_1, b_2\}$ é uma base reduzida.

O método termina ao fim de um número finito de iterações, porque a norma de b_1 vai decrescendo em cada passo e existe apenas um número finito de pontos de norma inferior ao vector inicial b_1.

Vamos agora mostrar que b_1 é o vector mais curto do reticulado.
Seja $u = xb_1 + yb_2 \neq 0$, com $x, y \in \mathbb{Z}$. Se $y = 0$ então $\|u\| = |x|\|b_1\| \geq \|b_1\|$.
Se $y \neq 0$, então

$$\|u\|^2 = \langle xb_2 + yb_1, xb_2 + yb_1\rangle$$

$$\begin{aligned} &= x^2\|b_2\|^2 + 2xy\langle b_1, b_2\rangle + y^2\|b_1\|^2 \\ &\geq (x^2+y^2)\|b_1\|^2 + 2xy\langle b_1, b_2\rangle \\ &\geq (x^2+y^2)\|b_1\|^2 - 2|x||y||\langle b_1, b_2\rangle| \\ &= \left(x^2+y^2-2|x||y|\left|\frac{\langle b_1, b_2\rangle}{\|b_1\|^2}\right|\right)\|b_1\|^2 \\ &\geq (x^2+y^2-|x||y|)\|b_1\|^2. \end{aligned} \tag{2.6}$$

Se $|x| \geq |y|$, então (2.6) é igual a

$$(|x|(|x|-|y|)+y^2)\|b_1\|^2 \geq \|b_1\|^2.$$

Se $|x| < |y|$, então (2.6) é igual a

$$(|y|(|y|-|x|)+x^2)\|b_1\|^2 \geq \|b_1\|^2.$$

■

Exemplo 2.2.1. Apliquemos o algoritmo de Redução de Gauss em $\mathbb{R}^2$ às colunas da matriz

$$A = \begin{pmatrix} 1 & 0 \\ 9677 & -12319 \end{pmatrix},$$

com $b_1 = (1, 9677)$ e $b_2 = (0, -12319)$. Na primeira iteração fazemos

$$v = b_2 - \lfloor \mu \rceil b_1,$$

com μ definido por (2.3). Neste caso

$$\lfloor \mu \rceil = -1,$$

e por isso,

$$v = b_2 + b_1 = (1, -2642).$$

Como

$$\|v\| \leq \|b_1\|,$$

fazemos uma nova iteração com $b_1 \leftarrow v = (1, -2642)$ e $b_2 \leftarrow b_1 = (1, 9677)$. Na segunda iteração fazemos novamente

$$v = b_2 - \lfloor \mu \rceil b_1,$$

com μ definido por (2.3). Neste caso

$$\lfloor \mu \rceil = -4,$$

e por isso,

$$v = b_2 + 4b_1 = (5, -891).$$

Como

$$\|v\| \leq \|b_1\|,$$

fazemos uma nova iteração com $b_1 \leftarrow v = (5, -891)$ e $b_2 \leftarrow b_1 = (1, -2642)$. Na terceira iteração fazemos novamente

$$v = b_2 - \lfloor \mu \rceil b_1,$$

com μ definido por (2.3). Neste caso

$$\lfloor \mu \rceil = 3,$$

e por isso,

$$v = b_2 - 3b_1 = (-14, 31).$$

Como

$$\|v\| \leq \|b_1\|,$$

fazemos uma nova iteração com $b_1 \leftarrow v = (-14, 31)$ e $b_2 \leftarrow b_1 = (5, -891)$. Na quarta iteração fazemos novamente

$$v = b_2 - \lfloor \mu \rceil b_1,$$

com μ definido por (2.3). Neste caso

$$\lfloor \mu \rceil = -24,$$

e por isso,

$$v = b_2 + 24b_1 = (-331, -147).$$

Como

$$\|v\| \geq \|b_1\|,$$

o algoritmo pára com $b_1 \leftarrow v = (-14, 31)$ e $b_2 \leftarrow b_1 = (-331, -147)$. O vector

$$b_1 = (-14, 31),$$

é o vector mais curto do reticulado gerado pelas colunas da matriz A.

De acordo com [2] página 6, o algoritmo de Redução de Gauss em $\mathbb{R}^2$ não é polinomial. Para garantirmos o seu fim em tempo polinomial deve-se substituir a condição "$\|b_1\| \geq \|b_2\|$" por

$$\|b_1\| \geq (1 - \epsilon)\|b_2\|,$$

onde ϵ é um número real, devidamente escolhido, inferior a um. Neste caso não garantimos que b_1 é o vector de norma mínima de $\mathcal{L}$, mas sabemos que o algoritmo termina ao fim de $\log_{(1-\epsilon)} \|b_1\|$ iterações, com o vector b_1 "razoavelmente"curto.

2.3 Método de redução de base

Seja $A = \begin{pmatrix} b_1 & | \ldots | & b_n \end{pmatrix}$ uma matriz não singular de ordem n e seja $\mathcal{L}$ o reticulado gerado pelas suas colunas. De acordo com a Desigualdade de Hadamard, que estudámos no Teorema 1.3.4, tem-se

$$\det \mathcal{L} \leq \|b_1\|_2 \ldots \|b_n\|_2,$$

com igualdade se os vectores $b_1, \ldots, b_n$ são ortogonais. Naturalmente, esse limite superior para $\det \mathcal{L}$ depende da escolha da base. Se $\mathcal{L}$ possui uma base ortogonal, então esse limite superior é o melhor possível, mas nem todo o reticulado possui uma base ortogonal. Como encontrar, então, um "bom" limite superior para $\det \mathcal{L}$?

Hermite [1850] (cf. [14] pág. 67) mostrou que para cada n existe um número $c\,(n)$ tal que todo o reticulado $\mathcal{L} \in \mathbb{R}^n$, de dimensão completa, tem uma base $\{b_1, \ldots, b_n\}$ que verifica

$$\|b_1\|_2 \ \ldots \ \|b_n\|_2 \leq c\,(n) \det \mathcal{L}. \tag{2.7}$$

Hermite mostrou que (2.7) verifica-se com

$$c\,(n) = \left(\frac{4}{3}\right)^{n(n-1)/4}. \tag{2.8}$$

Minkowski [1896] (cf. [14] pág. 67) melhorou esse resultado ao mostrar que (2.7) é verificado com

$$c\,(n) = \frac{2^n}{\mathcal{V}(n)}, \tag{2.9}$$

onde $\mathcal{V}(n) = \frac{\pi^{n/2}}{\Gamma(1+n/2)}$ representa o volume da bola unitária de dimensão n e Γ é a função Gama. Contudo, não se conhece nenhum algoritmo polinomial que encontre uma base $\{b_1, \ldots, b_n\}$ que satisfaça (2.7) para $c(n)$ definido por (2.8) ou (2.9).

Lovász [1982](cf. [14] pág. 68) definiu um algoritmo polinomial que permite encontrar uma base que satisfaz (2.7), com $c(n)$ definido por

$$c\,(n) = 2^{n(n-1)/4}. \tag{2.10}$$

Esse algoritmo, que apresentamos na Figura 2.5, é usualmente descrito como **Algoritmo LLL** (Lenstra, Lenstra e Lovász [8]) ou **Método de redução de base**.

Input: $n \in \mathbb{N}$ e $D \in \mathbb{Q}^{n\times n}$ simétrica, definida positiva.

Output: $B \in \mathbb{Z}^{n\times n}$ cujas colunas formam uma base reduzida de $\mathbb{Z}^n$.

Inicialização: $B \leftarrow I \in \mathbb{Z}^{n\times n}$, matriz identidade de $\mathbb{Z}^{n\times n}$.

Iteração:

Passo 1: Aplicar o processo de ortogonalização de Gram-Shmidt sobre as colunas de B, relativamente ao produto interno $\langle x, y\rangle = x^\top Dy$.
No final $B = CV$ onde $V = [\lambda_{ji}]_{n\times n}$ é uma matriz triangular superior com diagonal unitária e C é uma matriz com colunas ortonais

Passo 2: Efectuar eliminação de Gauss sobre V usando exclusivamente multiplicadores inteiros, de modo a obter uma matriz triangular superior com diagonal unitária e cujos elementos não diagonais satisfazem $|\lambda_{ji}| \leq 1/2$ para todo $j < i$.
Faça-se $B \leftarrow BE$ e $V \leftarrow VE$, sendo E o produto das correspondentes matrizes elementares.

Passo 3: Seja i o menor i tal que $\|b_i^*\|^2 > 2\|b_{i+1}^*\|^2$, sendo b_i^* a i-ésima coluna de C. Se não existe, então, parar. Caso contrário, faz-se $B \leftarrow BP_{i+1,i}$ sendo $P_{i+1,i}$ uma matriz de permutação e voltar ao *Passo 1*.

Figura 2.5: Algoritmo LLL

Passamos a apresentar algumas explicações sobre o modo como os passos do Algoritmo LLL, descrito na Figura 2.5, são implementados.

- *Passo 1*: Seja D uma matriz quadrada de ordem n, definida positiva de números racionais. Sem perda de generalidade, podemos supor que D é uma matriz de inteiros, caso contrário multiplique-se D por um

inteiro apropriado (e.g., o produto dos denominadores dos elementos de D).

No final do primeiro passo, a matriz C é constituída pelas colunas $b_1^*, \ldots, b_n^*$, que resultam da OGS aplicada à base $b_1, \ldots, b_n$ e relativamente ao produto interno $\langle x, y \rangle \equiv x^\top D y$. Pela Observação 1.3.1, para cada $i = 1, 2, \ldots, n$, tem-se

$$b_i^* = b_i - B_{i-1} \left(B_{i-1}^T D B_{i-1} \right)^{-1} B_{i-1}^T D b_i, \tag{2.11}$$

em que B_{i-1} é a matriz com vectores coluna $b_1, \ldots, b_{i-1}$. Em particular, para cada $i = 1, \ldots, n$, b_i^* é o único vector que satisfaz

$$\begin{aligned} & b_i^* - b_i \in \mathcal{S}\{b_1, \ldots, b_{i-1}\}, \\ & \langle b_i^*, b_j \rangle = 0,\ j = 1, 2, \ldots, i-1. \end{aligned}$$

Além disso, para cada i, tem-se

$$b_i^* \ = \ b_i - \lambda_{1i} b_1^* - \ldots - \lambda_{i-1,i} b_{i-1}^*,$$

para $\lambda_{1i}, \ldots, \lambda_{i-1,i} \in \mathbb{R}$. Recorrendo à linguagem matricial e de acordo com o Teorema 1.3.2, a matriz $B = \begin{pmatrix} b_1 & | \ldots | & b_n \end{pmatrix}$ escreve-se de forma única

$$B = CV,$$

onde $C = \begin{pmatrix} b_1^* & | \ldots | & b_n^* \end{pmatrix}$ é a matriz com vectores coluna $b_1^*, \ldots, b_n^*$ resultantes do processo de OGS, e V é uma matriz triangular superior com elementos iguais a um na diagonal principal. Portanto

$$b_i = \lambda_{1i} b_1^* + \ldots + \lambda_{i-1,i} b_{i-1}^* + b_i^* \,,$$

para $\lambda_{1i}, \ldots, \lambda_{i-1,i} \in \mathbb{R}$ que são os componentes acima do elemento

diagonal de V na posição (i,i), ou seja,

$$\left(\begin{array}{c|c|c|c} b_1 & \dots & b_n \end{array} \right) = \left(\begin{array}{c|c|c|c} b_1^* & \dots & b_n^* \end{array} \right) \begin{pmatrix} 1 & \lambda_{12} & \dots & \lambda_{1n} \\ 0 & 1 & \dots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \quad (2.12)$$

Exemplo do Passo 1: Considere-se a matriz D simétrica definida positiva

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

e a matriz identidade $B = I \in \mathbb{R}^{3\times 3}$. Para este exemplo, (2.12) resulta no seguinte

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 1 & \frac{1}{2} & -\frac{2}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}}_{V}. \quad (2.13)$$

- *Passo 2*: O passo 2 é implementado do seguinte modo: para $i = 2, \dots, n$ e para cada $j = i-1, i-2, \dots, 1$

$$V \leftarrow V E_{ji} \left(- \lfloor \lambda_{ji} \rceil\right).$$

No final, e após efectuarmos $B \leftarrow BE$, sendo E o produto de matrizes elementares, temos uma decomposição $B = CV$, onde $V = [\lambda_{ji}]_{n\times n}$ é uma matriz triangular superior, com diagonal unitária, que satisfaz, para cada $i > j$, a seguinte propriedade

$$|\lambda_{ji}| \leq \frac{1}{2}. \quad (2.14)$$

Note-se que, pelo Corolário 1.3.3, quando o processo de Ortogonalização de Gram-Schmidt é aplicado às novas colunas de B, a matriz C é a mesma.

Exemplo do Passo 2: Vejamos agora um exemplo de implementação do segundo passo da matriz V em (2.13): primeiro, faça-se $i = 2$, $j = 1$

$$V^{(0)}E_{12}(-1) = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{2}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{2}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} = V^{(1)}.$$

Segundo, faça-se $i = 3$, $j = 2$

$$V^{(1)}E_{23}(1) = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{2}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{7}{6} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} = V^{(2)}.$$

Terceiro para $i = 3$, $j = 1$

$$V^{(2)}E_{13}(1) = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{7}{6} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{6} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} = V^{(3)}.$$

Então, $V \longleftarrow V^{(3)}$, $E \longleftarrow E_{12}(-1)E_{23}(1)E_{13}(1)$ e

$$B \longleftarrow BE = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

No final deste passo, (2.12) resulta no seguinte

$$\underbrace{\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{6} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}}_{V}. \qquad (2.15)$$

- *Passo 3* Neste terceiro e último passo, deverá escolher-se, para $i = 1, \ldots, n-1$, um índice i de modo que

$$\|b_i^*\|^2 > 2\|b_{i+1}^*\|^2. \tag{2.16}$$

Se não existir um tal i, então

$$\|b_i^*\|^2 \leq 2\|b_{i+1}^*\|^2, \qquad i = 1, 2, \ldots, n-1$$

e portanto, $\{b_1, b_2, \ldots, b_n\}$ é uma base reduzida - ver página 22.

Exemplo do Passo 3: Continuando o exemplo anterior, determinamos $\|b_i^*\|^2$, $i = 1, 2, 3$, relativamente ao produto interno $\langle x, y\rangle = x^\top Dy$, em que b_i^* representa a i-ésima coluna de C em (2.15). Neste caso

$$\|b_1^*\|^2 = b_1^{*T} D b_1^* = 2;$$
$$\|b_2^*\|^2 = b_2^{*T} D b_2^* = 3/2;$$
$$\|b_3^*\|^2 = b_3^{*T} D b_3^* = 44/9.$$

Portanto, as colunas de B em (2.15) formam uma base reduzida para o reticulado $\mathbb{Z}^3$.

No próximo teorema mostramos que o algoritmo da Figura 2.5 termina ao fim de um número finito de operações.

Teorema 2.3.1. *O algoritmo da Figura 2.5 termina no máximo ao fim de*

$$n^2(log_2 n + log_2 T) log_{\frac{4}{3}} 2$$

iterações, sendo T o máximo valor absoluto de todos os elementos de D. No final, obtém-se uma base reduzida que satisfaz $\|b_1\| \ldots \|b_n\| \leq 2^{n(n-1)/4}\sqrt{\det\, D}$, para $\|x\| = \sqrt{x^T D x}$.

Demonstração: Primeiro, mostramos que $|\det B|$ permanece inalterado durante todo o algoritmo. No final do *Passo 2*,

$$\det(BE) = \det(B)\det(E) = \det(B)$$

uma vez que, sendo E o produto de matrizes elementares, $\det(E) = 1$. No final do *Passo 3*,

$$\det(BP_{i,i+1}) = -\det(B).$$

Fica assim demonstrado que $|\det B|$ permanece sempre igual a um durante todo o algoritmo.

Agora mostramos que, em cada iteração e no final do *Passo 1*, tem-se

$$\|b_1^*\|^2 \dots \|b_i^*\|^2 = \det\left(B_i^\top D B_i\right). \tag{2.17}$$

Seja $C_i = \begin{pmatrix} b_1^* & | \dots | & b_i^* \end{pmatrix}$ e V_i a matriz triangular superior com diagonal unitária tal que

$$C_i V_i = B_i.$$

Como $b_i^* \perp b_j^*$, $i \neq j$, então a matriz $C_i^T D C_i$ é diagonal, com elementos diagonais iguais a

$$b_1^{*T} D b_1^*, \dots, b_i^{*T} D b_i^*.$$

Por isso,

$$\det(C_i^T D C_i) = (b_1^{*T} D b_1^*) \dots (b_i^{*T} D b_i^*) = \|b_1^*\|^2 \dots \|b_i^*\|^2. \tag{2.18}$$

Então,

$$\det(B_i^T D B_i) \;=\; \det(V_i^T C_i^T D C_i V_i) = \det(C_i^T D C_i), \tag{2.19}$$

pelo que, (2.17) decorre de (2.18) e (2.19). Em particular, no final do *Passo 1* de cada iteração

$$\|b_1^*\|^2 \dots \|b_n^*\|^2 = \det(B_n^T D B_n) = \det D. \tag{2.20}$$

Agora, consideremos a seguinte função

$$f(b_1, \ldots, b_n) = \det(B_1^T D B_1) \det(B_2^T D B_2) \ldots \det(B_n^T D B_n),$$

que no final do *Passo 1* de cada iteração é igual a

$$\|b_1^*\|^2 (\|b_1^*\|^2 \|b_2^*\|^2) \ldots (\|b_1^*\|^2 \ldots \|b_n^*\|^2). \tag{2.21}$$

No final do *Passo 2*, a sequência que resulta da OGS associada às colunas da matriz BE continua a ser $b_1^*, b_2^*, \ldots, b_n^*$. Por isso f permanece igual a (2.21) também no final do *Passo 2*.

Suponhamos que, no *Passo 3* de uma dada iteração, as colunas de b_i e b_{i+1} são trocadas. Então, para

$$\tilde{B} \equiv \left(\begin{array}{c|c|c} \tilde{b_1} & \ldots & \tilde{b_n} \end{array} \right) = BP_{i,i+1},$$

tem-se

$$\tilde{B_k} \equiv \begin{cases} \left(\begin{array}{c|c|c} b_1 & \ldots & b_k \end{array} \right) = B_k & \text{se } k < i \\ \left(\begin{array}{c|c|c|c} b_1 & \ldots & b_{i-1} & b_{i+1} \end{array} \right) & \text{se } k = i \\ \left(\begin{array}{c|c|c|c|c} b_1 & \ldots & b_{i-1} & b_{i+1} & b_i \end{array} \right) = B_{i+1} P_{i,i+1}^k & \text{se } k = i+1 \\ \left(\begin{array}{c|c|c|c|c} b_1 & \ldots & b_{i-1} & b_i & \ldots & b_k \end{array} \right) = B_k P_{i,i+1}^k & \text{se } k > i+1 \end{cases}$$

onde $P_{i,i+1}^k$ denota uma matriz de permutação $k \times k$. Por isso,

$$\det(\tilde{B_k^T} D \tilde{B_k}) = \det(B_k^T D B_k), \quad k = 1, \ldots, i-1$$

e

$$\begin{aligned} \det(\tilde{B_k^T} D \tilde{B_k}) &= \det({P_{i,i+1}^k}^T B_k^T D B_k P_{i,i+1}^k) \\ &= \det(B_k^T D B_k), \quad k = i+1, \ldots, n. \end{aligned}$$

Seguidamente mostramos que

$$\det(\tilde{B_i^T} D \tilde{B_i}) < \frac{3}{4} \det(B_i^T D B_i). \tag{2.22}$$

Seja $\{\tilde{b}_1^*, \ldots, \tilde{b}_n^*\}$ a sequência que resulta da OGS associada à base ordenada $\{\tilde{b}_1, \ldots, \tilde{b}_n\}$. Como $\tilde{b_j} = b_j$, $j = 1, \ldots, i-1$, então $\tilde{b_j^*} = b_j^*$, $j = 1, \ldots, i-1$. Além disso,

$$\begin{aligned}
\tilde{b_i^*} &= \tilde{b_i} - \sum_{j=1}^{i-1} \frac{\langle \tilde{b_i}, \tilde{b_j^*} \rangle}{\|\tilde{b_j^*}\|^2} \tilde{b_j^*} \\
&= b_{i+1} - \sum_{j=1}^{i-1} \frac{\langle b_{i+1}, \tilde{b_j^*} \rangle}{\|\tilde{b_j^*}\|^2} b_j^* \\
&= b_{i+1} - \sum_{j=1}^{i-1} \lambda_{j,i+1} b_j^* \\
&= \left(b_{i+1} - \sum_{j=1}^{i} \lambda_{j,i+1} b_j^* \right) + \lambda_{i,i+1} b_i^* \\
&= b_{i+1}^* + \lambda_{i,i+1} b_i^*. \qquad (2.23)
\end{aligned}$$

Então,

$$\frac{\det(\tilde{B_i^T} D \tilde{B_i})}{\det(B_i^T D B_i)} = \frac{\|\tilde{b_1^*}\|^2 \ldots \|\tilde{b_i^*}\|^2}{\|b_1^*\|^2 \ldots \|b_i^*\|^2} = \frac{\|\tilde{b_i^*}\|^2}{\|b_i^*\|^2}. \qquad (2.24)$$

Por (2.23),

$$\begin{aligned}
\frac{\|\tilde{b_i^*}\|^2}{\|b_i^*\|^2} &= \frac{\|b_{i+1}^* + \lambda_{i,i+1} b_i^*\|^2}{\|b_i^*\|^2} \\
&= \frac{\|b_{i+1}^*\|^2 + (\lambda_{i,i+1})^2 \|b_i^*\|^2}{\|b_i^*\|^2} \\
&< \frac{1}{2} + \frac{1}{4} = \frac{3}{4}, \qquad (2.25)
\end{aligned}$$

pelo que (2.22) decorre de (2.24) e de (2.25).

Portanto, sempre que o algoritmo não termina, o valor de $f(b_1, \ldots, b_n)$ decresce pelo menos 3/4. No início do algoritmo B é a base canónica, por isso

$$f(b_1, \ldots, b_n) = \det(D_{11}) \ldots \det(D_{nn}),$$

onde D_{ii} representa a submatriz principal de D constituída pelas primeiras i linhas e colunas de D. Se T é o maior valor absoluto dos elementos de D então, pela Desigualdade de Hadamard

$$\det(D_{ii}) \leq \|D_{ii}^1\|_2 \dots \|D_{ii}^i\|_2 \leq (\sqrt{i}T)^i \leq (nT)^n, \tag{2.26}$$

onde D_{ii}^k denota a k-ésima coluna de D_{ii}. Por isso tem-se

$$f(b_1, \dots, b_n) \leq (nT)^{n^2}.$$

Então, ao fim de k iterações tem-se

$$1 \leq f(b_1, \dots, b_n) < \left(\frac{3}{4}\right)^k (nT)^{n^2},$$

sendo que a primeira desigualdade resulta do facto de $f(b_1, \dots, b_n)$ ser um inteiro positivo e por $b_1, \dots, b_n$ serem vectores linearmente independentes, com D uma matriz de inteiros. Como $\lim_{k \to +\infty} \left(\frac{3}{4}\right)^k (nT)^{n^2} = 0$, então existe k tal que

$$\left(\frac{3}{4}\right)^k (nT)^{n^2} < 1,$$

o que equivale a

$$k > log_{\frac{4}{3}}(nT)^{n^2} = log_{\frac{4}{3}}((nT)^{n^2}) = n^2 \left(log_2 n + log_2 T\right) log_{\frac{4}{3}} 2.$$

No final do Algoritmo LLL obtemos uma base $\{b_1, \dots, b_n\}$ reduzida de $\mathbb{Z}^n$. Isto significa que para cada $k = 1, 2, \dots, n$

$$\begin{aligned} \|b_k\|^2 &= \|b_k^* + \sum_{i=1}^{k-1} \lambda_i^k b_i^*\|^2 \\ &= \|b_k^*\|^2 + \sum_{i=1}^{k-1} \|\lambda_i^k b_i^*\|^2 \\ &= \|b_k^*\|^2 + \sum_{i=1}^{k-1} (\lambda_i^k)^2 \|b_i^*\|^2 \end{aligned}$$

$$\begin{aligned} &\leq \|b_k^*\|^2 + \sum_{i=1}^{k-1} \frac{1}{4}\|b_i^*\|^2 \\ &\leq \|b_k^*\|^2 + \sum_{i=1}^{k-1} \frac{1}{2^2} 2^{k-i}\|b_k^*\|^2 \\ &= \|b_k^*\|^2 \left(1 + \frac{1}{4}2^{k-1}\right) \\ &\leq 2^{k-1}\|b_k^*\|^2, \end{aligned}$$

pelo que

$$\begin{aligned} \|b_1\|^2 \dots \|b_n\|^2 &\leq \prod_{k=1}^{n} \left(2^{k-1}\|b_k^*\|^2\right) \\ &= \left(\prod_{k=1}^{n} 2^{k-1}\right) \left(\|b_1^*\|^2 \dots \|b_n^*\|^2\right) \\ &= 2^{n(n-1)/2}\|b_1^*\|^2 \dots \|b_n^*\|^2, \end{aligned}$$

por isso,

$$\|b_1\| \dots \|b_n\| \leq 2^{n(n-1)/4}\sqrt{\det\, D}$$

e portanto, a base satisfaz o limite de Lováz definido por (2.7) e (2.10). ■

Para concluir que o algoritmo funciona em tempo polinomial, precisamos mostrar que os números racionais gerados durante o Algoritmo LLL não se tornam demasiado grandes em tamanho. Pelo teorema anterior, o Algoritmo LLL requer um número de iterações máximo que é um polinómio do tamanho da instância que se quer resolver. Recorde o conceito de tamanho:

Definição 2.3.1 ([14]). *O **tamanho de um racional** α é o número de bits necessário para o representar no computador. No caso particular de $\alpha = p/q$, com p, q inteiros primos entre si e $A = [a_{ij}]_{m\times n} \in \mathbb{Q}^{m\times n}$ então*

*i) O **tamanho de** α é $1 + \lceil \log_2(|p| + 1)\rceil + \lceil \log_2(|q| + 1)\rceil$;*

ii) O ***tamanho de*** *A é* $mn + \sum_{i,j} tamanho\, de\, a_{ij}$

Teorema 2.3.2. *O tamanho de todos os números gerados pelo Algoritmo da Figura 2.5 é limitado polinomialmente pelo tamanho da matriz D.*

Demonstração: Seja $D \in \mathbb{Z}^{n\times n}$ simétrica, definida positiva e T o maior valor absoluto dos elementos de D. Vamos primeiro mostrar que:
i) O tamanho dos numeradores e denominadores de b_i^*, $i = 1, \ldots, n$, são majorados por um polinómio em $\log_2 n$ e $\log_2 T$.
As matrizes $\{B_i, i = 1, \ldots, n\}$ e D são matrizes de inteiros. Por isso as matrizes

$$\det\left(B_{i-1}^T D B_{i-1}\right)\left(B_{i-1}^T D B_{i-1}\right)^{-1} = \mathrm{adj}\left(B_{i-1}^T D B_{i-1}\right), \quad i = 1, \ldots, n,$$

também são matrizes de inteiros. Atendendo a que

$$b_i^* = b_i - B_{i-1}\left(B_{i-1}^T D B_{i-1}\right)^{-1} B_{i-1}^T D b_i, \quad i = 1, \ldots, n,$$

concluímos que

$$\det(B_{i-1}^T D B_{i-1}) b_i^*, \quad i = 1, \ldots, n,$$

são vectores de inteiros. Daqui concluímos, por (2.26), que

$$\text{denominadores de } b_i^* \leq \det(B_{i-1}^T D B_{i-1}) \leq (nT)^n.$$

Isto implica que

$$\begin{aligned} \|b_i^*\|_2 &= \sqrt{\sum_{j=1}^{n} (b_{ji}^*)^2} \\ &\geq \sqrt{\left(\frac{1}{(nT)^n}\right)^2} = (nT)^{-n}. \end{aligned} \tag{2.27}$$

Seja λ o mais pequeno valor próprio de D, então

$$\lambda^{-1} = \text{maior valor próprio de } D^{-1} \leq \mathrm{tr}\left(D^{-1}\right) \leq n(nT)^n. \tag{2.28}$$

Então, de acordo com (2.26), (2.27) e (2.28),

$$\begin{aligned}
\|b_i^*\|_2^2 &\leq \lambda^{-1}\|b_i^*\|^2 = \lambda^{-1}\det(B_i^T D B_i)\|b_i^*\|^{-2}\ldots\|b_{i-1}^*\|^{-2} \\
&\leq \lambda^{-n}\det(B_i^T D B_i)\|b_i^*\|_2^{-2}\ \ldots\ \|b_{i-1}^*\|_2^{-2} \\
&\leq n^n (nT)^{n^2}(nT)^{2n^2} \\
&\leq (nT)^{3n^2+2n}\,.
\end{aligned}$$

Portanto, o tamanho dos numeradores e denominadores de b_i^* são limitados polinomialmente pelo tamanho de D.

Agora vamos mostrar que:
ii) O tamanhos dos números de b_i, $i = 1, \ldots, n$ são também majorados por um polinómio em $\log_2 n$ e $\log_2 T$.
Após a aplicação do Passo 2, de acordo com (2.14), obtemos

$$\begin{aligned}
\|b_i\|_2 &= \|\lambda_{1i}b_1^* + \ldots + \lambda_{i-1,i}b_{i-1}^* + b_i^*\|_2 \\
&\leq \|b_1^*\|_2 + \ldots + \|b_i^*\|_2 \\
&\leq n\,(nT)^{3n^2+2n}\,.
\end{aligned}$$

Como os elementos de cada b_i são inteiros, o tamanho de cada b_i é polinomialmente limitado pelo tamanho de D.

■

De acordo com o Teorema 2.3.1 e o Teorema 2.3.2 , conclui-se que o Algortimo LLL funciona em tempo polinomial.
No próximo teorema, mostramos como determinar uma base reduzida para um outro reticulado $\mathcal{L}$ diferente de $\mathbb{Z}^n$.

Teorema 2.3.3. *Seja $A \in \mathbb{Q}^{n\times k}$, uma matriz de característica completa por colunas. Existe um algoritmo polinomial que determina uma base reduzida*

$\{b_1, \ldots, b_k\}$, para o reticulado $\mathcal{L}$ gerado pelas colunas de A, que satisfaz

$$\|b_1\|_2 \ldots \|b_k\|_2 \leq 2^{k(k-1)/4}\sqrt{\det(A^T A)}, \tag{2.29}$$

onde $\|x\|_2 = \sqrt{x^T x}$.

Demonstração: Aplicando o algoritmo de redução de base da Figura 2.5 à matriz $D = A^T A$, é possível encontrar uma base $b_1, \ldots, b_k$ para $\mathbb{Z}^k$ com

$$\begin{aligned} \|Ab_1\|_2 \ldots \|Ab_k\|_2 &= \sqrt{b_1^T D b_1} \ldots \sqrt{b_k^T D b_k} \\ &= \|b_1\| \ldots \|b_k\| \\ &\leq 2^{k(k-1)/4}\sqrt{\det D} \\ &= 2^{k(k-1)/4}\sqrt{\det A^T A}. \end{aligned}$$

Como é simples de verificar, os vectores $Ab_1, \ldots, Ab_k$ formam uma base reduzida para o reticulado $\mathcal{L}$ gerado pelas colunas de A. De facto, para $i, j \in \{1, \ldots, n\}$, $i \neq j$,

$$\langle Ab_i^*, Ab_j^* \rangle = b_i^{*T} A^T A b_j^* = 0,$$

e

$$\|Ab_i^*\|_2^2 = \|b_i^*\|^2 \leq 2\|b_{i+1}^*\|^2 = 2\|Ab_i^*\|_2^2.$$

■

Note-se que, fazendo $n = k$ no Teorema 2.3.3, o reticulado gerado pelas colunas de A tem dimensão completa e neste caso (2.29) pode ser escrito por

$$\|b_1\| \ \ldots \ \|b_n\| \leq 2^{n(n-1)/4}|\det A|.$$

Além disso, no algoritmo descrito na Figura 2.5 e no Teorema 2.3.3 são usadas apenas operações elementares por colunas. Por isso e pelo Lema 2.1.1, o output do Algoritmo LLL, aplicado à matriz A, pode resumir-se a uma matriz de inteiros unimodular U tal que $B = AU$.

2.4 Exemplos

Nesta secção, indicamos dois exemplos nos quais determinamos uma base reduzida para o reticulado gerado pelas colunas de uma matriz. No primeiro indicamos todos os cálculos efectuados; no segundo usamos um software específico.

Exemplo 2.4.1. Consideremos a matriz

$$A = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}, \tag{2.30}$$

e seja $\mathcal{L}$ o reticulado gerado pelos vectores

$$u_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \text{ e } u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

De acordo com o Teorema 2.3.3, vamos encontrar uma base $\{b_1, b_2\}$ para o reticulado $\mathbb{Z}^2$ usando a matriz

$$D = A^T A = \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix}. \tag{2.31}$$

- Entrada: A matriz D de (2.31), que é definida positiva, e os vectores da base canónica de $\mathbb{Z}^2$

$$b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ e } b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

- Passo 1: De acordo com o produto interno referido no Exemplo 1.1.2, determinam-se os vectores ortogonais tal como vimos em (2.11). Desta forma obtém-se

$$\begin{aligned} b_1^* &= b_1, \\ b_2^* &= b_2 - \frac{b_1^T D b_2}{b_1 D b_1} b_1. \end{aligned}$$

Como

$$\begin{aligned} b_1^T D b_2 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 5, \\ b_1^T D b_1 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 17, \end{aligned}$$

então

$$\begin{aligned} b_1^* &= b_1, \\ b_2^* &= b_2 - \frac{5}{17} b_1^*. \end{aligned}$$

Por isso

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{5}{17} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{17} \\ 0 & 1 \end{pmatrix} = CV.$$

- Passo 2:

$$b_2 \longleftarrow b_2 - \left\lfloor \frac{5}{17} \right\rceil b_1 = b_2.$$

- Passo 3: Calculam-se

$$\begin{aligned} \|b_1^*\|^2 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 17\,, \\ \|b_2^*\|^2 &= \begin{pmatrix} -\frac{5}{17} & 1 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} -\frac{5}{17} \\ 1 \end{pmatrix} = \frac{9}{17}. \end{aligned}$$

Como $\|b_1^*\|^2 > 2\|b_2^*\|^2$, então trocam-se as colunas da matriz $\begin{pmatrix} b_1 & b_2 \end{pmatrix}$

$$b_1 \longleftarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad e \quad b_2 \longleftarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

e volta-se ao primeiro passo.

- Passo 1: Aplica-se, novamente, o processo de ortogonalização de Gram-Schmidt aos vectores

$$b_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad e \quad b_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Obtém-se

$$\begin{aligned} b_1^* &= b_1, \\ b_2^* &= b_2 - \frac{b_1^T D b_2}{b_1 D b_1} b_1. \end{aligned}$$

Como

$$\begin{aligned} b_1^T D b_2 &= \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 5, \\ b_1^T D b_1 &= \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2, \end{aligned}$$

então

$$\begin{aligned} b_1^* &= b_1, \\ b_2^* &= b_2 - \frac{5}{2} b_1^*. \end{aligned}$$

Por isso

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{5}{2} \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{2} \\ 0 & 1 \end{pmatrix} = CV.$$

- Passo 2:

$$b_2 \longleftarrow b_2 - \left\lfloor \frac{5}{2} \right\rceil b_1 = b_2 - 3b_1 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Por isso

$$B = \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{5}{2} \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} = CV.$$

- Passo 3: Calculam-se

$$\|b_1^*\|^2 = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2,$$

$$\|b_2^*\|^2 = \begin{pmatrix} 1 & -\frac{5}{2} \end{pmatrix} \begin{pmatrix} 17 & 5 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -\frac{5}{2} \end{pmatrix} = \frac{9}{2} \geq \frac{\|b_1^*\|^2}{2}.$$

  Logo, não existe i que satisfaça (2.16) e o algoritmo pára.

- Saída:

$$b_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad e \quad b_2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

  Os vectores b_1 e b_2 formam uma base reduzida para o reticulado $\mathbb{Z}^2$ e de acordo com o Teorema 2.3.3 os vectores

$$Ab_1 = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$Ab_2 = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

  formam uma base reduzida do reticulado gerado pelas colunas de A.

No próximo exemplo, utilizamos o software NTL (Number Theory Library) disponível http://www.shoup.net/ntl/ que inclui uma implementação do Algoritmo LLL.

Exemplo 2.4.2. Veja-se, na Figura 2.6, uma sessão do NTL aplicada ao exemplo anterior.

Neste caso a matriz que resulta de A é a matriz

$$B_2 = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$$

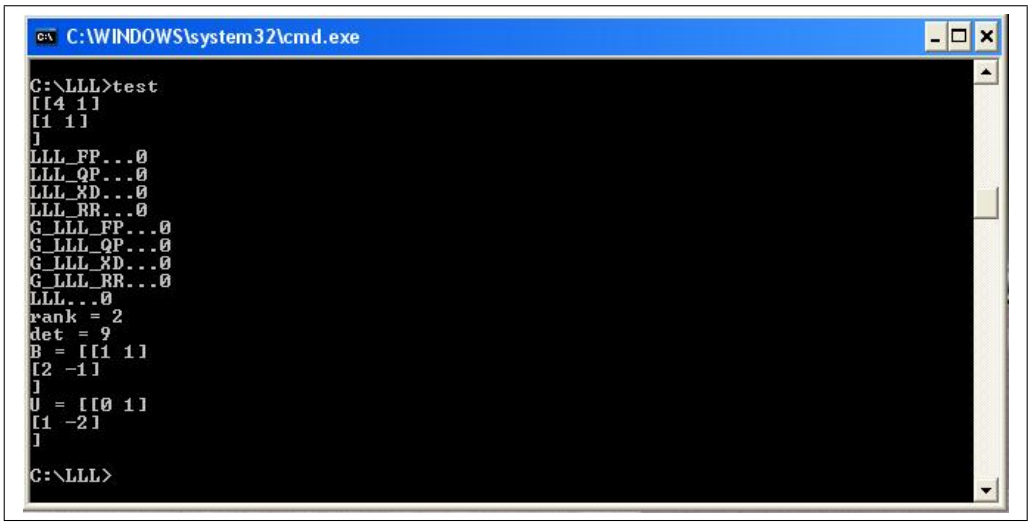

Figura 2.6: NTL aplicado ao Exemplo 2.4.1

que também é uma base reduzida para o reticulado gerado pelas colunas de A. Note-se que

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}}_{B_2} = \underbrace{\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}}_{U},$$

e que

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}}_{B_1} = \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}}_{B_2} \underbrace{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}}_{U'},$$

ou seja, as colunas das matrizes A, B_1 e B_2 geram o mesmo reticulado e as colunas de B_1 e B_2 formam, respectivamente, uma base reduzida para o reticulado gerado pelas colunas de A.

Capítulo 3

Aplicações

Neste capítulo, estudamos algumas aplicações do Algoritmo LLL. Na primeira secção, estudamos o *Problema do Vector mais Curto* (*Shortest Vector Problem - SVP*) de um reticulado. Mostramos que existe um algoritmo polinomial, que permite determinar um limite superior para o vector mais curto de um reticulado, considerando para tal o vector de norma mínima da base construída pelo Teorema 2.3.3.

Na segunda secção, estudamos o *Problema do Vector mais Próximo* (*Closest Vector Problem - SVP*) de um reticulado. Para tal, determinamos um algoritmo polinomial para encontrar uma aproximação para o vector w de um reticulado $\mathcal{L} \subset \mathbb{Q}^n$, mais próximo de um vector $v \in \mathbb{Q}^n$ arbitrariamente fixo.

Na terceira secção, estudamos uma aproximação muito fraca em tempo polinomial para o problema da *Aproximação Diofantina simultânea*. Este problema consiste em, para um dado vector de racionais a, encontrar um vector b de racionais, de denominador comum, o mais próximo possível de a.

Na quarta secção, veremos como determinar a Forma Normal de Hermite de uma matriz não singular, com recurso ao Método de redução de base estudado no Teorema 2.3.3.

Na última secção, propomos um método para estudar a viabilidade de alguns problemas da Programação Inteira.

3.1 O Problema do Vector mais Curto

O **Problema do Vector mais Curto** (*Shortest Vector Problem - SVP*) é o seguinte: "dado um reticulado $\mathcal{L} \in \mathbb{R}^n$, determinar o vector não nulo de norma mínima de $\mathcal{L}$". Este problema é $\mathcal{NP}$ - *completo* (Van Emde Boas [1981]). Um resultado clássico de Minkowski (c.f. [14] pág. 71) estabelece um limite superior para a norma euclidiana de um vector não nulo de um reticulado:

Teorema 3.1.1. *Seja $\mathcal{L} \in \mathbb{R}^n$ um reticulado de dimensão completa. Então existe $b \in L \setminus \{0\}$, tal que:*

$$\|b\| \leq 2 \left(\frac{\det \mathcal{L}}{\mathcal{V}(n)} \right)^{\frac{1}{n}} \tag{3.1}$$

onde $\mathcal{V}(n)$ representa o volume da bola unitária de dimensão n.

Demonstração: Seja $r = \frac{1}{2} \min\{\|b\| : b \in L \backslash \{0\}\}$. Então, para todo $x, y \in L$ com $x \neq y$ tem-se $\|x - y\| > r$. Por isso

$$B^0(x, r) \cap B^0(y, r) = \emptyset$$

onde $B^0(x, r)$ representa a bola aberta de centro x e raio r. A medida de $\bigcup_{x \in L} B^0(x, r)$ em $\mathbb{R}^n$ é inferior ou igual à medida de $\bigcup_{x \in L} (x + \pi)$. Como estes conjuntos são disjuntos e têm volume constante, então

$$\frac{r^n \mathcal{V}(n)}{\det \mathcal{L}} \leq 1, \tag{3.2}$$

que é a razão entre o volume de $B^0(x, r)$ e $(x + \pi)$, para um qualquer $x \in \mathcal{L}$, donde decorre (3.1). ∎

No entanto, ainda não foi encontrado nenhum algoritmo polinomial que determine um b que satisfaça (3.1). O Método de redução de base permite-nos encontrar, em tempo polinomial, um limite superior para a norma desse vector e uma aproximação para o vector mais curto.

Teorema 3.1.2. *Seja A uma matriz não singular de racionais de ordem n. Então, existe um algoritmo polinomial que permite encontrar um vector não nulo b que pertence ao reticulado gerado pelas colunas de A, tal que*

$$\|b\|_2 \leq 2^{(n-1)/4}(\det \mathcal{L})^{1/n} \tag{3.3}$$

Demonstração: Seja $\{b_1, \ldots, b_n\}$ uma base reduzida determinada pelo algoritmo da Figura 2.5 para $D = A^T A$. Então,

$$\|Ab_1\|_2 = \|b_1\| = \|b_1^*\|.$$

Se a base é reduzida então, por (1.10),

$$\|b_1^*\| \leq 2^{(k-1)/2}\|b_k^*\|, \quad k = 1, \ldots, n.$$

Então,

$$\begin{aligned}
\|Ab_1\|_2 &= \left(\prod_{k=1}^{n} \|b_1^*\|\right)^{1/n} \\
&\leq \left(\prod_{k=1}^{n} 2^{(k-1)/2}\|b_k^*\|\right)^{1/n} \\
&= \left(2^{n(n-1)/4}\prod_{k=1}^{n} \|b_k^*\|\right)^{1/n} \\
&= \left(2^{n(n-1)/4}\det \mathcal{L}\right)^{1/n} \\
&= 2^{(n-1)/4}(\det \mathcal{L})^{1/n}. \qquad \blacksquare
\end{aligned}$$

Nesta dissertação sempre que um vector b de um reticulado $\mathcal{L}$ satisfaça (3.3) considera-se que b é um **vector curto** de $\mathcal{L}$.

Exemplo 3.1.1. Recorrendo ao software NTL, podemos determinar uma aproximação para o vector mais curto do reticulado gerado pelas colunas da matriz

$$A = \begin{pmatrix} -19 & 21 & 22 \\ -20 & 22 & 23 \\ -10 & 11 & 12 \end{pmatrix}.$$

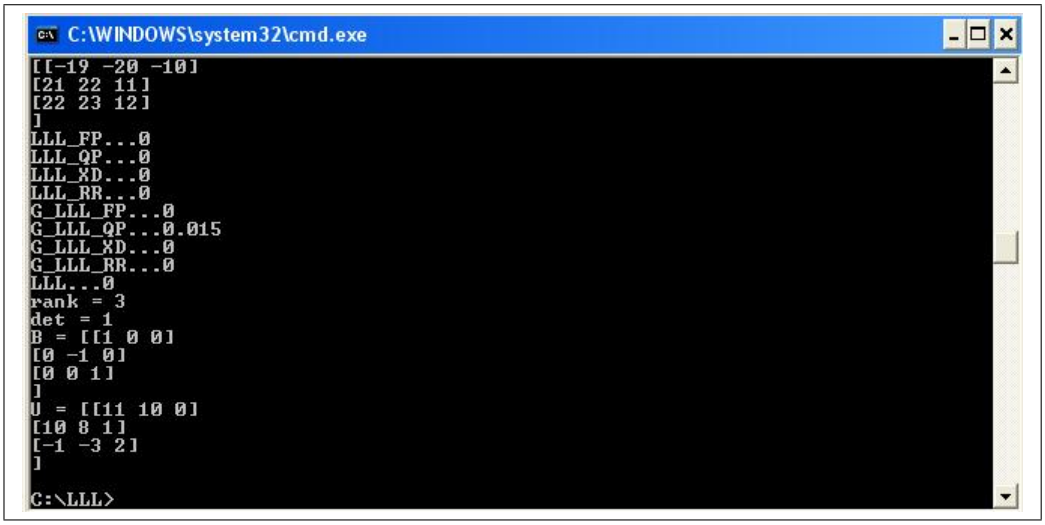

Figura 3.1: NTL aplicado à matriz A

De acordo com a Figura 3.1, tem-se

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} -19 & 21 & 22 \\ -20 & 22 & 23 \\ -10 & 11 & 12 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 11 & 10 & -1 \\ 10 & 8 & -3 \\ 0 & 1 & 2 \end{pmatrix}}_{U},$$

em que as colunas de B são uma base reduzida para o reticulado $\mathcal{L}$ gerado pelas colunas de A. Neste caso a aproximação é exacta e o vector mais curto do reticulado $\mathcal{L}$ é

$$b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

3.2 O Problema do Vector mais Próximo

O **Problema do Vector mais Próximo** (*Closest Vector Problem - CVP*) é o seguinte: "dado $b \in \mathbb{R}^n$, determinar o vector w pertencente a um reticulado $\mathcal{L}$ cuja distância a b é mínima". De acordo com [4] página 141, o Problema do Vector mais Próximo é $\mathcal{NP}-completo$ (Van Emde Boas [1981]). O Algoritmo LLL, pode também ser usado para aproximar o vector mais próximo. O próximo resultado deve-se a Babai (1986) e foi retirado de [4] pág. 149.

Teorema 3.2.1. *Seja $\mathcal{L}$ um reticulado de base $a_1, a_2, \ldots, a_n \in \mathbb{Q}^n$ e seja $b \in \mathbb{Q}^n$. Existe um algoritmo polinomial que permite encontrar $w \in \mathcal{L}$ tal que*

$$\|b - w\|_2 \leq 2^{n/2} \min \{\|b - v\|_2 : v \in \mathcal{L}\} \tag{3.4}$$

Demonstração: Seja $\{b_1, \ldots, b_n\}$ a base reduzida que resulta de $\{a_1, \ldots, a_n\}$ por aplicação do Teorema 2.3.3. Seja $\{b_1^*, \ldots, b_n^*\}$ a base de $\mathbb{Q}^n$ que resulta de $\{b_1, \ldots, b_n\}$ pela OGS. Recorde que, para $j = 1, \ldots, n$,

$$b_j = \sum_{i=1}^{j-1} \alpha_{ij} b_i^* + b_j^*.$$

Então, para $b \in \mathbb{Q}^n$, tem-se

$$b = \sum_{i=1}^{n} \lambda_i^0 b_i^*.$$

Podemos, em primeiro, subtrair a b o vector $\lfloor \lambda_n^0 \rceil$ de modo a obter

$$b - \lfloor \lambda_n^0 \rceil b_n = \sum_{i=1}^{n} \lambda_i^1 b_i^*, \tag{3.5}$$

com $|\lambda_n^1| \leq 1/2$. Seguidamente subtraímos $\lfloor \lambda_{n-1}^1 \rceil b_{n-1}$ em (3.5) e assim sucessivamente para os restantes índices $n-2, \ldots, 2, 1$ de modo a obtermos

$w \in \mathcal{L}$, tal que

$$b - w = \sum_{i=1}^{n} \lambda_i b_i^* \tag{3.6}$$

com $|\lambda_i| \leq 1/2$, para $i = 1, \ldots, n$. Resta mostrar que w satisfaz (3.4). Para tal, considere-se um outro vector $v \in \mathcal{L}$ em que

$$b - v = \sum_{i=1}^{n} \mu_i b_i^* \, .$$

Seja k o maior índice tal que $\mu_k \neq \lambda_k$, então

$$v - w = \sum_{i=1}^{k} (\lambda_i - \mu_i) b_i^* \in \mathcal{L} \, ,$$

ou seja, $\lambda_k - \mu_k$ é um inteiro não nulo, particularmente $|\lambda_i - \mu_i| \geq 1$ e desta forma $|\mu_k| \geq 1/2$. Então,

$$\|b - v\|_2^2 \geq \sum_{i=k+1}^{n} \mu_i^2 \|b_i^*\|_2^2 + \frac{1}{4}\|b_k^*\|_2^2 = \sum_{i=k+1}^{n} \lambda_i^2 \|b_i^*\|_2^2 + \frac{1}{4}\|b_k^*\|_2^2,$$

como $\{b_1, \ldots, b_n\}$ é uma base reduzida, por (1.11), tem-se

$$\begin{aligned}
\|w - b\|_2^2 &\leq \sum_{i=k+1}^{n} \lambda_i^2 \|b_i^*\|_2^2 + \frac{1}{4} \sum_{i=1}^{k} \|b_i^*\|_2^2 \\
&\leq \sum_{i=k+1}^{n} \lambda_i^2 \|b_i^*\|_2^2 + \frac{1}{4}\|b_k^*\|_2^2 \sum_{i=1}^{k} 2^{k-i} \\
&= \sum_{i=k+1}^{n} \lambda_i^2 \|b_i^*\|_2^2 + \frac{1}{4}\|b_k^*\|_2^2 \left(2^k - 1\right) \\
&< 2^k \|v - b\|_2^2 \leq 2^n \|v - b\|_2^2.
\end{aligned}$$

Portanto, o vector w de (3.6) satisfaz (3.4) e é uma aproximação do vector $x \in \mathcal{L}$ mais próximo de $b \in \mathbb{Q}^n$.

■

Exemplo 3.2.1. Seja

$$b = \begin{pmatrix} 73,6 \\ 50,4 \\ 43,7 \end{pmatrix}.$$

Recorrendo ao software NTL, podemos determinar uma aproximação para o vector mais próximo do reticulado gerado pelas colunas da matriz

$$A = \begin{pmatrix} -19 & 21 & 22 \\ -20 & 22 & 23 \\ -10 & 11 & 12 \end{pmatrix}.$$

Como vimos na Figura 3.1, tem-se

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} -19 & 21 & 22 \\ -20 & 22 & 23 \\ -10 & 11 & 12 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 11 & 10 & -1 \\ 10 & 8 & -3 \\ 0 & 1 & 2 \end{pmatrix}}_{U},$$

em que as colunas de B são uma base reduzida para o reticulado $\mathcal{L}$ gerado pelas colunas de A. Então, seguindo a demonstração do Teorema 3.2.1, tem-se que a base $\{b_1^*, b_2^*, b_3^*\}$ que resulta da OGS aplicada às colunas de B são ainda as colunas de B e tem-se

$$b = 73,6b_1^* - 50,4b_2^* + 43,7b_3^*.$$

Portanto, uma aproximação para o vector w do reticulado $\mathcal{L}$ mais próximo de b é

$$w = \begin{pmatrix} \lfloor 73,6 \rceil \\ \lfloor -50,4 \rceil \\ \lfloor 43,7 \rceil \end{pmatrix} = \begin{pmatrix} 74 \\ -50 \\ 44 \end{pmatrix}.$$

Neste caso, a aproximação é exacta, ou seja, o vector w é o vector do reticulado $\mathcal{L}$ mais próximo de b.

3.3 Aproximação Diofantina simultânea

Outra aplicação (c.f. [14] pág. 73) do Algoritmo LLL é a **Aproximação Diofantina simultânea**. O problema de *Aproximação Diofantina simultânea*

é o seguinte: "Sejam $\alpha_1, \ldots, \alpha_n$ números reais, M um número natural e $\varepsilon > 0$ um racional. Determinar números inteiros $p_1, \ldots, p_n, q$ tal que $|\alpha_i - \frac{p_i}{q}| \leq \frac{\varepsilon}{q}$ e $1 \leq q \leq M$".

O problema está bem definido, pois Dirichlet (1842 cf. [14] pág. 72-73) provou que se $\alpha_1, \ldots, \alpha_n, \varepsilon$ forem números reais e $0 < \varepsilon < 1$, então existem números inteiros $p_1, \ldots, p_n$ e q, tal que

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{\varepsilon}{q} \quad \text{para} \quad i = 1, \ldots, n \quad \text{e} \quad 1 \leq q \leq \epsilon^{-n}.$$

A demonstração é uma generalização da demonstração para $n = 1$ e pode ser consultada em [15] pág. 18-19.

No caso em que $n = 1$, o problema tem solução em tempo polinomial através do método das fracções contínuas. Quando $n > 1$, Lagarias (1982) (c.f. [14] pág. 74) demonstrou que este problema é $\mathcal{NP} - completo$.

No entanto (cf. [14] pág. 73) uma aproximação em tempo polinomial pode ser encontrada usando o Teorema 3.1.2.

Teorema 3.3.1 (Aproximação Diofantina simultânea)**.** *Seja $n \in \mathbb{N}$ tal que $n(n+1)$ é múltiplo de 4. Para qualquer vector $a = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Q}^n$ e racional ε com $0 < \varepsilon < 1$, existe um algoritmo polinomial que determina um vector de inteiros $p = (p_1, \ldots, p_n)$ e um inteiro q, tal que*

$$\left\| a - \frac{p}{q} \right\|_2 < \frac{\varepsilon}{q}, \quad e \quad 1 \leq q \leq 2^{n(n+1)/4} \varepsilon^{-n}$$

Demonstração: Seja $\mathcal{L}$ o reticulado gerado pelas colunas da seguinte matriz:

$$A = \begin{pmatrix} 1 & 0 & \ldots & 0 & -\alpha_1 \\ 0 & 1 & \ldots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & 1 & -\alpha_n \\ 0 & 0 & \ldots & 0 & 2^{-n(n+1)/4}\varepsilon^{n+1} \end{pmatrix},$$

Então, $\det \mathcal{L} = |\det A| = 2^{-n(n+1)/4}\varepsilon^{n+1}$. Aplicando o Teorema 3.1.2 à matriz A encontramos um vector não nulo $b = (\beta_1, \ldots, \beta_{n+1})^T$ de $\mathcal{L}$ que satisfaz

$$\|b\|_2 \leq 2^{n/4}(2^{-n(n+1)/4}\varepsilon^{n+1})^{1/(n+1)} = \varepsilon. \tag{3.7}$$

Como $b \in \mathcal{L}$, existem escalares inteiros $(p_1, \ldots, p_n, q)$, únicos porque A é invertível, tal que

$$b = \begin{pmatrix} 1 & 0 & \ldots & 0 & -\alpha_1 \\ 0 & 1 & \ldots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & 1 & -\alpha_n \\ 0 & 0 & \ldots & 0 & 2^{-n(n+1)/4}\varepsilon^{n+1} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \\ q \end{pmatrix}. \tag{3.8}$$

Começamos por provar que $q \neq 0$. De facto, se $q = 0$ então $b = (p_1, \ldots, p_n, 0)$ e como $p_1, \ldots, p_n$ são números inteiros e $b \neq 0$ teríamos $\|b\|_2 \geq 1 > \varepsilon$.
Sem perda de generalidade, supomos que $q \geq 1$. Caso contrário, substitui-se b por $-b$ em (3.7). Então,

$$q = \beta_{n+1}2^{n(n+1)/4}\varepsilon^{-(n+1)},$$

pois de (3.8),

$$\beta_{n+1} = q.2^{-n(n+1)/4}\varepsilon^{n+1},$$

pelo que, atendendo a que $\beta_{n+1} \leq \|b\| \leq \varepsilon$, concluímos que

$$q \leq 2^{n(n+1)/4}\varepsilon^{-n}.$$

Além disso, porque $\beta_{n+1} \neq 0$,

$$\left\| a - \frac{1}{q}.p \right\|_2 = \frac{1}{q}\|qa - p\|_2$$

$$
\begin{aligned}
&= \frac{1}{q}\sqrt{\sum_{i=1}^{n}\left(q\alpha_i - p_i\right)^2} \\
&= \frac{1}{q}\sqrt{\sum_{i=1}^{n}\left(\beta_i\right)^2} \\
&< \frac{1}{q}\|b\| \leq \frac{\varepsilon}{q}
\end{aligned}
$$

■

A restrição "$n(n+1)$ é múltiplo de 4" poderá ser retirada mas isso teria de ser averiguado. Sem essa restrição a demonstração, em [14] pág. 73, encontra-se incorrecta.

3.4 Forma Normal de Hermite

Outra aplicação do método de redução de base (c.f. [14] pág. 74) é a determinação da Forma Normal de Hermite de uma matriz em tempo polinomial.

Definição 3.4.1. *Seja A uma matriz de racionais. Diz-se que A está na* ***Forma Normal de Hermite*** *se $A = [B\,0]$, onde B é uma matriz triangular inferior, não negativa e tal que cada uma das suas linhas possui o maior dos elementos unicamente na diagonal.*

Consideremos A uma matriz não singular de inteiros de ordem n, para

$$M = \left\lceil 2^{n(n-1)/4} \left|\det A\right| \right\rceil,$$

seja C a matriz que se obtém de A multiplicando a i-ésima linha de A por M^{n-i}, para $i = 1, \ldots, n$. Note-se que

$$\det(C) = M^{n(n-1)/2} \det(A).$$

Pelo Teorema 2.3.3 podemos determinar, em tempo polinomial, uma base $b_1, \ldots, b_n$ para o reticulado gerado pelas colunas de C, de modo que

$$\|b_1\| \ldots \|b_n\| \leq 2^{n(n-1)/4} M^{n(n-1)/2} |\det A|. \tag{3.9}$$

Agora vamos mostrar que os vectores $b_1, \ldots, b_n$ podem ser reordenados de modo a que a matriz $\begin{pmatrix} b_{1'} & | \ldots | & b_{n'} \end{pmatrix}$ seja triangular inferior.

Através da reordenação podemos assumir que a j'-ésima coordenada de $b_{j'}$ é não nula, para $j' = 1, \ldots, n$, pois caso contrário os vectores $b_1, \ldots, b_n$ são linearmente dependentes. Então, a j'-ésima coordenada de $b_{j'}$ é pelo menos $M^{n-j'}$ em valor absoluto, e portanto,

$$\|b_{j'}\| \geq M^{n-j'}.$$

Suponhamos que a i-ésima coordenada de $b_{j'}$ é não nula, para algum i tal $1 \leq i < j' \leq n$, então

$$\|b_{j'}\| > M^{n-i} \geq M M^{n-j'},$$

e tem-se que

$$\|b_1\| \ldots \|b_n\| > \left(\prod_{j'=1}^{n} M^{n-j'} \right) M \geq 2^{n(n-1)/4} M^{n(n-1)/2} |\det A|,$$

o que contradiz (3.9). Portanto, os vectores $b_1, \ldots, b_n$ podem ser reordenados de modo a que a matriz $\begin{pmatrix} b'_1 & | \ldots | & b'_n \end{pmatrix}$ seja triangular inferior.
Desta forma, a matriz pode ser facilmente modificada por operações elementares por colunas para a Forma Normal de Hermite. Dividindo a j-ésima linha por M^{n-j}, para cada $j = 1, \ldots, n$, encontramos a Forma Normal de Hermite da matriz A.

Exemplo 3.4.1. Consideremos a matriz

$$A = \begin{pmatrix} 5 & 4 & 2 \\ 8 & 1 & 6 \\ 3 & 9 & 7 \end{pmatrix},$$

com

$$|\det A| = 249.$$

Determinamos

$$M = \left\lceil 2^{3(3-1)/4} 249 \right\rceil,$$

e

$$M^{3-1} = 497025,$$
$$M^{3-2} = 705,$$
$$M^{3-3} = 1.$$

Seguidamente, multiplicamos j-ésima linha de A por M^{n-j}, para $j = 1, 2, 3$. Desta forma obtemos

$$C = \begin{pmatrix} 2485125 & 1988100 & 994050 \\ 5640 & 705 & 4230 \\ 3 & 9 & 7 \end{pmatrix}.$$

Pelo Teorema 2.3.3, determinamos uma matriz

$$\begin{pmatrix} 0 & 0 & -497025 \\ 0 & 705 & 0 \\ 249 & 91 & -104 \end{pmatrix}, \tag{3.10}$$

cujas colunas constituem uma base reduzida para o reticulado gerado pelas colunas de C - veja-se a sua aplicação no NTL na Figura 3.2.

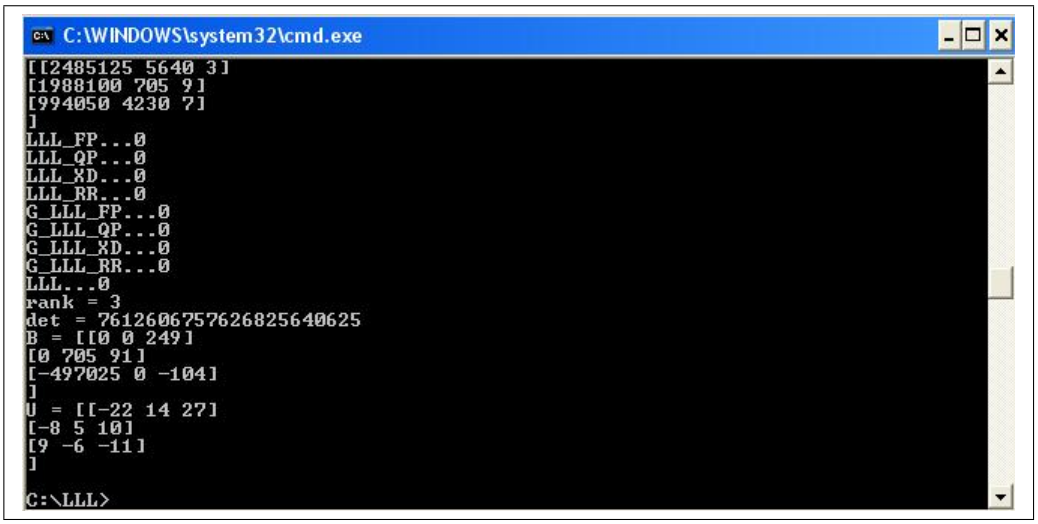

Figura 3.2: NTL aplicado à matriz C

Como se pode verificar, a matriz (3.10) pode ser facilmente modificada, por operações elementares por colunas, de maneira a obter

$$C' = \begin{pmatrix} 497025 & 0 & 0 \\ 0 & 705 & 0 \\ 104 & 91 & 249 \end{pmatrix}.$$

que é uma matriz triangular inferior não negativa. Para obtermos a Forma Normal de Hermite da matriz A, basta dividir a j-ésima linha de C' por M^{n-j}, para $j = 1, 2, 3$. Desta forma obtemos

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 104 & 91 & 249 \end{pmatrix}.$$

3.5 Programação Inteira

Nesta secção, propomos um método para verificar se a solução de alguns problemas da Programação Inteira (PI) é o conjunto vazio (cf. [7]), substituindo

o problema

$$\begin{aligned} b' \leq Ax \leq b \\ x \in \mathbb{Z}^n \,, \end{aligned} \tag{3.11}$$

por

$$\begin{aligned} b' \leq (AU)y \leq b \\ y \in \mathbb{Z}^n \,. \end{aligned} \tag{3.12}$$

A matriz U é uma matriz de inteiros e unimodular, gerada pelo Método de redução de base, de modo a que as colunas de AU sejam "curtas"e próximas da ortogonalidade. Desta forma, consideremos os seguintes poliedros

$$P = \{x : b' \leq Ax \leq b\} \,;$$

$$\tilde{P} = \{y : b' \leq (AU)y \leq b\}.$$

Existe uma correspondência biunívoca entre

$$P \cap \mathbb{Z}^n \quad \text{e} \quad \tilde{P} \cap \mathbb{Z}^n,$$

dada por

$$Uy = x,$$

em que A é uma matriz $n \times k$, U é uma matriz de inteiros unimodular $k \times k$ e b, b' são n-vectores. Através do Método de redução de base fazemos o seguinte: determinamos uma matriz de inteiros e unimodular U, de acordo com o Teorema 2.3.3, que torna as colunas de AU pequenas e próximas da ortogonalidade. Desta forma substituímos o problema $P \cap \mathbb{Z}^n$ por $\tilde{P} \cap \mathbb{Z}^n$. A redução de base das colunas de A é claramente uma transformação válida; a

correspondente relação entre os pontos (inteiros) em $P \cap \mathbb{Z}^n$ e $\tilde{P} \cap \mathbb{Z}^n$ é dada por

$$Uy = x \quad \text{e} \quad y = U^{-1}x\,,$$

onde a unimodularidade de U implica que existe a matriz U^{-1} de inteiros. Note-se que a dimensão do problema não se altera. A esta técnica chama-se ***reformulação do espaço das colunas***.

Exemplo 3.5.1. Considere-se a resolução do seguinte problema $P \cap \mathbb{Z}^n$:

$$106 \leq 21x_1 + 19x_2 \leq 113$$
$$0 \leq x_1, x_2 \leq 6$$
$$x_1, x_2 \in \mathbb{Z}$$

cujo espaço solução, na sua relaxação linear, está descrito à esquerda na Figura 3.4. Se aplicarmos um processo de ramificação (branch-and-bound) é necessário verificar pelo menos 6 pontos que podem ser solução do problema até verificarmos que não existe solução inteira. Se aplicarmos o Método de redução de base, pelo NTL - ver Figura 3.3, às colunas da matriz

$$A = \begin{pmatrix} 21 & 19 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

obtemos

$$AU = \begin{pmatrix} -2 & 7 \\ -1 & -6 \\ 1 & 7 \end{pmatrix} \quad , \quad U = \begin{pmatrix} -1 & -6 \\ 1 & 7 \end{pmatrix}.$$

Então, a respectiva reformulação é o seguinte problema $\tilde{P} \cap \mathbb{Z}^n$:

$$106 \leq -2y_1 + 7y_2 \leq 113$$

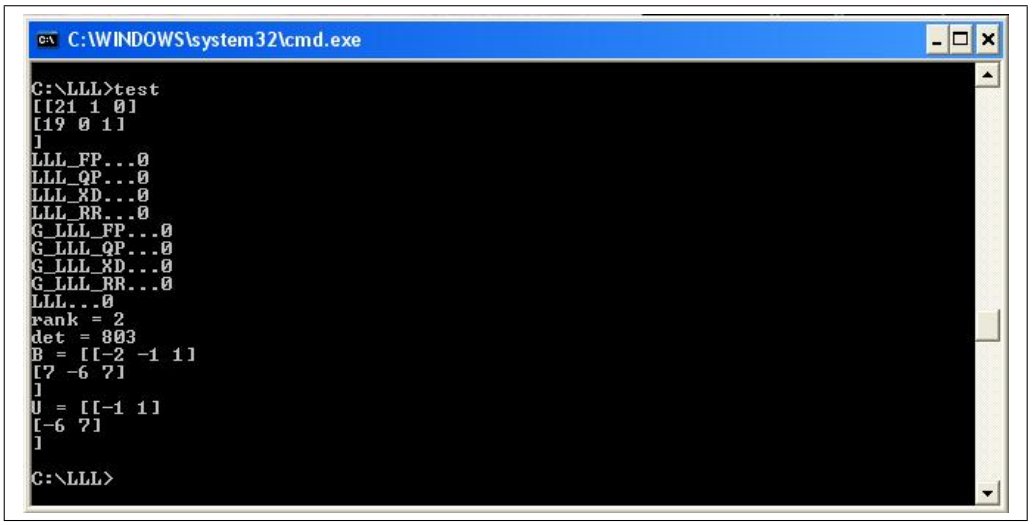

Figura 3.3: NTL aplicado à matriz A

$$\begin{aligned} 0 \leq -y_1 - 6y_2 \leq 6 \\ 0 \leq y_1 + 7y_2 \leq 6 \\ y_1 \ , \ y_2 \in \mathbb{Z}, \end{aligned} \tag{3.13}$$

cujo espaço solução, na sua relaxação linear, está descrito à direita na Figura 3.4.

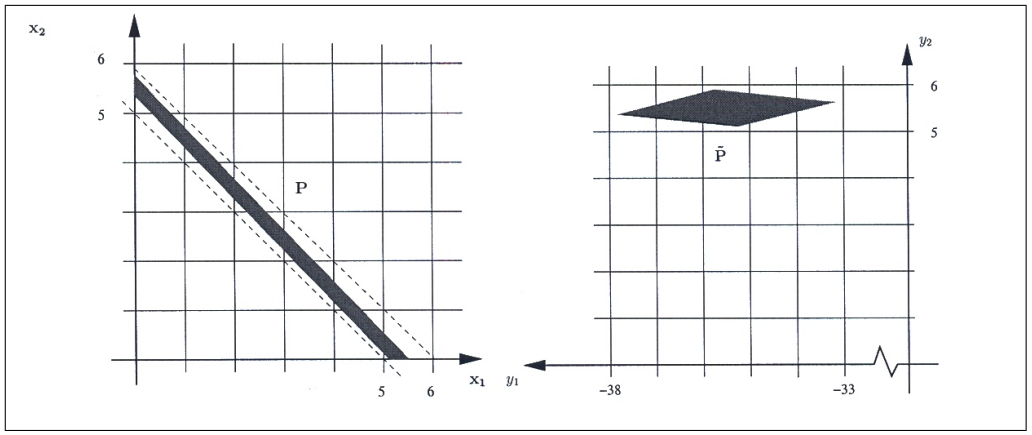

Figura 3.4: Poliedro antes e depois da reformulação

Desta forma, aplicando o processo de ramificação (branch-and-bound) a par-

tir de y_1, existem quatro pontos que podem ser solução do problema. No entanto, a restrição linear de (3.13) implica que

$$5,04 \leq y_2 \leq 5,94\,,$$

ou seja, a ramificação (branch-and-bound) a partir de y_2 permite imediatamente verificar $\tilde{P} \cap \mathbb{Z}^n$ é vazio e consequentemente também $P \cap \mathbb{Z}^n$, como podemos ver na segunda imagem da figura 3.4

Exemplo 3.5.2. Neste exemplo, apresentamos uma versão simplificada do que foi proposto por Jeroslow em [6]. Seja n um inteiro positivo ímpar. O problema $P \cap \mathbb{Z}^n$

$$\begin{gathered} n - \tfrac{1}{2} \leq 2\sum_{i=1}^{n} x_i \leq n + \frac{1}{2} \\ 0 \leq x \leq e \\ x \in \mathbb{Z}^n\,, \end{gathered} \tag{3.14}$$

não tem solução inteira, uma vez que o único inteiro no intervalo $\left[n - \frac{1}{2}\ ,\ n + \frac{1}{2}\right]$ é o próprio n (e representa um vector de uns).

Se aplicarmos um processo de ramificação (branch-and-bound) a cada uma das x_i variáveis, teremos de enumerar pelo menos $2^{(n-1)/2}$ pontos. De facto, supondo que fixamos no máximo $(n-1)/2$ variáveis a zero ou um, o dobro da soma destas variáveis é no máximo $n-1$, enquanto que a soma dos coeficientes das variáveis livres é no mínimo $n+1$. Desta forma, pelo menos uma das variáveis livres poderá assumir um valor não inteiro.

Aplicando o Método de redução de base, temos

$$A = \begin{pmatrix} 2e_{1\times n} \\ I_n \end{pmatrix} \quad , \quad U = \begin{pmatrix} I_{n-1} & 0_{(n-1)\times 1} \\ -e_{1\times(n-1)} & 1 \end{pmatrix}\,,$$

$$AU = \begin{pmatrix} 0_{1\times(n-1)} & 2 \\ I_{n-1} & 0_{(n-1)\times 1} \\ -e_{1\times(n-1)} & 1 \end{pmatrix},$$

cuja reformulação é

$$\begin{gathered} n - \tfrac{1}{2} \leq 2y_n \leq n + \tfrac{1}{2} \\ 0 \leq y_1, \ldots, y_{n-1} \leq 1 \\ 0 \leq -\sum_{i=1}^{n-1} y_i + y_n \leq 1 \\ y \in \mathbb{Z}^n. \end{gathered} \tag{3.15}$$

A primeira restrição de (3.15) é equivalente a

$$\frac{(n+1)}{2} - \frac{3}{4} \leq y_n \leq \frac{(n+1)}{2} - \frac{1}{4} \tag{3.16}$$

Como $\frac{(n+1)}{2}$ é um inteiro não existe nenhum inteiro y_n com limites definidos por (3.16). Logo, a ramificação (branch-and-bound) a partir de y_n implica imediatamente a impossibilidade de (3.15) e consequentemente de (3.14).

Capítulo 4

"Ataque"ao RSA

A criptografia é a ciência que se ocupa das comunicações secretas, comunicações em que há um emissor e um receptor e se pretende que nenhum terceiro tenha acesso à informação transmitida. Para atingir este objectivo, a informação ou mensagem é cifrada, isto é, substituída por outra conforme uma regra pré-estabelecida.
Na primeira secção deste capítulo descrevemos o sistema criptográfico RSA (Rivest, Shamir e Adleman), um dos sistemas de chave pública mais conhecidos. Veremos que neste sistema existe uma assimetria entre os processos de cifragem e de decifragem, isto é, saber cifrar uma mensagem utilizando este método não significa que se consiga depois decifrá-la.
Na última secção estudamos um processo de quebra do código RSA, utilizando para tal o Algoritmo LLL.

Um dos mais antigos e rudimentares sistemas de criptografia foi usado por Júlio César. Este consiste em substituir cada letra do alfabeto pela letra situada três posições à frente, sendo a antepenúltima, a penúltima e a última letras substituídas, respectivamente pela primeira, segunda e terceira. Assim

para o alfabeto inglês tem-se a seguinte correspondência

A B C D E...I J K L M...V W X Y Z

D E F G H...L M N O P...Y Z A B C

O **Código de César** pode ser descrito usando a teoria das congruências. O primeiro passo consiste em substituir a mensagem por um único número, através de uma correspondência fixa, por exemplo

A	B	...	W	X	Y	Z
01	02	...	23	24	25	26

. (4.1)

Em seguida, substitui-se p, número associado a uma letra de mensagem, por $c \in \mathbb{Z}_{26}$ tal que

$$c \equiv (p+3)(\text{mod } 26)\,,$$

para descodificar a mensagem, usa-se

$$p \equiv (c+23)(\text{mod } 26)\,,$$

uma vez que 23 é o simétrico de 3 em $\mathbb{Z}_{26}$ e finalmente recorre-se à tabela (4.1). Trata-se de um sistema muito simples mas extremamente inseguro.

Nos códigos ditos "convencionais" tais como o código de César existe uma chave secreta conhecida pelo emissor e pelo receptor.
Os sistemas de "chave pública" diferem dos sistemas "convencionais" pelo facto de usarem *duas chaves*; uma para codificar e outra para descodificar. Embora as chaves efectuem operações inversas e portanto, estejam relacionadas, não há um método de computação eficiente que permita deduzir a chave de descodificação a partir da chave de codificação.

4.1 O sistema criptográfico RSA

Os sistemas criptográficos utilizados na vida real são bem mais complexos que o descrito anteriorormente. Em 1977, Rivest, Shamir e Adleman [12] criaram um sistema criptográfico de chave pública que usa conhecimentos da Teoria dos Números, nomeadamente o Teorema de Euler (cf. [11] pág. 26). Vamos desvendar nesta secção a matemática que está por detrás deste sistema.

Descrição do sistema criptográfico RSA

Escolhem-se dois primos distintos, p e q, suficientemente grandes (*e.g.*, cerca de 200 dígitos cada um) e calcula-se $n = pq$.

Escolhe-se um inteiro positivo e primo com $\varphi(n)$ onde φ é a função de Euler (note-se que $\varphi(n) = (p-1)(q-1)$). Este número e pode ser seleccionado no conjunto $\{1, 2, \ldots, (p-1)(q-1)-1\}$. Verificar se um dado número é primo com $(p-1)(q-1)$ pode ser feito através do Algoritmo de Euclides. Com o auxílio de um computador essa tarefa é trivial.

O receptor torna público o par (n, e), denominado **chave pública**, mas não os primos p e q; estes podem até ser esquecidos, pois não são necessários para nenhuma operação, servem apenas para iniciar o sistema.

O processo de codificação, por parte do emissor, começa com a conversão da mensagem num número a através de um "alfabeto digital" no qual cada letra, sinal de pontuação ou algarismo é substituído por um inteiro com dois dígitos, por exemplo representando cada letra da mensagem pelo seu código ASCII. Um procedimento "standard" consiste em usar a correspondência

$$\begin{aligned} &\text{A} = 01,\ \text{B} = 02, \ldots, \text{Z} = 26,\ , = 27,\ . = 28,\ ? = 29 \\ &0 = 30,\ 1 = 31, \ldots, 9 = 39,\ 00 = \text{espaço entre palavras} \end{aligned} \tag{4.2}$$

Caso $a > n$ divide-se a em blocos de dígitos $a_1, \ldots, a_i$ tais que $a_j < n$ para $j = 1, \ldots, i$ e cada um destes blocos será codificado separadamente.
Suponha-se então que $a < n$. O emissor, conhecedor da chave pública (n, e), converte a no número $b \in \mathbb{Z}_n$ tal que

$$a^e \equiv b(\mathrm{mod}\, n)\,, \tag{4.3}$$

onde e é denominado o **expoente de codificação** e (4.3) **fórmula de codificação**.
O receptor, conhecedor não só de (n, e) mas também de $\varphi(n)$, determina o **expoente de descodificação**, $d \in \mathbb{Z}_{\varphi(n)}$ tal que

$$ed \equiv 1(\mathrm{mod}\, \varphi(n))\,. \tag{4.4}$$

Tal d existe e é único, porque, sendo e e $\varphi(n)$ primos entre si, a congruência linear

$$ex \equiv 1(\mathrm{mod}\, \varphi(n)) \tag{4.5}$$

tem uma e uma só solução em $\mathbb{Z}_{\varphi(n)}$, que pode ser determinada através da resolução da equação Diofantina

$$ex + \varphi(n)y = 1.$$

Note-se que o expoente de descodificação d requer o conhecimento simultâneo de e e de $\varphi(n) = (p-1)(q-1)$.
Para descodificar a mensagem, o receptor vai determinar $a \in \mathbb{Z}_n$ tal que

$$b^d \equiv a(\mathrm{mod}\, n)\,, \tag{4.6}$$

denominada **fórmula de descodificação** e em seguida recorre ao "alfabeto digital" (4.2).
Resta justificar (4.6), o que será feito na seguinte proposição.

Proposição 4.1.1. *Sejam p e q dois números primos distintos e a um natural tal que $a < n = pq$. Se e e d são números naturais tais que $ed \equiv 1(\mathrm{mod}\,\varphi(n))$, então $a^{ed} \equiv a(\mathrm{mod}\,n)$.*

Demonstração: Por hipótese $ed \equiv 1(\mathrm{mod}\,\varphi(n))$, então tem-se

$$ed = 1 + t\varphi(n),$$

para certo inteiro não negativo t. Se $a < n = pq$ com p, q primos distintos, então

$$m.d.c(a,n) = 1 \quad \text{ou} \quad m.d.c(a,n) = p \quad \text{ou} \quad m.d.c(a,n) = q\,.$$

Se $m.d.c(a,n) = 1$, pelo Teorema de Euler

$$a^{\varphi(n)} \equiv 1(\mathrm{mod}\,n)\,, \tag{4.7}$$

Vem então, de acordo com [11] pág. 21 e (4.7)

$$a^{ed} = a^{1+t\varphi(n)} = a.a^{t\varphi(n)} = a.(a^{\varphi(n)})^t \equiv a(\mathrm{mod}\,n)$$

Se $m.d.c(a,n) = p$ então pelo Teorema de Fermat (caso particular do Teorema de Euler), tem-se

$$a^{q-1} \equiv 1(\mathrm{mod}\,q)\,, \tag{4.8}$$

então, de acordo com [11] pág. 21 e (4.8), tem-se

$$a^{kj} = a^{1+t\varphi(n)} = a.a^{(p-1)(q-1)t} = a.(a^{(q-1)})^{(p-1)t} \equiv a(\mathrm{mod}\,q)\,.$$

Por outro lado, como $p|a$ tem-se

$$a \equiv 0(\mathrm{mod}\,p)\,, \tag{4.9}$$

logo, por [11] pág. 21, obtemos

$$a^{ed} \equiv 0(\mathrm{mod}\,p) \quad \text{e} \quad 0 \equiv a(\mathrm{mod}\,p)\,,$$

portanto, de acordo com a propriedade transitiva das congruências

$$a^{ed} \equiv a(\mathrm{mod}\, p)\,.$$

Então, como $m.d.c(p,q) = 1$ vem

$$a^{ed} \equiv a(\mathrm{mod}\, pq = n)\,.$$

Se $m.d.c(a,n) = q$ a justificação é idêntica à do caso anterior, basta trocar os papeis de p e q.

■

Nesta demonstração considerámos a um número natural tal que $a < n$, diferente da que está em [11] pág. 56-57, em que se considera a um inteiro primo com n.

Então, de acordo com a Proposição 4.1.1, $a^{ed} \equiv a(\mathrm{mod}\, n)$, mas como $a^e \equiv b(\mathrm{mod}\, n)$, tem-se

$$b^d \equiv a(\mathrm{mod}\, n)\,.$$

Ou seja: para reaver a a partir de b, o receptor apenas tem de calcular b^d módulo n. Mais precisamente, a mensagem decifrada a é o resto da divisão de b^d por n.

O que é interessante nisto é que o receptor divulga n e e, isto é, divulga publicamente a *chave de cifragem* para as mensagens que lhe são enviadas. Mas não divulga d, a chave necessária para a decifragem.

Para isto fazer sentido é necessário que d seja muito difícil de calcular para outra pessoa que não o receptor. Tudo depende de uma boa escolha do n. Deverão escolher-se dois números primos p e q muito grandes e fazer $n = pq$. O número n é tornado público mas o factores p e q são mantidos secretos pelo receptor. Como vimos, para calcular a chave necessária para a decifragem d é necessário conhecer $\varphi(n) = (p-1)(q-1)$ e, portanto, é necessário conhecer

p e q. Ora, se p e q forem muito grandes é muito difícil, mesmo usando computadores poderosos, obter os factores p e q a partir do seu produto n. Há assim uma grande assimetria entre entre os processos de cifragem e de decifragem. Saber cifrar uma mensagem utilizando este método não significa que se consiga decifrá-la.

Pode justificar-se a segurança do sistema RSA de outra forma: se se puder quebrar este sistema, então pode utilizar-se o mesmo procedimento para encontrar factores primos de n. O problema de factorização foi estudado por inúmeros matemáticos e nenhum método eficiente foi encontrado, o que torna a segurança dos sistema RSA bastante provável.

Exemplo 4.1.1. Sejam $p = 29$ e $q = 53$ dois números primos, eventualmente demasiado pequenos para garantir segurança, então,

$$n = p \times q = 29 \times 53 = 1537\,,$$

e

$$\varphi(n) = (p-1)(q-1) = 28 \times 52 = 1456\,,$$

e escolha-se como expoente de codificação, por exemplo $e = 47$ uma vez que

$$m.d.c(47, 1456) = 1\,.$$

O expoente de descodificação d obtém-se resolvendo a congruência linear

$$47d \equiv 1 (\mathrm{mod}\, 1456)\,,$$

ou seja $d = 31$, no conjunto $\{0, 1, \ldots, 1455\}$. Então a chave pública será

$$(1537, 47).$$

Se quisermos codificar a mensagem

`NO WAY`

obtém-se, usando a tabela referida em (4.2),

$$a = 141500230125\,.$$

Considere-se este número dividido, por exemplo, em quatro blocos de três dígitos. Efectuemos a codificação do primeiro bloco $a = 141$. Como

$$141^{47} \equiv 658 (\mathrm{mod}\, 1537)\,,$$

portanto os três primeiros dígitos da mensagem codificada são 658. Procede-se de modo idêntico para os outros blocos; a mensagem codificada, na forma digital, é

$$658 \quad 1408 \quad 1250 \quad 1252\,.$$

Vamos agora proceder à descodificação do primeiro bloco, usando o expoente de descodificação (não público) $d = 31$, ou seja

$$658^{31} \equiv 141 (\mathrm{mod}\, 1537)\,.$$

A descodificação dos outros blocos faz-se de modo análogo. Conhecido o a, recorre-se ao "alfabeto digital"definido em (4.2).

■

4.2 "Ataque"ao RSA

Veremos agora de que forma é possível quebrar o código de encriptação RSA. O ataque que aqui descrevemos baseia-se no pressuposto de que o expoente público de encriptação e é relativamente pequeno e os primos p e q escolhidos para determinar o valor $n = pq$ estão também relativamente próximos, nomeadamente,

$$e < \varphi(n)\,, \tag{4.10}$$

e

$$\frac{1}{2}n^{1/2} < p < n^{1/2} < q < 2n^{1/2}\,, \tag{4.11}$$

Vamos assumir também que, para algum $\delta < 1$,

$$d = n^{\delta}.$$

Veremos agora que, nas condições descritas, é possível quebrar o código RSA. Para tal considere-se que o emissor pretende enviar ao receptor o número a, ou seja utiliza a chave pública (n, e) para encriptar a da seguinte forma

$$a^e \equiv b \,(\mathrm{mod}\, n)\,,$$

e envia a mensagem b, que é interceptada. Pretender-se-á descobrir o número a sem o conhecimento da chave de desencriptação d e dos primos p e q. Considere-se a função de Euler

$$\varphi(n) = (p-1)(q-1) = n - p - q + 1 = n - \lambda, \tag{4.12}$$

com $\lambda = p + q - 1$. Por (4.11) tem-se

$$\lambda = n - \varphi(n) < 3n^{1/2}. \tag{4.13}$$

Considere-se a equação

$$ed = 1 + k\varphi(n)\,, \tag{4.14}$$

cujos valores desconhecidos são, neste momento, os inteiros positivos d, k e $\varphi(n)$. Note-se que, de (4.10) conclui-se que

$$k < d = n^{\delta}.$$

A equação (4.14) é equivalente a

$$ed - kn = 1 - k\lambda, \tag{4.15}$$

o que pode ser escrito na forma matricial

$$\begin{pmatrix} 1 & 0 \\ e & -n \end{pmatrix} \begin{pmatrix} d \\ k \end{pmatrix} = \begin{pmatrix} d \\ 1-k\lambda \end{pmatrix},$$

o que equivale a

$$\underbrace{\begin{pmatrix} \lfloor n^{1/2} \rfloor & 0 \\ e & -n \end{pmatrix}}_{A} \begin{pmatrix} d \\ k \end{pmatrix} = \underbrace{\begin{pmatrix} d\lfloor n^{1/2} \rfloor \\ 1-k\lambda \end{pmatrix}}_{v_1}, \tag{4.16}$$

ou seja, o vector v_1 é uma combinação linear inteira das colunas de A. Portanto, v_1 pertence ao reticulado $\mathcal{L}_1$ gerado pelas colunas de A.
De acordo com o Teorema 3.1.2 o vector (não nulo) mais curto do reticulado $\mathcal{L}_1$, em $\mathbb{R}^2$, satisfaz

$$\|v\| \leq 2^{1/4}(\det \mathcal{L}_1)^{1/2} < \sqrt{10}(\det \mathcal{L}_1)^{1/2}. \tag{4.17}$$

Então, sabendo que

$$\begin{aligned} \|v_1\| &\leq \sqrt{nd^2 + (1-k\lambda)^2} \\ &< \sqrt{nd^2 + (k\lambda)^2} \\ &\leq \sqrt{n^{1+2\delta} + 9n^{1+2\delta}} \\ &= \sqrt{10}\; n^{\delta+1/2}, \end{aligned}$$

e ainda que

$$\sqrt{10}(\det \mathcal{L}_1)^{1/2} \leq \sqrt{10}(n^{3/2})^{1/2} = \sqrt{10}n^{3/4},$$

então para

$$\delta < \frac{1}{4} \tag{4.18}$$

tem-se que

$$d\lfloor n^{1/2} \rfloor \leq \|v_1\| < \sqrt{10}n^{3/4} \approx \sqrt{10}(\det \mathcal{L}_1)^{1/2}, \tag{4.19}$$

Seguidamente podemos usar o Algoritmo LLL de acordo com o Teorema 3.1.2 de forma a tentar encontrar v_1.

Para ilustrar esta situação, mostramos um exemplo em que é possível quebrar o código RSA. Para tal e de forma a encontrar uma chave pública que satisfaça (4.18) recorremos ao software Cryptool. Esta ferramenta pode ser encontrada em http://www.cryptool.org/ e permite trabalhar com alguns conceitos de criptografia, entre os quais o código RSA. Utilizamos também o software NTL.

Exemplo 4.2.1. Considere-se a chave pública

$$(n, e) = (12319, 9677)\,,$$

gerada pelo software Cryptool, em consonância com (4.18), conforme podemos ver nas Figuras 4.1 e 4.2.

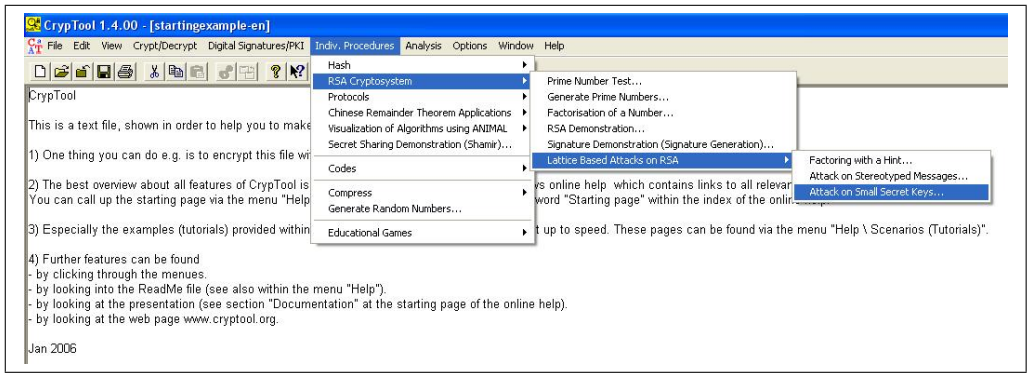

Figura 4.1: Ambiente de trabalho do Cryptool

Vamos tentar descobrir d tal que $ed \equiv 1(\mathrm{mod}\varphi(n))$. Para tal considere-se (4.16), ou seja, considere-se o reticulado $\mathcal{L}_1$ gerado pelas colunas da matriz

$$A = \begin{pmatrix} 111 & 0 \\ 9677 & -12319 \end{pmatrix},$$

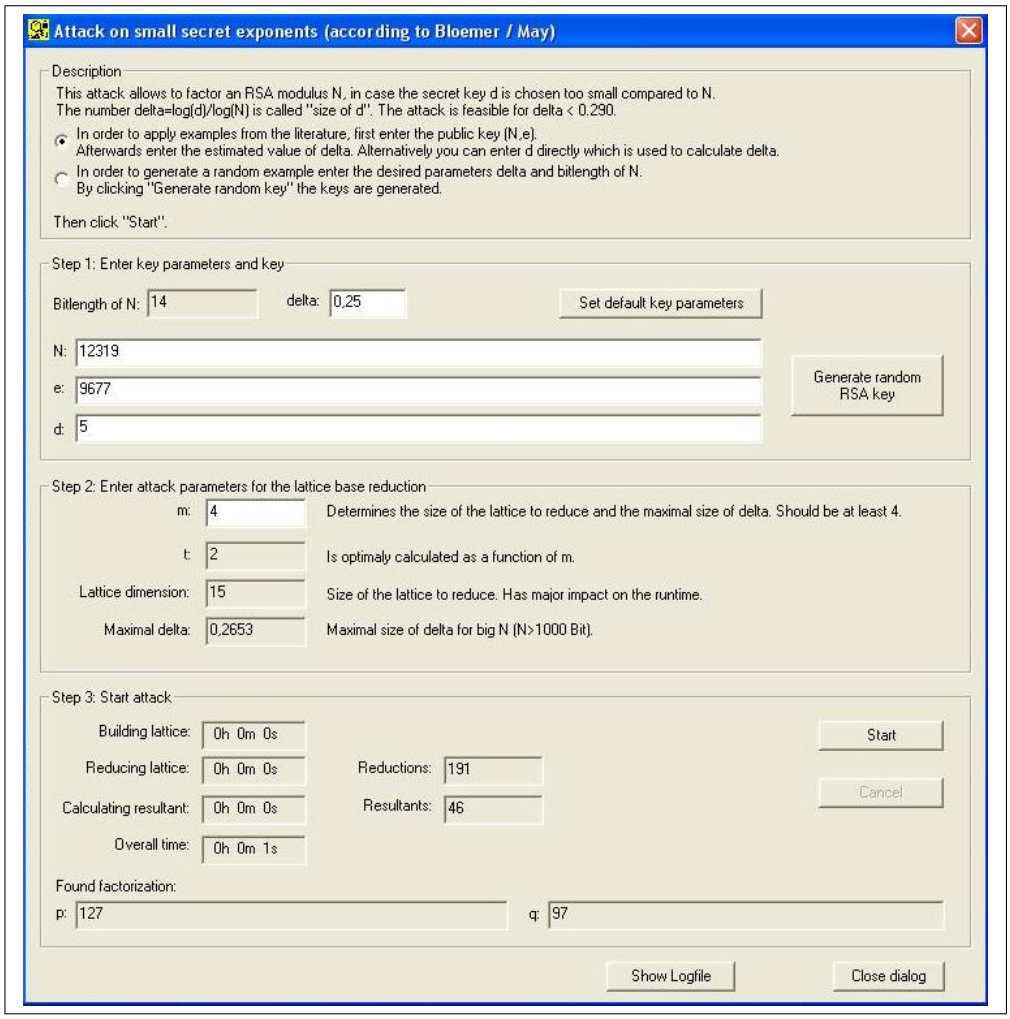

Figura 4.2: Ataque ao RSA no Cryptool

e um pequeno vector de $\mathcal{L}_1$ determinado pelo Teorema 3.1.2. Com o auxílio do NTL (ver Figura 4.3) determinamos uma base reduzida de $\mathcal{L}_1$ constituída pelos vectores coluna da matriz

$$B = \begin{pmatrix} 555 & 990 \\ -891 & 860 \end{pmatrix}, \tag{4.20}$$

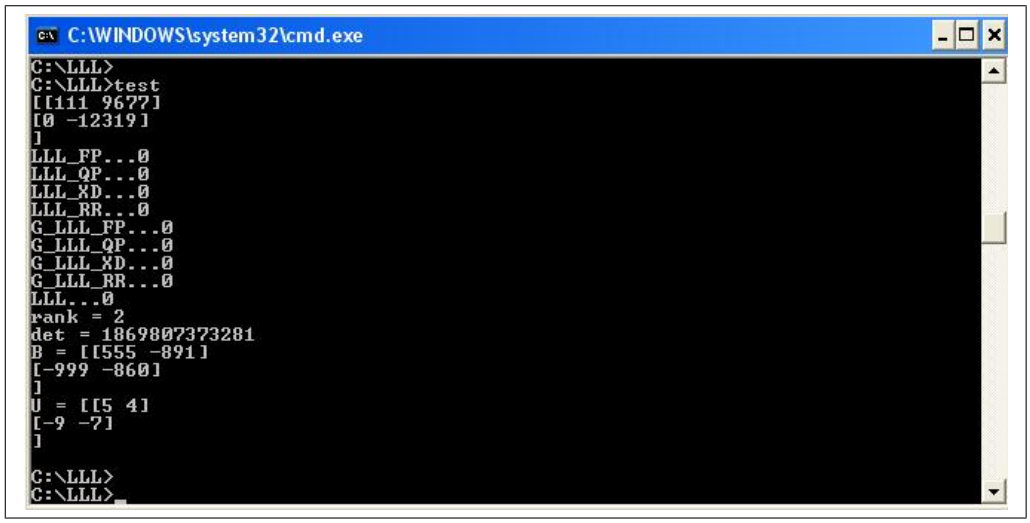

Figura 4.3: Executável do NTL para o Algoritmo LLL

então um pequeno vector de $\mathcal{L}_1$ que satisfaz (4.17) é

$$v = \begin{pmatrix} 555 \\ -891 \end{pmatrix}.$$

Considere-se agora a igualdade

$$\begin{aligned} \begin{pmatrix} d\lfloor n^{1/2} \rfloor \\ 1 - k\lambda \end{pmatrix} &= \begin{pmatrix} 555 \\ -891 \end{pmatrix} \\ \iff \begin{pmatrix} 110d \\ 1 - k\lambda \end{pmatrix} &= \begin{pmatrix} 555 \\ -891 \end{pmatrix} \end{aligned} \tag{4.21}$$

e tentamos descobrir os inteiros positivos d, k e λ que satisfaçam (4.12), (4.13), (4.14) e primos p e q de modo que $n = pq$. De (4.21) tem-se

$$\begin{cases} d &= 5 \\ k\lambda &= 892 \end{cases},$$

assim por (4.15) obtém-se o sistema

$$\begin{cases} k\lambda &= 892 \\ k(12319 - \lambda) &= 9677 \times 5 - 1 \end{cases},$$

cuja solução é

$$\begin{cases} k &= 4 \\ \lambda &= 223 \end{cases}.$$

Finalmente testamos a factorização de $n = 12319$ através do sistema

$$\begin{cases} 12319 &= pq \\ 223 &= p + q - 1 \end{cases},$$

cuja solução é

$$\begin{cases} p &= 127 \\ q &= 97 \end{cases}.$$

Conseguimos factorizar o número $n = 127 \times 97$ com o auxílio do Algoritmo LLL e desta forma o expoente de desencriptação $d = 5$ está correcto. Caso contrário tentaríamos os valores mais próximos como $d = 2, 3, 4, 6, 7, \ldots$, desde que d satisfaça (4.19), na tentativa de encontrar a factorização de n.

Bibliografia

[1] Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[2] Matthew C. Cary. *Lattice Basis Reduction Algorithms and Applications*. University of Washington, Washington, Department of Computer Science, 2002. Disponível em http://www.cs.washington.edu/homes/cary/.

[3] C.F. Gauss. Disquisitiones arithmeticae, article 171. Edição Inglesa, trans. by A.A. Clarke, Springer - Verlag.

[4] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1988.

[5] Alexander D. Healy. *Lattice Basis Reduction and Public-Key Cryptography*. Cambridge, Massachusetts, 2002. Mathematics and Computer Science in partial fulfillment of the honors requirements for the degree of Bachelor of Arts.

[6] R. G. Jeroslow. Trivial integer programs unsolvable by branch-and-bound. *Math. Programming*, 6:105–109, 1974.

[7] Bala Krishnamoorthy and Gabor Pataki. Column basis reduction and decomposable knapsack problems. http://www.ie.buffalo.edu/soda-ext2.pdf.

[8] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[9] Carl Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.).

[10] Paulo R. Pinto. *Resumo das Aulas Teóricas de Álgebra Linear*. Departamento de Matemática, Seccção de àlgebra e Análise, Lisboa, Instituto Superior Técnico, 2005. Disponível em http://www.math.ist.utl.pt/˜ppinto/al0506/ALteoricas.pdf.

[11] João Filipe Queiró. *Teoria de Números*. Departamento de Matemática, FCTUC, Coimbra, Universidade de Coimbra, 2002. Disponível em http://www.mat.uc.pt/˜jfqueiro/TN.pdf.

[12] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Comm. ACM*, 21(2):120–126, 1978.

[13] Ana Paula Santana, João Filipe Queiró. *Álgebra Linear e Geometria Analítica*. Departamento de Matemática, FCTUC, Coimbra, Universidade de Coimbra, 2003. Disponível em http://www.mat.uc.pt/%7Ejfqueiro/ALGA2003.pdf.

[14] Alexander Schrijver. *Theory of linear and integer programming*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986. A Wiley-Interscience Publication.

[15] João L.C. Soares. *Problemas Diofantinos*. Departamento de Matemática, FCTUC, Coimbra, Universidade de Coimbra, 2006.

[16] Joachim von zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge University Press, Cambridge, segunda edição, 2003.