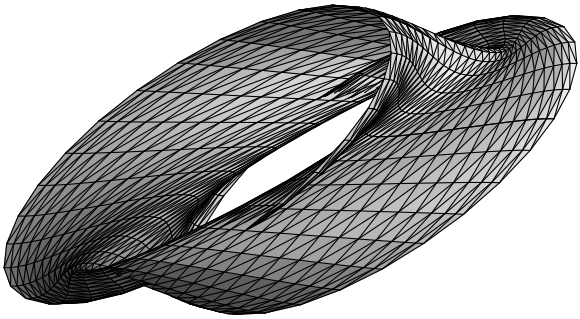
ISSN 1980-024X

Boletim de Iniciação Científica em Matemática – BICMat



Volume IX
Outubro de 2012
Departamento de Matemática
IGCE – Rio Claro

unesp

ISSN 1980-024X

Boletim de Iniciação Científica em Matemática – BICMat

Realização

Conselho de Curso de Graduação em Matemática

Departamento de Matemática

IGCE – Unesp – Rio Claro

Editorial

O Boletim de Iniciação Científica em Matemática – BICMat é uma publicação que se destina a difundir prioritariamente trabalhos de iniciação científica em Matemática que fazem parte de projetos desenvolvidos por alunos do Curso de Graduação em Matemática do IGCE – Unesp – Rio Claro. Eventualmente trabalhos de Iniciação Científica realizados em outras instituições poderão também ser publicados neste Boletim.

O BICMat foi criado em 1998 e nessa época foram publicados dois volumes; o primeiro no ano de criação e o segundo em 2000.

Considerando a importância da Iniciação Científica para o graduando, e o sempre crescente número de projetos desta natureza desenvolvidos em nossa instituição, resolvemos reativar a publicação do BICMat, com ISSN 1980-024X.

Destacamos que a autoria dos trabalhos apresentados no BICMat é dos alunos. O orientador figura apenas como responsável científico.

Este Boletim também está aberto à divulgação de trabalhos que não sejam frutos de projetos de iniciação científica, mas que sejam de interesse dos alunos do curso de graduação em Matemática. Estes trabalhos serão selecionados pelos Editores.

Este número estará disponibilizado eletronicamente na página do Departamento de Matemática no endereço

`www.rc.unesp.br/igce/matematica`

SUMÁRIO

Códigos de Huffman

Adriele Ribeiro dos Santos[1]
Orientador(a): Prof. Dr. Henrique Lazari

Resumo: O código de Huffman é o modelo clássico de um código sem ruído ótimo, com significativa importância teórica e em aplicações práticas de métodos de compressão de dados. O presente trabalho consiste de uma demonstração detalhada da otimalidade do código de Huffman que não é comumente encontrada na literatura sobre códigos. Essa pesquisa foi realizada em materiais bibliográficos e na internet. O objetivo é de apresentar uma descrição sucinta de codificação de fonte, entropia e a demonstração da otimalidade do código de Huffman como apresentada em ASH. Esse é um resultado relevante sobre codificação de fonte que pode ser apresentado em nível de iniciação cientifica.

Palavras-chave: álgebra; telecomunicações; computação

Definição 1. Seja $S = \{s_1, \ldots, s_n\}$ uma fonte de informação com probabilidades $p(s_1) = p_1, \ldots, p(s_n) = p_n$ de ocorrências das mensagens. A informação fornecida pela ocorrência do evento s_i é $I(p_i) = k \cdot \log_r(p_i)$ para algum $k \in N$. Como $p_i \leq 1$ e desejamos que $I(p_i) \geq 0$, tomamos $k = -1$ e obtemos:

$$I(p_i) = -\log_r(p_i) = \log_r(\frac{1}{p_i})$$

para alguma base fixada de logaritmos, é comum tomar base $r = 2$.

Definição 2. Entropia: Seja $S = \{s_1, \ldots, s_n\}$ uma fonte de informação como acima. A entropia $H_r(S)$ de S na base r, é dada por:

$$H_r(S) = \sum_{i=0}^{n} p_i \cdot \log_r(\frac{1}{p_i})$$

[1]Bolsista PET SESu/MEC

que é a média (ponderada pelas probabilidades dos símbolos) de informação recebida por todos os símbolos do conjunto, sendo que a média de informação recebida por cada símbolo s_i é dado por:

$$p_i \cdot I(s_i) = p_i \cdot \log(\frac{1}{p_i}).$$

Proposição 3. *Seja $S = \{s_1, \ldots, s_n\}$ uma fonte de informação com probabilidades $p(s_i) = p_i$, para $i = 1, \ldots, n$ então $H(S) \leq \log(n)$. Se $p_i = \frac{1}{n}$ para todo i ou seja, se a distribuição de probabilidades é uniforme, então:*

$$H(S) = \sum_{i=1}^{n} \frac{1}{n} \log(n) = \log(n)$$

de onde resulta o

Corolário 4. *Seja $S = \{s_1, \ldots, s_n\}$ uma fonte de informação, então para qualquer distribuição de probabilidades dos símbolos da fonte, a entropia da mesma é limitada superiormente pela entropia da distribuição uniforme.*

Definição 5. Seja S um conjunto de símbolos formando um alfabeto, digamos, $S = \{s_1, \ldots, s_n\}$. Um código sobre S é uma função do conjunto consistindo de todas as sequências possíveis de elementos de S no conjunto de todas as sequências de algum outro alfabeto $X = \{x_1, \ldots, x_r\}$. Chamamos S de alfabeto da fonte e X de alfabeto-código. As sequências de elementos de X são chamadas de palavras-códigos.

Definição 6. Seja $S = \{s_1, \ldots, s_n\}$ uma fonte de informação. O comprimento médio de palavras-códigos ou comprimento médio ponderado de um código C é:

$$AV(C, f) = \sum_{i=1}^{n} p_i \cdot m_i$$

onde $m_i = comp(f(s_i))$ é o comprimento (número de caracteres ou dígitos) da palavra-código $f(s_i)$ e p_i é a probabilidade.

Definição 7. Se todas as palavras-códigos de um código C tem o mesmo comprimento, dizemos que C é um código de bloco ou um código de comprimento fixo. Se C contêm palavras-código de comprimentos diferentes então dizemos que C é um código de comprimento variável de palavras.

Definição 8. Dizemos que um código C é unicamente decifrável se cada palavra-código corresponde no máximo a uma mensagem.

Definição 9. Um código é dito instantâneo se cada palavra-código em cada sequência de palavras-código pode ser decodificada tão logo seja recebida (lida da esquerda para a direita).

Se um código é instantâneo então ele é unicamente decifrável, mas a recíproca nem sempre é verdadeira.

1 Códigos de Huffman

O código de Huffman é um código de comprimento variável e tal que os comprimentos das palavras-código são inversamente proporcionais às probabilidades de ocorrência dos símbolos correspondentes da fonte.

Definição 10. Um código ótimo é um código unicamente decodificável (decifrável).

Os dois teoremas seguintes são da referência [1].

Teorema 11. *Existe um código instantâneo com comprimentos de palavras* $m_1, \ldots, m_n$ *se e somente se,* $\sum_{i=1}^{n} \frac{1}{2^{m_i}} \leq 1$.

Teorema 12. *Se um código é unicamente decifrável então:* $\sum_{i=1}^{n} \frac{1}{2^{m_i}} \leq 1$.

Lema 13. *Se C é um código ótimo na classe dos códigos instantâneos com probabilidades $p_1, \ldots, p_n$ então C é ótimo na classe dos códigos unicamente decifráveis.*

Prova: Se C' é unicamente decifrável e tem comprimento médio menor que C, sejam $u'_1, \ldots, u'_n$ os comprimentos de C' então $\sum_{i=1}^{n} \frac{1}{2^{m_i}} \leq 1$ e existe um código instantâneo C'' com o menor comprimento médio de C' contra a hipótese de C. ■

Lema 14. *Dado C um código instantâneo, com comprimentos $m_1, \ldots, m_k$ e comprimento médio m associado com o conjunto de probabilidades $p_1, \ldots, p_k$. Suponha que os símbolos sejam arranjados em ordem decrescente de probabilidades e que um grupo de símbolos com a mesma probabilidade é arranjado em ordem crescente de comprimento das palavras. Então se C é um código ótimo dentro da classe dos códigos instantâneos, C deve satisfazer as seguintes propriedades:*

(a) *$p_i > p_k$ então $m_i < m_k$;*

(b) *$m_{k-1} = m_k$;*

(c) *entre as palavras-códigos de comprimento m_k, deverá existir pelo menos duas palavras-código diferindo apenas no último dígito sendo que os demais são iguais.*

Prova: (a) se $p_i > p_k$ e $m_i < m_k$ construímos C' trocando w_i e w_k, então se m' é o comprimento médio de C':

$$m' - m = p_i \cdot m_k + p_k \cdot m_i - (p_i \cdot m_i + p_k \cdot m_k) = (p_i - p_k) \cdot (m_k - m_i) < 0.$$

(b) se $p_{n-1} > p_n$ então por (a) $m_{n-1} \leq m_n$ se $p_{n-1} = p_n$ então $m_{n-1} \leq m_n$ por hipótese. Assim $m_{n-1} \leq m_n$ e se $m_{n-1} < m_n$ então como w_{n-1} não é prefixo de w_n eliminamos o último dígito de w_n e obtemos um código melhor.

(c) se não acontece, eliminamos o último dígito das duas e obtemos um código melhor. ■

O processo de codificação de Huffman pode ser descrito como se segue abaixo: Combinamos dois símbolos mais improváveis do alfabeto em um

único símbolo cuja probabilidade é a soma das probabilidades dos anteriores. Passamos então a codificar um novo alfabeto com um símbolo a menos, e repetimos este procedimento até restar somente dois símbolos que serão codificados com 0 e 1. Retornamos pelos passos anteriores, colocando novos valores 0 e 1 se o símbolo foi obtido da soma de dois símbolos anteriores e não acrescentando nada caso contrário. Ao chegar na fonte original, estes dígitos nos fornecem a codificação de Huffman.

O Método de Huffman com as notações dos lemas acima combinamos os dois últimos símbolos x_{n-1} e x_n em um símbolo equivalente $x_{n,n-1}$ com probabilidade $p_n + p_{n-1}$ vamos supor que em algum momento construímos um código ótimo C_2 para o novo conjunto de símbolos. Construímos um novo código C_1 para o conjunto original de símbolos do seguinte modo: as palavras-código de $x_1, \ldots, x_{n-2}$ continuam as mesmas de C_2 e as de x_{n-1} e x_n no obtidos adicionado respectivamente 0 e 1 a $w_{n,n-1}$ associado a $x_{n,n-1}$ em C_2.

Afirmação: C_1 é ótimo para $p_1, \ldots, p_k$. Se C_1 não é ótimo, seja C_1' ótimo para $x_1, \ldots, x_k$ com palavras-código $w_1', \ldots, w_k'$ com comprimentos $m_1', \ldots, m_k'$ então pelo lema $m_{k-1}' = m_k'$ e pelo menos duas palavras de comprimento m_k' coincidem exceto pelo último dígito. (w_{k-1}' e w_k') combinamos x_{k-1} e x_k para construir um código C_2' associado a $x_{k,k-1}$ a palavra w_k' (ou w_{k-1}') com o último dígito removido. Vamos ver que: $m_2' \leq m_2$ então C_2 é ótimo. $m_2' = p_1 \cdot m_1' + \cdots + p_{k-2} \cdot m_{k-2}' + (p_{k-1} + p_k) \cdot (m_{k-1}' - 1)$ $= p_1 \cdot m_1' + \cdots + p_{k-2} \cdot m_{k-2}' + p_{k-1} \cdot m_{k-1}' + p_k \cdot m_k' - (p_{k-1} + p_k)$ $\leq p_1 \cdot m_1 + \ldots + p_{k-2} \cdot m_{k-2} + p_{k-1} \cdot m_{k-1} + p_k \cdot m_k$ e como $m_{k-1} = m_k$ obtemos: $m_2' \leq p_1 \cdot m_1 + \cdots + p_{k-2} \cdot m_{k-2} + (p_{k-1} + p_k) \cdot m_{k-1} - (p_{k-1} + p_k) = m_2$.

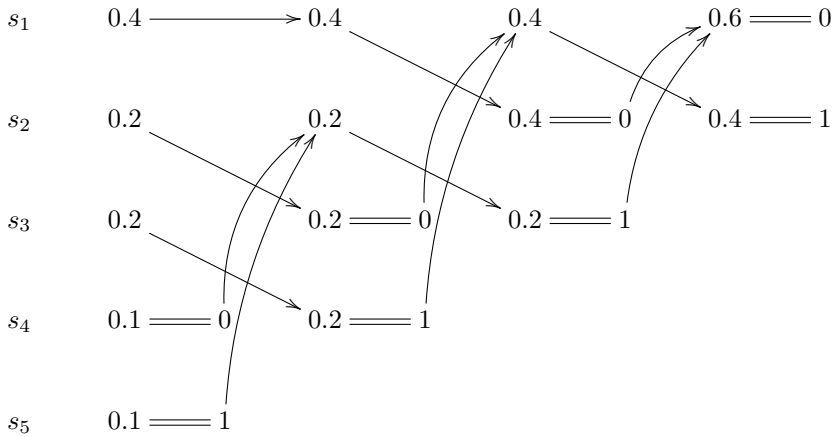
Logo C_2 é ótimo.

Exemplo 15. Vamos considerar a fonte de informação $S = \{s_1, s_2, s_3, s_4, s_5\}$ com perfil de probabilidades

$$p(s_1) = 0.4, \quad p(s_2) = 0.2, \quad p(s_3) = 0.2, \quad p(s_4) = 0.1, \quad p(s_5) = 0.1.$$

Repetindo neste caso particular a construção feita na demonstração da

otimalidade do Código de Huffman, montaremos a árvore de codificação de Huffman e a partir da atribuição de valores binários nos diversos níveis da árvore, percorrendo o caminho inverso construiremos a codificação da fonte de informação S.



Desse modo, $s_1 = 00$, $s_2 = 10$, $s_3 = 11$, $s_4 = 010$, $s_5 = 110$ Assim podemos calcular a entropia e comprimento médio:

$$H_2(S) = 0.4 \cdot \log(\frac{1}{0.4}) + 0.2 \cdot \log(\frac{1}{0.2}) + 0.2 \cdot \log(\frac{1}{0.2}) + 0.1 \cdot \log(\frac{1}{0.1}) + 0.1 \cdot \log(\frac{1}{0.1}) =$$
$$\frac{2}{5} \cdot \log(\frac{1}{\frac{2}{5}}) + \frac{1}{5} \cdot \log(\frac{1}{\frac{1}{5}}) + \frac{1}{5} \cdot \log(\frac{1}{\frac{1}{5}}) + \frac{1}{10} \cdot \log(\frac{1}{\frac{1}{10}}) + \frac{1}{10} \cdot \log(\frac{1}{\frac{1}{10}}) =$$
$$\frac{2}{5} \cdot \log(\frac{5}{2}) + \frac{1}{5} \cdot \log(5) + \frac{1}{5} \cdot \log(5) + \frac{1}{10} \cdot \log(10) + \frac{1}{10} \cdot \log(10) =$$
$$= \log(5) - \frac{1}{5}.$$

$$AV(C, f) = 0.4 \cdot 2 + 0.2 \cdot 2 + 0.2 \cdot 2 + 0.1 \cdot 3 + 0.1 \cdot 3 = \frac{22}{10} = 2.2.$$

Agradecimentos: PET SESu/MEC

Abstract: The Huffman code is the classic model of a code without great noise, with significant theoretical importance and practical applications of methods of data compression. This work consists of a detailed demonstration of the optimality of the Huffman code that is not commonly found in the literature about codes. This survey was conducted in bibliographic materials and the Internet. The goal is to present a brief description of source coding, entropy and the demonstration of optimality of the Huffman code as presented in ASH. This is an important result on source coding that can be presented in scientific initiation level.

Keywords: algebra; telecommunications; computation

Referências Bibliográficas

[1] Ash, R.B., *Information Theory*, Interscience New York, N.Y., 1965.

[2] Palazzo Junior, R., *Transmissão de Dados*, Campinas - Unicamp, 1999.

Conjunto de Cantor e Propriedades

Cristiano dos Santos
Orientador(a): Profa. Dra. Marta Cilene Gadotti

Resumo: Neste trabalho apresentamos um breve relato sobre a vida de Georg Cantor e suas contribuições para o desenvolvimento da matemática; introduzimos o importante conjunto de Cantor e suas propriedades que estão relacionadas ao conceito de medida nula e cardinalidade.

"Ninguém nos expulsará do paraíso que Cantor criou para nós." (Hilbert)

Palavras-chave: conjunto de Cantor; não enumerável; medida nula

1 Georg Cantor (1845–1918)

Cantor nasceu em S. Peterburgo, mas a maior parte de sua vida ele passou na Alemanha. Georg se interessou fortemente pelos argumentos sutis dos teólogos medievais sobre a continuidade e o infinito, e isso contribuiu para que não quisesse seguir uma carreira mundana em engenharia, como seu pai sugeria. Em seus estudos de Zurique, Göttingen e Berlim o jovem consequentemente concentrou-se em filosofia, física e matemática, programa que parece ter estimulado sua enorme imaginação matemática. Doutorou-se em Berlim em 1967 com uma tese sobre teoria dos números, mas suas primeiras publicações mostram atração pela análise. Suas con-

tribuições mais originais centram-se na provocativa palavra "infinito". Os incríveis resultados de Cantor o levaram a estabelecer a teoria dos conjuntos como uma disciplina matemática completamente desenvolvida, ramo que em meados do século vinte teria efeitos profundos sobre o ensino da matemática. Cantor estava entre os matemáticos mais notáveis, e certamente mais originais, de sua época; no entanto não conseguiu uma posição profissional de primeiro plano. Passou maior parte de sua carreira na Universidade de Halle, pequena escola sem reputação. Em 1884 Cantor sofreu o primeiro dos esgotamentos nervosos que viriam a reaparecer durante os trinta anos restantes de sua vida. Acessos de depressão às vezes o levavam a duvidar de sua própria obra. Quase no fim ele obteve o reconhecimento de suas realizações, mas sua morte em 1918 numa instituição para doentes mentais em Halle faz lembrar que o gênio e a loucura às vezes estão relacionados de perto.

2 Definições

Definição 1 (Conjunto totalmente desconexo). Dizemos que um conjunto $A \subset \mathbb{R}$ é totalmente desconexo quando para quaisquer $x, y \in A$ e $x < y$, existir $z \notin A$ tal que $x < z < y$.

Definição 2 (Medida nula). Um conjunto $A \subset \mathbb{R}$ tem medida nula, quando dado $\varepsilon > 0$ existir uma família enumerável de intervalos abertos I_j tais que: (a) $A \subset \bigcup_{j=1}^{\infty} I_j$, (b) $\sum_{j=1}^{\infty} \ell(I_j) < \varepsilon$, onde $\ell(I_j) = b_j - a_j$ é o comprimento de $I_j = (a_j, b_j)$.

2.1 Cardinalidade

Dois conjuntos finitos X, Y tem o mesmo número de elementos se, e somente se, existe uma correspondência um-a-um $f : X \longrightarrow Y$. Embora dizer que dois conjuntos "tem o mesmo número de elementos" não se aplique para o caso em que X, Y são infinitos, parece natural pensar que dois con-

juntos infinitos, que estejam em correspondência um-a-um, tem o mesmo "tamanho". Formalizaremos esta intuição neste texto.

Definição 3. Dizemos que dois conjuntos são equipotentes, fato denotado por $X \sim Y$ quando existir uma correspondência um-a-um $f : X \longrightarrow Y$.

Observação 4. A relação $\sim$, "ser equipotente a", entre conjuntos é uma relação de equivalência.

Para tratar do "tamanho" do conjunto de Cantor, é preciso introduzir os números cardinais e algumas propriedades desses números, que serão dadas na forma de definições.

Definição 5. Cada conjunto A está associado a um número cardinal, denotado por $\text{card}(A)$, e para cada número cardinal a, existe um conjunto A com $\text{card}(A) = a$.

Definição 6. $\text{card}(A) = 0$ se, e somente se, $A = \emptyset$.

Definição 7. Se A é um conjunto finito não vazio, isto é, $A \sim \{1, 2, \ldots, k\}$ para algum $k \in \mathbb{N}$, então $\text{card}(A) = k$.

Definição 8. Para quaisquer dois conjuntos A e B, $\text{card}(A) = \text{card}(B)$ se, e somente se, $A \sim B$.

Definição 9. Sejam A e B conjuntos. Então dizemos que $\text{card}(A) < \text{card}(B)$, quando A é equipotente a um subconjunto de B, mas o conjunto B não é equipotente a nenhum subconjunto de A.

Definição 10. Se A e B são conjuntos não vazios, dizemos que:

- $\text{card}(A) \leqslant \text{card}(B)$, se existe $f : X \longrightarrow Y$ injetora.
- $\text{card}(A) \geqslant \text{card}(B)$, se existe $f : X \longrightarrow Y$ sobrejetora.

Exemplo 11. $\text{card}(\mathbb{N}) < \text{card}(\mathbb{R})$.

De fato, como o conjunto $\mathbb{N}$ é um subconjunto de $\mathbb{R}$, $\mathbb{N}$ é equipotente a um subconjunto de $\mathbb{R}$, $\mathbb{N} \sim \mathbb{N} \subset \mathbb{R}$, mas sabemos que o conjunto infinito $\mathbb{R}$ é não enumerável. Portanto pela Definição 9, temos $\text{card}(\mathbb{N}) < \text{card}(\mathbb{R})$.

Definição 12. Sejam a e b números cardinais. A soma cardinal $a + b$, é o número cardinal $\text{card}(A \cup B)$, em que A e B são conjuntos disjuntos tais que $\text{card}(A) = a$ e $\text{card}(B) = b$.

Segundo Georg Cantor, os símbolos $\aleph_0$ (*leia-se álefe zero*) e c tem sido usados para denotar, respectivamente, o número cardinal de um conjunto enumerável e o número cardinal do *continuum* significando o conjunto dos números reais. Em outras palavras, $\text{card}(\mathbb{N}) = \aleph_0$ e $\text{card}(\mathbb{R}) = c$.

Exemplo 13. $\aleph_0 + \aleph_0 = \aleph_0$.

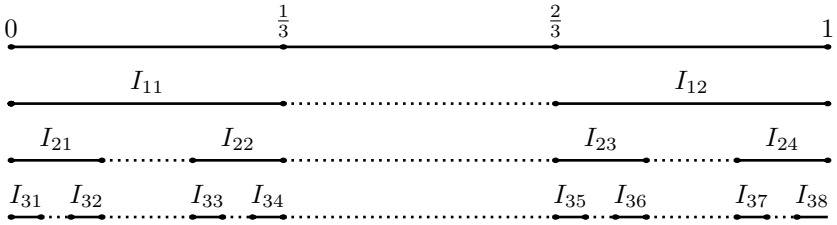
Para verificar esse fato, considere $\mathbb{N}_p$ e $\mathbb{N}_i$, o conjunto de números naturais pares e de números naturais ímpares, respectivamente. Então, $\mathbb{N}_p$ e $\mathbb{N}_i$ são subconjuntos de $\mathbb{N}$ enumeráveis e disjuntos, e a união deles é $\mathbb{N}$. Consequentemente, pela Definição 12, temos:

$$\aleph_0 + \aleph_0 = \text{card}(\mathbb{N}_p) + \text{card}(\mathbb{N}_i) = \text{card}(\mathbb{N}_p \cup \mathbb{N}_i) = \text{card}(\mathbb{N}) = \aleph_0.$$

3 O Conjunto de Cantor

O conjunto de Cantor é construído por etapas da seguinte maneira: dividimos o intervalo $[0, 1]$ em três partes iguais e removemos o intervalo aberto do meio, $J_{11} = (\frac{1}{3}, \frac{2}{3})$. Isto nos deixa com dois intervalos fechados, I_{11} e I_{12}; em cada um destes repetimos a mesma operação, removendo os intervalos (abertos) do meio J_{21} e J_{22}. Isto nos deixa com quatro intervalos fechados, $I_{21}, I_{22}, I_{23}, I_{24}$. Assim prosseguimos indefinidamente(como segue na figura abaixo). O conjunto C de Cantor é o conjunto dos pontos não removidos. É claro então que C é o conjunto que obtemos ao removermos do intervalo $[0, 1]$ o conjunto J, união dos intervalos J_{rs},

$$J = \cup\{J_{rs} : \ r = 1, 2, \ldots, \ s = 1, \ldots, 2^{r-1}\}.$$



A observação a seguir será útil na prova de que C é desconexo; e fornece uma caracterização dos elementos de C, quando usamos a base 3.

Observação 14 (Expansão Ternária). Os pontos do conjunto de Cantor têm uma caracterização em termos de sua representação em base 3. Dado $x \in [0,1]$, representar x na base 3 (isto é escrever uma expansão ternária de x) significa escrever $x = 0, x_1x_2x_3\ldots$ onde cada um dos dígitos x_n é igual a $0, 1$ ou 2 de tal modo que:

$$x = \frac{x_1}{3} + \frac{x_2}{3^2} + \cdots + \frac{x_n}{3^n} + \cdots$$

Não é difícil provar, usando cada etapa da construção do conjunto de Cantor que os elementos de C na base 3 são compostos por 0 e 2. Os extremos também são escritos dessa forma, como por exemplo $\frac{1}{3} = 0,0222\ldots$, pois

$$\frac{1}{3} = \frac{0}{3} + \frac{2}{9} + \frac{2}{27} + \cdots = \sum_{j=0}^{\infty} \frac{2}{3^{j+2}} = \frac{2}{9} \cdot \sum_{j=0}^{\infty} \frac{1}{3^j} = \frac{2}{9} \cdot \frac{1}{1-\frac{1}{3}} = \frac{1}{3}.$$

4 Propriedades

A seguir serão apresentadas e demonstradas as propriedades que o conjunto de Cantor possui.

(1) C é compacto.

Prova: De fato, pois $C \subset [0,1]$ e é fechado pois J é aberto. ∎

(2) C não tem pontos isolados.

Prova: Mostremos que todos os pontos de C são de acumulação, nenhum sendo isolado. Seja primeiro um ponto p que seja extremo de um dos intervalos removidos. De fato, basta observar que os intervalos que restam depois da n-ésima operação de remover intervalos tem todos o mesmo comprimento $\frac{1}{3^n}$. Assim, dado qualquer $\varepsilon > 0$, tomamos n suficientemente grande para que $\frac{1}{3^n} < \varepsilon$, logo $V_\varepsilon(p)$ certamente conterá o extremo de algum intervalo removido na n-ésima operação. Seja agora $p \in C$ um ponto que não seja extremo de nenhum intervalo J_{rs}. Dado qualquer $\varepsilon > 0$, existe algum ponto de C no intervalo $]p, p+\varepsilon[$, do contrário este intervalo estaria todo contido num dos intervalos removidos e p só poderia ser extremo de um intervalo J_{rs}, o que contraria a hipótese. Portanto p é ponto de acumulação à direita de C. Analogamente prova-se que p é ponto de acumulação à esquerda. ∎

(3) C é não enumerável.

Prova: De fato, se fosse, seus pontos seriam os elementos de uma sequência (c_n). Seja I_1 um intervalo fechado, de comprimento menor do que 1, contendo uma infinidade de pontos de C, mas não c_1; seja $I_2 \subset I_1$ outro intervalo fechado, de comprimento menor do que $\frac{1}{2}$, contendo uma infinidade de pontos de C, mas não c_2 (isto é possível porque, como provamos, todo ponto de C é de acumulação). Prosseguindo assim indefinidamente, obtemos uma sequência de intervalos fechados e encaixados $I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots$, com o comprimento de I_n tendendo a zero quando $n \to \infty$. Seja p a interseção desses intervalos. É claro que p é ponto de acumulação de C, pois qualquer vizinhança $V_\varepsilon(p)$ certamente conterá I_n a partir de certo índice $n = N$, logo conterá infinitos elementos de C. Mas C é fechado, portanto $p \in C$. Por outro lado, $p \neq c_n$ para todo n, pois $c_n \notin I_n$. Isto contradiz a hipótese inicial de que todos os elementos de C estão numa sequência (c_n). ∎

(4) C é totalmente desconexo.

Prova: Dados $x, y \in C$ com $x < y$ (análogo $y < x$). Pela representação ternária, $x = \sum x_j 3^{-j}$ e $y = \sum y_j 3^{-j}$. Como $x \neq y$, $\exists n \in \mathbb{N}$ tal que $x_j = y_j,\ j \in \{1, 2, \ldots, n-1\}$ e $x_n < y_n$. Considere $z = \sum z_j 3^{-j}$, com $z_j = x_j = y_j;\ j \in \{1, 2, \ldots, n-1\}$ e $z_n = 1$. Logo $z \notin C$, pela Observação 14, e $x < z < y$, portanto C é totalmente desconexo (Definição 1). ■

(5) C tem medida nula.

Prova: Mostremos que C tem medida nula, ou seja, $m(C) = 0$. Observe que no primeiro passo removemos um intevalo de comprimento $\frac{1}{3}$ do intervalo $[0, 1]$, depois removemos dois intervalos de comprimento $\frac{1}{3^2}$ e em seguida quatro intervalos de comprimento $\frac{1}{3^3}$ e assim por diante. Como $C \subset [0, 1]$ então:

$$m(C) = 1 - \sum_{j=0}^{\infty} \frac{2^j}{3^{j+1}} = 1 - \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} = 0.$$

■

(6) $\text{card}(C) = c$**.**

Prova: Vale observar que para provar esta propriedade basta mostrar que existe uma função $f : C \longrightarrow [0, 1]$ sobrejetora (Definição 10) já que $C \subset [0, 1]$. Seja $x \in C$ então pela Observação 14 $x = \sum a_j 3^{-j}$ onde $a_j = 0$ ou 2 para todo j. Seja $f(x) = \sum b_j 2^{-j}$ onde $b_j = \frac{a_j}{2}$. A série que define $f(x)$ é a expansão na base dois de um número em $[0, 1]$, e qualquer número em $[0, 1]$ pode ser obtido desta forma. Portanto f é sobrejetora o que implica $\text{card}(C) = c$. ■

5 Função de Cantor

Vamos examinar a função f da demonstração anterior. Primeiramente vemos que se $x, y \in C$ e $x < y$, então $f(x) < f(y)$ exceto para os pontos

extremos dos intervalos retirados. Estendemos f a uma função definida em $[0,1]$ fazendo-a constante em cada um dos intervalos retirados e com valor igual ao valor do extremo do intervalo. Esta função estendida é crescente e como a sua imagem é todo intervalo $[0,1]$ ela não pode ter saltos de descontinuidade e é portanto contínua. f é chamada a *função de Cantor.*

Agradecimentos: Agradeço a minha orientadora, Profa. Dra. Marta Cilene Gadotti, a todo apoio e dedicação oferecido neste trabalho.

Abstract: In this work we present a short historical report about Georg Cantor's life and his contributions for mathematical developing; we introduce the Cantor's set and its properties which that are related to the concept of zero measure and cardinal number.

"No one shall expel us from the paradise that Cantor has created for us." (Hilbert)

Keywords: set of Cantor; not enumerable; zero measure

Referências Bibliográficas

[1] Ávila, G., *Introdução a Análise Matemática*. São Paulo: Edgar Blücher, 1993.

[2] Barreto, A.C., *Tópicos de Análise*. Rio de Janeiro: 8º Colóquio Brasileiro de Matemática, 1971.

[3] Boyer, C.B., *História da Matemática*. São Paulo: Edgar Blücher, 1996.

[4] Carvalho, A.N., *Notas de Aula - SMA 5826 - Análise I.* ICMC - USP, São Carlos, SP, 2007.

[5] Sampaio, J.C.V., *Introdução à Teoria dos Conjuntos* `http://www.dm.ufscar.br/~sampaio/itc.html`, acesso em 02/08/2012.

Da "Teoria" dos Números à Prática

Diego Marques Mesquita[1]

Orientador(a): Prof. Dr. Romulo Campos Lins

Resumo: A Teoria dos Números é uma área da Matemática de grande interesse na Graduação. Por um lado, em seus aspectos mais elementares, ela se relaciona com muito do que é estudado de Matemática na educação básica, o que permite comparar o tratamento da Matemática escolar com o tratamento dado na Matemática "superior". Boa parte do "espírito" da Teoria dos Números reside na solução de problemas, solução que muitas vezes depende de se perceber padrões e testar conjecturas. Não tem como falar sobre teoria dos números sem falar de números primos. Na primeira parte apresentarei dois problemas envolvendo números primos, um deles que pode levar à reflexão sobre a distribuição destes nos naturais. Na segunda parte apresento e discuto um exemplo de aplicação da Teoria dos Números, o Cadastro de Pessoas Físicas (CPF), um número que é considerado "o" número de identificação das pessoas, mais do que RG ou outros. Com isso, pretendo combinar os dois pontos que não podem ficar desvinculados: a teoria e a prática.

Palavras-chave: teoria dos números; CPF; números não-livres de quadrado; números compostos

1 Sequências de números compostos

Definição 1. Chamamos de número composto todo número inteiro ≥ 0 que possui mais de dois divisores naturais.

[1]Bolsista do Programa de Educação Tutorial–MEC/SESu

Uma forma simples de determinar uma sequência utilizando o fatorial

Vamos mostrar que existem sequências de inteiros positivos e consecutivos, de tamanho arbitrário, e nas quais todos os números são compostos.

O fatorial de um número natural $n > 0$ é definido por, $n! = \prod_{i=1}^{n} i$. Claramente $n!$ é divisível por todos os inteiros de 2 a n.

E a partir de um fatorial podemos gerar uma sequência de números compostos. O raciocínio é bem simples: já que $n!$ tem fator comum com o 2, se fizermos a soma $n!+2$ obteremos um número que também é divisível por 2 (já que se $2 \mid n!$ e $2 \mid 2$ então $2 \mid (n!+2)$) e é maior que 2, logo é composto. Com o mesmo argumento, $3 \mid (n!+3), 4 \mid (n!+4), 5 \mid (n!+5), \ldots, n \mid (n!+n)$.

Com isso, acabamos de construir uma sequência de números que são compostos, porém construímos uma que contém $(n-1)$ números compostos. Logo para determinar uma sequência de n números compostos basta tomarmos a sequência:

$$(n+1)!+2, (n+1)!+3, (n+1)!+4, (n+1)!+5, \ldots, (n+1)!+n, (n+1)!+(n+1)$$

Perceba que determinamos uma sequência de n compostos, mas não sabemos se esta a que começa no menor inteiro positivo possível. Tome como exemplo a sequência com quatro números compostos:

$$(4+1)!+2, (4+1)!+3, (4+1)!+4, (4+1)!+5 = 122, 123, 124, 125$$

Sequência que obviamente é de números compostos, mas não é a menor. Tome como exemplo a sequência: $32, 33, 34, 35$. Também é uma sequência de números compostos e é menor que a anterior.

Este problema nos faz refletir sobre a distribuição dos números primos, já que conseguimos achar sequências de números compostos de tamanhos arbitrários, embora o conjunto dos números primos seja infinito.

Apresentarei agora uma sequência bem peculiar, que envolve números que não são livres de quadrado, ou seja, números que possuem pelo menos

um divisor que é um quadrado perfeito. A função de Möbius está relacionada a esta classificação dos inteiros positivos.

Vamos agora mostrar que, de maneira semelhante aos compostos, existem sequências de números que não são livres de quadrado, e de comprimento qualquer que queiramos. Utilizaremos o

Teorema 2 (Teorema Chinês do Resto). *Sejam $m_1, m_2, \ldots, m_k$ inteiros maiores que 1, primos entre si dois a dois, isto é, $(m_i, m_j) = 1$ se $i \neq j$. Neste caso, o sistema de congruências*

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ x \equiv a_3 \pmod{m_3} \\ \quad \vdots \\ x \equiv a_k \pmod{m_k} \end{cases}$$

sempre tem solução.

Prova: A demonstração desse teorema pode ser feita via uma *construção.*

Consideremos a solução $S = S_1 + S_2 + \cdots + S_k$. O fato de a representarmos como soma de k parcelas já expressa nossa intenção de que cada S_i seja $\equiv a_i \pmod{m_i}$ e $\equiv 0 \pmod{m_j}$ para $j \neq i$. A construção é a seguinte:

Definimos M_i, $1 \leq i \leq k$:

$$M_i = \frac{\prod_{t=1}^{k} m_t}{m_i}$$

Assim, M_i é divisível por todo $m_j, j \neq i$. Além disto, como os m's são primos dois a dois,

$$(M_i, m_i) = 1$$

de modo que M_i tem inverso multiplicativo M_i', $\pmod{m_i}$. Agora definimos

$$S_i = a_i M_i M_i'$$

para $1 \leq i \leq k$, e, finalmente,

$$\begin{aligned} S &= S_1 + S_2 + \cdots + S_k \\ &= a_1 M_1 M_1' + a_2 M_2 M_2' + \cdots + a_k M_k M_k'. \end{aligned}$$

É simples verificar que

$$a_i M_i M_i' \equiv a_i \times 0 \times M_i' \equiv 0 \pmod{m_j}$$

para $j \neq i$ (já que $m_j \mid M_i$), e que

$$a_i M_i M_i' \equiv a_i \times 1 \equiv a_i \pmod{m_i}$$

já que M_i' é o inverso multiplicativo de $M_i \pmod{m_i}$. ∎

Assim, determinemos, como exemplo, uma sequência com três números não-livres de quadrado.

$$\begin{cases} x & \equiv 0 \pmod{2^2} \\ x-1 & \equiv 0 \pmod{3^2} \\ x-2 & \equiv 0 \pmod{5^2} \end{cases}$$

$$M_1 = \frac{\prod_{t=1}^{3} m_t}{2^2} = 225, \qquad M_2 = \frac{\prod_{t=1}^{3} m_t}{3^2} = 100, \qquad M_3 = \frac{\prod_{t=1}^{3} m_t}{5^2} = 36.$$

E por conseguinte,

$$\begin{aligned} 225 M_1' &\equiv 1 \pmod 4 \Rightarrow M_1' \equiv 1 \pmod 4 \\ 100 M_2' &\equiv 1 \pmod 9 \Rightarrow M_2' \equiv 1 \pmod 9 \\ 36 M_3' &\equiv 1 \pmod{25} \Rightarrow M_3' \equiv 16 \pmod{25} \end{aligned}$$

Logo a solução do sistema é dado por:

$$S \equiv 0 \times 225 \times 1 + 1 \times 100 \times 1 + 2 \times 36 \times 16 \pmod{900} \Rightarrow$$
$$S \equiv 1252 \equiv 352 \pmod{900}$$

Portanto uma sequência de 3 números não-livres de quadrado pode ser

$$352, 351, 350$$

Assim como acontecia com as sequências de compostos, esta demonstração de existência não garante que vamos encontrar a "menor" sequência de números livres de quadrado. Na verdade, este é um problema em aberto! Computacionalmente, foram encontradas as menores destas sequências para tamanhos até 18, mas não acima disto (`www.marmet.org/louis/sqfgap`).

Note que também poderíamos determinar uma sequência de números compostos utilizando o *Teorema Chinês do Resto*, bastava tomar o sistema com módulos primos.

2 O Cadastro de Pessoa Física – CPF

O número de CPF de uma pessoa é constituído de 11 dígitos, sendo um primeiro bloco com 9 algarismos e um segundo com dois algarismos, que são dígitos de controle ou verificação.

O nono dígito (da esquerda para direita) corresponde a região do brasil onde o CPF foi feito. A saber:

0 : Rio Grande do Sul; 1 : Distrito Federal, Goiás, Mato Grosso, Mato Grosso do Sul e Tocantins; 2 : Amazonas, Pará, Roraima, Amapá, Acre e Rondônia; 3 : Ceará, Maranhão e Piauí; 4 : Paraíba, Pernambuco, Alagoas e Rio Grande do Norte; 5 : Bahia e Sergipe; 6 : Minas Gerais; 7 : Rio de Janeiro e Espírito Santo; 8 : São Paulo; 9 : Paraná e Santa Catarina;

O décimo dígito (que é o primeiro dígito verificador) é o resultado de uma congruência módulo 11 de um número obtido por um algoritmo com o primeiro bloco de algarismos.

Seja $a_1a_2a_3a_4a_5a_6a_7a_8a_9$ a sequência formada pelos 9 dígitos, devemos multiplicá-los, nessa ordem, pela base $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ e somar os produtos obtidos. Se a soma obtida é S, então o décimo dígito a_{10} é dado por $S - a_{10} \equiv 0 \pmod{11}$. Note que tal número será o próprio resto da divisão

por 11. A determinação do segundo dígito verificador é feita de modo similar, sendo que agora acrescentamos o décimo dígito e usamos uma base de multiplicação de 0 a 9. Caso a_{10} ou a_{11} forem iguais a 10 consideraremos como dígito o número 0.

Assim, o CPF é constituído por oito "números aleatórios", um número prefixado (região) e dois obtidos dos outros através da congruência.

Achemos os dígitos de controle deste CPF:

2	3	5	3	4	3	1	0	4
1	2	3	4	5	6	7	8	9

Efetuando as operações correspondentes, teremos:

$$2\times 1+3\times 2+5\times 3+3\times 4+4\times 5+3\times 6+1\times 7+0\times 8+4\times 9=116.$$
$$116\equiv 6 \pmod{11}$$

Obtemos agora o segundo dígito:

$$2\times 0+3\times 1+5\times 2+3\times 3+4\times 4+3\times 5+1\times 6+0\times 7+4\times 8+6\times 9=145.$$
$$145\equiv 2 \pmod{11}$$

Logo o CPF completo é 235.343.104–62.

Discussão

Mas por que multiplicar por esta base? Ou por que multiplicar por uma base, não poderíamos apenas somar e fazer a congruência? Por que

são dois dígitos verificadores? Não poderia ser apenas um? Por que a soma dos 11 dígitos resulta num número de dois dígitos iguais?

Repare que não necessariamente precisamos multiplicar pela base $\{1,2,3,4,5,6,7,8,9\}$, poderíamos multiplicar por $\{10,9,8,7,6,5,4,3,2\}$, efetuar a soma e fazer a congruência módulo 11. Caso o resto da divisão seja menor que 2, o nosso primeiro dígito verificador se torna 0 (zero), caso contrário subtrai-se o valor obtido de 11. E para calcular o segundo dígito utiliza-se a base $\{11,10,9,8,7,6,5,4,3,2\}$, e repete-se o processo.

Utiliza-se os multiplicadores e a congruência para ser possível detectar os erros de digitação mais frequentes, que são o erro em um único dígito (DS) e a inversão de dígitos adjacentes (IDA), que representam cerca de 85% dos erros de digitação, segundo levantamentos estatísticos. A soma simples, com congruência módulo 11 já permite detectar DS. De fato, já que o valor mínimo que um dígito pode ter é 0 e o máximo é 9, variando nove unidades, fazendo a congruência mod 11 detectar DS. Agora se errássemos dois dígitos, apenas a congruência não detectaria, pois poderíamos ter variado os dois dígitos de forma que o número continuasse congruente a zero mod 11.

A importância dos multiplicadores se deve, além de detectar o erro DS, na detecção dos IDAs, ou seja, quando ocorre troca de dois dígitos adjacentes, sendo que se houvesse apenas a soma dos dígitos (sem os multiplicadores) não seria detectado o erro. De fato, suponha que o dígito a_i está sendo multiplicado por m, logo o dígito a_{i+1} é multiplicado por $(m+1)$. Queremos saber quando não é detectado o erro ao trocar a_i por a_{i+1}, em outras palavras,

$$\begin{aligned} ma_i + (m+1)a_{i+1} \equiv ma_{i+1} + (m+1)a_i \pmod{11} \Rightarrow \\ ma_i + ma_{i+1} + a_{i+1} \equiv ma_i + ma_{i+1} + a_i \pmod{11} \Rightarrow \\ a_{i+1} \equiv a_i \pmod{11} \Rightarrow a_{i+1} = a_i, \end{aligned}$$

já que ambos são menores que onze.

O que faz todo sentido, já que para não ser detectado o "erro" os dois

dígitos devem ser iguais! Observe que o uso dos multiplicadores detecta até mesmo se houver a troca do nono dígito com o primeiro dígito verificador.

Se apenas o primeiro dígito verificador já sustenta a validade dos 9 anteriores, qual seria a "utilidade" do segundo dígito?

Em buscas não achei nenhuma fonte que afirmasse com certeza qual o propósito do segundo dígito verificador, mas durante as discussões com meu orientador fizemos uma reflexão sobre esse item. Concluímos que um motivo seria dar integralidade ao primeiro dígito verificador, já que o mesmo tem a importância fundamental de validar os outros, e seu erro pode impedir a validação do CPF. Mas esta é apenas uma conjectura nossa; o ISBN, número que identifica livros, usa o mesmo esquema de multiplicadores do CPF, mas com apenas um dígito verificador.

Uma curiosidade que a maioria das pessoas conhecem sobre o CPF é que a soma de seus dígitos resulta em um número de dois dígitos iguais, ou seja, um número de 0 a 100 que é múltiplo de 11. É fácil mostrar a validade dessa curiosidade por meio da construção do CPF.

De fato, seja $a_1\,a_2\,a_3\,a_4\,a_5\,a_6\,a_7\,a_8\,a_9\,a_{10}\,a_{11}$ os dígitos do CPF, tal que $a_{10} = a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 + 6a_6 + 7a_7 + 8a_8 + 9a_9 \pmod{11}$ e $a_{11} = a_2 + 2a_3 + 3a_4 + 4a_5 + 5a_6 + 6a_7 + 7a_8 + 8a_9 + 9a_{10} \pmod{11}$

Verifiquemos como ocorre a soma dos dígitos na congruência módulo 11.

$$\begin{gathered}
a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + a_{10} + a_{11} \pmod{11} \Rightarrow \\
a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + \\
(a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 + 6a_6 + 7a_7 + 8a_8 + 9a_9) + \\
(a_2 + 2a_3 + 3a_4 + 4a_5 + 5a_6 + 6a_7 + 7a_8 + 8a_9 + 9a_{10}) \pmod{11} = \\
= (a_1 + a_1 + 9a_1) + (a_2 + 2a_2 + a_2 + 18a_2) + (a_3 + 3a_3 + 2a_3 + 27a_3) + \\
(a_4 + 4a_4 + 3a_4 + 36a_7) + (a_5 + 5a_5 + 4a_5 + 45a_5) + (a_6 + 6a_6 + 5a_6 + 54a_6) + \\
(a_7 + 7a_7 + 6a_7 + 63a_7) + (a_8 + 8a_8 + 7a_8 + 72a_8) + (a_9 + 9a_9 + 8a_9 + 81a_9) \pmod{11} \\
= 11a_1 + 22a_2 + 33a_3 + 44a_4 + 55a_5 + 66a_6 + 77a_7 + 88 + a_8 + 99a_9 \equiv 0 \pmod{11}
\end{gathered}$$

Comprovando, então, a validade da característica do CPF.

Agradecimentos: Agradeço, primeiramente, ao meu orientador Prof. Dr. Romulo Campos Lins por ter trabalhado com tanto afinco, por ter sido um dos responsáveis pela minha evolução no conhecimento matemático nesses dois anos de parceria e também, além de tudo, por ser um ótimo amigo. Agradeço aos meus familiares e à minha namorada Andressa que sempre estão me incentivando e me fazendo feliz.

Abstract: Number Theory is an area of mathematics that is of great interest to undergraduate students. On the one hand, in its most basic aspects, it relates to a lot of the mathematics that is studied in basic school, allowing one to compare the treatment of school mathematics with the treatment given in college mathematics. Much of the "heart" of number theory lies in the solution of problems, solution that often depends on detecting patterns and testing conjectures. One cannot talk about number theory without talking about prime numbers. On the first part I present two problems involving prime numbers, which may lead to an initial reflection about the distribution of prime numbers. On the second part I present and discuss an example of application of number theory, the *Cadastro de Pessoas Físicas* (CPF), an identification number for Brazilian citizens (similarly to the Social Security number in the USA). My intention is to bring together two aspects which should not be left separated: theory and practice.

Keywords: number theory; CPF; square-free number; composite number

Referências Bibliográficas

[1] Vinogradov, I.M., *Elements of number theory*. NY, Dover Publications, 1954.

[2] Santos, J.P. de Oliveira, *Introdução à Teoria dos Números*. Rio de Janeiro, IMPA, CNPq, 2000.

[3] de Sá, I.P., *Aritmética modular e algumas de suas aplicações*. Rio de Janeiro, UNIFESO.

O autovetor de $ 25.000.000.000

Givanildo Donizeti de Melo[1] e Márcia Ritchielle da Silva[1]
Orientador(a): Profa. Dra. Marta Cilene Gadotti

Resumo: Neste artigo apresentamos o método utilizado pelo *Google* a fim de ordenar os *sites* gerados a partir de uma pesquisa. A ideia central da ordenação é atribuir pontos para um *site* dado, segundo à quantidade de *links* feitos de outras páginas para este *site*. A página apresentada pelo *Google* em primeiro lugar é aquela que possui maior pontuação. Para a construção do algoritmo utilizamos alguns conceitos básicos de Álgebra Linear, tais como autovetor e autovalor.

Palavras-chave: *Google*; álgebra linear; classificação de páginas; autovetor

1 Introdução

Quando um determinado assunto é inserido no *Google*, nota-se que os *sites* relacionados a esta pesquisa estão dispostos em uma certa ordem. A ordenação desses *sites* é dada segundo uma pontuação que está vinculada a quantidade de *links* feitos entre as páginas da *web*, isto é, quanto mais *links* uma página receber das outras páginas, maior será sua pontuação. Para definir o índice de importância de uma dada página na *web*, tratamos de escrever o problema na forma matricial e estudamos os autovalores e as multiplicidades geométricas.

Na seção 2 apresentamos as definições necessárias para o entendimento deste trabalho; na seção 3 introduzimos o algoritmo no caso da multiplicidade geométrica do autovalor ser 1; nas seções 4 e 5, tratamos do caso em que a multiplicidade geométrica é maior que 1 e por fim, na seção 6, apresentamos um resultado sobre a construção do autovetor numa *web* com

[1]Bolsista do Programa de Educação Tutorial (PET)–SESu/MEC

muitas páginas. Este trabalho é baseado na referência [2].

2 Preliminares

Nesta seção definimos alguns conceitos básicos de Álgebra Linear, que serão utilizados na construção do algoritmo de classificação das páginas.

Definição 1. Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear. Um escalar $\lambda \in \mathbb{R}$ é chamado autovalor de T, se existir um vetor não-nulo $v \in \mathbb{R}^n$, para o qual

$$T(v) = \lambda v.$$

Todo vetor não nulo que satisfaz essa relação é chamado um *autovetor* de T associado ao autovalor λ. O conjunto $V_\lambda(T) = \{v \in \mathbb{R}^n; T(v) = \lambda v\}$ é um subespaço de $\mathbb{R}^n$, chamado autoespaço associado a λ. A dimensão de $V_\lambda(T)$ é chamada multiplicidade geométrica do autovalor λ.

Observação 2. A definição 1 pode ser dada para uma matriz quadrada A, já que sempre podemos construir a matriz da transformação. Os conceitos e resultados sobre a matriz da transformação podem ser encontrados em [1].

Definição 3. Se $A = (a_{ij})$, $a_{ij} \in \mathbb{R}$ é uma matriz $n \times n$ e I é a matriz identidade, definimos o polinômio característico de A como sendo:

$$P_A(\lambda) = \det(A - \lambda I).$$

Então os autovalores de A são os valores de λ que satisfazem $P_A(\lambda) = 0$. E os autovetores associados a um certo autovalor λ, são os vetores $v \in \mathbb{R}^n$ que satisfazem $(A - \lambda I)v = 0$.

Definição 4. Uma matriz quadrada é uma matriz coluna estocástica quando todas as sua entradas são não negativas e a soma das entradas em cada coluna é 1.

Definição 5. Uma matriz $M = (M_{ij})$ quadrada de ordem n é positiva se

$$M_{ij} > 0, \ \forall i, j \in \{1, 2, \ldots, n\}.$$

3 Algoritmo

O *Google* organiza os *sites* segundo uma pontuação e a ideia central na atribuição de pontos para um *site* dado, está vinculada à quantidade de *links* feitos de outras páginas para este *site*.

Chamamos esta pontuação de índice de importância, que será um número real não negativo.

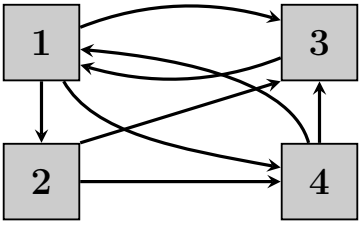
Notação: x_k (índice de importância da página k).

Este valor é definido por:

$$x_k = \sum_{j \in L_k} \frac{x_j}{n_j}, \tag{1.1}$$

onde L_k é o conjunto dos *sites* que possuem ligações para a página k e n_j é o número de *links* que saem da página j.

Exemplo 6. De acordo com uma pesquisa sobre um determinado tema inserido no *Google*, obtemos a seguinte *web*:



Os quadrados numerados representam as páginas da *web* e as setas os links entre elas.

Veremos qual destas páginas o *Google* classificará como a mais importante, segundo (1.1).

Os conjuntos dos links que chegam em cada página k, $k \in \{1,2,3,4\}$, são:

$$L_1 = \{3,4\}; \qquad L_2 = \{1\}; \qquad L_3 = \{1,2,4\}; \qquad L_4 = \{1,2\}.$$

E o número de links que saem de cada página j, $j \in \{1,2,3,4\}$, são:

$$n_1 = 3; \qquad n_2 = 2; \qquad n_3 = 1; \qquad n_4 = 2.$$

Usando a igualdade (1.1), obtemos:

$$\begin{cases} x_1 &= x_3 + \dfrac{x_4}{2} \\ x_2 &= \dfrac{x_1}{3} \\ x_3 &= \dfrac{x_1}{3} + \dfrac{x_2}{2} + \dfrac{x_4}{2} \\ x_4 &= \dfrac{x_1}{3} + \dfrac{x_2}{2} \end{cases}$$

Essas equações lineares podem ser escritas na forma $Ax = x$:

$$\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

onde A é uma matriz quadrada de ordem 4 e $x = (x_1, x_2, x_3, x_4)$ é o autovetor associado ao autovalor 1 da matriz A. Cada componente x_i deste vetor corresponde ao índice de importância da página i ($i \in \{1,2,3,4\}$). Isso transforma o problema de classificação de páginas em encontrar o autovetor x associado ao autovalor 1 da matriz A.

De $Ax = x$ segue que $Ax - Ix = 0 \Leftrightarrow (A - I)(x) = 0 \Leftrightarrow$

$$\begin{pmatrix} -1 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & -1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolvendo este sistema temos,

$$x_2 = \frac{x_1}{3}, \qquad x_3 = \frac{3x_1}{4}, \qquad x_4 = \frac{x_1}{2}.$$

Assim,

$$V_1(A) = \left\{ \left(x_1, \frac{x_1}{3}, \frac{3x_1}{4}, \frac{x_1}{2} \right), x_1 \in \mathbb{R} \right\} \Leftrightarrow V_1(A) = [(12, 4, 9, 6)],$$

onde $[(12, 4, 9, 6)]$ indica o espaço gerado pelo vetor $(12, 4, 9, 6)$.

Portanto, todos os autovetores da matriz A que estão associados ao autovalor 1 pertencem ao conjunto

$$\{\alpha\,(12, 4, 9, 6)\,;\ \text{com } \alpha \in \mathbb{R}\}.$$

Normalizando o vetor gerador segundo a norma

$$\|(x_1, \ldots, x_n)\| = |x_1| + \cdots + |x_n|,$$

obtemos

$$\left(\frac{12}{31}, \frac{4}{31}, \frac{9}{31}, \frac{6}{31} \right).$$

Cada entrada x_i deste autovetor representa o índice de importância do *site* i. Assim, listamos abaixo a classificação de páginas dada pelo *Google*, segundo a pontuação encontrada neste exemplo.

1º) A página 1 com índice de importância $\frac{12}{31} = 0.387$;

2º) A página 3 com índice de importância $\frac{4}{31} = 0.290$;

3º) A página 4 com índice de importância $\frac{9}{31} = 0.194$;

4º) A página 2 com índice de importância $\frac{6}{31} = 0.129$.

Portanto, a página 1 é a mais importante, logo, esta será a primeira página listada na pesquisa.

Note que a matriz A é uma matriz coluna estocástica e o autovetor encontrado está associado ao autovalor 1. Esta matriz sempre terá autovalor 1? A proposição a seguir responderá esta pergunta.

Proposição 7. *Toda matriz coluna estocástica tem* 1 *como autovalor.*

Prova: Seja $A = (a_{ij})_{n\times n}$ uma matriz coluna estocástica e considere o vetor $w = (1, 1, \ldots, 1) \in \mathbb{R}^n$.

Como A é uma matriz coluna estocástica, segue que

$$\sum_{i=1}^{n} a_{ij} = 1, \ \forall j = 1, \ldots, n \tag{1.2}$$

Lembremos que A e sua transposta A^t tem os mesmos autovalores, pois

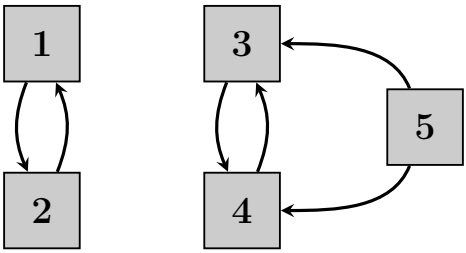
$$\det(A^t - \lambda I) = \det(A - \lambda I).$$

Desse modo, basta provar que $\lambda = 1$ é autovalor de A^t associado ao autovetor w, isto é, $A^t w = w$. Como

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \Leftrightarrow \begin{cases} a_{11} + \cdots + a_{n1} &= 1 \\ a_{12} + \cdots + a_{n2} &= 1 \\ \vdots & \vdots \\ a_{1n} + \cdots + a_{nn} &= 1 \end{cases}$$

segue de (1.2) que 1 é autovalor para A^t. Logo, é também para A. ∎

Exemplo 8. Considere, agora, a *web* com 5 páginas:



Analogamente ao Exemplo 6, obtemos a matriz

$$A' = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0,5 \\ 0 & 0 & 1 & 0 & 0,5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

A solução da igualdade $A'x = x \Leftrightarrow (A' - I)(x) = 0$, é

$$x_1 = x_2; \qquad x_3 = x_4; \qquad x_5 = 0.$$

Logo,

$$V_1(A') = \{(x_1, x_1, x_3, x_3, 0), x_1, x_3 \in \mathbb{R}\} \Leftrightarrow$$
$$V_1(A') = [(1, 1, 0, 0, 0), (0, 0, 1, 1, 0)].$$

Note que a dimensão deste autoespaço $V_1(A')$ é 2. E agora? Qual vetor devemos escolher para saber qual é a página mais importante?

A seguir veremos um algoritmo eficaz para a escolha do autovetor, cujas coordenadas fornecem a ordenação das páginas, no caso em que a multiplicidade geométrica do autoespaço é maior que 1.

4 A Modificação da matriz A quando $\dim V_1(A) > 1$

Seja S uma matriz $n \times n$ com todas as entradas iguais a $\frac{1}{n}$, em que n é o número de páginas da *web*. Segue que a matriz S é coluna estocástica, ou seja, a soma das entradas de cada coluna é igual a 1.

Proposição 9. *Seja $V_1(S)$ o autoespaço associado ao autovalor 1 da matriz S. Este espaço vetorial é unidimensional.*

Prova: Considere a igualdade $Sx = x$, onde $x = (x_1, \ldots, x_n)$, assim

$$\begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow$$

$$\begin{cases} \frac{1}{n}x_1 + \cdots + \frac{1}{n}x_n &= x_1 \\ \vdots & \vdots \\ \frac{1}{n}x_1 + \cdots + \frac{1}{n}x_n &= x_n \end{cases} \Rightarrow$$

$$\begin{cases} \frac{1}{n}(x_1 + \cdots + x_n) &= x_1 \\ \vdots & \vdots \\ \frac{1}{n}(x_1 + \cdots + x_n) &= x_n \end{cases} \Rightarrow$$

$$\begin{cases} x_1 + \cdots + x_n &= nx_1 \\ \vdots & \vdots \\ x_1 + \cdots + x_n &= nx_n \end{cases} \Rightarrow$$

$$x_1 = \cdots = x_n.$$

Logo $V_1(S) = [(1, \ldots, 1)]$. Portanto $V_1(S)$ é unidimensional. ∎

A partir das matrizes A e S, definimos a matriz M como sendo a média ponderada entre A e S, isto é:

$$M = (1 - m)A + mS, \tag{1.3}$$

onde $m \in [0, 1]$. Note que se $m = 0$, voltamos ao problema inicial $M = A$. Por outro lado, se $m = 1$ temos que $M = S$. Assim todas as páginas da *web* serão classificadas com o mesmo índice de importância, já que o autoespaço $V_1(S) = [(1, \ldots, 1)]$.

O valor m originalmente usado pelo *Google* é $0,15$.[2]

[2] A justificativa para o valor de $m = 0,15$ está nas referências [3, 4].

Lema 10. *A matriz M dada em* (1.3) *é uma matriz coluna estocástica.*

Prova: Seja L uma coluna qualquer de M.

$$L = \begin{pmatrix} (1-m)a_{1j} + m\dfrac{1}{n} \\ \vdots \\ (1-m)a_{nj} + m\dfrac{1}{n} \end{pmatrix}$$

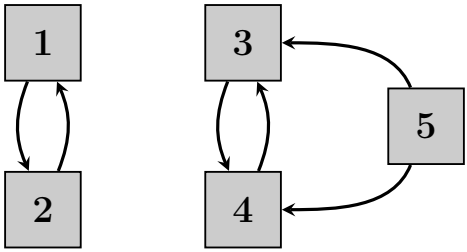
Então,

$$\sum_{i=1}^{n}\left[(1-m)a_{ij} + m\frac{1}{n}\right] = \sum_{i=1}^{n} a_{ij} - \sum_{i=1}^{n} ma_{ij} + \sum_{i=1}^{n} m\frac{1}{n} =$$

$$= 1 - m\sum_{i=1}^{n} a_{ij} + m\sum_{i=1}^{n}\frac{1}{n} = 1 - m + m = 1.$$

Portanto, M é uma matriz coluna estocástica. ∎

Exemplo 11. Considere agora a seguinte *web*:



Repetindo o procedimento do exemplo 6, obtemos a matriz A':

$$A' = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0,5 \\ 0 & 0 & 1 & 0 & 0,5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Neste caso S é a matriz:

$$S = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

Utilizando $m = 0,15$ podemos calcular $M = (1-m)A + mS$, logo

$$M = \begin{pmatrix} 0,03 & 0,88 & 0,03 & 0,03 & 0,03 \\ 0,88 & 0,03 & 0,03 & 0,03 & 0,03 \\ 0,03 & 0,03 & 0,03 & 0,88 & 0,455 \\ 0,03 & 0,03 & 0,88 & 0,03 & 0,455 \\ 0,03 & 0,03 & 0,03 & 0,03 & 0,03 \end{pmatrix}.$$

Resolvendo $Mx = x$, encontramos o autovetor

$$x = (0.2, 0.2, 0.285, 0.285, 0.03)\,.$$

Portanto, o *Google* apresentará as páginas desta *web* na seguinte ordem:
1º) As páginas 3 e 4 com índice de importância 0.285;
2º) As páginas 1 e 2 com índice de importância 0.2;
3º) A página 5 com índice de importância 0.03.

5 Análise da matriz M

O objetivo desta seção é demonstrar que a dimensão do autoespaço $V_1(M)$ é 1, lembrando que a matriz M é a média ponderada entre as matrizes A e S, que definimos na seção anterior. Na prova desse resultado utilizamos duas proposições que serão enunciadas e provadas a seguir.

Proposição 12. *Se $M = (M_{ij})$ é uma matriz coluna estocástica positiva, então qualquer autovetor em $V_1(M)$ tem todas as componentes positivas ou todas as componentes negativas.*

Prova: Vamos provar por contradição. Temos pela desigualdade triangular que

$$\left|\sum_i y_i\right| \leqslant \sum_i |y_i|\,,$$

para todo y_i real. Quando os y_i são de sinais mistos, a desigualdade acima é estritamente menor.

Seja $x \in V_1(M)$, x contém elementos de sinais mistos. Como $Mx = x$ temos $x_i = \sum_{j=1}^{n} M_{ij}x_j$, onde $M_{ij}x_j$ somados são de sinais mistos (desde que $M_{ij} > 0$). Então,

$$|x_i| = \left|\sum_{j=1}^{n} M_{ij}x_j\right| < \sum_{j=1}^{n} M_{ij}|x_j| \Rightarrow$$

$$\Rightarrow \sum_{i=1}^{n} |x_i| < \sum_{i=1}^{n}\left(\sum_{j=1}^{n} M_{ij}|x_j|\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} M_{ij}\right)|x_j| = \sum_{j=1}^{n}|x_j|,$$

na última igualdade usamos o fato de M ser uma matriz coluna estocástica, isto é, $\sum_{i=1}^{n} M_{ij} = 1$ para $j \in 1, \ldots, n$. Note que $\sum_{i=1}^{n}|x_i| < \sum_{j=1}^{n}|x_j|$ é uma contradição.

Portanto, x não pode ter componentes positivas e negativas. ∎

Observação 13. Seja $x = (x_1, \ldots, x_n) \in V_1(M)$. Pela Proposição 12 temos que x não possui componentes de sinais mistos. Se $x_i \geqslant 0$ para todo i (nem todos os x_i são nulos) então $x_i > 0$, segue imediatamente de $x_i = \sum_{j=1}^{n} M_{ij}x_j$ e $M_{ij} > 0$. Do mesmo modo $x_i \leqslant 0$ para todo i implica que cada $x_i < 0$.

Proposição 14. *Sejam v e w vetores linearmente independentes em $\mathbb{R}^n$ ($n \geqslant 2$). Para alguns valores $\alpha, \beta \in \mathbb{R}$, não todos nulos, o vetor $x = \alpha v + \beta w$ tem componentes positivas e negativas.*

Prova: Por hipótese, $\{v, w\}$ é L.I., logo $v \neq 0$ e $w \neq 0$.

Seja $d = \sum_{i=1}^{n} v_i$. Se $d = 0$ segue que $v_1 + \cdots + v_n = 0$, isto é, v contém componentes de sinais mistos. Escolhendo $\alpha = 1$ e $\beta = 0$, implica $v = x$. Portanto, x tem componentes positivas e negativas, e o resultado está provado.

Agora, se $d \neq 0$, considere $\alpha = -\frac{1}{d}\sum_{i=1}^{n} w_i$ e $\beta = 1$, e assim

$$x = \left(-\frac{1}{d}\sum_{i=1}^{n} w_i\right) v + w.$$

Logo, as componentes de $x = (x_1, \ldots, x_n)$ satisfazem:

$$\left\{\begin{array}{rcl} x_1 &=& \left(-\frac{1}{d}\sum_{i=1}^{n} w_i\right) v_1 + w_1 \\ x_2 &=& \left(-\frac{1}{d}\sum_{i=1}^{n} w_i\right) v_2 + w_2 \\ &\vdots& \\ x_n &=& \left(-\frac{1}{d}\sum_{i=1}^{n} w_i\right) v_n + w_n \end{array}\right. \Rightarrow$$

$$\sum_{i=1}^{n} x_i = \left(-\frac{1}{d}\sum_{i=1}^{n} w_i\right)(v_1+\cdots+v_n)+(w_1+\cdots+w_n) = -\sum_{i=1}^{n} w_i + \sum_{i=1}^{n} w_i = 0.$$

Portanto, x possui componentes de sinais mistos. ∎

Agora possuímos ferramentas para demonstrar o resultado abaixo.

Teorema 15. *Se M é uma matriz coluna estocástica positiva, então $V_1(M)$ é unidimensional.*

Prova: Suponhamos, por contradição, que existam dois autovetores linearmente independente $v, w \in V_1(M)$. Seja $x \in V_1(M)$ não nulo, tal que $x = \alpha v + \beta w$, onde $\alpha, \beta \in \mathbb{R}$ são escolhidos segundo a Proposição 14. Assim x tem componentes positivas e negativas, isto é, x possui componentes de sinais mistos. Mas isso é uma contradição, pois pela Proposição 12, temos que, qualquer vetor em $V_1(M)$ tem todas as suas componentes positivas ou negativas.

Portanto $V_1(M)$ não contém dois vetores L.I.. Logo $V_1(M)$ é unidimensional. ■

6 Como encontrar o autovetor em uma *web* com muitas páginas?

Atualmente, a *web* contém ao menos oito milhões de páginas. Por isso calcular os autovetores utilizando o método apresentado torna-se inviável.

A ideia básica para calcular o autovetor q de $ 25.000.000.000 da matriz M numa *web* com n páginas é gerar uma sequência de vetores que converge para o mesmo.

Para isto, seja $u_0 \in \mathbb{R}^n$ um vetor positivo qualquer fixado com $\|u_0\| = 1$ e considere a seguinte sequência $(u_k)_{k\in\mathbb{N}}$ de vetores no $\mathbb{R}^n$, definida por

$$\begin{array}{ccccc} u_1 & = & Mu_0 & & \\ u_2 & = & Mu_1 & = & M^2u_0 \\ u_3 & = & Mu_2 & = & M^3u_0 \\ \vdots & & \vdots & & \vdots \\ u_k & = & Mu_{k-1} & = & M^ku_0 \\ \vdots & & \vdots & & \vdots \end{array}$$

Lembremos que a norma $\|\cdot\|$ de um vetor $v = (v_1, \ldots, v_n)$ em $\mathbb{R}^n$ é definida pela norma da soma, isto é,

$$\|v\| = \sum_{i=1}^{n} |v_i| = |v_1| + \cdots + |v_n|.$$

Proposição 16. *Sejam M uma matriz coluna estocástica e*

$$V = \{v \in \mathbb{R}^n; \sum_{j=1}^{n} v_j = 0\},$$

que é um subespaço do $\mathbb{R}^n$. Então $Mv \in V$ e $\|Mv\| \leqslant c\|v\|$ para qualquer $v \in V$, onde

$$c = \max_{1 \leqslant j \leqslant n} |1 - 2 \min_{1 \leqslant i \leqslant n} M_{ij}| < 1.$$

Prova: Para provar que $Mv \in V$ considere $w = Mv$, então $w_i = \sum_{j=1}^{n} M_{ij} v_j$ e

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} v_j = \sum_{j=1}^{n} v_j \left(\sum_{i=1}^{n} M_{ij} \right) = \sum_{j=1}^{n} v_j = 0.$$

Note que $\sum_{i=1}^{n} M_{ij} = 1$, pois a matriz M é coluna estocástica. Portanto, $w = Mv \in V$.

Mostremos agora que $\|Mv\| \leqslant c\|v\|$, note que

$$\|w\| = \sum_{i=1}^{n} e_i w_i = \sum_{i=1}^{n} e_i \left(\sum_{j=1}^{n} M_{ij} v_j \right),$$

onde $e_i = \operatorname{sgn}(w_i)$ (sgn significa o sinal de w_i), e que os e_i não são todos de um mesmo sinal, uma vez que $\sum_{i=1}^{n} w_i = 0$ (a menos que $w = 0$). E temos também

$$\|w\| = \sum_{j=1}^{n} v_j \left(\sum_{i=1}^{n} e_i M_{ij} \right) = \sum_{j=1}^{n} a_j v_j, \tag{1.4}$$

onde $a_j = \sum_{i=1}^{n} e_i M_{ij}$. Como e_i é de sinal misto e $\sum_{i=1}^{n} M_{ij} = 1$ com $0 < M_{ij} < 1$, vemos que

$$-1 < -1 + 2 \min_{1 \leqslant i \leqslant n} M_{ij} \leqslant a_j \leqslant 1 - 2 \min_{1 \leqslant i \leqslant n} M_{ij} < 1.$$

Pois, lembremos que

$$M_{1j} + M_{2j} + \cdots + M_{nj} = 1 \tag{1.5}$$

e $0 < M_{ij} < 1$, então

$$a_j = \sum_{i=1}^{n} e_i M_{ij} = -\left(M_{i_1 j} + \cdots + M_{i_k j}\right) + M_{i_{k+1} j} + \cdots + M_{i_n j}, \tag{1.6}$$

neste caso suponhamos que k elementos do somatório acima sejam de sinais negativos e $n-k$ elementos sejam positivos e por isso introduzimos os índices $i_1, \ldots, i_k, i_{k+1}, \ldots, i_n$. Usando (1.5) na igualdade (1.6) temos

$$a_j = -\left(M_{i_1 j} + \cdots + M_{i_k j}\right) + \left(1 - \left(M_{i_1 j} + \ldots + M_{i_k j}\right)\right) \Rightarrow$$

$$a_j = -2\left(M_{i_1 j} + \cdots + M_{i_k j}\right) + 1 \Rightarrow$$

$$a_j \leqslant -2 \min_{1 \leqslant i \leqslant n} M_{ij} + 1 < 1.$$

Novamente, usando

$$a_j = e_1 M_{1j} + \cdots + e_n M_{nj} = -\left(M_{i_1 j} + \cdots + M_{i_k j}\right) + M_{i_{k+1} j} + \cdots + M_{i_n j}$$

por (1.5), segue que

$$a_j = -\left(1 - \left(M_{i_{k+1} j} + \cdots + M_{i_n j}\right)\right) + \left(M_{i_{k+1} j} + \cdots + M_{i_n j}\right) \Rightarrow$$

$$\Rightarrow a_j = -1 + 2\left(M_{i_{k+1} j} + \ldots + M_{i_n j}\right) \Rightarrow$$

$$\Rightarrow a_j \geqslant -1 + 2(n-k) \min_{1 \leqslant i \leqslant n} M_{ij} \geqslant -1 + 2 \min_{1 \leqslant i \leqslant n} M_{ij} > -1.$$

Assim,

$$|a_j| \leqslant |1 - 2 \min_{1 \leqslant i \leqslant n} M_{ij}| < 1, \forall\ j = 1, \ldots, n$$

Seja $c = \max_{1 \leqslant j \leqslant n} |1 - 2 \min_{1 \leqslant i \leqslant n} M_{ij}|$. Observe que $c < 1$ e $|a_j| \leqslant c$ para todo j. Usando (1.4), temos

$$\|Mv\| = \|w\| = \sum_{j=1}^{n} a_j v_j = \left|\sum_{j=1}^{n} a_j v_j\right| \leqslant \sum_{j=1}^{n} |a_j||v_j| \leqslant c \sum_{j=1}^{n} |v_j| = c\|v\|,$$

o que prova a proposição. ∎

Os resultados anteriores são ferramentas necessárias na demonstração do seguinte teorema, que fornece o autovetor quando tratamos de uma *web* com muitas páginas.

Teorema 17. *Cada matriz coluna estocástica positiva M, tem um único vetor q com componentes positivas de tal forma que $Mq = q$ com $\|q\| = 1$. Então o vetor q pode ser calculado como $q = \lim_{k\to\infty} M^k u_0$, para qualquer vetor inicial u_0 com componentes positivas tais que $\|u_0\| = 1$.*

Prova: Pelo Lema 15 a dimensão de $V_1(M)$ é 1, isto é, $V_1(M) = [p]$.

Se $p > 0$, então escolha $q = \frac{p}{\|p\|}$. Agora, se $p < 0$, tome $q = -\frac{p}{\|p\|}$.

Note que $q > 0$, isto é, $q_i > 0$ para todo $i \in \{1, \ldots, n\}$ e

$$\|q\| = |q_1| + \cdots + |q_n| = q_1 + \cdots + q_n = 1.$$

Seja $u_0 \in \mathbb{R}^n$ arbitrário com todas as componentes positivas, tais que $\|u_0\| = 1$. Podemos escrever $u_0 = q + v$, onde $v \in V$ (V é um subespaço de $\mathbb{R}^n$ definido por $V = \{v \in \mathbb{R}^n / v_1 + \cdots + v_n = 0\}$ como na Proposição 16).

Tem-se que $M^k u_0 = M^k q + M^k v = q + M^k v$, o que implica

$$M^k u_0 - q = M^k v. \tag{1.7}$$

Usando indução sobre k e a Proposição 16 é fácil ver que $\|M^k v\| \leqslant c^k \|v\|$ para $0 \leqslant c < 1$. Logo,

$$0 < \|M^k v\| \leqslant c^k \|v\| \Rightarrow 0 = \lim_{k\to\infty} 0 \leqslant \lim_{k\to\infty} \|M^k v\| \leqslant \lim_{k\to\infty} c^k \|v\| = 0.$$

Portanto, $\lim_{k\to\infty} \|M^k v\| = 0 \Rightarrow \lim_{k\to\infty} M^k v = 0$. De (1.7), concluímos que $\lim_{k\to\infty} M^k u_0 = q$. ∎

Note que o vetor q do Teorema 17 é o autovetor de $ 25.000.000.000, cujas as componentes indicam a pontuação de cada *site* da *web*. Dessa forma podemos descrever a ordenação das páginas no *Google*.

7 Conclusões

A relevância desta pesquisa consiste em seu caráter multidisciplinar, sendo portanto acessível aos alunos de outros cursos, como as Engenharias, Ciências da Computação, Física, etc.

Trabalhando com a referência [2], tivemos a oportunidade de estudar uma aplicação relativamente simples da Matemática, mas muito bonita, em um assunto interessante e importante, sobre o *site* de pesquisa utilizado pelo mundo todo, que é o *Google*.

Agradecimentos: Agradecemos a nossa orientadora Marta Cilene Gadotti e ao auxílio financeiro do Programa de Educação Tutorial (PET)–SESu/MEC.

Abstract: In this paper we present the method used by *Google* to sort the *sites* generated from a research through importance score. The central idea is to assign points for a given *site*, according to the amount of *links* made from other pages on this *site*. The page that appears first in *Google*, is the one that has highest importance score. To construct the algorithm will use some basic concepts of linear algebra, such as eigenvector and eigenvalue.

Keywords: *Google*; linear algebra; ordination of *sites*; eigenvector

Referências Bibliográficas

[1] Poole, D., *Álgebra Linear*, editora Thomson, 2004.

[2] Bryan, K.; Leise, T., *The* $25.000.000.000 *Eigenvector. The Linear Algebra Behind Google.*

[3] Langville, A.N.; Meyer, C.D., *Deeper inside PageRank*, Internet Math., 1 (2005).

[4] Moler, C., *The world's largest matrix computation*, `www.mathworks.com/company/newsletters/news_notes/clevescorner/oct02_cleve.html` (acessado em 05 de setembro de 2012).

Módulos Livres e Produtos Tensoriais

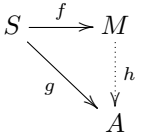
Gustavo Cazzeri Innocencio Figueiredo[1]
Orientador(a): Profa. Dra. Alice Kimie Miwa Libardi

Resumo: A Álgebra Homológica foi criada originalmente para tratar de conceitos ligados à Topologia Algébrica. No presente trabalho pretendemos introduzir algumas ideias básicas da teoria dos módulos livres e dos produtos tensoriais que são diretamente aplicados no estudo da Álgebra Homológica.

Palavras-chave: módulos sobre anéis; módulos livres; produtos tensoriais; álgebra homológica

1 Módulos livres

Definição 1. Sejam M um R-módulo, S um conjunto e $f : S \to M$ uma função. Dizemos que "M é um *módulo livre* sobre S com f" se, e somente se, para todo R-módulo A e toda função $g : S \to A$, existe um único homomorfismo $h : M \to A$ tal que $h \circ f = g$, isto é, tal que o diagrama abaixo comuta.



Exemplo 2. O R-módulo trivial $M = \{0_M\}$ é um módulo livre sobre o conjunto $\varnothing$ com a função $\varnothing : \varnothing \to \{0_M\}$. De fato, se A é um R-módulo, então existe apenas um homomorfismo $h : \{0_M\} \to A$. Claro que $h(0_M) = 0_A$. Assim, para todo R-módulo A e toda função $g : \varnothing \to A$, segue que

[1]Bolsista FAPESP 2012/04716–8

$g = \varnothing$ e existe um único homomorfismo $h : \{0_M\} \to A$. Naturalmente, temos que $h \circ \varnothing = \varnothing = g$.

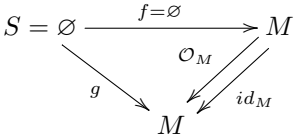
Proposição 3. *Sejam M um R-módulo livre sobre o conjunto S com a função $f : S \to M$. Temos que*

(i) *Se $S = \varnothing$, então $f = \varnothing$ e $M = \{0_M\}$;*

(ii) *f é injetora;*

(iii) *$\langle im(f)\rangle = M$.*

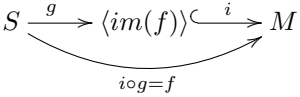
Prova: **(i)** Como $f : \varnothing \to M$, então $f = \varnothing$. Suponha que $M \neq \{0_M\}$ e sejam $g : \varnothing \to M$ e $x_0 \in M$, com $x_0 \neq 0_M$. Então, $g = \varnothing$. Considere as funções $id_M : M \to M$ e $\mathcal{O}_M : M \to M$ tais que $id_M(x) = x$ e $\mathcal{O}_M(x) = 0_M, \forall x \in M$. Então, id_M é a função identidade de M e $\mathcal{O}_M$ é a função nula de M, ambas homomorfismos. Segue que $id_M(x_0) = x_0 \neq 0_M = \mathcal{O}_M(x_0)$ e, portanto, $id_M \neq \mathcal{O}_M$. Porém, temos que $id_M \circ f = id_M \circ \varnothing = \varnothing = g = \varnothing = \mathcal{O}_M \circ \varnothing = \mathcal{O}_M \circ f$. Dessa forma, mostramos que existem um R-módulo (o próprio M) e uma função $g : \varnothing \to M$ tais que existem dois homomorfismos distintos (id_M e $\mathcal{O}_M$) que satisfazem a comutatividade do diagrama abaixo, ou seja, M não é livre sobre $\varnothing$, um absurdo com nossa hipótese. Logo, só pode ser $M = \{0_M\}$.

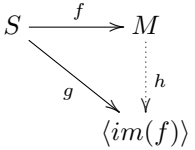


(ii) Se $S = \varnothing$, então $f = \varnothing$, que é injetora. Se S é unitário, então f é constante, que é injetora também. Suponha que existam $s_0, s_1 \in S$, com $s_0 \neq s_1$ e $f(s_0) = f(s_1)$. Considere o anel R como sendo um R-módulo sobre si mesmo. Como R é um anel comutativo com unidade e $1 \neq 0$, tome a função $g : S \to R$ tal que $g(s_1) = 1$ e $g(s) = 0, \forall s \in S \setminus \{s_1\}$. Assim, $g(s_0) = 0$. Como M é livre, temos que existe um único homomorfismo $h : M \to R$ tal que $h \circ f = g$. Portanto, $0 = g(s_0) = (h \circ f)(s_0) =$

$h\big(f(s_0)\big) = h\big(f(s_1)\big) = (h \circ f)(s_1) = g(s_1) = 1$, um absurdo. Logo, $\forall s_0, s_1 \in S$, se $s_0 \neq s_1$, então $f(s_0) \neq f(s_1)$, ou seja, f é injetora.

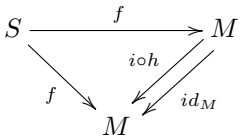
(iii) Como $im(f) \subset \langle im(f)\rangle \subset M$, seja $g : S \to \langle im(f)\rangle$ tal que $g(s) = f(s)$, $\forall s \in S$. Na teoria de conjuntos, a função g definida, vista como conjunto de pares ordenados, é a própria f. Porém, como *morfismo* da categoria de conjuntos e funções, $\big(dom(g), cod(g), g\big) = \big(S, \langle im(f)\rangle, f\big)$ é igual a $\big(dom(f), cod(f), f\big) = (S, M, f)$ se, e somente se, $\langle im(f)\rangle = M$, que é exatamente o que queremos mostrar. Para tanto, seja $i : \langle im(f)\rangle \hookrightarrow M$ a inclusão, isto é, $i(x) = x$, $\forall x \in \langle im(f)\rangle$. É claro que $i \circ g = f$, pois $(i \circ g)(s) = i\big(g(s)\big) = i\big(f(s)\big) = f(s)$, $\forall s \in S$. Temos que $dom(i) = \langle im(f)\rangle = im(i) \subset cod(i) = M$. Ficamos com os diagramas





Como $S \xrightarrow{\ g\ } \langle im(f)\rangle \leqslant M$ e M é livre sobre S, existe um único homomorfismo $h : M \to \langle im(f)\rangle$ tal que $h \circ f = g$.

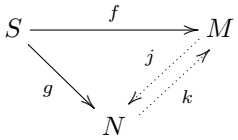
Como i e h são homomorfismos, então $i \circ h : M \to M$ também é homomorfismo. Como já vimos, a função identidade $id_M : M \to M$ também é um homomorfismo. Daí, temos o diagrama



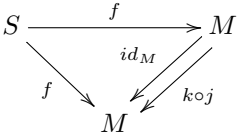
É claro que $id_M \circ f = f$. Temos também que $(i \circ h) \circ f = i \circ (h \circ f) = i \circ g = f$. Portanto, $i \circ h = id_M$. Como id_M é sobrejetora, decorre que i também é sobrejetora. Logo, $\langle im(f)\rangle = im(i) = cod(i) = M$, como queríamos demonstrar. ∎

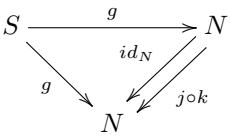
Teorema 4 (Unicidade). *Sejam S um conjunto, M um R-módulo livre sobre S com $f : S \to M$ e N um R-módulo livre sobre S com $g : S \to N$. Então, $N \cong M$ e existe um único isomorfismo $h : M \to N$ tal que $h \circ f = g$.*

Prova: Por hipótese, existem únicos homomorfismos $j : M \to N$ e $k : N \to M$ tais que $j \circ f = g$ e $k \circ g = f$.



Dessa forma, ficamos com os diagramas





É claro que id_M, id_N, $j \circ k$ e $k \circ j$ são homomorfismos e que $id_M \circ f = f$ e $id_N \circ g = g$. Temos também que $(k \circ j) \circ f = k \circ (j \circ f) = k \circ g = f$ e que $(j \circ k) \circ g = j \circ (k \circ g) = j \circ f = g$. Como M é livre, $k \circ j = id_M$. Como N é livre, $j \circ k = id_N$. Portanto, j e k são isomorfismos, com $k = j^{-1}$. Basta tomar $h = j$. ∎

Sejam S um conjunto, R um anel e M um R-módulo. Para todos $f, g \in M^S$ e todo $r \in R$, definimos $f + g : S \to M$ e $r \cdot f = rf : S \to M$ de modo que $(f + g)(s) = f(s) + g(s)$ e $(r \cdot f)(s) = r \cdot f(s)$, $\forall s \in S$. Também definimos $\alpha : M^S \times M^S \to M^S$ e $\mu : R \times M^S \to M^S$ tais que $\alpha(f, g) = f + g$ e $\mu(r, f) = r \cdot f$, $\forall f, g \in M^S$, $\forall r \in R$. Pode-se mostrar facilmente que $\big(M^S, (R, +, \cdot), \alpha, \mu\big)$ é um $(R, +, \cdot)$-módulo à esquerda. Diremos simplesmente que M^S é um R-módulo. Se $S = \varnothing$, então $M^S = \{\varnothing\}$ e $0_{M^S} = \varnothing$. Se $S \neq \varnothing$, então $\varnothing \notin M^S$ e o zero de M^S é a função constante $0_{M^S} = S \times \{0_M\}$, isto é, $0_{M^S} = \mathcal{O} : S \to M$ é tal que $\mathcal{O}(s) = 0_M$, $\forall s \in S$.

Montamos, dessa mesma maneira, o R-módulo R^S, considerando o anel R como sendo um módulo sobre si mesmo.

No mesmo contexto, sejam S um conjunto, R um anel e M um R-módulo. Considere o R-módulo M^S e seja o conjunto

$$M_S = \big\{f \in M^S : (\exists E \subset S)\big(E \text{ é finito e } f[S \setminus E] \subset \{0_M\}\big)\big\},$$

isto é,

$$M_S = \big\{f \in M^S : S \setminus f^{-1}[\{0_M\}] \text{ é finito}\big\}.$$

Os elementos de M_S são as funções definidas em S, que assumem valores em M, os quais são diferentes de 0_M para, no máximo, um número finito de elementos de S. É claro que, se S é finito, então $M_S = M^S$, pois $S \setminus f^{-1}[\{0_M\}]$ é finito, $\forall f \in M^S$. Note também que, se existe $f_0 \in M_S$ tal que $0_M \notin im(f_0)$, então S é finito.

Se $S = \varnothing$, então $M_S = M^S = \{\varnothing\} \Rightarrow M_S \neq \varnothing$. Se $S \neq \varnothing$, seja $f_0 : S \to M$ tal que $f_0(s) = 0_M$, $\forall s \in S$, isto é, f_0 é a função nula (constante) $f_0 = S \times \{0_M\} \in M^S$. Assim, $f_0^{-1}[\{0_M\}] = S \Rightarrow S \setminus f_0^{-1}[\{0_M\}] = S \setminus S = \varnothing$ é finito $\Rightarrow f_0 \in M_S \Rightarrow M_S \neq \varnothing$. Segue que $M_S \neq \varnothing$, qualquer que seja o conjunto S.

Vamos mostrar que M_S é um R-submódulo de M^S. Sejam $f, g \in M_S$. Então, existem $F \subset S$ e $G \subset S$ finitos tais que $f[S \setminus F] \subset \{0_M\}$ e $g[S \setminus G] \subset \{0_M\}$. Assim, $F \cup G \subset S$, $F \cup G$ é finito e, $\forall y \in M$, se $y \in (f+g)[S \setminus (F \cup G)]$, então existe $s \in S \setminus (F \cup G)$ tal que $y = (f+g)(s) = f(s) + g(s)$. Mas, $S \setminus (F \cup G) = (S \setminus F) \setminus G \subset S \setminus G \Rightarrow s \in S \setminus G \Rightarrow g(s) \in g[S \setminus G] \subset \{0_M\} \Rightarrow g(s) = 0_M$. Analogamente, $S \setminus (F \cup G) = (S \setminus G) \setminus F \subset S \setminus F \Rightarrow s \in S \setminus F \Rightarrow f(s) \in f[S \setminus F] \subset \{0_M\} \Rightarrow f(s) = 0_M$. Dessa forma, $y = f(s) + g(s) = 0_M + 0_M = 0_M \Rightarrow y \in \{0_M\}$. Portanto, $(f+g)[S \setminus (F \cup G)] \subset \{0_M\}$ e, consequentemente, $f + g \in M_S$. Seja também $r \in R$. Daí, $\forall z \in M$, se $z \in (r \cdot f)[S \setminus F]$, então, existe $s \in S \setminus F$ tal que $z = (r \cdot f)(s) = r \cdot f(s)$. Mas, $f(s) \subset f[S \setminus F] \subset \{0_M\} \Rightarrow f(s) = 0_M \Rightarrow z = r \cdot f(s) = r \cdot 0_M = 0_M \Rightarrow z \in \{0_M\}$. Portanto, $(r \cdot f)[S \setminus F] \subset \{0_M\}$ e, consequentemente, $r \cdot f \in M_S$. Como f, g e r são arbitrários e $M_S \neq \varnothing$, segue que $M_S \leqslant M^S$, como queríamos.

Em particular, temos que $R_S \leqslant R^S$.

Seja $S \neq \varnothing$. Note que, para todo $\phi \in R^S$, todo R-módulo M e toda função $g : S \to M$, temos que $\phi(s)\cdot g(s) \in M$. Além disso, se $\phi \in R_S$, existe um conjunto finito $E_\phi \subset S$ tal que $\phi[S\backslash E_\phi] \subset \{0\}$. Como $S \neq \varnothing$, temos que $\varnothing \notin R_S$ e, se $E_\phi = \varnothing$, então $\phi = \mathcal{O}$, em que $\mathcal{O} : S \to R$ é tal que $\mathcal{O}(s) = 0$, $\forall s \in S$, isto é, $\mathcal{O} = S\times\{0\}$. Nesse caso, temos que $\phi(s)\cdot g(s) = 0\cdot g(s) = 0_M$, $\forall s \in S$. Se $E_\phi \neq \varnothing$, existe $n \in \mathbb{N}$ tal que $E_\phi = \{s_1, \ldots, s_n\}$. Segue que, $\forall j \in \{1, \ldots, n\}$, $\phi(s_j)$ pode ser zero ou não, mas $\phi(s) = 0$, se s não for nenhum dos s_j. Denotaremos $\sum_{j=1}^{n} \phi(s_j)g(s_j)$ por '$\sum_{s\in E_\phi} \phi(s)g(s)$' ou por '$\sum_{s\in S} \phi(s)g(s)$', sem fazer menção ao conjunto E_ϕ. Dessa forma, mesmo a função nula $\mathcal{O} = S \times \{0\}$ pode ter um conjunto $E_\mathcal{O} = \{s_0\} \subset S$ finito e não-vazio (no caso, unitário) tal que $\mathcal{O}[S \setminus E_\mathcal{O}] \subset \{0\}$ e podemos escolher qualquer elemento $s_0 \in S$, dado que $S \neq \varnothing$, pois não é obrigatório que seja $\mathcal{O}(s_0) \neq 0$ e, no caso, temos justamente que $\mathcal{O}(s_0) = 0$. Portanto, não iremos considerar o caso de $E_\phi = \varnothing$, qualquer que seja $\phi \in R_S$. A diferença que queremos salientar é que, por essas considerações, se $\phi = \mathcal{O}$ ou $g[E_\phi] = \{0_M\}$, então $\sum_{s\in S} \phi(s)g(s) = 0_M$.

Além disso, para qualquer subconjunto finito $F = \{t_1, \ldots, t_k\} \subset S$, se $E_\phi \subset F$, então $\sum_{t\in F} \phi(t)g(t) = \sum_{j=1}^{k} \phi(t_j)g(t_j) = \sum_{s\in E_\phi} \phi(s)g(s) = \sum_{s\in S} \phi(s)g(s)$, pois, $\forall j \in \{1, \ldots, k\}$, se $t_j \notin E_\phi$, então $\phi(t_j)g(t_j) = 0 \cdot g(t_j) = 0_M$ não contribui na soma. Por meio de cálculos e desenvolvimento das somas, podemos concluir $\sum_{s\in S} (\phi+\psi)(s)g(s) = \sum_{s\in S} \phi(s)g(s) + \sum_{s\in S} \psi(s)g(s)$. Isso será de crucial importância em nosso próximo

Teorema 5 (Existência). *Para todo conjunto S e todo anel R comutativo com unidade, existem um R-módulo M e uma função $f : S \to M$ tais que M é livre sobre S com f.*

Prova: Considere R como sendo um R-módulo. Se $S = \varnothing$, então $f = \varnothing$ e $\{0\}$ é um R-módulo livre sobre S com f, como mostrado no exemplo 2. Se $S \neq \varnothing$, tome o R-módulo R_S. Seja $f : S \to R^S$ tal que, $\forall s \in S$,

$f(s) : S \to R$ seja dada, $\forall t \in S$, por

$$[f(s)](t) = \begin{cases} 1, & \text{se } t = s; \\ 0, & \text{se } t \neq s. \end{cases}$$

Primeiro vamos mostrar que $im(f) \subset R_S$. Para tanto, seja $\phi \in im(f)$. Então, existe $s_0 \in S$ tal que $\phi = f(s_0)$. Assim, ϕ é uma função $\phi : S \to R$ tal que

$$\phi(t) = [f(s_0)](t) = \begin{cases} 1, & \text{se } t = s_0; \\ 0, & \text{se } t \neq s_0. \end{cases}$$

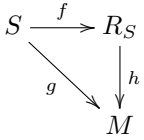
Sejam $E = \{s_0\}$ e $y \in \phi[S \setminus E]$. Então, E é finito (unitário) e $y = \phi(t)$, para algum $t \in S \setminus E = S \setminus \{s_0\} \Rightarrow t \neq s_0 \Rightarrow y = \phi(t) = 0 \in \{0\}$. Como y é qualquer, $\phi[S \setminus E] \subset \{0\}$. Portanto, $\phi \in R_S$. Como ϕ é arbitrário, temos que $im(f) \subset R_S$.

Dessa maneira, f é uma função do tipo $f : S \to R_S$. Vamos mostrar que R_S é um R-módulo livre sobre S com f.

Sejam M um R-módulo, $g : S \to M$ e $h : R_S \to M$ tais que, $\forall \phi \in R_S$, $h(\phi) = \sum\limits_{s \in S} \phi(s)g(s)$. Vamos mostrar que h é um homomorfismo.

Sejam $\phi, \psi \in R_S$ e $r \in R$. Temos que $h(\phi + \psi) = \sum\limits_{s \in S} (\phi + \psi)(s)g(s) = \sum\limits_{s \in S} \phi(s)g(s) + \sum\limits_{s \in S} \psi(s)g(s) = h(\phi) + h(\psi)$ e que $h(r\phi) = \sum\limits_{s \in S} (r\phi)(s)g(s) = r \sum\limits_{s \in S} \phi(s)g(s) = rh(\phi)$.

Temos também que $(h \circ f)(s) = h\big(f(s)\big) = \sum\limits_{t \in S} [f(s)](t) \cdot g(t) = [f(s)](s) \cdot g(s) = 1 \cdot g(s) = g(s)$, $\forall s \in S$. Daí, $h \circ f = g$, ou seja, o diagrama abaixo comuta.



Seja $w : R_S \to M$ um homomorfismo tal que $w \circ f = g$. Note que, $\forall \phi \in R_S$, temos que $\left[\sum\limits_{s \in S} \phi(s) \cdot f(s)\right](t) = \sum\limits_{s \in S} [\phi(s) \cdot f(s)](t) = \sum\limits_{s \in S} \phi(s) \cdot [f(s)](t) =$

$\phi(t) \cdot [f(t)](t) = \phi(t) \cdot 1 = \phi(t)$, $\forall t \in S$. Portanto, $\phi = \sum_{t \in S} \phi(s) \cdot f(s)$. Daí, $\forall \phi \in R_S$, ficamos com $w(\phi) = w\left(\sum_{s \in S} \phi(s) \cdot f(s)\right) = \sum_{s \in S} w\big(\phi(s) \cdot f(s)\big) = \sum_{s \in S} \phi(s) \cdot w\big(f(s)\big) = \sum_{s \in S} \phi(s) \cdot (w \circ f)(s) = \sum_{s \in S} \phi(s) \cdot g(s) = h(\phi)$. Assim, $w = \phi$. Daí, $h : R_S \to M$ é o único homomorfismo tal que $h \circ f = g$.

Como M e g são arbitrários, R_S é um R-módulo livre sobre S com f. ■

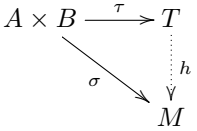
2 Produtos tensoriais

Definição 6. Sejam A, B e T R-módulos e $\tau : A \times B \to T$ uma função. Dizemos que "τ é *bilinear*" se, e somente se,

(i) $\tau(r_1 \cdot a_1 + r_2 \cdot a_2, b) = r_1 \cdot \tau(a_1, b) + r_2 \cdot \tau(a_2, b)$, $\forall a_1, a_2 \in A$, $\forall b \in B$, $\forall r_1, r_2 \in R$;

(ii) $\tau(a, s_1 \cdot b_1 + s_2 \cdot b_2) = s_1 \cdot \tau(a, b_1) + s_2 \cdot \tau(a, b_2)$, $\forall a \in A$, $\forall b_1, b_2 \in B$, $\forall s_1, s_2 \in R$.

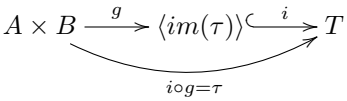
Definição 7. Sejam A, B e T R-módulos e $\tau : A \times B \to T$ uma função bilinear. Dizemos que "T é um *produto tensorial* de A e B com τ" se, e somente se, para todo R-módulo M e toda função bilinear $\sigma : A \times B \to M$, existe um único homomorfismo $h : T \to M$ tal que $h \circ \tau = \sigma$, isto é, tal que o diagrama abaixo comuta.



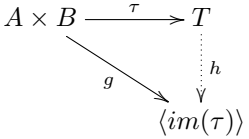
Proposição 8. *Se um R-módulo T é um produto tensorial dos R-módulos A e B com τ, então $\langle im(\tau) \rangle = T$.*

Prova: Como $im(\tau) \subset \langle im(\tau) \rangle \subset T$, seja $g : A \times B \to \langle im(\tau) \rangle$ tal que $g(a, b) = \tau(a, b)$, $\forall a \in A$, $\forall b \in B$. Seja também $i : \langle im(\tau) \rangle \hookrightarrow T$ a inclusão, isto é, $i(t) = t$, $\forall t \in \langle im(f) \rangle$. É claro que $i \circ g = f$, pois

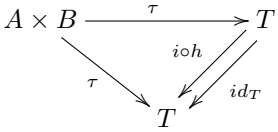
$(i \circ g)(a,b) = i\big(g(a,b)\big) = i\big(\tau(a,b)\big) = \tau(a,b)$, $\forall a \in A$, $\forall b \in B$. Temos que $dom(i) = \langle im(\tau)\rangle = im(i) \subset cod(i) = T$. Ficamos com o diagrama



Como $\langle im(\tau)\rangle \leqslant T$, g é bilinear e T é um produto tensorial de A e B com τ, existe um único homomorfismo $h : T \to \langle im(\tau)\rangle$ tal que $h \circ \tau = g$, isto é, tal que o diagrama abaixo comuta.



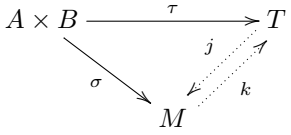
Como i e h são homomorfismos, então $i \circ h : T \to T$ também é homomorfismo. Como já vimos, a função identidade $id_T : T \to T$ também é um homomorfismo. Daí, temos o diagrama



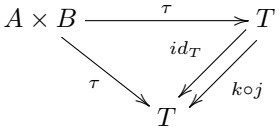
É claro que $id_T \circ \tau = \tau$. Temos também que $(i \circ h) \circ \tau = i \circ (h \circ \tau) = i \circ g = \tau$. Portanto, $i \circ h = id_T$. Como id_T é sobrejetora, decorre que i também é sobrejetora. Logo, $\langle im(\tau)\rangle = im(i) = cod(i) = T$, como queríamos demonstrar. ∎

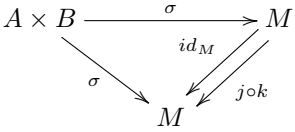
Teorema 9 (Unicidade). *Se T é um produto tensorial de A e B com τ e M é um produto tensorial de A e B com σ, então $M \cong T$ e existe um único isomorfismo $\varphi : T \to M$ tal que $\varphi \circ \tau = \sigma$.*

Prova: Por hipótese, existem únicos homomorfismos $j : T \to M$ e $k : M \to T$ tais que $j \circ \tau = \sigma$ e $k \circ \sigma = \tau$.



Dessa forma, ficamos com os diagramas





É claro que id_T, id_M, $j \circ k$ e $k \circ j$ são homomorfismos e que $id_T \circ \tau = \tau$ e $id_M \circ \sigma = \sigma$. Temos também que $(k \circ j) \circ \tau = k \circ (j \circ \tau) = k \circ \sigma = \tau$ e que $(j \circ k) \circ \sigma = j \circ (k \circ \sigma) = j \circ \tau = \sigma$. Como T é um produto tensorial de A e B com τ, então $k \circ j = id_T$. Como M é um produto tensorial de A e B com σ, então $j \circ k = id_M$. Portanto, j e k são isomorfismos, com $k = j^{-1}$. Basta tomar $\varphi = j$. ∎

Para nosso próximo objetivo, vamos precisar de uma construção que explicitaremos a seguir: Sejam F e M R-módulos, $G \leqslant F$, $T = \frac{F}{G}$, $p : F \to T$ a projeção canônica e $j : F \to M$ um homomorfismo tal que $j[G] \subset \{0_M\}$, ou seja, tal que $G \subset ker(j)$. Assim, ficamos com $G \leqslant ker(j) \leqslant F$. Considere o conjunto

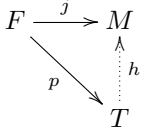
$$H = \Big\{\nu \subset T \times M : (\forall X \in T)(\forall x \in F)(\forall y \in M) \\ \big[(X, y) \in \nu \text{ e } X = x + G \Rightarrow y = j(x)\big]\Big\}.$$

Para cada $x \in F$ seja $\nu_x = \big\{\big(x + G, j(x)\big)\big\}$. Temos que $\nu_x \in H$, $\forall x \in F$. De fato, seja $x \in F$. Temos que $x + G \in T$ e $j(x) \in M \Rightarrow \big(x + G, j(x)\big) \in T \times M \Rightarrow \nu_x \subset T \times M$. Também, $\forall X \in T$, $\forall x \in F$, $\forall y \in M$, não pode acontecer que $(X, y) \in \nu_x$ e $X = x + G$ e $y \neq j(x)$. Portanto, se $(X, y) \in \nu_x$ e $X = x + G$, então $y = j(x)$. Isso mostra que $\nu_x \in H$. Como x é qualquer, o resultado segue.

Vamos mostrar que $h = \cup H$ é uma função de T em M.

Primeiramente, note que $dom(h) = T$. De fato, $\forall X \in T$, existe $x \in F$ tal que $X = x+G \Rightarrow \big(X, j(x)\big) = \big(x+G, j(x)\big) \in \nu_x \in H \Rightarrow \Rightarrow \big(X, j(x)\big) \in \cup H = h \Rightarrow X \in dom(h)$. Assim, $T \subset dom(h) \Rightarrow dom(h) = T$. Sejam $(X, y), (X, v) \in h$. Então, existem $\nu, \theta \in H$ tais que $(X, y) \in \nu$ e $(X, v) \in \theta$. Como $(X, y) \in \nu \subset T \times M$, existe $x \in F$ tal que $X = x + G \Rightarrow y = j(x)$. Como $(X, v) \in \theta \subset T \times M$, existe $u \in F$ tal que $X = u + G \Rightarrow v = j(u)$. Assim, $x + G = X = u + G \Rightarrow x - u \in G \subset ker(j) \Rightarrow j(x) - j(u) = j(x - u) = 0_M \Rightarrow v = j(u) = j(x) = y$. Como X, y e v são quaisquer, h é uma função $h : T \to M$ tal que $h(x + G) = j(x)$, $\forall x \in F$.

Note que $(h \circ p)(x) = h\big(p(x)\big) = h(x + G) = j(x)$, $\forall x \in F$. Assim, $h \circ p = j$, isto é, o diagrama abaixo comuta.



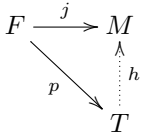
Na realidade, h é a única função que faz comutar o diagrama. De fato, seja $w : T \to M$ tal que $w \circ p = j$. Então, $w(x + G) = w\big(p(x)\big) = (w \circ p)(x) = j(x) = h(x + G)$, $\forall x \in F$. Daí, $w = h$.

Temos que h é um homomorfismo. Com efeito, $\forall x, y \in F$, $\forall r \in R$, temos que $h\big((x + G) + (y + G)\big) = h\big((x + y) + G\big) = j(x + y) = j(x) + j(y) = h(x + G) + h(y + G)$ e $h\big(r(x + G)\big) = h\big((rx) + G\big) = j(rx) = rj(x) = rh(x + G)$. Além disso, $im(h) = im(j)$. De fato, $\forall y \in M$, temos que $y \in im(h) \Leftrightarrow$ existe $x \in F$ tal que $y = h(x + G) = j(x) \Leftrightarrow y \in im(j)$.

Por fim, temos que $ker(h) = p[ker(j)]$. Com efeito, seja $X \in ker(h)$. Então, $X \in T \Rightarrow$ existe $x \in F$ tal que $X = x+G = p(x)$ e $j(x) = h(x+G) = h(X) = 0_M \Rightarrow x \in ker(j)$. Daí, $X \in p[ker(j)]$. Como X é qualquer, $ker(h) \subset p[ker(j)]$. Seja $Y \in p[ker(j)]$. Então, existe $y \in ker(j)$ tal que $Y = p(y) = y + G$. Assim, $h(Y) = h(y + G) = j(y) = 0_M \Rightarrow Y \in ker(h)$. Como Y é arbitrário, $p[ker(j)] \subset ker(h)$. Logo, $ker(h) = p[ker(j)]$.

Dessa forma, acabamos de demonstrar a seguinte

Proposição 10. *Sejam F e M R-módulos, $G \leqslant F$, $T = \frac{F}{G}$, $p : F \to T$ a projeção canônica e $j : F \to M$ um homomorfismo tal que $G \subset ker(j)$. Então, existe uma única função $h : T \to M$ tal que $h \circ p = j$, ou seja, tal que o diagrama abaixo comuta.*



Temos também que h é um homomorfismo, com $h(x + G) = j(x)$, $\forall x \in F$. Ademais, $im(h) = im(j)$ e $ker(h) = p[ker(j)]$.

Teorema 11. *Para quaisquer R-módulos A e B, existem algum R-módulo T e alguma função bilinear $\tau : A \times B \to T$ tais que T é um produto tensorial de A e B com τ.*

Prova: Pelo teorema 5, existem um R-módulo F e uma função $g : A \times B \to F$ tais que F é livre sobre o conjunto $A \times B$ com g. Sejam

$$I = \big\{g(r_1a_1 + r_2a_2, b) - [r_1g(a_1, b) + r_2g(a_2, b)] \in F : a_1, a_2 \in A, \\ b \in B \text{ e } r_1, r_2 \in R\big\},$$

$$J = \big\{g(a, s_1b_1 + s_2b_2) - [s_1g(a, b_1) + s_2g(a, b_2)] \in F : a \in A, \\ b_1, b_2 \in B \text{ e } s_1, s_2 \in R\big\},$$

$G = \langle I \cup J\rangle \leqslant F$, $T = \frac{F}{G}$, $p : F \to T$ a projeção canônica e $\tau = p \circ g$.

Sabemos que $I \subset (I \cup J) \subset \langle I \cup J\rangle = G$ e $J \subset (I \cup J) \subset \langle I \cup J\rangle = G$. Portanto, para todos $a, a_1, a_2 \in A$, todos $b, b_1, b_2 \in B$ e todos $r_1, r_2, s_1, s_2 \in R$, temos que $g(r_1a_1+r_2a_2, b)-[r_1g(a_1, b)+r_2g(a_2, b)] \in I \subset G$ e $g(a, s_1b_1+s_2b_2) - [s_1g(a, b_1) + s_2g(a, b_2)] \in J \subset G$. Dessa forma, $G = \big\{g(r_1a_1 + r_2a_2, b)-[r_1g(a_1, b)+r_2g(a_2, b)]\big\}+G \Rightarrow [g(r_1a_1+r_2a_2, b)]+G = [r_1g(a_1, b)+r_2g(a_2, b)]+G = \big\{[r_1g(a_1, b)]+G\big\}+\big\{[r_2g(a_2, b)]+G\big\} = r_1\big\{[g(a_1, b)]+G\big\}+$

$r_2\big\{[g(a_2,b)]+G\big\}$ e $G=\big\{g(a,s_1b_1+s_2b_2)-[s_1g(a,b_1)+s_2g(a,b_2)]\big\}+G \Rightarrow$ $[g(a,s_1b_1+s_2b_2)]+G=[s_1g(a,b_1)+s_2g(a,b_2)]+G=\big\{[s_1g(a,b_1)]+G\big\}+$ $\big\{[s_2g(a,b_2)]+G\big\}=s_1\big\{[g(a,b_1)]+G\big\}+s_2\big\{[g(a,b_2)]+G\big\}$.

Pelas considerações do parágrafo acima, $\forall a,a_1,a_2\in A$, $\forall b,b_1,b_2\in B$, $\forall r_1,r_2,s_1,s_2\in R$, temos que

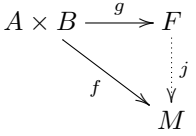
$$\begin{aligned}\tau(r_1a_1+r_2a_2,b)=(p\circ g)(r_1a_1+r_2a_2,b)=p\big(g(r_1a_1+r_2a_2,b)\big)=\\ [g(r_1a_1+r_2a_2,b)]+G=r_1\big\{[g(a_1,b)]+G\big\}+r_2\big\{[g(a_2,b)]+G\big\}=\\ r_1p\big(g(a_1,b)\big)+r_2p\big(g(a_2,b)\big)=r_1(p\circ g)(a_1,b)+r_2(p\circ g)(a_2,b)=\\ r_1\tau(a_1,b)+r_2\tau(a_2,b)\end{aligned}$$

e que

$$\begin{aligned}\tau(a,s_1b_1+s_2b_2)=(p\circ g)(a,s_1b_1+s_2b_2)=p\big(g(a,s_1b_1+s_2b_2)\big)=\\ [g(a,s_1b_1+s_2b_2)]+G=s_1\big\{[g(a,b_1)]+G\big\}+s_2\big\{[g(a,b_2)]+G\big\}=\\ s_1p\big(g(a,b_1)\big)+s_2p\big(g(a,b_2)\big)=s_1(p\circ g)(a,b_1)+s_2(p\circ g)(a,b_2)=\\ s_1\tau(a,b_1)+s_2\tau(a,b_2).\end{aligned}$$

Dessa forma, τ é bilinear.

Vamos mostrar que T é um produto tensorial de A e B com τ. Para tanto, sejam M um R-módulo e $f:A\times B\to M$ uma função bilinear. Como F é um R-módulo livre sobre $A\times B$ com g, existe um único homomorfismo $j:F\to M$ tal que $j\circ g=f$, ou seja, tal que o diagrama abaixo comuta.

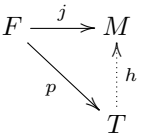


Seja $x\in I$. Então, existem $r_1,r_2\in R$, $a_1,a_2\in A$ e $b\in B$ tais que $x=$ $g(r_1a_1+r_2a_2,b)-r_1g(a_1,b)-r_2g(a_2,b)$. Assim, $j(x)=j\big(g(r_1a_1+r_2a_2,b)-$ $r_1g(a_1,b)-r_2g(a_2,b)\big)=j\big(g(r_1a_1+r_2a_2,b)\big)-j\big(r_1g(a_1,b)\big)-j\big(r_2g(a_2,b)\big)=$ $(j\circ g)(r_1a_1+r_2a_2,b)-r_1j\big(g(a_1,b)\big)-r_2j\big(g(a_2,b)\big)=f(r_1a_1+r_2a_2,b)-$

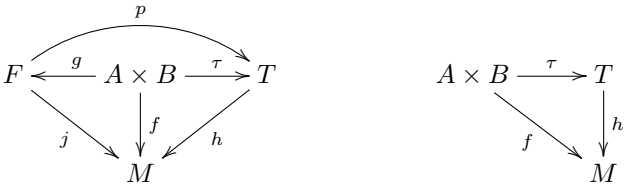
$r_1(j \circ g)(a_1, b) - r_2(j \circ g)(a_2, b) = r_1 f(a_1, b) + r_2 f(a_2, b) - r_1 f(a_1, b) - r_2 f(a_2, b) = 0_M \Rightarrow x \in ker(j)$. Como x é qualquer, $I \subset ker(j)$.

Seja $x \in J$. Então, existem $s_1, s_2 \in R$, $a \in A$ e $b_1, b_2 \in B$ tais que $x = g(a, s_1 b_1 + s_2 b_2) - s_1 g(a, b_1) - s_2 g(a, b_2)$. Assim, $j(x) = j\big(g(a, s_1 b_1 + s_2 b_2) - s_1 g(a, b_1) - s_2 g(a, b_2)\big) = j\big(g(a, s_1 b_1 + s_2 b_2)\big) - j\big(s_1 g(a, b_1)\big) - j\big(s_2 g(a, b_2)\big) = (j \circ g)(a, s_1 b_1 + s_2 b_2) - s_1 j\big(g(a, b_1)\big) - s_2 j\big(g(a, b_2)\big) = f(a, s_1 b_1 + s_2 b_2) - s_1(j \circ g)(a, b_1) - s_2(j \circ g)(a, b_2) = s_1 f(a, b_1) + s_2 f(a, b_2) - s_1 f(a, b_1) - s_2 f(a, b_2) = 0_M \Rightarrow x \in ker(j)$. Como x é qualquer, $J \subset ker(j)$.

Portanto, temos que $I \cup J \subset ker(j) \Rightarrow G = \langle I \cup J \rangle \leqslant ker(j)$. Pela proposição 10, temos que existe uma única função $h : T \to M$ tal que $h \circ p = j$, ou seja, tal que o diagrama abaixo comuta.

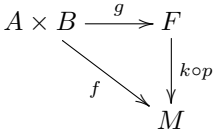


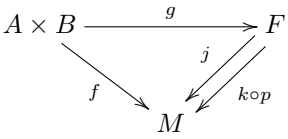
Temos também que h é um homomorfismo, com $h(x + G) = j(x)$, $\forall x \in F$. Ficamos com os diagramas



Em especial, o diagrama à direita comuta. De fato, $h \circ \tau = h \circ (p \circ g) = (h \circ p) \circ g = j \circ g = f$.

Falta então mostrar que h é a única com essa propriedade. Para isso, seja $k : T \to M$ um homomorfismo tal que $k \circ \tau = f$. Então, $f = k \circ \tau = k \circ (p \circ g) = (k \circ p) \circ g$. Assim, temos um homomorfismo $k \circ p : F \to M$ tal que o diagrama abaixo à esquerda comuta.





Mas F é livre sobre $A \times B$ com g e, como já foi dito j é o único homomorfismo tal que $j \circ g = f$. Portanto, $k \circ p = j$.

Por outro lado, como mostrado na proposição 10, h é a única função tal que $h \circ p = j$. Logo, $k = h$, como queríamos demonstrar. ∎

Dados R-módulos A e B, podem existir infinitos produtos tensoriais de A e B. Porém, como mostrado no teorema da unicidade, todos eles são isomorfos entre si e, no que diz respeito à sua estrutura, são o mesmo. Portanto, cometendo um abuso de linguagem, vamos denotar um produto tensorial qualquer de A e B por '$A \otimes B$'. Vamos nos dar a liberdade de dizer que, se A e B são R-módulos, então existem um R-módulo $A \otimes B$, chamado de "o produto tensorial de A e B", e uma função bilinear $\tau : A \times B \to A \otimes B$, chamada de "a *aplicação tensorial* de $A \otimes B$", tais que, para todo R-módulo M e toda função bilinear $\sigma : A \times B \to M$ existe um único R-homomorfismo $h : T \to M$ tal que $h \circ \tau = \sigma$.

Para todo $a \in A$ e todo $b \in B$, vamos denotar a imagem $\tau(a, b)$ por '$a \otimes b$'. Como $\langle im(\tau) \rangle = A \otimes B$, um gerador típico de $A \otimes B$ é da forma $\tau(a, b) = a \otimes b$, para algum $a \in A$ e algum $b \in B$. Como τ é bilinear, temos que $(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b$, $a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2$ e $(ra) \otimes b = r(a \otimes b) = a \otimes (rb)$, $\forall a, a_1, a_2 \in A$, $\forall b, b_1, b_2 \in B$, $\forall r \in R$. Portanto, para todo $t \in A \otimes B$, existem $(a'_1, b_1), \ldots, (a'_n, b_n) \in A \times B$ e $r_1, \ldots, r_n \in R$ tais que $t = \sum_{j=1}^{n} r_j \tau(a'_j, b_j) = \sum_{j=1}^{n} r_j (a'_j \otimes b_j) = \sum_{j=1}^{n} (r_j a'_j) \otimes b_j$. Como $a_j = r_j a'_j \in A$, $\forall j \in \{1, \ldots, n\}$, então temos que, para todo $t \in A \otimes B$, existem $a_1, \ldots, a_n \in A$ e $b_1, \ldots, b_n \in B$ tais que $t = \sum_{j=1}^{n} a_j \otimes b_j$.

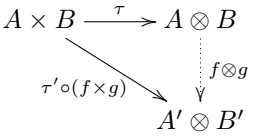
Temos também que $a \otimes 0_B = a \otimes (0 \cdot 0_B) = 0 \cdot (a \otimes 0_B) = 0_{A \otimes B} = 0 \cdot (0_A \otimes b) = (0 \cdot 0_A) \otimes b = 0_A \otimes b$, $\forall a \in A$, $\forall b \in B$.

Uma outra construção é importante. Sejam A, A', B e B' R-módulos, com aplicações tensoriais $\tau : A \times B \to A \otimes B$ e $\tau' : A' \times B' \to A' \otimes B'$. Sejam também homomorfismos $f : A \to A'$ e $g : B \to B'$. Definimos o produto cartesiano de f com g como sendo a função $f \times g : A \times B \to A' \times B'$ tal que $(f \times g)(a, b) = \big(f(a), g(b)\big)$, $\forall a \in A$, $\forall b \in B$. Temos que $f \times g$ é um homomorfismo. De fato, $\forall (a_1, b_1), (a_2, b_2) \in A \times B$, $\forall r \in R$, temos que $(f \times g)\big((a_1, b_1) + (a_2, b_2)\big) = (f \times g)(a_1 + a_2, b_1 + b_2) = \big(f(a_1 + a_2), g(b_1 + b_2)\big) = \big(f(a_1) + f(a_2), g(b_1) + g(b_2)\big) = \big(f(a_1), g(b_1)\big) + \big(f(a_2), g(b_2)\big) = (f \times g)(a_1, b_1) + (f \times g)(a_2, b_2)$ e que $(f \times g)\big(r(a_1, b_1)\big) = (f \times g)(ra_1, rb_1) = \big(f(ra_1), g(rb_1)\big) = \big(rf(a_1), rg(b_1)\big) = r\big(f(a_1), g(b_1)\big) = r(f \times g)(a_1, b_1)$. Temos também que $\tau' \circ (f \times g)$ é bilinear. Com efeito, $[\tau' \circ (f \times g)](r_1 a_1 + r_2 a_2, b) = \tau'\big((f \times g)(r_1 a_1 + r_2 a_2, b)\big) = \tau'\big(f(r_1 a_1 + r_2 a_2), g(b)\big) = \tau'\big(r_1 f(a_1) + r_2 f(a_2), g(b)\big) = r_1 \tau'\big(f(a_1), g(b)\big) + r_2 \tau'\big(f(a_2), g(b)\big) = r_1 \tau'\big((f \times g)(a_1, b)\big) + r_2 \tau'\big((f \times g)(a_2, b)\big) = r_1 [\tau' \circ (f \times g)](a_1, b) + r_2 [\tau' \circ (f \times g)](a_2, b)$ e $[\tau' \circ (f \times g)](a, s_1 b_1 + s_2 b_2) = \tau'\big((f \times g)(a, s_1 b_1 + s_2 b_2)\big) = \tau'\big(f(a), g(s_1 b_1 + s_2 b_2)\big) = \tau'\big(f(a), s_1 g(b_1) + s_2 g(b_2)\big) = s_1 \tau'\big(f(a), g(b_1)\big) + s_2 \tau'\big(f(a), g(b_2)\big) = s_1 \tau'\big((f \times g)(a, b_1)\big) + s_2 \tau'\big((f \times g)(a, b_2)\big) = s_1 [\tau' \circ (f \times g)](a, b_1) + s_2 [\tau' \circ (f \times g)](a, b_2)$, $\forall a, a_1, a_2 \in A$, $\forall b, b_1, b_2 \in B$, $\forall r_1, r_2, s_1, s_2 \in R$.

Dessa forma, por definição, existe um único homomorfismo

$$f \otimes g : A \otimes B \to A' \otimes B'$$

tal que $(f \otimes g) \circ \tau = \tau' \circ (f \times g)$, isto é, tal que o diagrama abaixo comuta.



Assim, o homomorfismo $f \otimes g$ é o único pelo qual o quadrado abaixo é

comutativo.

$$\begin{array}{ccc} A\times B & \xrightarrow{\ \tau\ } & A\otimes B \\ {\scriptstyle f\times g}\downarrow & & \downarrow{\scriptstyle f\otimes g} \\ A'\times B' & \xrightarrow{\ \tau'\ } & A'\otimes B' \end{array}$$

O homomorfismo $f\otimes g$ é chamado de "o produto tensorial de f com g". Temos que $(f\otimes g)(a\otimes b) = (f\otimes g)\big(\tau(a,b)\big) = [(f\otimes g)\circ\tau](a,b) = [\tau'\circ(f\times g)](a,b) = \tau'\big((f\times g)(a,b)\big) = \tau'\big(f(a),g(b)\big) = f(a)\otimes g(b)$, $\forall a\in A$, $\forall b\in B$.

Uma propriedade importante desses homomorfismos é a seguinte

Proposição 12. *Sejam A, A', A'', B, B' e B'' R-módulos, com aplicações tensoriais $\tau : A\times B\to A\otimes B$, $\tau' : A'\times B'\to A'\otimes B'$ e $\tau'' : A''\times B''\to A''\otimes B''$. Sejam também os homomorfismos $f : A\to A'$, $f' : A'\to A''$, $g : B\to B'$ e $g' : B'\to B''$. Então, temos que*

$$(f'\circ f)\otimes(g'\circ g) = (f'\otimes g')\circ(f\otimes g).$$

Prova: Temos o seguinte diagrama

$$\begin{array}{ccc} A\times B & \xrightarrow{\ \tau\ } & A\otimes B \\ {\scriptstyle f\times g}\downarrow & & \downarrow{\scriptstyle f\otimes g} \\ A'\times B' & \xrightarrow{\ \tau'\ } & A'\otimes B' \\ {\scriptstyle f'\times g'}\downarrow & & \downarrow{\scriptstyle f'\otimes g'} \\ A''\times B'' & \xrightarrow{\ \tau''\ } & A''\otimes B'' \end{array}$$

em que $(f\otimes g)\circ\tau = \tau'\circ(f\times g)$ e $(f'\otimes g')\circ\tau' = \tau''\circ(f'\times g')$.

Também temos o seguinte diagrama

$$\begin{array}{ccc} A\times B & \xrightarrow{\ \tau\ } & A\otimes B \\ {\scriptstyle (f'\times g')\circ(f\times g)}\downarrow & & \downarrow{\scriptstyle (f'\otimes g')\circ(f\otimes g)} \\ A''\times B'' & \xrightarrow{\ \tau''\ } & A''\otimes B'' \end{array}$$

em que $(f'\times g')\circ(f\times g)$ e $(f'\otimes g')\circ(f\otimes g)$ são homomorfismos.

Por outro lado, temos também o homomorfismo

$$(f' \circ f) \times (g' \circ g) : A \times B \to A'' \times B''.$$

Portanto, existe um único homomorfismo $(f' \circ f) \otimes (g' \circ g) : A \otimes B \to A'' \otimes B''$ tal que $[(f' \circ f) \otimes (g' \circ g)] \circ \tau = \tau'' \circ [(f' \circ f) \times (g' \circ g)]$, isto é, tal que o diagrama abaixo comuta.

$$\begin{array}{ccc} A \times B & \xrightarrow{\tau} & A \otimes B \\ {\scriptstyle (f' \circ f) \times (g' \circ g)} \downarrow & & \downarrow {\scriptstyle (f' \circ f) \otimes (g' \circ g)} \\ A'' \times B'' & \xrightarrow{\tau''} & A'' \otimes B'' \end{array}$$

Vamos mostrar que $(f' \circ f) \times (g' \circ g) = (f' \times g') \circ (f \times g)$. De fato, $\forall (a, b) \in A \times B$, temos $[(f' \circ f) \times (g' \circ g)](a, b) = \big((f' \circ f)(a), (g' \circ g)(b)\big) = \big(f'(f(a)), g'(g(b))\big) = (f' \times g')\big(f(a), g(b)\big) = (f' \times g')\big((f \times g)(a, b)\big) = [(f' \times g') \circ (f \times g)](a, b)$. Como $f' \circ f$ e $g' \circ g$ são homomorfismos e τ'' é bilinear, temos que $\tau'' \circ [(f' \circ f) \times (g' \circ g)]$ é bilinear.

Seja $\sigma = \tau'' \circ [(f' \circ f) \times (g' \circ g)] = \tau'' \circ (f' \times g') \circ (f \times g)$. Temos que σ é bilinear e que $[(f' \otimes g') \circ (f \otimes g)] \circ \tau = (f' \otimes g') \circ [(f \otimes g) \circ \tau] = (f' \otimes g') \circ [\tau' \circ (f \times g)] = [(f' \otimes g') \circ \tau'] \circ (f \times g) = [\tau'' \circ (f' \times g')] \circ (f \times g) = \tau'' \circ (f' \times g') \circ (f \times g) = \sigma$.

Portanto, o homomorfismo $j = (f' \otimes g') \circ (f \otimes g)$ é tal que $j \circ \tau = \sigma$, isto é, tal que o diagrama abaixo comuta.

$$\begin{array}{ccc} A \times B & \xrightarrow{\tau} & A \otimes B \\ & {\scriptstyle \sigma} \searrow & \downarrow {\scriptstyle j} \\ & & A'' \otimes B'' \end{array}$$

Como $(f' \circ f) \otimes (g' \circ g)$ é o único com esta propriedade, temos que $(f' \circ f) \otimes (g' \circ g) = j = (f' \otimes g') \circ (f \otimes g)$, como queríamos demonstrar. ■

Exemplo 13. Sejam A e B R-módulos, $id_A : A \to A$ e $id_B : B \to B$ os homomorfismos identidade de A e B, respectivamente, $\tau : A \times B \to A \otimes B$ uma função tensorial de $A \otimes B$ e $id_{A \otimes B} : A \otimes B \to A \otimes B$ o homomorfismo

identidade de $A \otimes B$. É claro que $id_A \times id_B : A \times B \to A \times B$ é um homomorfismo e, assim, $\tau \circ (id_A \times id_B) : A \times B \to A \otimes B$ é bilinear. Temos o seguinte diagrama.

$$\begin{array}{ccc} A \times B & \xrightarrow{\tau} & A \otimes B \\ {\scriptstyle id_A \times id_B}\downarrow & & \downarrow\downarrow {\scriptstyle id_A \otimes id_B} \\ A \times B & \xrightarrow{\tau} & A \otimes B \end{array} \qquad {\scriptstyle id_{A\otimes B}}$$

Em que $id_A \otimes id_B$ é o único homomorfismo tal que o quadrado acima é comutativo, isto é, $(id_A \otimes id_B) \circ \tau = \tau \circ (id_A \times id_B)$. Como $id_{A\otimes B}$ é a função identidade, temos que $id_{A\otimes B} \circ \tau = \tau$. Também, para todo $(a, b) \in A \times B$, segue que $[\tau \circ (id_A \times id_B)](a, b) = \tau\big(id_A(a), id_B(b)\big) = \tau(a, b)$. Assim, $\tau \circ (id_A \times id_B) = \tau = id_{A\otimes B} \circ \tau$. Ou seja, $id_{A\otimes B}$ também faz com que o quadrado acima seja comutativo. Logo, $id_{A\otimes B} = id_A \otimes id_B$.

Abstract: Homological Algebra was originally created to deal with concepts from Algebraic Topology. In this work we will introduce some basic ideas of the theory of free modules and tensor products wich are directly applied on the study of Homological Algebra.

Keywords: modules over rings; free modules; tensor products; homological algebra

Referências Bibliográficas

[1] Hu, S.T., *Introduction to Homological Algebra*, Holden-Day, 1968.

[2] Vick, J.W., *Homology Theory* An Introduction to Algebraic Topology, Graduate Texts in Mathematics No. 145, Springer-Verlag, 1994.

[3] Weibel, C.A., *An Introduction to Homological Algebra*, Cambridge University Press, 1997.

[4] Hilton, P.J.; Stammbach U., *A Course in Homological Algebra*, Graduate Texts in Mathematics No. 4, Springer-Verlag, 1997.

[5] Gelfand, S.I.; Manin, Y.I., *Methods of Homological Algebra*, Springer-Verlag, 2003.

Presentação de Grupos e o Teorema de Tietze

Pablo Gonzalez Pagotto[1]

Orientador(a): Profa. Dra. Alice Kimie Miwa Libardi

Resumo: Em nosso trabalho apresentaremos os conceitos de *grupos livres* e *presentação de grupos*. Estes conceitos são indispensáveis para o estudo do grupo fundamental de nós. Também apresentaremos e demonstraremos o *Teorema de Tiezte* de acordo com o qual toda equivalência de presentação entre presentações finitas pode ser decomposta em um número finito de equivalências de Tietze.

Palavras-chave: grupos livres; presentação de grupos; teorema de Tietze

1 Grupos Livres

Seja $\mathscr{A}$ um conjunto de cardinalidade $\#\mathscr{A}$ dado.

Chamaremos $\mathscr{A}$ de *alfabeto* e seus elementos de *letras*. Diremos que uma *sílaba* é um símbolo a^n, onde $a \in \mathscr{A}$ e $n \in \mathbb{Z}$. Diremos que uma *palavra* é uma sequência ordenada e finita de sílabas. e.g. $b^{12}b^{-3}a^{32}b^0c^{-2}h^{23}$. Em uma palavra as sílabas são escritas uma após a outra na forma de um produto formal.

Existe uma única palavra que não possui sílabas, denominada *palavra vazia* e é denotada pelo símbolo 1. Por brevidade denotaremos a^1 por a.

No conjunto $W[\mathscr{A}]$, de todas as palavras formadas pelo alfabeto $\mathscr{A}$, está definida a multiplicação natural, que consiste em concatenar duas palavras.

Definição 1 (Contração/Expansão elementar do tipo I). Se uma palavra u tem a forma $w_1a^0w_2$, onde w_1, w_2 são palavras, dizemos que a palavra $v = w_1w_2$ foi obtida de u por uma contração elementar do tipo I ou que u foi obtida de v por uma expansão elementar do tipo I.

[1]Bolsista FAPESP – Processo 2011/00377–1

Definição 2 (Contração/Expansão elementar do tipo II). Se uma palavra u tem a forma $w_1 a^p a^q w_2$, onde w_1, w_2 são palavras, dizemos que a palavra $v = w_1 a^{p+q} w_2$ foi obtida de u por uma contração elementar do tipo II ou que u foi obtido de v por uma expansão elementar do tipo II.

Definamos a relação $\sim$ em $W[\mathscr{A}]$ por, $w_1 \sim w_2$ se, e somente se, w_1 pode ser obtido de w_2 por uma sequência finita de contrações/expansões dos tipos I ou II. Esta relação é de equivalência e denotaremos por $[u]$ a classe de equivalência da palavra u e por $F[\mathscr{A}]$ o conjunto das classes de equivalência de palavras.

Proposição 3. *O conjunto $F[\mathscr{A}]$ com o operador $. : F[\mathscr{A}] \times F[\mathscr{A}] \to F[\mathscr{A}]$ dado por $([u],[v]) \to [uv]$ é um grupo.*

O grupo $F[\mathscr{A}]$ é chamado *o grupo livre no alfabeto $\mathscr{A}$*.

Proposição 4. *Sejam G um grupo e $E \subset G$ um subconjunto. Então $[E] = \bigcap_{\substack{E \subset K \\ K < G}} K$ é um subgrupo de G e será denominado o subgrupo gerado por E em G.*

Proposição 5. *Sejam G um grupo e $E \subset G$ um subconjunto não vazio. Então $[E] = \{{g_1}^{n_1} {g_2}^{n_2} \cdots {g_k}^{n_k}; g_i \in E, n_i \in \mathbb{Z}\}$.*

Definição 6. Seja $E \subset G$ um subconjunto de G. Dizemos que E é um conjunto de elementos geradores de G se $[E] = G$.

Definição 7. Seja G um grupo. Se E é um subconjunto de G tal que $[E] = G$ e toda função $\phi : E \to H$ pode ser estendida para um homomorfismo $\phi' : G \to H$, onde H é um grupo qualquer, dizemos que E é uma base livre de G e o grupo G é dito livre.

Exemplo 8. Se $G = \{1\}$ então $\emptyset$ é uma base livre de G pois, $[\emptyset] = \{1\} = G$ e dado um grupo H e uma função $\phi : \emptyset \to H$ temos que o homomorfismo

$\phi' : G \to H$, $\phi'(1) = 1_H$ é uma extensão de ϕ. Portanto G é um grupo livre.

Proposição 9. *Sejam G e H grupos. Se $E \subset G$, $[E] = G$ e $\phi : E \to H$ pode ser estendida para um homomorfismo $\phi' : G \to H$ então tal extensão é única.*

Prova: Suponhamos que o homomorfismo $\psi : G \to H$ seja uma extensão de ϕ e mostremos que $\psi = \phi'$. De fato, como ψ e ϕ' são extensões de ϕ segue que $\forall x \in E, \psi(x) = \phi(x) = \phi'(x)$.

Seja agora $x \in G$ qualquer. Como $[E] = G$ segue que

$$x = \prod_{i=1}^{n} {a_i}^{n_i} \quad , a_i \in E,\ n_i \in \mathbb{Z}$$

Logo,

$$\psi(x) = \psi\left(\prod_{i=1}^{n} {a_i}^{n_i}\right) = \prod_{i=1}^{n} \psi(a_i)^{n_i} = \prod_{i=1}^{n} \phi(a_i)^{n_i}$$
$$= \prod_{i=1}^{n} \phi'(a_i)^{n_i} = \phi'\left(\prod_{i=1}^{n} {a_i}^{n_i}\right) = \phi'(x)$$

Portanto $\psi = \phi'$. ∎

Teorema 10. *Um grupo é livre se, e somente se, é isomorfo a $F[\mathscr{A}]$ para algum $\mathscr{A}$.*

Prova: Primeiramente mostremos que $F[\mathscr{A}]$ é um grupo livre, mais especificamente, que $[\mathscr{A}] = \{[u] \in F[\mathscr{A}]; u \in \mathscr{A}\}$ é uma base livre de $F[\mathscr{A}]$. Para este fim considere um grupo H e uma função $\phi : [\mathscr{A}] \to H$. Defina $\phi' : F[\mathscr{A}] \to H$ pondo

$$\phi'\left(\prod_{i=1}^{n} [a_i]^{n_i}\right) = \prod_{i=1}^{n} \phi([a_i])^{k_i} \quad , [a_i] \in [\mathscr{A}],\ n_i \in \mathbb{Z}.$$

Pela definição de $F[a]$ e como $[\mathscr{A}]$ é um conjuntos de geradores de $F[\mathscr{A}]$, segue que todo elemento de $F[\mathscr{A}]$ pode ser escrito como um produto de

finitos elementos de $[\mathscr{A}]$ e assim ϕ' define um homomorfismo. Portanto $[\mathscr{A}]$ é uma base livre de $F[\mathscr{A}]$.

Agora se G é um grupo isomorfo a $F[\mathscr{A}]$ por isomorfismo λ, então $E = \lambda^{-1}([\mathscr{A}])$ é uma base livre de G e portanto G é um grupo livre. De fato, sejam H um grupo qualquer e uma função $\psi : E \to H$. Como λ é um isomorfismo então $\lambda^{-1}\psi : [\mathscr{A}] \to H$ se estende ao homomorfismo $(\lambda^{-1}\psi)' : F[\mathscr{A}] \to H$ e portanto $\psi' = (\lambda^{-1}\psi)'\lambda : G \to H$ é um homomorfismo e também uma extensão de ψ.

Reciprocamente, sejam G um grupo livre e E uma base livre deste. Seja $F[\mathscr{A}]$ o grupo livre no alfabeto $\mathscr{A}$ cuja cardinalidade é a mesma que a de E, isto é, $\#\mathscr{A} = \#E$. Seja $\kappa : E \to [\mathscr{A}]$ uma função bijetora (esta existe pois, como $\#\mathscr{A} = \#E$ segue por definição que existe uma bijeção $j : \mathscr{A} \to E$. Mas, a projeção $\rho : \mathscr{A} \to [\mathscr{A}], a \to [a]$ é uma bijeção e portanto $i = j\rho^{-1} : [\mathscr{A}] \to E$ é uma bijeção). Como E é base uma livre, a correspondência κ se estende ao homomorfismo $\phi : G \to F[\mathscr{A}]$. Da mesma forma, $[\mathscr{A}]$ é uma base livre de $F[\mathscr{A}]$ e portanto a correspondência κ^{-1} se estende ao homomorfismo $\mu : F[\mathscr{A}] \to G$.

Os homomorfismo compostos $\phi\mu : F[\mathscr{A}] \to F[\mathscr{A}]$ e $\mu\phi : G \to G$ são extensões das funções $\kappa\kappa^1$ e $\kappa^{-1}\kappa$ respectivamente. Como essas últimas são funções identidade, se estendem aos automorfismos identidade de $F[\mathscr{A}]$ e G respectivamente. Segue da unicidade de tais extensões que $\phi\mu$ e $\mu\phi$ são automorfismos identidade. Portanto ϕ e μ são isomorfismos. ∎

Corolário 11. *O grupo livre $F[\mathscr{A}]$ é de fato livre.*

Corolário 12 (da demonstração do Teorema 10). *Sejam G e D grupos livres. Se as bases livres E e Q de G e D, respectivamente, tem a mesma cardinalidade então G e D são isomorfos.*

2 Presentação

Definição 13. Um elemento f de um grupo arbitrário Q é dito uma consequência dos elementos $\{f_i\}_{i\in\Lambda}$ em Q se todo homomorfismo $\psi : Q \to H$, H grupo, tal que $\{f_i\}_{i\in\Lambda} \subset \ker\psi$ satisfaz $\psi(f) = 1$.

Proposição 14. *Um elemento f de um grupo é uma consequência dos elementos $\{f_i\}_{i\in\Lambda} \in Q$ se, e somente se, todo subgrupo normal de Q que contem os elementos $\{f_i\}_{i\in\Lambda}$ também contem f.*

Definição 15. O conjunto $\mathcal{C}_{\{f_i\}_{i\in\Lambda}} = \{f \in Q; f \text{ é consequência de } \{f_i\}_{i\in\Lambda}\}$ é denominado a consequência de $\{f_i\}_{i\in\Lambda}$.

Teorema 16. *Seja Q um grupo. A consequência dos elementos $\{f_i\}_{i\in\Lambda}$ é a interseccção de todos os subgrupos normais de Q que contêm os elementos $\{f_i\}_{i\in\Lambda}$, simbolicamente,*

$$\mathcal{C}_{\{f_i\}_{i\in\Lambda}} = \bigcap_{\substack{K\lhd Q \\ K\supset\{f_i\}_{i\in\Lambda}}} K$$

Prova: Por simplicidade, denotemos por V o conjunto:

$$\bigcap_{\substack{K\lhd Q \\ K\supset\{f_i\}_{i\in\Lambda}}} K.$$

Seja $b \in \mathcal{C}_{\{f_i\}_{i\in\Lambda}}$. Então,pela proposição 14, para todo subgrupo normal K de Q, se $\{f_i\}_{i\in\Lambda} \subset K$ então $b \in K$, logo $b \in V$. Reciprocamente, se $b \in V$ então para todo subgrupo normal K de Q tal que $\{f_i\}_{i\in\Lambda} \subset K$ temos que $b \in K$ e portanto $b \in \mathcal{C}_{\{f_i\}_{i\in\Lambda}}$. ∎

Como a intersecção de subgrupos normais é um subgrupo normal, o teorema anterior nos dá que a consequência dos elementos $\{f_i\}_{i\in\Lambda}$ é o "menor" subgrupo normal que contem $\{f_i\}_{i\in\Lambda}$.

Teorema 17. *Sejam $\{g_i\}_{i\in\Lambda} \subset G$, G um grupo, e $\phi : G \to H$ um epimorfismo. Então ϕ leva a consequência de $\{g_i\}_{i\in\Lambda}$ na consequência de $\phi(\{g_i\}_{i\in\Lambda})$.*

Seja F um grupo livre com uma base livre E que é suposta suficientemente grande de tal modo que contenha tantos elementos básicos quanto se deseja, denominado conjunto subjacente de geradores.

Definição 18. Uma presentação de grupo, denotada por $(X : \mathbf{r})$, é um objeto que consiste de um subconjunto $X \subset E$ e um subconjunto $\mathbf{r}$ do subgrupo $F(X)$ gerado em F por X. O conjunto X é denominado conjunto de geradores da presentação e o conjunto $\mathbf{r}$ é chamado de conjunto dos relatores da presentação.

É importante observar que $F(X)$ é um grupo livre e que X é uma base livre já que $F(X)$ é isomorfo a $F[X]$ no alfabeto X. De fato, se $X = \emptyset$ é imediato. Se $X \neq \emptyset$ temos, pela proposição 5,

$$F(X) = \{x_1^{n_1} \dots x_\ell^{n_\ell}; x_i \in X,\, n_i \in \mathbb{Z}\}.$$

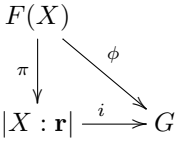
Seja $\mu : F(X) \to F[X]$ o homomorfismo dado por $\mu(x_i) = [x_i]$. Claramente μ determina um isomorfismo de grupos.

Definição 19. O grupo com presentação $(X : \mathbf{r})$ é o grupo quociente $|X : \mathbf{r}| = F(X)/\mathcal{C}_{\mathbf{r}}$, onde $\mathcal{C}_{\mathbf{r}}$ é a consequência de $\mathbf{r}$ em $F(X)$.

Definição 20. A presentação de um grupo G consiste de uma presentação de grupo $(X : \mathbf{r})$ e um isomorfismo $i : |X : \mathbf{r}| \to G$.

Pelo *primeiro teorema de isomorfismo*[2], todo epimorfismo $\phi : F(X) \to G$ cujo kernel é $\mathcal{C}_{\mathbf{r}}$ determina uma presentação de G. Reciprocamente, toda presentação de G determina tal epimorfismo ϕ. Como mostra o diagrama comutativo abaixo:

[2]Sejam G e H grupos, e seja $\varphi : G \to H$ um epimorfismo. Então o grupo quociente $G/\ker\varphi$ é isomorfo a H.



Se necessário escreveremos $(X : \mathbf{r})_\phi$ para indicar que $(X : \mathbf{r})$ é a presentação do grupo G com respeito ao homomorfismo ϕ. Introduziremos o conceito de relação, já que sua manipulação é mais simples do que a de relatores.

Definição 21. Pela fórmula $u = v$ diremos que $u \equiv v \pmod{\mathcal{C}_\mathbf{r}}$, ou seja, $uv^{-1} \in \mathcal{C}_\mathbf{r}$.

Exemplo 22. Se a e b comutam, então é fácil ver que $(ab)^2 = 1$ implica $a^2b^2 = 1$ mas, não é tão simples ver que a^2b^2 é uma consequência dos relatores $aba^{-1}b^{-1}$ e $(ab)^2$. De fato, $a^2b^2 = b^{-1}(aba^{-1}b^{-1})^{-1}bb^{-1}(ab)^2b$.

Observação 23. $(X : \ )$, isto é, quando $\mathbf{r} = \emptyset$, é a presentação do grupo livre $F(X)$.

$(\ : \ )$, isto é, quando $X = \emptyset$ e $\mathbf{r} = \emptyset$, é a presentação especialmente simples do grupo trivial.

Definição 24. Seja $(X : \mathbf{r})$ uma presentação.

(*i*) Se X é finito, dizemos que $(X : \mathbf{r})$ é finitamente gerado.

(*ii*) Se $\mathbf{r}$ é finito, dizemos que $(X : \mathbf{r})$ é finitamente relacionado.

(*iii*) Se X e $\mathbf{r}$ são finitos, dizemos que $(X : \mathbf{r})$ é uma presentação finita.

(*iv*) Um grupo é dito finitamente gerado se possui uma presentação finitamente gerada.

(*v*) Um grupo é dito finitamente relacionado se possui uma presentação finitamente relacionada.

(*vi*) Um grupo é dito finitamente presentado se possui uma presentação finita.

3 Tipos de Presentação

Um grupo pode possuir muitas presentações e isso pode dificultar a determinação de quando duas presentações são isomorfas. Para tal daremos algumas condições necessárias para que duas presentações sejam isomorfas.

Definição 25. Uma aplicação de presentações $f\colon (X\colon \mathbf{r}) \to (Y\colon \mathbf{s})$ consiste de duas presentações $(X\colon \mathbf{r})$ e $(Y\colon \mathbf{s})$ e um homomorfismo $f\colon F(X) \to F(Y)$ tal que $f(\mathbf{r}) \subset \mathcal{C}_{\mathbf{s}}$.

Proposição 26. *Toda aplicação de presentação* $f : (X : \mathbf{r}) \to (Y : \mathbf{s})$ *determina unicamente um homomorfismo* $f_* : |X : \mathbf{r}| \to |Y : \mathbf{s}|$ *satisfazendo* $f_*\pi_X = \pi_Y f$, *onde* $\pi_X : F(X) \to |X : \mathbf{r}|$ *e* $\pi_Y : F(Y) \to |Y : \mathbf{s}|$ *são as projeções canônicas.*

Prova: Considere o diagrama:

$$\begin{array}{ccc} F(X) & \xrightarrow{\ f\ } & F(Y) \\ {\scriptstyle \pi_X}\downarrow & & \downarrow{\scriptstyle \pi_Y} \\ |X : \mathbf{r}| & \xrightarrow[f_*]{} & |Y : \mathbf{s}| \end{array}$$

Para cada elemento $x \in F(X)$ defina $f_*(x+\mathcal{C}_{\mathbf{r}}) = f(x)+\mathcal{C}_{\mathbf{s}}$. Primeiramente mostremos que f_* está bem definida.

Sejam $u, v \in |X : \mathbf{r}|$ tais que $u = v$, então $u = u' + \mathcal{C}_{\mathbf{r}}$ e $v = v' + \mathcal{C}_{\mathbf{r}}$ com $u'v'^{-1} \in \mathcal{C}_{\mathbf{r}}$. Como $\mathbf{r} \in \ker \pi_Y f$ segue que $\pi_Y f(u'v'^{-1}) = 1$ e portanto $f(u'v'^{-1}) = f(u')f(v')^{-1} \in \mathcal{C}_{\mathbf{r}}$, por conseguinte $f(u') + \mathcal{C}_{\mathbf{s}} = f(v') + \mathcal{C}_{\mathbf{s}}$, ou seja, $f_*(u) = f_*(v)$.

Provemos agora a unicidade. Sejam $\alpha + \mathcal{C}_{\mathbf{r}} \in |X\colon \mathbf{r}|$ qualquer e

$$k\colon |X\colon \mathbf{r}| \to |Y\colon \mathbf{s}|$$

um homomorfismo satisfazendo $k\pi_X = \pi_Y f$. Então

$$\begin{aligned} f_*(\alpha + \mathcal{C}_{\mathbf{r}}) &= f_*(\pi_X(\alpha)) = (f_*\pi_X)(\alpha) \\ &= (\pi_Y f)(\alpha) = (k\pi_X)(\alpha) = k(\alpha + \mathcal{C}_{\mathbf{r}}). \end{aligned}$$

Portanto $f_* = k$ e assim f_* é única. ∎

Definição 27. Sejam $f : (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $g : (Y : \mathbf{s}) \to (Z : \mathbf{t})$ aplicações de presentação. Então a aplicação composta $gf : (X : \mathbf{r}) \to (Z : \mathbf{t})$ consiste de $(X : \mathbf{r})$, $(Z : \mathbf{t})$ e do homomorfismo $gf : F(X) \to F(Z)$

Proposição 28. (i) *A composta de aplicações de presentação é uma aplicação de presentação.*

(ii) *A composição de presentações é associativa.*

(iii) *Sejam* $1\colon (X\colon \mathbf{r}) \to (X\colon \mathbf{r})$, $f\colon (X\colon \mathbf{r}) \to (Y\colon \mathbf{s})$ *e* $g\colon (Y\colon \mathbf{s}) \to (Z\colon \mathbf{t})$ *aplicações de presentação. Então* $1_* = 1$ *e* $(gf)_* = g_*f_*$.

Prova: (i) Basta mostrarmos que $gf(\mathbf{r}) \subset \mathcal{C}_{\mathbf{t}}$. De fato, como g é aplicação de presentação temos que $g(\mathbf{s}) \subset \mathcal{C}_{\mathbf{t}}$. Agora, de $g(\mathbf{s}) \subset \mathcal{C}_{\mathbf{t}}$ temos que $\mathbf{s} \subset \ker \pi_Z g$, pela definição de consequência segue que $\mathcal{C}_{\mathbf{s}} \subset \ker \pi_Z g$, ou seja, $g(\mathcal{C}_{\mathbf{s}}) \subset \ker \pi_Z = \mathcal{C}_{\mathbf{t}}$. Portanto temos $f(\mathbf{r}) \subset \mathcal{C}_{\mathbf{s}}$ e $g(\mathcal{C}_{\mathbf{s}}) \subset \mathcal{C}_{\mathbf{t}}$, isto é, $gf(\mathbf{r}) \subset \mathcal{C}_{\mathbf{t}}$.
(ii) Decorre imediatamente da associatividade de homomorfismos.
(iii) É fácil ver que $1 : (X : \mathbf{r}) \to (X : \mathbf{r})$ é uma aplicação de presentação e como $\pi_X 1_* = 1\pi_X$ segue o resultado. Mostremos agora que $(gf)_* = g_*f_*$.

$$\begin{array}{ccccc} F(X) & \xrightarrow{f} & F(Y) & \xrightarrow{g} & F(Z) \\ \downarrow{\scriptstyle \pi_X} & & \downarrow{\scriptstyle \pi_Y} & & \downarrow{\scriptstyle \pi_Z} \\ |X : \mathbf{r}| & \xrightarrow[f_*]{} & |Y : \mathbf{s}| & \xrightarrow[g_*]{} & |Z : \mathbf{t}| \end{array}$$

Para isto basta notarmos que $f_*\pi_X = \pi_Y f$ e $g_*\pi_Y = \pi_Z g$ donde $g_*f_*\pi_X = g_*\pi_Y f = \pi_Z gf$. ∎

Definição 29. Duas aplicações de presentação $f_1, f_2 : (X : \mathbf{r}) \to (Y : \mathbf{s})$, são ditas homotópicas, e denotaremos $f_1 \simeq f_2$, se para todo $z \in X$, o elemento $f_1(z)f_2(z)^{-1}$ pertence a consequência de $\mathbf{s}$ ou equivalentemente, $\pi_Y f_1(z) = \pi_Y f_2(z)$.

Proposição 30. *Duas aplicações de presentação $f_1, f_2 : (X : \mathbf{r}) \to (Y : \mathbf{s})$ são homotópicas se, e só se, $f_{1_*} = f_{2_*}$.*

Prova: Se $f_1 \simeq f_2$ então, pela definição de aplicações homotópicas, temos $\pi_Y f_1 = \pi_Y f_2$ o que implica $f_{1_*}\pi_X = f_{2_*}\pi_X$. Como π_X é sobrejeção segue que $f_{1_*} = f_{2_*}$. Agora, se $f_{1_*} = f_{2_*}$ então $\pi_Y f_1 = f_{1_*}\pi_X = f_{2_*}\pi_X = \pi_Y f_2$ e portanto $f_1 \simeq f_2$. ■

Corolário 31. *Sejam $f_1, f_2 : (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $g_1, g_2 : (Y : \mathbf{s}) \to (Z : \mathbf{t})$ aplicações de presentação tais que $f_1 \simeq f_2$ e $g_1 \simeq g_2$. então $g_1 f_1 \simeq g_2 f_2$.*

Prova: De $f_1 \simeq f_2$ e $g_1 \simeq g_2$ segue que $f_{1_*} = f_{2_*}$ e $g_{1_*} = g_{2_*}$ donde, $g_{1_*} f_{1_*} = g_{2_*} f_{2_*}$ e portanto $g_1 f_1 \simeq g_2 f_2$. ■

Teorema 32. *Para cada homomorfismo $\Theta : |X : \mathbf{r}| \to |Y : \mathbf{s}|$, existe uma aplicação de presentação $f : (X : \mathbf{r}) \to (Y : \mathbf{s})$ tal que $f_* = \Theta$. Além disso, esta é única a menos de homotopia.*

Prova: Considere o diagrama:

$$\begin{array}{ccc} F(X) & \overset{f}{\rightarrowtail} & F(Y) \\ {\scriptstyle \pi_X}\downarrow & & \downarrow{\scriptstyle \pi_Y} \\ |X : \mathbf{r}| & \underset{\Theta}{\longrightarrow} & |Y : \mathbf{s}| \end{array}$$

Como π_X e π_Y são sobrejeções podemos assinalar a cada $x \in X$ um elemento $f(x) \in F(Y)$ de tal forma que $\pi_Y f(x) = \Theta\pi_X(x)$. Como $F(X)$ é um grupo livre com base X, a função $f : X \to F(Y)$ se estende ao homomorfismo $f : F(X) \to F(Y)$ com $\pi_Y f = \Theta\pi_X$.

Temos também que $f(\mathbf{r}) \subset \mathcal{C}_{\mathbf{s}}$ pois, como $\mathbf{r} \subset \mathcal{C}_{\mathbf{r}} = \ker \pi_X$ segue que $\pi_Y f(\mathbf{r}) = \Theta\pi_X(\mathbf{r}) = \{1\}$. Portanto f é uma aplicação de presentação satisfazendo $f_* = \Theta$. A unicidade decorre do corolário 30. ■

Vemos assim que as classes de homotopia de aplicações de presentação estão em correspondência biunívoca com os homomorfismo entre os grupos presentados. Além disso tal correspondência preserva composições.

Definição 33. Duas presentações $(X : \mathbf{r})$ e $(Y : \mathbf{s})$ são ditas do mesmo tipo se existem aplicações $f : (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $g : (Y : \mathbf{s}) \to (X : \mathbf{r})$ tais que $gf \simeq 1 : F(X) \to F(X)$ e $fg \simeq 1 : F(Y) \to F(Y)$. O par de aplicações (f, g) (ou cada uma separadamente) é chamado equivalência de presentação.

Teorema 34. *Duas presentações são do mesmo tipo se, e somente se, seus grupos são isomorfos.*

Prova: Sejam $f : (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $g : (Y : \mathbf{s}) \to (X : \mathbf{r})$ aplicações de presentação. Se (f, g) é uma equivalência de presentação então

$$g_* f_* = (gf)_* = 1_* = 1$$
$$f_* g_* = (fg)_* = 1_* = 1$$

Portanto f_* é um isomorfismo entre $|X : \mathbf{r}|$ e $|Y : \mathbf{s}|$.

Reciprocamente, se $\Theta : |X : \mathbf{r}| \to |Y : \mathbf{s}|$ é um isomorfismo o teorema 32 nos dá aplicações de presentação f e g tais que $f_* = \Theta$ e $g_* = \Theta^{-1}$. Então

$$(gf)_* = g_* f_* = \Theta^{-1}\Theta = 1 = 1_* \Rightarrow gf \simeq 1$$
$$(fg)_* = f_* g_* = \Theta\Theta^{-1} = 1 = 1_* \Rightarrow fg \simeq 1$$

Portanto (f, g) é uma equivalência de presentação. ■

4 Equivalências de Tietze

Entre as equivalências de presentação merecem destaque as equivalências de Tietze (I, I') e (II, II') as quais serão agora consideradas.

Definição 35. Sejam $(X : \mathbf{r})$ uma presentação e $\zeta \in \mathcal{C}_{\mathbf{r}}$. Considere a presentação $(Y : \mathbf{s})$ onde, $Y = X$ e $\mathbf{s} = \mathbf{r} \cup \{\zeta\}$. Definimos a equivalência de Tietze I e I' por $(X : \mathbf{r})$, $(Y : \mathbf{s})$ e o automorfismo identidade $I : F(X) \to F(Y)$ e $(X : \mathbf{r})$, $(Y : \mathbf{s})$ e o automorfismo identidade $I' : F(Y) \to F(X)$, respectivamente.

Definição 36. Sejam $(X : \mathbf{r})$ uma presentação, $\zeta \in E - X$ e $\xi \in F(X)$ quaisquer. Considere a presentação $(Y : \mathbf{s})$ onde, $Y = X \cup \zeta$ e $\mathbf{s} = \mathbf{r} \cup \{\zeta\xi^{-1}\}$. Definimos a equivalência de Tietze II por $(X : \mathbf{r})$, $(Y : \mathbf{s})$ e o homomorfismo $II : F(X) \to F(Y)$ dado por $II(x) = x, \forall x \in X$. Definimos também a equivalência II' por $(X : \mathbf{r})$, $(Y : \mathbf{s})$ e o homomorfismo $II' : F(Y) \to F(X)$ dado por $II'(x) = x, \forall x \in X$ e $II'(\zeta) = \xi$.

Definição 37. Seja $\phi : X \to Y$ um homomorfismo. Dizemos que ϕ é uma retração se $Y \subset X$ e $\phi(x) = x, \forall x \in Y$.

Proposição 38. *Os pares (I, I') e (II, II') constituem equivalências de presentação.*

Prova: Demonstremos primeiramente para o par (I, I'). Note que $\mathcal{C}_{\mathbf{r}} = \mathcal{C}_{\mathbf{s}}$ pois, como $\mathbf{r} \subset \mathbf{s}$ segue que $\mathcal{C}_{\mathbf{r}} \subset \mathcal{C}_{\mathbf{s}}$. Para mostrar a outra inclusão sejam $\phi : F(X) \to H$, H grupo, um homomorfismo qualquer e $a \in \mathcal{C}_{\mathbf{s}}$. Se $\mathbf{r} \subset \ker\phi$ temos que $\zeta \in \ker\phi$ logo, $\mathbf{s} = \mathbf{r} \cup \{\zeta\} \subset \ker\phi$ e portanto $\phi(a) = 1$. Como ϕ é arbitrário segue que $a \in \mathcal{C}_{\mathbf{r}}$ e assim $\mathcal{C}_{\mathbf{r}} \supset \mathcal{C}_{\mathbf{s}}$. Por conseguinte temos que $I(\mathbf{r}) = \mathbf{r} \subset \mathcal{C}_{\mathbf{s}}$ e $I'(\mathbf{s}) = \mathbf{s} \subset \mathcal{C}_{\mathbf{r}}$ e portanto $I : (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $I' : (Y : \mathbf{s}) \to (X : \mathbf{r})$ são aplicações de presentação. Como os homomorfismos I e I' são automorfismos identidade segue que o par (I, I') é uma equivalência de presentação.

Mostremos agora para o par $(\mathit{II}, \mathit{II}')$. Como II é o homomorfismo inclusão segue que $\mathit{II}(\mathbf{r}) = \mathbf{r} \subset \mathcal{C}_{\mathbf{s}}$ e assim $\mathit{II} : (X : \mathbf{r}) \to (Y : \mathbf{s})$ é uma aplicação de presentação. Agora II' leva $\mathbf{s}$ em $\mathbf{r} \cup \{1\} = \mathbf{r}$ pois, $\mathit{II}'(\mathbf{r}) = \mathbf{r}$ e $\mathit{II}'(\zeta\xi^{-1}) = 1$ logo, $\mathit{II}'(\mathbf{s}) \subset \mathcal{C}_{\mathbf{r}}$ e $\mathit{II}' : (Y : \mathbf{s}) \to (X : \mathbf{r})$ é uma aplicação de presentação.

Como $\mathit{II}'\mathit{II}(x) = x, \forall x \in X$ segue que $\mathit{II}'\mathit{II} : F(X) \to F(X)$ é a identidade. Agora, para todo $x \in X$ temos $\mathit{II}\mathit{II}'(x)x^{-1} = 1$ e $\mathit{II}\mathit{II}'(\zeta)\zeta^{-1} = \mathit{II}(\xi)\zeta^{-1} = \xi\zeta^{-1} = (\zeta\xi^{-1})^{-1}$ que pertence a consequência de $\mathbf{s}$, ou seja, $\mathit{II}\mathit{II}' \simeq 1$. Portanto o par $(\mathit{II}, \mathit{II}')$ é uma equivalência de presentação. ■

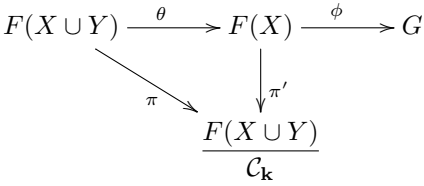
Note que $\mathit{II} : F(X) \to F(X)$ é o homomorfismo inclusão e $\mathit{II}' : F(Y) \to F(Y)$ é uma retração.

Lema 39. *Sejam X e Y subconjuntos disjuntos de E e θ uma retração de $F(X \cup Y)$ sobre $F(X)$. Seja também $(X : \mathbf{r})_\phi$ uma presentação de um grupo G. Então o kernel do homomorfismo $\phi\theta : F(X \cup Y) \to G$ é a consequência, $\mathcal{C}_{\mathbf{k}}$, de $\mathbf{k} = \mathbf{r} \cup \{y\theta(y)^{-1}, y \in Y\}$.*

Prova: Claramente, $\phi\theta(r) = \phi(r) = 1$ para qualquer $r \in \mathbf{r}$. Como θ é uma retração, $\theta^2 = \theta$, e portanto

$$\begin{aligned}\phi\theta(y\theta(y)^{-1}) &= \phi(\theta(y)\theta(\theta(y)^{-1})) = \phi(\theta(y)\theta^2(y)^{-1}) \\ &= \phi(\theta(y)\theta(y)^{-1}) = \phi(1) = 1.\end{aligned}$$

Portanto $\mathbf{k} \subset \ker\phi\theta$ e assim, $\mathcal{C}_{\mathbf{k}}$ está contido no $\ker\phi\theta$. Para mostrar que $\ker\phi\theta \subset \mathcal{C}_{\mathbf{k}}$, considere a projeção canônica $\pi : F(X \cup Y) \to F(X \cup Y)/\mathcal{C}_{\mathbf{k}}$ e sua restrição $\pi' = \pi|F(X)$.



Temos então que $\pi'\theta(x) = \pi'(x) = \pi(x)$ para $x \in X$. Além disso, para $y \in Y$ temos

$$\pi(y)\pi'\theta(y)^{-1} = \pi(y)\pi\theta(y)^{-1} = \pi(y\theta(y)^{-1}) = 1.$$

Portanto $\pi'\theta(y) = \pi(y)$ para todo $y \in X \cup Y$. Isso mostra que $\pi'\theta = \pi$.

Suponha agora que $u \in F(X \cup Y)$ seja tal que $u \in \ker \phi\theta$. Então

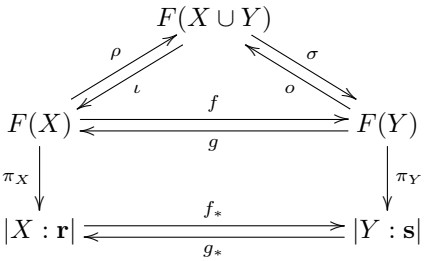
$$\pi(u\theta(u)^{-1}) = \pi'\theta(u\theta(u)^{-1}) = \pi(\theta(u)\theta(u)^{-1}) = \pi(1) = 1,$$

e portanto $u\theta(u)^{-1} \in \mathcal{C}_{\mathbf{k}}$. Mas $\phi\theta(u) = 1$, logo $\theta(u) \in \ker \phi = \mathcal{C}_{\mathbf{r}}$ e portanto $\theta(u) \in \mathcal{C}_{\mathbf{k}}$. Concluimos então que $u = u\theta(u)\theta(u)^{-1} \in \mathcal{C}_{\mathbf{k}}$. ∎

Teorema 40 (de Tietze). *Sejam $f\colon (X : \mathbf{r}) \to (Y : \mathbf{s})$ e $g\colon (Y : \mathbf{s}) \to (X : \mathbf{r})$ aplicações de presentação. Se (f, g) é uma equivalência de presentação e as presentações $(X : \mathbf{r})$ e $(Y : \mathbf{s})$ são finitas então existe uma sequência finita $(T_1, T_1'), (T_2, T_2'), \ldots, (T_\ell, T_\ell')$ de equivalências de Tietze tais que $f = T_1 T_2 \cdots T_\ell$ e $g = T_\ell' T_{\ell-1}' \cdots T_1'$.*

Prova: Primeiramente analisemos o caso $X \cap Y = \emptyset$. Considere o seguinte diagrama, onde ι e o são inclusões e ρ e σ são retrações definidas de forma que

$$\rho(y) = g(y), \forall y \in Y \quad \sigma(x) = f(x), \forall x \in X.$$



Temos que $(X : \mathbf{r}) \underset{\iota}{\overset{\rho}{\rightleftarrows}} (X \cup Y : \mathbf{t})$ onde $\mathbf{t} = \mathbf{r} \cup \{y\rho(y)^{-1}; y \in Y\}$, é uma equivalência de presentação. De fato, como ι é o homomorfismo

inclusão temos que $\iota(\mathbf{r}) = \mathbf{r} \subset \mathcal{C}_{\mathbf{t}}$ e como ρ é uma contração segue que $\mathbf{t}$ é levado em $\mathbf{r} \cup \{1\}$ logo ι e ρ são aplicações de presentação. Agora $\rho\iota : F(X) \to F(X)$ é o automorfismo identidade, além disso $\iota\rho(x)x^{-1} = 1$, $\forall x \in X$ e $\iota\rho(y)y^{-1} = \iota g(y)y^{-1} = g(y)y^{-1} = \rho(y)y^{-1} = (y\rho(y)^{-1})$ portanto $\rho\iota = 1$ e $\iota\rho \simeq 1$, ou seja, (ι, ρ) é uma equivalência de presentação.

Vemos que a equivalência (ι, ρ) pode ser fatorada em m equivalências de Tietze *II* $(T_1, T_1'), (T_2, T_2'), \ldots, (T_m, T_m')$, onde m é o número de elementos do conjunto Y. Sendo $Y = \{y_1, y_2, \ldots, y_m\}$ podemos fazer a fatoração como mostra o diagrama abaixo:

$$
\begin{array}{c}
(X : \mathbf{r}) \\
{\scriptstyle T_1'}\uparrow\downarrow{\scriptstyle T_1} \\
(X \cup \{y_1\} : \mathbf{r} \cup \{y_1\rho(y_1)^{-1}\}) \\
{\scriptstyle T_2'}\uparrow\downarrow{\scriptstyle T_2} \\
(X \cup \{y_1, y_2\} : \mathbf{r} \cup \{y_1\rho(y_1)^{-1}, y_2\rho(y_2)^{-1}\}) \\
\vdots \\
((X \cup Y) - \{y_m\} : \mathbf{t} - \{y_m\rho(y_m)^{-1}\}) \\
{\scriptstyle T_m'}\uparrow\downarrow{\scriptstyle T_m} \\
(X \cup Y : \mathbf{t})
\end{array}
$$

Então $\iota = T_m T_{m-1} \cdots T_1$ e $\rho = T_1' T_2' \cdots T_m'$.

De modo análogo a equivalência

$$(Y : \mathbf{s}) \underset{\sigma}{\overset{o}{\rightleftarrows}} (X \cup Y : \mathbf{w}),$$

onde $\mathbf{w} = \mathbf{s} \cup \{x\rho(x)^{-1}; x \in X\}$ pode ser fatorada em n equivalências de Tietze *II* (S_1, S_1'), $(S_2, S_2'), \ldots, (S_m, S_m')$, onde n é o número de elementos do conjunto X. Então $o = S_n S_{n-1} \cdots S_1$ e $\sigma = S_1' S_2' \cdots S_n'$. Agora, o lema 39 nos dá que $\ker \pi_X\rho$ é $\mathcal{C}_{\mathbf{t}}$. Mas $\pi_X\rho = g_*\pi_Y\sigma$ e $\pi_Y\sigma(\mathbf{w}) = 1$, já que σ leva $\mathbf{w}$ em $\mathbf{s} \cup \{1\} = \mathbf{s}$. Portanto, $\mathbf{w} \subset \mathcal{C}_{\mathbf{t}}$. O mesmo argumento mostra que $\mathbf{t} \subset \mathcal{C}_{\mathbf{w}}$.

Logo as equivalências de presentação

$$(X \cup Y : \mathbf{t}) \underset{\alpha'}{\overset{\alpha}{\rightleftarrows}} (X \cup Y : \mathbf{w} \cup \mathbf{t}) \underset{\beta'}{\overset{\beta}{\rightleftarrows}} (X \cup Y : \mathbf{w}),$$

decorrentes dos automorfismos identidade podem ser fatoradas em equivalências de Tietze I $(U_1, U_1'), (U_2, U_2'), \ldots, (U_{q+n}, U_{q+n}')$ e $(V_1, V_1'), \ldots, (V_{p+m}, V_{p+m}')$, onde p e q são o número de elemetos de $\mathbf{r}$ e $\mathbf{s}$, respectivamente. Então

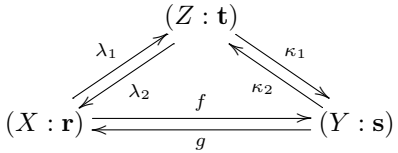
$$\begin{aligned} \alpha &= U_{q+n}U_{q+n-1}\cdots U_1, & \beta &= V_{p+m}V_{p+m-1}\cdots V_1, \\ \alpha' &= U_1'U_2'\cdots U_{q+n}', & \beta' &= V_1'V_2'\cdots V_{p+m}' \end{aligned}$$

e portanto

$$\begin{aligned} f &= \sigma\beta'\alpha\iota = S_1'S_2'\cdots S_n'V_1'V_2'\cdots V_{p+m}'U_{q+n}U_{q+n-1}\cdots U_1T_mT_{m-1}\cdots T_1 \\ g &= \rho\alpha'\beta o = T_1'T_2'\cdots T_m'U_1'U_2'\cdots U_{q+n}'V_{p+m}V_{p+m-1}\cdots V_1S_nS_{n-1}\cdots S_1 \end{aligned}$$

Analisemos agora o caso $X \cap Y \neq \emptyset$. Selecionamos um subconjunto Z de E tal que $Z \cap (X \cup Y) = \emptyset$ e que exista uma bijeção com X.

Essa bijeção induz um isomorfismo $\lambda_1 : F(X) \to F(Z)$ com inversa $\lambda_2 = \lambda_1^{-1} : F(Z) \to F(X)$. Sejam $\mathbf{t} = \lambda_1(\mathbf{r})$, $\kappa_1 = f\lambda_2$ e $\kappa_2 = \lambda_1 g$. Então $f = \kappa_1\lambda_1$ e $g = \lambda_2\kappa_2$.



Claramente (λ_1, λ_2) é uma equivalência de presentação. Queremos provar que (κ_1, κ_2) também é uma equivalência de presentação.

Temos que $\kappa_1(\mathbf{t}) = f\lambda_2(\mathbf{t}) = f(\mathbf{r}) \subset \mathcal{C}_\mathbf{s}$ e $\kappa_2(\mathbf{s}) = \lambda_1 g(\mathbf{t}) \subset \lambda_1(\mathcal{C}_\mathbf{r})$ mas, pelo teorema 17 tem-se que $\lambda_1(\mathcal{C}_\mathbf{r}) = \mathcal{C}_\mathbf{t}$, portanto κ_1 e κ_2 são aplicações de presentação. Além disso

$$\kappa_1\kappa_2 = f\lambda_2\lambda_1 g = fg \simeq 1$$
$$\kappa_2\kappa_1 = \lambda_1 g f \lambda_2 \simeq \lambda_1\lambda_2 = 1.$$

Portanto (κ_1, κ_2) é equivalência de presentação. Agora basta aplicar a primeira parte da demonstração aos pares de presentações $(X : \mathbf{r}), (Z : \mathbf{t})$ e $(Z : \mathbf{t}), (Y : \mathbf{s})$ e o resultado esta provado. ■

A importância do teorema de Tietze é que este reduz o problema de mostrar que uma dada aplicação de presentação de grupos depende somente do grupo presentado pela checagem de que ela permanece inalterada pelas operações de Tietze I e II.

Exemplo 41. Podemos utilizar as equivalências de Tietze para obter uma presentação a partir de outra. Desta forma mostremos que os grupos $|\{x, y, z\} : \{xyz = yzx\}|$ e $|\{x, y, a\} : \{xa = ax\}|$ são isomorfos.

$$\begin{array}{c}
(\{x,y,z\} : \{xyz(yzx)^{-1}\}) \\
\downarrow II \\
(\{x,y,z,a\} : \{xyz(yzx)^{-1}, a(yz)^{-1}\}) \\
\downarrow I \\
(\{x,y,z,a\} : \{xyz(yzx)^{-1}, a(yz)^{-1}, xa(ax)^{-1}\}) \\
\downarrow I' \\
(\{x,y,z,a\} : \{a(yz)^{-1}, xa(ax)^{-1}\}) \\
\downarrow I \\
(\{x,y,z,a\} : \{a(yz)^{-1}, xa(ax)^{-1}, z(y^{-1}a)^{-1}\}) \\
\downarrow I' \\
(\{x,y,z,a\} : \{xa(ax)^{-1}, z(y^{-1}a)^{-1}\}) \\
\downarrow II' \\
(\{x,y,a\} : \{xa(ax)^{-1}\})
\end{array}$$

Abstract: In this work we introduce the concepts of Free Group and Group Presentation. These concepts are indispensable to the study of the fundamental group of knots. We also present and prove the Tietze Theorem which say that every presentation equivalence between finite presentations can be factorized into a finite number of Tietze equivalences.

Keywords: free group; group presentation; Tietze theorem

Referências Bibliográficas

[1] Crowell, R.H.; Fox, R.H., *Introduction to Knot Theory*, Blaisdell Publishing Co., 1965.

[2] Jacobson, N., *Basic Algebra I*, W.H. Freeman and Co., 1985.

[3] Herstein, I.N., *Topics in Algebra*, John Wiley & Sons, 1975.

ISSN 1980-024X

Boletim de Iniciação Científica em Matemática – BICMat

Orientação aos autores

Ao redigir o material a ser divulgado o autor deve observar que o alvo principal é o aluno de graduação, devendo a redação ser clara e objetiva incentivando-o à leitura.

O trabalho deve ser enviado à Comissão Editorial, via e-mail, na linguagem LaTeX, usando a classe `bicmat`. Mais informações sobre a formatação do trabalho podem ser encontradas em `www.rc.unesp.br/igce/matematica/bicmat`, assim como o endereço para o envio do trabalho.

A responsabilidade de cada artigo é exclusiva do autor e respectivo orientador.