SOCIEDADE BRASILEIRA DE MATEMÁTICA
FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL

SÉZANI MORAIS GONÇALVES DE CARVALHO

MATRIZES, DETERMINANTES E POLINÔMIOS: Aplicações em códigos corretores de erros, como estratégia motivacional para o ensino de matemática.

PORTO VELHO
2014

SÉZANI MORAIS GONÇALVES DE CARVALHO

MATRIZES, DETERMINANTES E POLINÔMIOS: Aplicações em códigos corretores de erros, como estratégia motivacional para o ensino de matemática.

Trabalho de conclusão apresentado ao Mestrado Profissional em Matemática em Rede Nacional – PROFMAT no polo da Universidade Federal de Rondônia – UNIR, como requisito parcial para a obtenção do título de Mestre em Matemática Profissional.

Orientador: Prof. Dr. Tomás Daniel Menéndez Rodriguez.

Porto Velho
2014

FICHA CATALOGRÁFICA

BIBLIOTECA PROF. ROBERTO DUARTE PIRES

Carvalho, Sézani Morais Gonçalves de.

C3311m Matrizes, determinantes e polinômios: aplicação em códigos corretores de erros, como estratégia motivacional para o ensino de matemática / Sézani Morais Gonçalves de Carvalho. Porto Velho, Rondônia, 2014.

166 fl.; il.

Dissertação (Mestrado Profissional em Matemática) – Fundação Universidade Federal de Rondônia – UNIR.

Orientador: Prof. Tomás Daniel Menéndez Rodriguez

1.Matrizes. 2.Determinantes. 3.Polinômios. 4. Códigos corretores de erros. I. Rodriguez, Tomás Daniel Menéndez. II.Título.

CDU: 519.612

Bibliotecária responsável: Rejane Sales - CRB11/903

Sézani Morais Gonçalves de Carvalho

MATRIZES, DETERMINANTES E POLINÔMIOS: Aplicação em códigos corretores de erros, com estratégia motivacional para o ensino de matemática

Este trabalho foi julgado e aprovado para obtenção do titulo de Mestre em Matemática do Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, do Departamento de Matemática da Fundação Universidade Federal de Rondônia, Campus de Porto Velho - RO.

Porto Velho, 05 de dezembro de 2014

COMISSÃO EXAMINADORA

Prof. Dr. Tomás Daniel Menéndez Rodrigues
Orientador/Presidente
PROFMAT / UNIR

Prof. Dr. Marinaldo Felipe da Silva
PROFMAT / UNIR

Prof. Dr. José Ivan da Silva Ramos
PROFMAT / UFAC

Aos meus pais Dorival e Terezinha.
À minha amada filha Aízis.
Ao saudoso Professor Domingos dos Reis.
Ao grande amigo e professor Marinaldo.

AGRADECIMENTOS

Com muita sinceridade e bastante carinho, externo meus agradecimentos:

A Deus pela oportunidade de vitória a mim concedida, pois nos momentos mais difíceis dessa caminhada encontrei nele o conforto e a força necessária para prosseguir.

Ao meu orientador Prof. Dr. Tomás Daniel Menéndez Rodriguez por acreditar em mim desde o saudoso período da graduação.

Aos meus pais Dorival Gonçalves de Carvalho e Terezinha Morais de Carvalho por me incentivarem sempre nos estudos, além de me apoiarem.

À minha filha Aízis Morais de Carvalho por ser a razão pela qual sempre insisto em alcançar maiores conquistas.

Ao grande professor e estimado amigo Dr. Marinaldo Felipe por contribuir comigo e demais alunos com seu extraordinário conhecimento de matemática e com sua peculiaridade em ser sempre entusiasta.

À Querida professora e amiga Maria das Graças por ser uma fonte inspiradora em como ensinar e amar a matemática.

Aos professores e amigos Adeilton Costa, Flávio Batista Simão e Ronaldo Chaves pelas contribuições em minha formação além da boa amizade.

Aos meus grandes amigos e companheiros de academia que ao meu lado batalharam durante esses anos e foram incorporados à minha vida: Adalberto Carlos, Alisson, Francenildo, Francisco Sales, Gilson Caliani, Jean, José Inildo, Kleber Sales, Luci, Marizete Nink, Magno Martins, Vicente e, em especial ao jovem Guilherme, com quem aprendi muito, me diverti muito e construí uma sólida e eterna amizade.

A todos os que contribuíram com minha formação, aos quais mencionei acima, e aos colaboradores que por ventura deixei de mencionar, externo minha eterna gratidão.

RESUMO

Este trabalho consiste em um material de apoio aos professores de matemática atuantes nas séries finais do ensino médio, bem como para os alunos concluintes desse ciclo da educação básica, que desejem aprofundar seus conhecimentos.
Inicialmente, abordamos neste texto os fatores motivadores para a construção deste material de apoio. Em seguida apresentamos os conteúdos de matrizes, determinantes e polinômios, que estão presentes no currículo da disciplina de matemática no ensino médio. São apresentadas também as estruturas algébricas elementares que, embora não façam parte dos currículos de matemática na educação básica, aparecem parcialmente desde o ensino fundamental mesmo que de forma implícita nessa disciplina. Por fim, apresentamos as aplicações desses conteúdos matemáticos na teoria dos códigos corretores de erros, que é foco deste trabalho, além de um rol de atividades propostas sobre os conteúdos abordados.

Palavras Chave: Matrizes. Determinantes. Polinômios. Estruturas Algébricas. Códigos Corretores de Erros.

ABSTRACT

This work consists of a support material for teachers of mathematics acting in high school finals series, as well as for students graduating from this cycle of basic education, who aim to deepen their knowledge on the subject.
Initially, we discussed in this text the motivating factors for the construction of this support material. Then we present the contents of matrix, determinants and polynomials, which are present in the high school mathematics discipline curriculum. Also the elementary algebraic structures are presented which, although not part of the curriculum of mathematics in basic education, appear partially since elementary school, even if implicitly in this discipline. Finally, applications of these mathematic contents are presented in the theory of errors correcting codes, main focus of this work, besides a roster of proposed activities about the addressed contents.

Key words: Matrix, Determinants, Polynomials, Algebraic structures, Errors correcting codes.

LISTA DE FIGURAS

LISTA DE GRÁFICOS

SUMÁRIO

INTRODUÇÃO

Na atualidade, vários esforços têm sido realizados com o objetivo de proporcionar melhorias no ensino e aprendizagem de matemática, dentre eles podemos citar o programa Pacto Nacional pela Alfabetização na Idade Certa, do Governo Federal ou ainda programas das Secretarias Estaduais e Municipais de Educação de diversos Estados e Municípios brasileiros.

Embora existam medidas de diversas partes para proporcionar essa melhoria, temos visto ao longo dos anos que se trata de um processo demorado e complexo atingir esse objetivo, uma vez que pesquisas específicas realizadas nas várias etapas da educação básica, como exemplos a Provinha Brasil e o Sistema de Avaliação da Educação Básica (SAEB), tem mostrado.

O Índice de Desenvolvimento da Educação Básica (IDEB) mostra um avanço sutil no desenvolvimento da educação básica. Os dois últimos resultados do IDEB, a saber, dos anos de 2011 e 2013, apontam um discreto progresso, pois, do 1º ao 5º ano do ensino fundamental o índice aumentou de 5,0 para 5,2, enquanto que do 6º ao 9º ano do ensino fundamental, o índice aumentou de 4,1 para 4,2 e, no ensino médio o índice se manteve em 3,7.

Com base nos resultados apresentados acima, percebemos que à medida que avançamos para as séries finais da educação básica, dois fenômenos são observados:

1º - Os índices são menores;

2º - O progresso em cada etapa avaliada diminui à medida que avançamos aos anos finais, pois de 2011 para 2013, houve aumento de 0,2 pontos no índice do 1º ao 5º ano do ensino fundamental, 0,1 ponto do 6º ao 9º do ensino fundamental e não houve aumento no índice referente ao ensino médio.

Como componente presente nos currículos da educação básica, a matemática está inserida nesse contexto e seu ensino/aprendizagem tem participação no fracasso ou sucesso dos estudantes nessas etapas da educação.

É sabido que no desenvolvimento humano o sujeito, inicialmente adquire suas experiências a partir do concreto e em uma etapa posterior, decorrente das experiências adquiridas, atinge o estágio de abstração. Segundo Piaget, “Após os 11 ou 12 anos, o pensamento formal torna-se possível, isto é, as operações lógicas começam a ser transpostas do plano da manipulação para as ideias” (PIAGET, 1995, p 59). Com a matemática não é diferente: as experiências iniciais são adquiridas a partir do concreto e, em uma etapa posterior, à medida que vai avançando, a matemática vai se distanciando do concreto e sendo

imersa em um contexto abstrato e cada vez mais abstrato. Porém, cabe ressaltar que embora adquira status avançado de abstração, não deixa de ter aplicabilidade no mundo concreto, mesmo porque, grande parte dos avanços matemáticos existentes surgiu da necessidade de atender a alguma demanda do mundo concreto. Meyer *et al* evidencia que "os gregos desenvolveram o cálculo de área por que tinham de fazer as medições das terras do Nilo; os fenícios desenvolveram conceitos aritméticos de contabilidade porque eram comerciantes" (MEYER *et al*, 2011, p 25).

Nesse contexto, podemos ver nos Parâmetros Curriculares Nacionais o seguinte texto:

> A Matemática, por sua universalidade de quantificação e expressão, como linguagem portanto, ocupa uma posição singular. No Ensino Médio, quando nas ciências torna-se essencial uma construção abstrata mais elaborada, os instrumentos matemáticos são especialmente importantes. Mas não é só nesse sentido que a Matemática é fundamental. Possivelmente, não existe nenhuma atividade da vida contemporânea, da música à informática, do comércio à meteorologia, das engenharias às comunicações, em que a Matemática não compareça de maneira insubstituível para codificar, ordenar, quantificar e interpretar compassos, taxas, dosagens, coordenadas, tensões, frequências e quantas outras variáveis houver (PNC/Ensino Médio, p 9).

Sendo assim, a abstração é essencial para "o aprender" matemática, é impensável uma matemática que se alimente puramente do concreto, porém, mesmo que de maneira implícita, a matemática está presente nas atividades cotidianas, o que leva-nos a pensar em estratégias de ensino que apresentem as aplicabilidades da matemática no dia a dia. O ato de conhecer não deve estar puramente ligado ao "saber para que serve", mas quando apresentamos utilidades àquilo que ensinamos e pontes de ligação entre o abstrato e o concreto, ensinamos uma Matemática possivelmente mais capaz de despertar interesse aos estudantes, revelar identidades e afinidades e por consequência, construir um conhecimento mais sólido.

A proposta deste trabalho é a apresentação de alguns conteúdos do currículo escolar de matemática do ensino médio, dando atenção especial às demonstrações das propriedades e dos teoremas pertinentes, para em seguida, apresentar uma aplicação desses conteúdos em um contexto "extramatemático" ou "extraescolar". A escolha da aplicação na teoria dos códigos corretores de erros deu-se em virtude de este ser um assunto pouco conhecido ou discutido entre os jovens, porém, muito presente em recursos tecnológicos utilizados pelos mesmos, uma vez que as telecomunicações e os dispositivos de armazenamento presentes no nosso dia a dia, muito mais entre os jovens, não seriam confiáveis nem eficientes sem a utilização dessa teoria. Sendo assim, procuramos neste trabalho, apresentar os conceitos básicos dessa teoria,

com uma preocupação maior em atrair a atenção dos estudantes à disciplina de matemática, a partir do pressuposto do conhecimento de uma das suas vastas aplicações.

Cabe salientar que o material apresentado neste trabalho é primeiramente direcionado aos professores ou pessoas que tenham um conhecimento prévio de matemática além das operações fundamentais. Não é necessário que os alunos do ensino médio saibam demonstrar os teoremas apresentados no capítulo referente à teoria dos códigos corretores de erros, sendo mais interessante, a partir dos conceitos apresentados pelos professores sobre essa teoria, saberem operar com matrizes, determinantes, polinômios e, conhecendo suas propriedades, desenvolverem com mais habilidades os cálculos e argumentações relacionadas com esses assuntos.

1 MOTIVAÇÃO E ABORDAGEM DO TRABALHO

1.1 POR QUE ESTUDAR MATRIZES, DETERMINANTES E POLINÔMIOS?

O questionamento acima fez parte da minha vida em pelo menos duas ocasiões diferentes: a primeira enquanto eu ainda era aluno do ensino médio e presenciava meu professor destrinchar matrizes enormes de ordem 5 ou 6, determinando cada um dos seus elementos através de uma sentença que aparecia em função dos "is" e dos "jotas". Por vezes, recebia listas de exercícios nas quais volta e meia aparecia para ser calculado um determinante de uma matriz quadrada de ordem 5 através da regra de Laplace, ou ainda polinômios de graus elevados, dos quais tinha que determinar o quociente e o resto ou ainda encontrar as raízes reais. Na condição de aluno, resolvia essas atividades, mas não sabia para o que serviam. Por vezes imaginava que eram caprichos matemáticos que serviam simplesmente para treinar habilidades em multiplicar ou dividir números reais. Saí do ensino médio sem saber para o que serviam as matrizes, os determinantes e os polinômios. Por vezes encontrava alguma aplicação, porém, sempre dentro da própria matemática. Na segunda situação na qual deparei-me com o questionamento acima, anos já tinham passado e eu encontrava-me na posição de professor e ouvia dos meus alunos as mesmas indagações que no meu tempo de ensino médio, fazia a mim mesmo ou a amigos ou ainda ao próprio professor: para que servem as matrizes, determinantes e os polinômios? Quando eu terminar o ensino médio, aonde irei usar isso? Onde aplicarei esses conhecimentos no meu trabalho? A primeira pergunta seguramente sou capaz de responder, pois o objetivo deste trabalho, por si só traz a resposta. A segunda pergunta, se interpretada com um olhar matemático, também pode ser respondida: toda vez que um computador for utilizado, um telefone celular ou qualquer canal de comunicação, implicitamente estarão sendo usadas as matrizes, os polinômios e tantos outros conhecimentos matemáticos. Já a terceira pergunta não possui uma resposta formal, pois tal resposta está condicionada à atividade profissional que o estudante irá executar no futuro.

Certamente essas dúvidas não estão presentes somente nos conteúdos de matrizes, determinantes e polinômios, porém, tendo em vista a grande quantidade de cálculos que

geralmente são utilizados na resolução de problemas referentes a esses conteúdos, embora elementares, é plausível que com maior frequência ouçamos essas indagações ao ensiná-los.

Criar pontes de acesso entre os conteúdos matemáticos e as aplicações práticas certamente constitui uma estratégia para o ensino desta disciplina. O ato de "saber para que serve" pode ser motivador ao aluno e, caso alguém sonhe em ser engenheiro ou trabalhar com informática ou áreas afins, certamente terá subsídios para o seu direcionamento. Uma matemática que seja trabalhada de modo a associar os conteúdos estudados às aplicações nos fenômenos vivenciados pelos alunos é uma matemática contextualizada. A contextualização segundo Fogaça "é o ato de vincular o conhecimento à sua origem e à sua aplicação" (FOGAÇA, 2012). As ações pedagógicas no ensino da matemática devem apresentar preocupações com a contextualização. Reconhecemos que quanto mais abstrato for o conceito matemático a ser estudado, mais dificultosa será sua contextualização, porém, a abstração excessiva inerente a alguns conteúdos da matemática não constitui entrave algum em ações para que outros conteúdos, menos abstratos, sejam facilmente contextualizados e, por conseguinte tornem-se mais atrativos e mais facilmente compreendidos. Parece-nos que a teoria e a prática caminham em vias divergentes nas quais, à medida que progredimos nos conteúdos matemáticos presentes nos currículos escolares, mais distantes ficam a teoria e a prática. D'Ambrósio diz: "Do ponto de vista de motivação contextualizada, a matemática que se ensina hoje nas escolas é morta. Poderia ser tratada como fato histórico" (D' AMBROSIO, 2012, p 29). É perceptível esse distanciamento quando observamos os livros didáticos adotados pelas escolas de nível médio, nos quais são frequentes as listas de exercícios no fim de cada capítulo, nas quais aparecem: "calcule", "determine", "encontre" etc., sem nenhum elo entre os exercícios e as aplicações. Acerca da contextualização, Meyer *et al* dizem:

> A maioria das pessoas não consegue relacionar a Matemática nem com as outras ciências e muito menos com situações de seus cotidianos, porque foi criado um universo à parte, ou seja, para elas, a Matemática não está presente em outros contextos (MEYER *et al*, 2011, p 24).

Ou ainda, segundo Meyer *et al*, na educação básica, a matemática "chega para os alunos neutra e descontextualizada, com pouca ou nenhuma relação com a realidade de quem está na sala de aula: professores e alunos" (MEYER *et al*, 2011, p. 53).

Particularmente, reconhecemos a necessidade da resolução de exercícios do tipo "calcule", "determine" etc., porém, para um aprendizado consolidado de matemática, há necessidade de problemas que estimulem o pensar, que sirvam de ponte entre teoria e prática,

que suscitem o aluno à busca por respostas tendo como referência os fenômenos da vida extraescolar.

> A solução de problemas baseia-se na apresentação de situações abertas e sugestivas que exijam dos alunos uma atitude ativa ou um esforço para buscar suas próprias respostas, seu próprio conhecimento. O ensino baseado na solução de problemas pressupõe promover nos alunos o domínio de procedimentos, assim como a utilização dos conhecimentos disponíveis, para dar resposta a situações variáveis e diferentes (POZO; ECHEVERRÍA, 1988, p 9).

Trazer a realidade cotidiana para o interior de uma sala de aula pode representar um avanço pedagógico na educação matemática, pois, à medida que o aluno percebe a necessidade de aprender matemática para lidar com os fenômenos da vida real, por mais abstratos que sejam esses conteúdos, provavelmente melhor será o aprendizado. O interior de uma sala de aula de matemática deve conter as realidades vividas pelo aluno quando estão fora da escola, assim como no exterior da sala de aula o aluno deve vivenciar os conhecimentos matemáticos adquiridos. Para D'Ambrosio:

> Particularmente em matemática, parece que há uma fixação na ideia de haver necessidade de um conhecimento hierarquizado, em que cada degrau é galgado numa certa fase da vida, com atenção exclusiva durante horas de aula, como um canal de televisão que se sintoniza para as disciplinas e se desliga acabada a aula. Como se fossem duas realidades disjuntas, a da aula e a de fora da aula (D'AMBROSIO, 2012, p. 76).

A citação acima evidencia a disparidade existente entre o mundo dentro e o mundo fora da sala de aula. Um dos desafios da educação, em especial a matemática, é justamente colocar a seu favor a prática matemática vivenciada de forma explícita ou implícita pelos alunos em seu cotidiano.

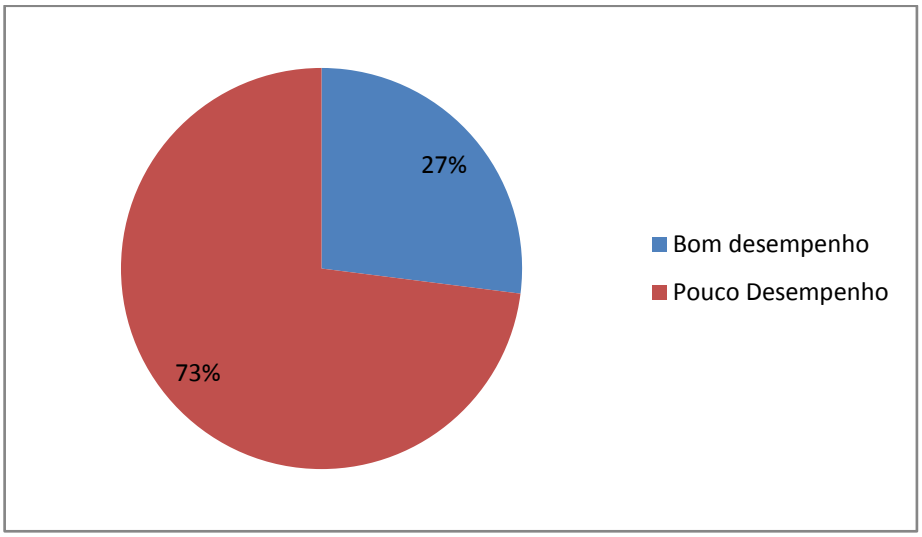
1.2 A MOTIVAÇÃO PARA A REALIZAÇÃO DO TRABALHO

Em relação ao exposto no tópico anterior, fomos movidos a estabelecer um canal de diálogo com os alunos de três turmas do 3º ano do ensino médio de uma escola pública de tempo integral no Município de Porto Velho-RO, para obter deles informações acerca do aprendizado dos conteúdos de matrizes, determinantes e polinômios. Os resultados obtidos são apresentados a seguir:

Questionados sobre terem estudado os conteúdos de matrizes, determinantes e polinômios, tivemos unanimidade em respostas afirmativas.

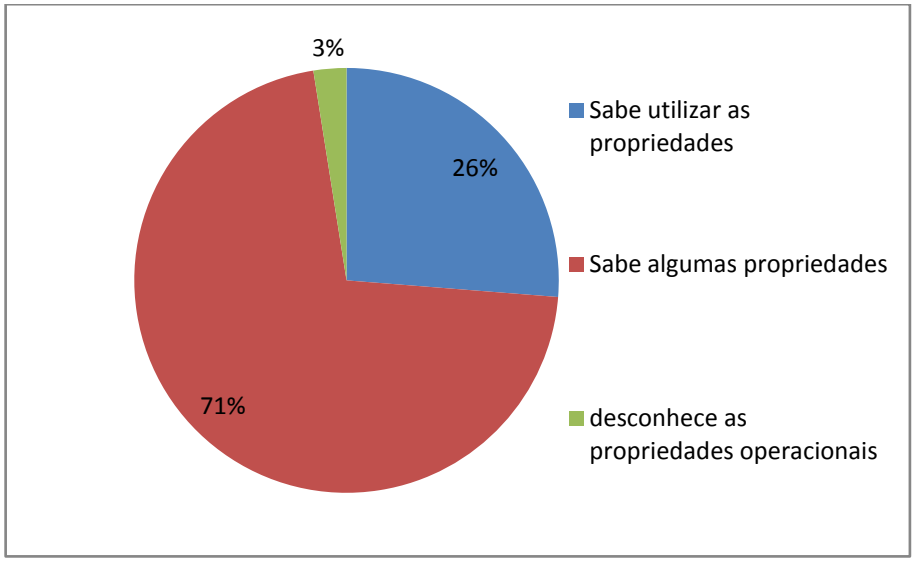
O Gráfico 1 apresenta os resultados obtido sobre o desempenho dos mesmos na resolução de atividades que envolvam as operações com matrizes:

Gráfico 1: Desempenho dos alunos em operações com matrizes



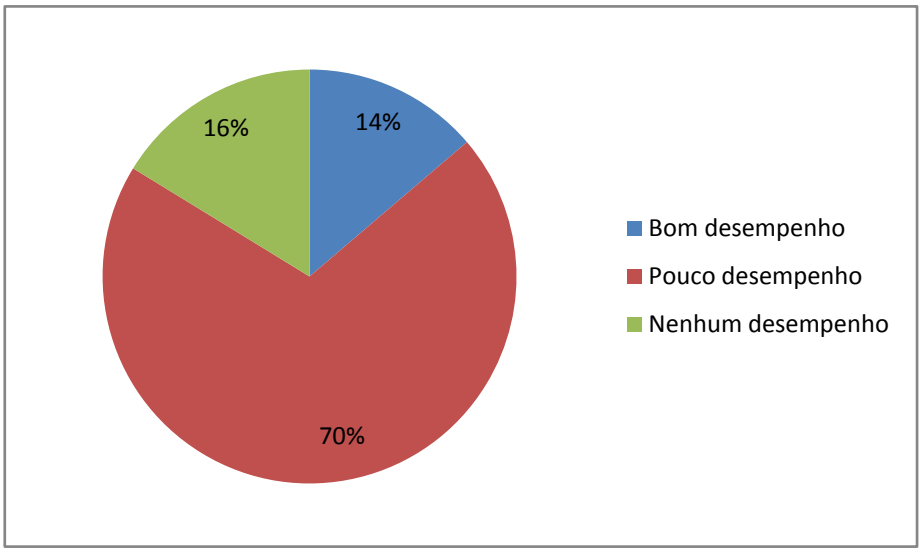
O Gráfico 2 apresenta os resultados obtidos acerca do conhecimento das propriedades operacionais das matrizes:

Gráfico 2: Conhecimento e aplicação das propriedades operacionais das matrizes



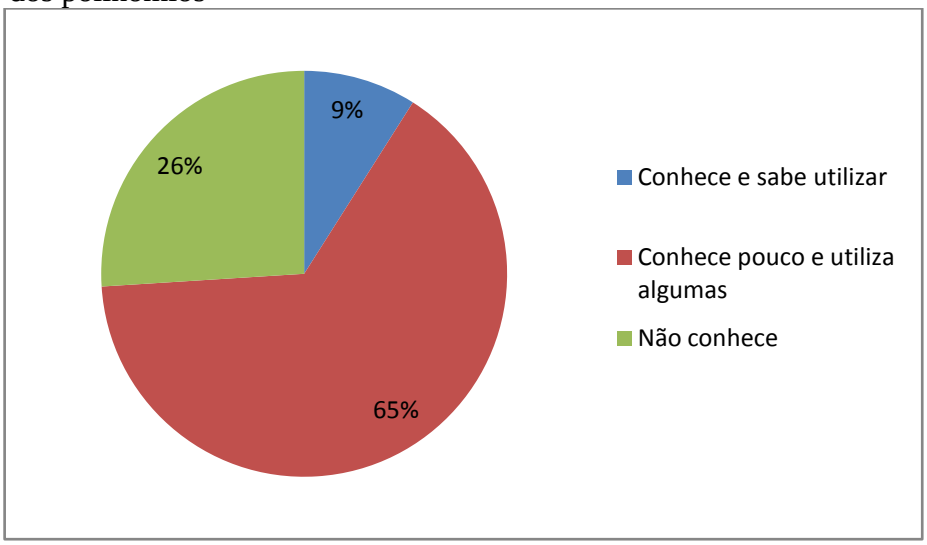
O Gráfico 3 apresenta os resultados obtido sobre o desempenho dos mesmos na resolução de atividades que envolvam as operações com polinômios:

Gráfico 3: Desempenho dos alunos em operações com polinômios



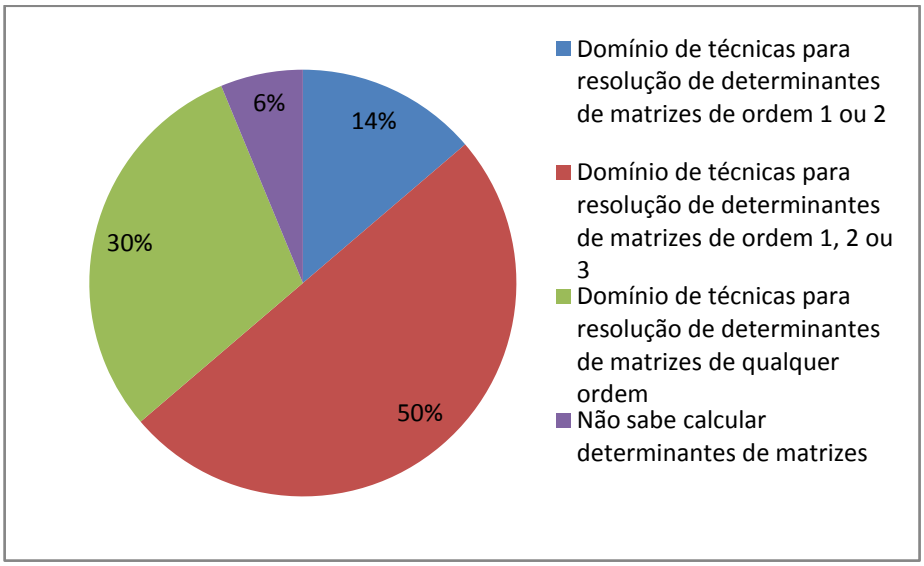
O Gráfico 4 apresenta os resultados obtidos acerca do conhecimento das propriedades operacionais dos polinômios:

Gráfico 4: Conhecimento e utilização das propriedades operacionais dos polinômios



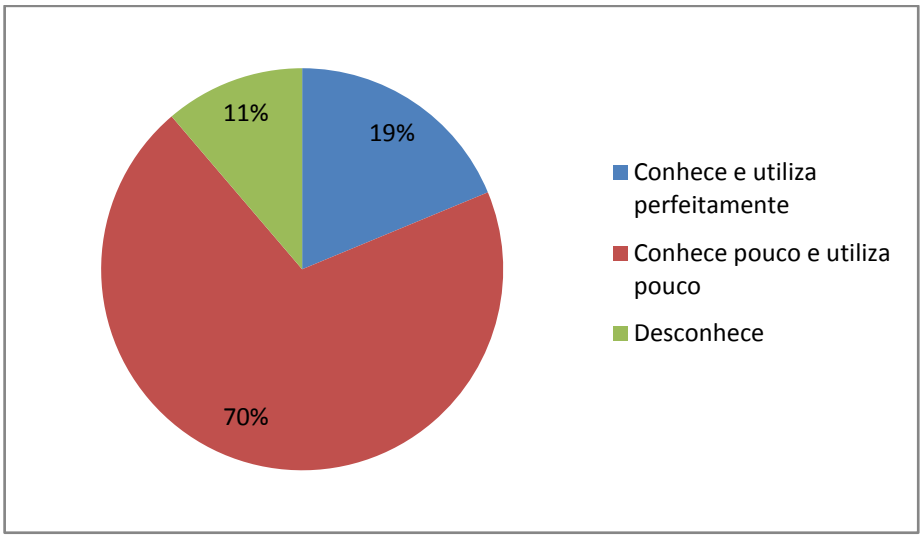
Em relação ao cálculo de determinantes de matrizes, questionados sobre as habilidades na resolução, as respostas obtidas são apresentadas no Gráfico 5:

Gráfico 5: Domínio de técnicas na resolução de determinantes de matrizes



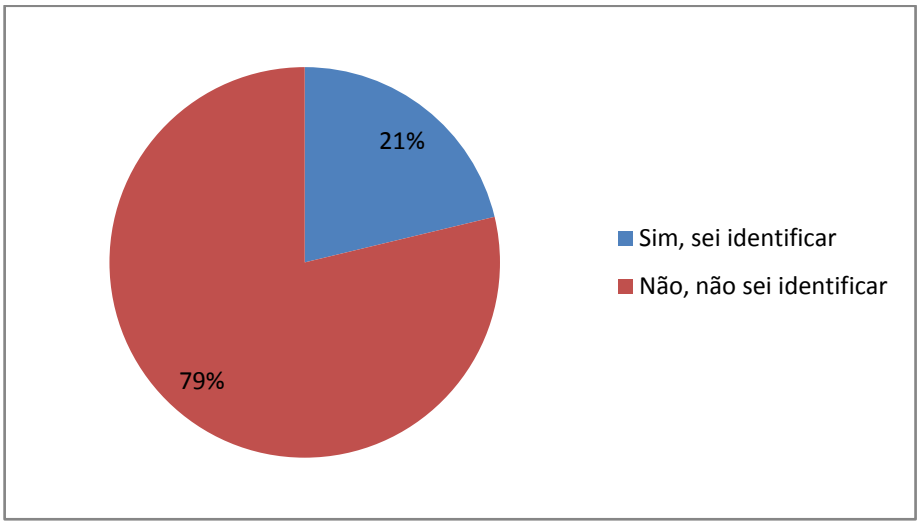
Em relação ao conhecer e saber utilizar as propriedades dos determinantes, obtivemos os resultados apresentados no Gráfico 6:

Gráfico 6: Conhecimento e utilização das propriedades dos determinantes



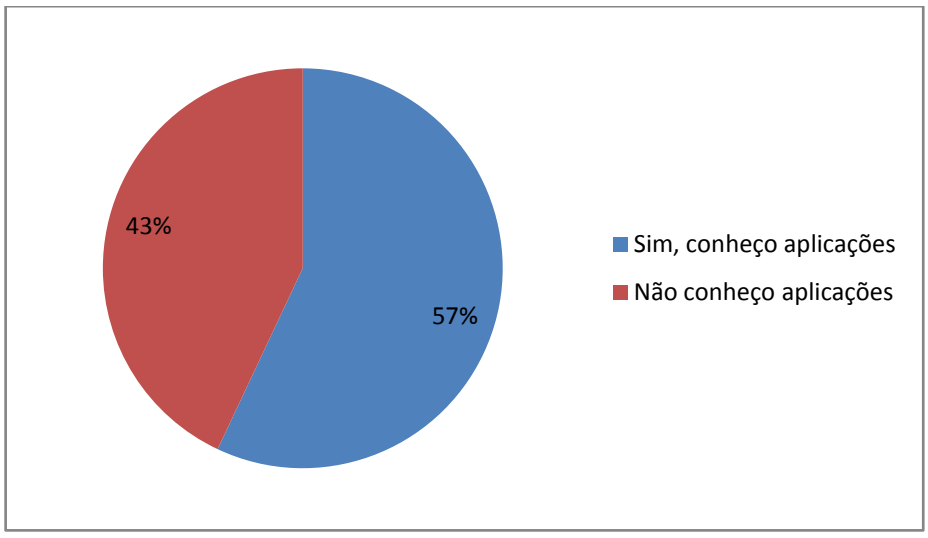
Perguntamos ainda aos alunos se os mesmos sabiam como identificar quando uma matriz é invertível. As respostas obtidas são apresentadas no Gráfico 7:

Gráfico 7: Habilidades na identificação da invertibilidade de uma matriz



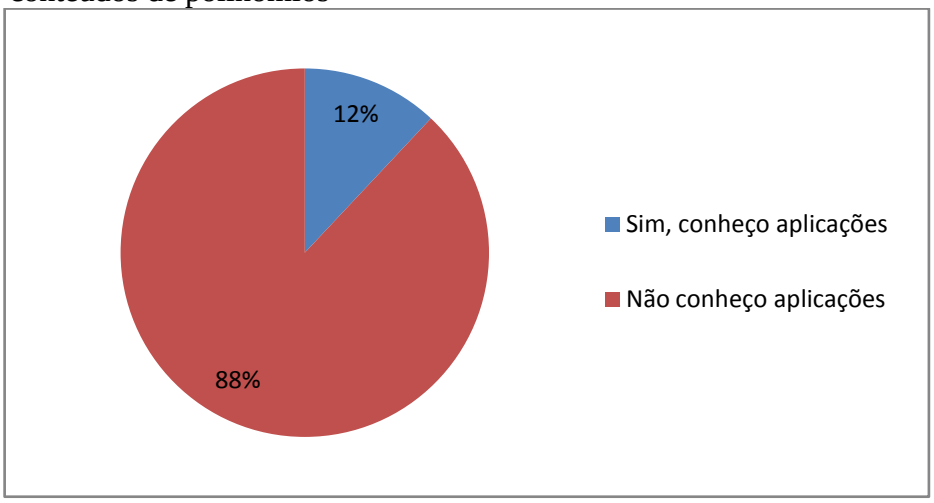
Perguntamos aos alunos se os mesmos conheciam alguma aplicação ou utilidade para as matrizes e os determinantes. Os resultados obtidos são apresentados no Gráfico 8:

Gráfico 8: Conhecimento dos alunos sobre a aplicabilidade dos conteúdos de matrizes e determinantes



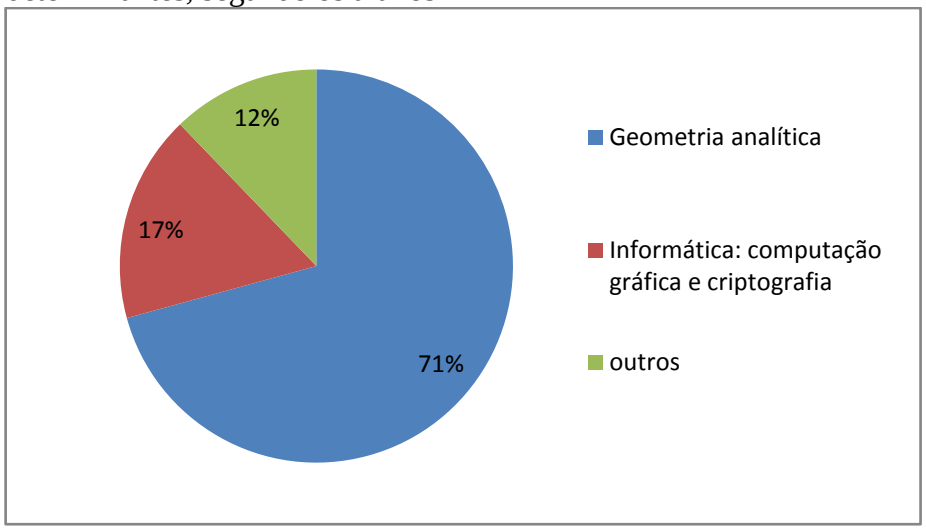
Perguntamos aos alunos se os mesmos conheciam alguma aplicação ou utilidade para o estudo dos polinômios. Os resultados obtidos são apresentados no Gráfico 9:

Gráfico 9: Conhecimento dos alunos sobre a aplicabilidade dos conteúdos de polinômios



Em virtude de 57% dos alunos terem respondido afirmativamente que conhecem aplicações para as matrizes e determinantes, solicitamos que fossem informadas as aplicações que os mesmo conhecem acerca desses conteúdos. Os resultados obtidos são apresentados no Gráfico 10:

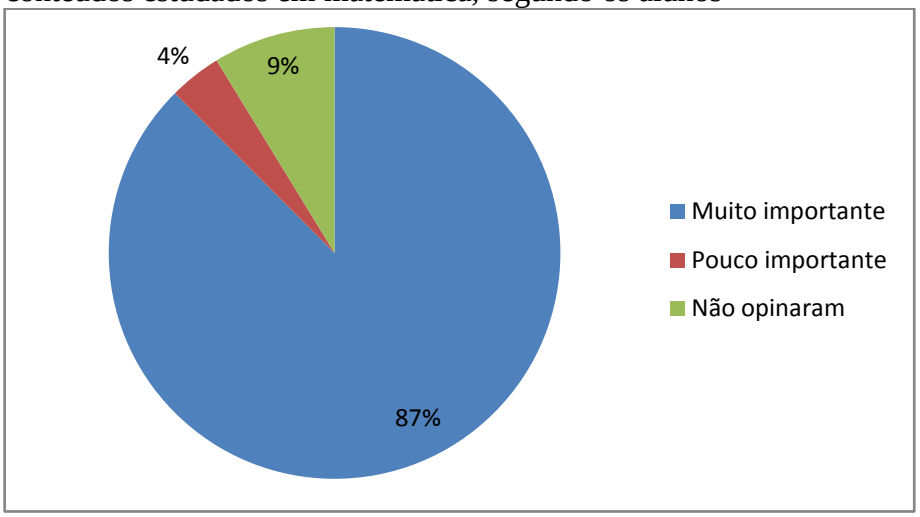
Gráfico 10: Áreas de aplicação dos conteúdos de matrizes e determinantes, segundo os alunos



É interessante observar que grande maioria dos alunos, mais precisamente 71% deles conhecem aplicações das matrizes e determinantes em geometria analítica, ou seja, conhecem uma aplicação da matemática dentro da própria matemática. Menos de 30% dos alunos já ouviram falar de alguma aplicação desses conteúdos em outra área do conhecimento ou em algum fenômeno.

Em um último questionamento, solicitamos aos alunos que opinassem a respeito da importância de conhecer a aplicabilidade dos conteúdos estudados em matemática, nas outras áreas de conhecimento bem como em situações do cotidiano. Os resultados obtidos são apresentados no Gráfico 11:

Gráfico 11: Grau de importância sobre conhecer a aplicabilidade dos conteúdos estudados em matemática, segundo os alunos



Os resultados obtidos através do diálogo com os alunos constituíram um fator motivador para a realização deste trabalho, uma vez que a proposta do mesmo é justamente atender parte da necessidade dos estudantes em conhecer as aplicações para os conteúdos estudados na disciplina de matemática no ensino médio.

1.3 ABORDAGEM DOS CONTEÚDOS DE MATRIZES, DETERMINANTES E POLINÔMIOS NO LIVRO DIDÁTICO.

Além de conter aplicações para os conteúdos de matrizes, determinantes e polinômios, o presente trabalho procura apresentar a demonstração da validade de cada uma das propriedades apresentadas e dos teoremas enunciados. Reconhecemos que a complexidade de alguma dessas demonstrações foge ao nível de conhecimento matemático praticado hoje em dia, em especial no ensino público, porém, mesmo assim julgamos necessário que essas demonstrações se fizessem presentes.

Verificamos que nos livros didáticos atualmente adotados nas escolas públicas, as demonstrações estão deixando de figurar, apenas as propriedades operacionais das matrizes, determinantes e dos polinômios são apresentadas. É sugerido ao aluno que verifique a

validade dessas propriedades através da análise de casos particulares. As análises matemáticas obtidas através de casos particulares são extremamente importantes, pois a partir dessas análises é possível que os alunos obtenham inferências e, por conseguinte, a capacidade de generalização, porém, o fato de uma propriedade ser verificada em casos particulares, pode não garantir a sua validade para uma infinidade de casos. Um exemplo bem simples disso consiste em um aluno que desconheça as propriedades operacionais das matrizes e deseje verificar se a multiplicação de matrizes goza da propriedade comutativa. Para tanto, escolhe ao acaso duas matrizes quadradas de ordem 2: $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$ e efetua as operações $A \cdot B$ e $B \cdot A$ e obtém os seguintes resultados:

$$A \cdot B = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3+0 & 2-1 \\ 0+0 & 0+2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$B \cdot A = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3+0 & -3+4 \\ 0+0 & 0+2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

O aluno observa que $A \cdot B = B \cdot A$. É levantada a suspeita de que a multiplicação de matrizes é comutativa.

Em uma nova tentativa, o aluno escolhe ao acaso duas outras matrizes, com finalidade de validar sua suspeita: $C = \begin{bmatrix} 5 & 2 \\ 0 & -3 \end{bmatrix}$ e $D = \begin{bmatrix} 2 & 1 \\ 0 & -2 \end{bmatrix}$ e efetua as operações $C \cdot D$ e $D \cdot C$ e obtém os seguintes resultados:

$$C \cdot D = \begin{bmatrix} 5 & 2 \\ 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 10+0 & 5-4 \\ 0+0 & 0+6 \end{bmatrix} = \begin{bmatrix} 10 & 1 \\ 0 & 6 \end{bmatrix}$$

$$D \cdot C = \begin{bmatrix} 2 & 1 \\ 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 \\ 0 & -3 \end{bmatrix} = \begin{bmatrix} 10+0 & 4-3 \\ 0+0 & 0+6 \end{bmatrix} = \begin{bmatrix} 10 & 1 \\ 0 & 6 \end{bmatrix}$$

Novamente os resultados obtidos são iguais, o que leva o aluno, tendo como base os casos particulares que analisou, a inferir que a multiplicação de matrizes goza da propriedade comutativa, generalizando esse resultado equivocadamente como veremos em 2.2.3. Portanto, embora a análise de casos particulares seja uma ferramenta útil na matemática para que os alunos busquem por padrões e façam conjecturas, essa ferramenta não pode ser utilizada como verdade absoluta. Em consequência disso, justifica-se a necessidade da presença das demonstrações das propriedades e teoremas pertinentes a cada assunto matemático abordado nos livros didáticos, bem como a prática dessa ação em sala de aula.

Outro ponto que observamos em alguns livros didáticos adotados pelas escolas é a abordagem dos determinantes como uma mera operação matemática a ser realizada com os elementos de uma matriz. Assim, os alunos desconhecem, por exemplo, que o determinante de uma matriz real é uma função com domínio no conjunto das matrizes reais quadradas e contradomínio no conjunto dos números reais, com isso, perdem a oportunidade de

associarem esse conteúdo, com outros conteúdos vistos anteriormente, como a caracterização de uma função como injetiva, sobrejetiva, bijetiva, existência da inversa ou composição de funções.

Cabe salientar ainda que a determinação da matriz inversa de uma matriz A quadrada de ordem n, seja por operações elementares sobre as linhas de uma matriz $[A|I_n]$ ou ainda através do produto da matriz adjunta de A pelo inverso multiplicativo do determinante da matriz A, vem perdendo espaço nos livros didáticos, tirando com isso, a oportunidade dos alunos aprenderem sobre esses conceitos que são fundamentais no estudo das matrizes.

Com relação aos conteúdos sobre polinômios, não foi encontrado texto algum adotado no ensino médio que trate sobre a interpolação de Lagrange.

1.4 A ESCOLHA DA APLICAÇÃO DAS MATRIZES, DETERMINANTES E POLINÔMIOS NOS CÓDIGOS CORRETORES DE ERROS

Diante dos resultados obtidos nos diálogos com os alunos das turmas de 3º ano do ensino médio, além das observações comentadas anteriormente acerca dos livros didáticos adotados pelas escolas, fomos motivados a elaborar um material que buscasse suprir as faltas de demonstrações da validade das propriedades enunciadas nesses livros, bem como a ausência dos principais teoremas de cada um desses conteúdos abordados, além da apresentação de alguma aplicação desses conteúdos em alguma área do conhecimento ou algum fenômeno do cotidiano dos alunos. Esse fato levou-nos aos códigos corretores de erros, uma vez que essa teoria é vastamente utilizada em meios de comunicação e equipamentos de armazenamento de informações que, frequentemente, são utilizados no nosso cotidiano, em especial por grande parte dos jovens que nos dias atuais fazem uso constante de recursos tecnológicos de comunicação e armazenamento tais como telefones celulares, *tablets*, computadores, entre outros.

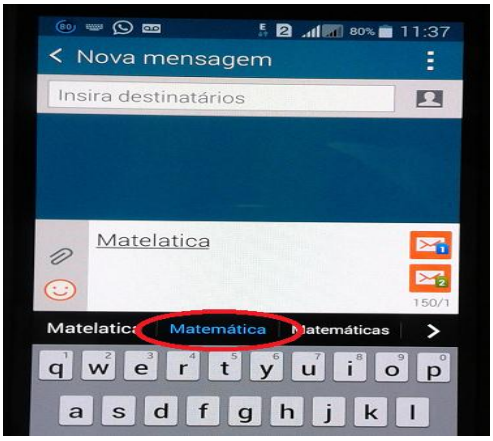
É perceptível nos dias atuais, que a sociedade, numa velocidade muito rápida, tem sido imersa em uma realidade digital. A tecnologia desenvolvida pelas engenharias tem avançado a passos rápidos e esses fatores suscitam às novas ações educacionais, capazes de aproveitar as novas tecnologias a favor de um ensino/aprendizagem com melhor qualidade e capaz de preparar o aluno para ser atuante no meio social. Para Henriques (2010):

> As mudanças sociais e o rápido desenvolvimento tecnológico que se têm verificado na sociedade conduzem a uma alteração nas suas necessidades e, consequentemente, nas competências que é preciso desenvolver nos alunos em áreas fundamentais

> como a da Matemática. Existe actualmente a convicção de que os alunos precisarão de um conjunto muito vasto de competências matemáticas para desempenhar, com eficiência, funções na sociedade actual. De acordo com diversos documentos de referência na área da educação matemática, ao nível do ensino básico e secundário [...], os alunos devem ser capazes de: (i) desenvolver uma profunda compreensão dos conceitos e princípios matemáticos; (ii) raciocinar com rigor e comunicar com clareza; (iii) reconhecer as aplicações matemáticas no mundo que os rodeia e enfrentar os problemas matemáticos com confiança; (iv) aprender a investigar, por si próprios, as ideias matemáticas; e (v) usar experiências e observações para formular conjecturas. (p 4)

Em virtude da teoria dos códigos corretores de erros estar inserida em grande parte dos recursos tecnológicos utilizados pelos alunos, optamos por trabalhar esse tema. Muitos de nós utilizamos recursos tecnológicos disponíveis na atualidade, sem darmos conta da matemática que existe por trás do bom funcionamento de cada um deles. Ao enviarmos uma mensagem no celular ou através de e-mail, por exemplo, o que nos dá garantia que o destinatário irá receber a mensagem tal qual a enviamos? O que garante a fidelidade entre a mensagem enviada e a recebida? Quem de nós ao digitar uma palavra errada em uma mensagem de celular não percebeu que o próprio equipamento sugere uma correção prévia, conforme a figura 1?

Figura 1: Telefone celular



Fonte: Foto retirada pelo autor

Encontrando nesses recursos tecnológicos utilizados pela sociedade atual a matemática necessária das matrizes, determinantes e polinômios aplicada nos códigos corretores de erros, vimos uma oportunidade útil de socializar esses conhecimentos e propiciar aos alunos uma forma diferenciada na abordagem dos assuntos estudados por eles.

Entre os conteúdos de matrizes, determinantes e polinômio, pertinentes ao currículo do ensino médio, e a teoria dos códigos corretores de erros, existe um elo que consiste no conhecimento das estruturas algébricas elementares. As estruturas algébricas elementares não pertencem ao rol de conteúdos presentes nos currículos de matemática da educação básica.

Nos livros didáticos do ensino fundamental, os conjuntos numéricos não são apresentados não como estruturas algébricas elementares, mas as suas propriedades, em geral, definem essas estruturas. Por exemplo, os livros do 7º ano do ensino fundamental apresentam o conjunto $\mathbb{Z}$ dos números inteiros, como sendo um conjunto no qual a adição está definida e goza das seguintes propriedades: comutatividade, associatividade, elemento neutro aditivo, elemento simétrico. A apresentação dessas propriedades no livro didático caracteriza o conjunto $\mathbb{Z}$ como sendo um grupo aditivo, ademais, por ser apresentada a propriedade comutativa, então, temos $\mathbb{Z}$ como um grupo abeliano. Ao introduzir a multiplicação no conjunto $\mathbb{Z}$ dos números inteiros, os livros didáticos apresentam as propriedades comutativa, associativa, elemento neutro multiplicativo e a distributividade em relação à adição, o que caracteriza $\mathbb{Z}$ como um anel ou, mais ainda, um anel comutativo com unidade. Posteriormente, mais precisamente quando se estudam equações em $\mathbb{Z}$, é apresentado aos alunos sentenças do tipo $3x = 0 \Rightarrow x = 0$, que, em outras palavras, significa que no conjunto $\mathbb{Z}$ não existem divisores próprios de zero, logo, $\mathbb{Z}$ é um domínio de integridade.

Quando o conjunto $\mathbb{Q}$ dos números racionais e $\mathbb{R}$ dos números reais são apresentados, além de serem mencionadas para $\mathbb{Q}$ e $\mathbb{R}$ todas as propriedades anteriormente enumeradas no conjunto $\mathbb{Z}$, é enunciado ainda que todo elemento não nulo desses conjuntos possui um inverso multiplicativo, o que define $\mathbb{Q}$ e $\mathbb{R}$ como corpos. Sendo assim, os conceitos de grupo, anéis, domínios de integridade e corpos, que constituem parte das estruturas algébricas elementares, vão sendo construídos implicitamente no aprendizado dos alunos. Portanto, não encaramos o "elo" das estruturas algébricas elementares como sendo um obstáculo para o acesso à teoria dos códigos corretores de erros e, em consequência disso, apresentamos esses conceitos neste trabalho.

2 MATRIZES

2.1 DEFINIÇÃO DE MATRIZES REAIS – ALGUNS CONCEITOS

Sendo $m, n \in \mathbb{N}$, definimos uma matriz real de ordem m por n como uma tabela formada por $m \cdot n$ elementos do conjunto $\mathbb{R}$ agrupados em m linhas e n colunas.

Ao elemento que ocupa a $i-$ésima linha e $j-$ésima coluna de uma matriz A, representamos por a_{ij}, com $1 \leq i \leq m$ e $1 \leq j \leq n$.

Cada elemento a_{ij} da matriz A é denominado *entrada da matriz*.

Exemplo:

$A = \begin{bmatrix} 2 & -1 & 0 \\ \sqrt{2} & \frac{1}{2} & -3 \end{bmatrix}$ é a representação de uma matriz de ordem 2×3. Observemos, por exemplo, que o elemento -3 ocupa a posição que corresponde à interseção da segunda linha com a terceira coluna, portanto $-3 = a_{23}$.

Uma matriz A de ordem m por n é genericamente representada por $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$ ou ainda $A = \left[a_{ij}\right]_{m \times n}$ ou, quando a ordem da matriz for conhecida, podemos representar apenas por $A = \left[a_{ij}\right]$.

À matriz cuja ordem seja $1 \times n$ denominamos *matriz linha* e à matriz cuja ordem seja $m \times 1$ denominamos *matriz coluna*.

Exemplos:

$A = \begin{bmatrix} -4 & \frac{\sqrt{3}}{2} & \frac{1}{9} \end{bmatrix}$ e $B = \begin{bmatrix} -\frac{3}{4} \\ 0 \\ \pi \end{bmatrix}$ são, respectivamente, *matriz linha* e *matriz coluna.*

À matriz $A = \left[a_{ij}\right]$ de ordem m por n, que possui $a_{ij} = 0$ para todo $i \in \{1, 2, \dots, m\}$ e todo $j \in \{1, 2, \dots, n\}$, denominamos $matriz\ nula$.

Dada uma matriz $A = \left[a_{ij}\right]$ de ordem m por n, definimos a *matriz oposta de A* como sendo a matriz $-A = \left[-a_{ij}\right]$, de mesma ordem de A.

Se em uma matriz $A = [a_{ij}]$ de ordem $m \times n$ tivermos $m = n$, então dizemos que A é uma matriz quadrada de ordem n.

Exemplo:

$A = \begin{bmatrix} 0 & -5 & \pi \\ -\frac{\sqrt{5}}{3} & \frac{1}{2} & 0 \\ e & -2 & 1 \end{bmatrix}$ é quadrada de ordem 3.

Em uma matriz quadrada $A = [a_{ij}]$, de ordem n, os elementos a_{ij}, com $i = j$ formam a *diagonal principal.*

À matriz $A = [a_{ij}]$, quadrada, de ordem n, onde $a_{ij} = 0$ quando $i \neq j$, denominamos *matriz diagonal*.

Exemplo:

$$A = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -\frac{2}{3} & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Uma matriz diagonal de ordem n, cujos elementos da diagonal principal forem todos iguais a 1 é denominada *matriz identidade de ordem n* e é representada por I_n.

Exemplo:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

À matriz $A = [a_{ij}]$, quadrada de ordem n, que possui os elementos $a_{ij} = 0$ quando $i < j$ (ou $i > j$) denominamos *matriz triangular inferior (ou matriz triangular superior).*

Exemplos:

$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -5 & 0 \\ 3 & \frac{1}{2} & -3 \end{bmatrix}$ e $B = \begin{bmatrix} -7 & 6 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{bmatrix}$ A e B são, respectivamente, matriz triangular inferior e matriz triangular superior, ambas de ordem 3.

O símbolo $\mathcal{M}(m, n)$ representará o conjunto de todas as matrizes de ordem m por n.

2.1.1 Igualdade de matrizes

Dadas duas matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$, pertencentes a $\mathcal{M}(m, n)$, ou seja, de mesma ordem, dizemos que A e B são iguais, ou ainda $A = B$, quando $a_{ij} = b_{ij}$ para todo $i \in \{1, 2, \dots, m\}$ e todo $j \in \{1, 2, \dots, n\}$.

2.2 OPERAÇÕES COM MATRIZES

2.2.1 Adição de matrizes

Definimos a operação de adição em $\mathcal{M}(m,n)$ como sendo uma função de $\mathcal{M}(m,n) \times \mathcal{M}(m,n)$ em $\mathcal{M}(m,n)$, que a cada par $(A,B) \in \mathcal{M}(m,n) \times \mathcal{M}(m,n)$ faz corresponder a matriz $A + B = C \in \mathcal{M}(m,n)$, de maneira que $a_{ij} + b_{ij} = c_{ij}$ para todo $i \in \{1,2,\dots,m\}$ e todo $j \in \{1,2,\dots,n\}$.

Exemplo:

Sejam as matrizes $A = \begin{bmatrix} -1 & 0 & 5 \\ 0 & 2 & -3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 4 & -3 \\ 2 & -1 & 4 \end{bmatrix}$ pertencentes a $\mathcal{M}(2,3)$, temos:

$$A + B = \begin{bmatrix} -1 & 0 & 5 \\ 0 & 2 & -3 \end{bmatrix} + \begin{bmatrix} 0 & 4 & -3 \\ 2 & -1 & 4 \end{bmatrix} = \begin{bmatrix} -1+0 & 0+4 & 5+(-3) \\ 0+2 & 2+(-1) & -3+4 \end{bmatrix} =$$

$$= \begin{bmatrix} -1 & 4 & 2 \\ 2 & 1 & 1 \end{bmatrix} = C \in \mathcal{M}(2,3)$$

Propriedades da adição de matrizes

Sejam A, B e C matrizes pertencentes a $\mathcal{M}(m,n)$, temos:

I) Propriedade associativa da adição: $A + (B + C) = (A + B) + C$

II) Propriedade comutativa da adição: $A + B = B + A$

III) Elemento Neutro da adição: $A + 0 = 0 + A = A$, onde 0 significa a matriz nula

IV) $A + (-A) = -A + A = 0$, onde $-A$ representa a matriz oposta de A.

Demonstrações:

I) Dadas $A = [a_{ij}]$, $B = [b_{ij}]$ e $C = [c_{ij}]$ matrizes pertencentes a $\mathcal{M}(m,n)$, temos:

$$A + (B + C) = [a_{ij}] + [b_{ij} + c_{ij}] = [a_{ij} + (b_{ij} + c_{ij})] = [(a_{ij} + b_{ij}) + c_{ij}] =$$

$$= [a_{ij} + b_{ij}] + [c_{ij}] = (A + B) + C$$ (utilizamos a associatividade da adição de números reais)

II) Dadas A e B matrizes pertencentes a $\mathcal{M}(m,n)$, temos:

$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$ (Utilizamos a comutatividade da adição de números reais)

III) Seja A uma matriz pertencente a $\mathcal{M}(m,n)$ e 0 a matriz nula de $\mathcal{M}(m,n)$, temos:

$A + 0 = [a_{ij}] + 0 = [a_{ij} + 0] = [a_{ij}] = A = [0 + a_{ij}] = 0 + [a_{ij}] = 0 + A$

IV) Seja A uma matriz pertencente a $\mathcal{M}(m,n)$ e $-A$ a sua matriz oposta. Temos:

$A + (-A) = [a_{ij}] + [-a_{ij}] = [a_{ij} + (-a_{ij})] = [a_{ij} - a_{ij}] = 0 = [-a_{ij} + a_{ij}] =$

$= [-a_{ij}] + [a_{ij}] = -A + A$

2.2.2 Multiplicação de um escalar real por uma matriz

Dada uma matriz $A = [a_{ij}]$ pertencente a $\mathcal{M}(m,n)$, definimos o produto da matriz A por um escalar $k \in \mathbb{R}$, como a matriz $kA = [ka_{ij}]$.

Exemplo:

Seja $A = \begin{bmatrix} -1 & 0 & 5 \\ 0 & 2 & -3 \end{bmatrix}$ e $k = -5$, temos:

$$-5A = -5.\begin{bmatrix} -1 & 0 & 5 \\ 0 & 2 & -3 \end{bmatrix} = \begin{bmatrix} -5.(-1) & -5.0 & -5.5 \\ -5.0 & -5.2 & -5(-3) \end{bmatrix} = \begin{bmatrix} 5 & 0 & -25 \\ 0 & -10 & 15 \end{bmatrix}$$

Propriedades da multiplicação de uma matriz por um escalar real

Sejam A e B matrizes pertencentes a $\mathcal{M}(m,n)$ e k_1 e k_2 escalares reais, temos:

I) $k_1.(A + B) = k_1.A + k_1.B$

II) $(k_1 + k_2).A = k_1.A + k_2.A$

III) $k_1.(k_2.A) = (k_1.k_2).A$

IV) $1A = A$

Demonstrações:

Sejam A e B matrizes pertencentes a $\mathcal{M}(m,n)$ e $k_1, k_2 \in \mathbb{R}$, temos:

I) $k_1 \cdot (A + B) = k_1 \cdot [a_{ij} + b_{ij}] = [k_1 \cdot (a_{ij} + b_{ij})] = [k_1 \cdot a_{ij} + k_1 \cdot b_{ij}] =$

$= [k_1 \cdot a_{ij}] + [k_1 \cdot b_{ij}] = k_1 \cdot [a_{ij}] + k_1 \cdot [b_{ij}] = k_1 \cdot A + k_1 \cdot B$

(utilizamos a distributividade da multiplicação em relação à adição de números reais)

II) $(k_1 + k_2).A = (k_1 + k_2).[a_{ij}] = [(k_1 + k_2).a_{ij}] = [k_1.a_{ij} + k_2.a_{ij}] =$

$= [k_1.a_{ij}] + [k_2.a_{ij}] = k_1 \cdot [a_{ij}] + k_2 \cdot [a_{ij}] = k_1 \cdot A + k_2 \cdot B$

(utilizamos a distributividade do produto em relação à adição de números reais)

III) $k_1.(k_2.A) = k_1 \cdot (k_2.[a_{ij}]) = k_1 \cdot [k_2 \cdot a_{ij}] = [k_1 \cdot (k_2 \cdot a_{ij})] = [(k_1 \cdot k_2).a_{ij}] =$

$= (k_1 \cdot k_2) \cdot [a_{ij}] = (k_1 \cdot k_2) \cdot A$ (utilizamos a associatividade da multiplicação de números reais)

IV) Sendo A uma matriz pertencente a $\mathcal{M}(m,n)$, como 1 é um escalar real, então o produto $1A$ está bem definido e $1A = 1 \cdot [a_{ij}] = [1 \cdot a_{ij}] = [a_{ij}] = A$

2.2.3 Multiplicação de matrizes

A multiplicação de matrizes acontece mediante a seguinte condição: para que exista a multiplicação entre duas matrizes A e B, é necessário que o número de colunas de A seja igual ao número de linhas de B, ou seja, $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{n\times p}$. Sendo C o produto $A \cdot B$, então a matriz C é de ordem m por p.

De acordo com a condição acima, temos que a multiplicação de matrizes quadradas de mesma ordem é sempre possível.

Passemos a definição formal da multiplicação de matrizes:

Sejam $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{n\times p}$ duas matrizes, definimos o produto $A \cdot B$ como sendo a matriz $C = [c_{ij}]_{m\times p}$ tal que $c_{ij} = \sum_{k=1}^{n}(a_{ik}.b_{kj})$ ou seja,

$c_{ij} = a_{i1}.b_{1j} + a_{i2}.b_{2j} + \cdots + a_{in}.b_{nj}$

Exemplo:

Sejam as matrizes $A = \begin{bmatrix} -1 & 1 & 3 \\ 2 & -2 & -4 \\ 5 & 0 & 0 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & -1 \\ -2 & 2 \\ 1 & 3 \end{bmatrix}$. Vemos que A é de ordem 3×3 e B de ordem 3×2, ou seja, o número de colunas da matriz A é igual ao número de linhas da matriz B, logo é possível o produto $A.B$

Seja $C = A.B$, temos:

$$C = \begin{bmatrix} -1 & 1 & 3 \\ 2 & -2 & -4 \\ 5 & 0 & 0 \end{bmatrix}.\begin{bmatrix} 0 & -1 \\ -2 & 2 \\ 1 & 3 \end{bmatrix} =$$

$$= \begin{bmatrix} -1.0 + 1.(-2) + 3.1 & -1.(-1) + 1.2 + 3.3 \\ 2.0 + (-2).(-2) + (-4).1 & 2.(-1) + (-2).2 + (-4).3 \\ 5.0 + 0.(-2) + 0.1 & 5.(-1) + 0.2 + 0.3 \end{bmatrix} =$$

$$= \begin{bmatrix} 0-2+3 & 1+2+9 \\ 0+4-4 & -2-4-12 \\ 0+0+0 & -5+0+0 \end{bmatrix} = \begin{bmatrix} 1 & 12 \\ 0 & -18 \\ 0 & -5 \end{bmatrix}$$

Propriedades da multiplicação de matrizes

Desde que as operações sejam possíveis, a multiplicação de matrizes goza das seguintes propriedades:

I) Distributividade à esquerda da multiplicação em relação à adição:

$A.(B+C) = A.B + A.C$

II) Distributividade à direita da multiplicação em relação à adição:

$(A+B).C = A.C + B.C$

III) Associatividade:

$(A.B).C = A.(B.C)$

IV) Considerando A uma matriz quadrada, temos $A.I = I.A = A$, onde I é o elemento neutro da multiplicação (matriz identidade).

Demonstrações:

I) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{n\times p}$ e $C = [c_{ij}]_{n\times p}$ matrizes quaisquer, temos:

$$A.(B+C) = \sum_{k=1}^{n} a_{ik} \cdot (b_{kj} + c_{kj}) = \sum_{k=1}^{n} (a_{ik} \cdot b_{kj} + a_{ik} \cdot c_{kj}) =$$

$$= \left(\sum_{k=1}^{n} a_{ik} \cdot b_{kj}\right) + \left(\sum_{k=1}^{n} a_{ik} \cdot c_{kj}\right) = A \cdot B + A \cdot C$$

II) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$ e $C = [c_{ij}]_{n\times p}$ matrizes quaisquer, temos:

$$(A+B) \cdot C = \sum_{k=1}^{n} (a_{kj} + b_{kj}) \cdot c_{ik} = \sum_{k=1}^{n} (a_{ik} \cdot c_{kj} + b_{ik} \cdot c_{kj}) =$$

$$= \left(\sum_{k=1}^{n} a_{ik} \cdot c_{kj}\right) + \left(\sum_{k=1}^{n} b_{ik} \cdot c_{kj}\right) = A \cdot C + B \cdot C$$

III) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{n\times p}$ e $C = [c_{ij}]_{p\times q}$ matrizes quaisquer, temos:

$$(A.B).C = \big((A.B).C\big)_{ij} = \sum_{k=1}^{p} (A \cdot B)_{ik} \cdot c_{kj} = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} a_{il} \cdot b_{lk}\right) \cdot c_{kj} =$$

$$= \sum_{l=1}^{n} a_{il} \cdot \left(\sum_{k=1}^{p} b_{lk} \cdot c_{kj}\right) = \sum_{l=1}^{n} a_{il} \cdot (B \cdot C)_{lj} = \big(A \cdot (B \cdot C)\big)_{ij} = A \cdot (B \cdot C)$$

IV) Seja $A = [a_{ij}]_n$ e I_n ($representaremos\ apenas\ por\ I$), temos:

$$A.I = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} . \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}.1 + a_{12}.0 + \cdots + a_{1n}.0 & \cdots & a_{11}.0 + a_{12}.0 + \cdots + a_{1n}.1 \\ \vdots & \ddots & \vdots \\ a_{n1}.1 + a_{n2}.0 + \cdots + a_{nn}.0 & \cdots & a_{n1}.0 + a_{n2}.0 + \cdots + a_{nn}.1 \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = A$$

De maneira análoga:

$$I.A = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}.\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} 1.a_{11} + 0.a_{21} + \cdots + a_{n1}.0 & \cdots & 1.a_{1n} + 0.a_{2n} + \cdots + 0.a_{nn} \\ \vdots & \ddots & \vdots \\ 0.a_{11} + 0.a_{21} + \cdots + 1.a_{n1}. & \cdots & 0.a_{1n} + 0.a_{2n} + \cdots + 1.a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = A$$

Portanto, $A.I = I.A = A$.

A multiplicação de matrizes, em geral não goza da propriedade comutativa. Ilustramos essa afirmação com um contra exemplo.

Sejam $A = \left[a_{ij}\right]_2$ e $B = \left[b_{ij}\right]_2$ tais que $A = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$, temos:

$$A.B = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}.\begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 0.2 + 1.1 & 0.1 + 1.5 \\ 2.2 + 3.1 & 2.1 + 3.5 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 7 & 17 \end{bmatrix}$$

e

$$B.A = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}.\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2.0 + 1.2 & 2.1 + 1.3 \\ 1.0 + 5.2 & 1.1 + 5.3 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 10 & 16 \end{bmatrix}$$

Portanto, temos $A.B \neq B.A$

2.2.4 Potenciação de matrizes

Definimos a potenciação de matrizes da seguinte forma:

Dada uma matriz $A = \left[a_{ij}\right]_n$, definimos, $A^0 = I_n$, $A^1 = A$ e $A^m = \underbrace{A.A.A.\cdots.A}_{m\ fatores}$

2.3 Transposta de uma matriz

Dada uma matriz $A = \left[a_{ij}\right]_{m\times n}$, definimos a matriz transposta de A como sendo a matriz $A^t = \left[a'_{ij}\right]_{n\times m}$ onde $a'_{ij} = a_{ji}$ para todo $i \in \{1, 2, \ldots, n\}$ e todo $j \in \{1, 2, \ldots, m\}$.

Exemplo:

Seja $A = \begin{bmatrix} 0 & -1 \\ -2 & 2 \\ 1 & 3 \end{bmatrix}$, por definição, a matriz transposta de A é $A^t = \begin{bmatrix} 0 & -2 & 1 \\ -1 & 2 & 3 \end{bmatrix}$.

Quando $A = A^t$, dizemos que A é uma *matriz simétrica* e quando $A = -A^t$, dizemos que A é uma *matriz antissimétrica*.

Propriedades da transposição de matrizes

I) $(A^t)^t = A$

II) $(A + B)^t = A^t + B^t$

III) $(k \cdot A)^t = k \cdot A^t, \quad \forall k \in \mathbb{R}$

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

Demonstrações:

I) Seja $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$, temos $(A^t)^t = \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}^t \right)^t = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}^t =$

$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = A$

II) Sejam $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ e $B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$, temos $(A + B)^t =$

$\left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \right)^t = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{11} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}^t =$

$= \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{m1} + b_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} + b_{11} & \cdots & a_{mn} + b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & \ddots & \vdots \\ b_{11} & \cdots & b_{mn} \end{bmatrix} =$

$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}^t + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}^t = A^t + B^t$

III) Sejam $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ e $k \in \mathbb{R}$, temos $(k \cdot A)^t = \left(k \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \right)^t =$

$= \begin{bmatrix} k \cdot a_{11} & \cdots & k \cdot a_{1n} \\ \vdots & \ddots & \vdots \\ k \cdot a_{m1} & \cdots & k \cdot a_{mn} \end{bmatrix}^t = \begin{bmatrix} k \cdot a_{11} & \cdots & k \cdot a_{m1} \\ \vdots & \ddots & \vdots \\ k \cdot a_{1n} & \cdots & k \cdot a_{mn} \end{bmatrix} = k \cdot \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} =$

$= k \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}^t = k \cdot A^t$

IV) Sejam $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ e $B = \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix}$, como o número de colunas de A é igual ao número de linhas de B, então existe o produto $A \cdot B$. Assim, $(A \cdot B)^t =$

$$= \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} . \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} \right)^{\boldsymbol{t}} =$$

$$= \begin{bmatrix} a_{11} \cdot b_{11} + \cdots + a_{1n} \cdot b_{n1} & \cdots & a_{11} \cdot b_{1p} + \cdots + a_{1n} \cdot b_{np} \\ \vdots & \ddots & \vdots \\ a_{m1} \cdot b_{11} + \cdots + a_{mn} \cdot b_{n1} & \cdots & a_{m1} \cdot b_{1p} + \cdots + a_{mn} \cdot b_{np} \end{bmatrix}^{t} =$$

$$= \begin{bmatrix} a_{11} \cdot b_{11} + \cdots + a_{1n} \cdot b_{n1} & \cdots & a_{m1} \cdot b_{11} + \cdots + a_{mn} \cdot b_{n1} \\ \vdots & \ddots & \vdots \\ a_{11} \cdot b_{1p} + \cdots + a_{1n} \cdot b_{np} & \cdots & a_{m1} \cdot b_{1p} + \cdots + a_{mn} \cdot b_{np} \end{bmatrix} =$$

$$= \begin{bmatrix} b_{11} \cdot a_{11} + \cdots + b_{n1} \cdot a_{1n} & \cdots & b_{11} \cdot a_{m1} + \cdots + b_{n1} . a_{mn} \\ \vdots & \ddots & \vdots \\ b_{1p} \cdot a_{11} + \cdots + b_{np} . a_{1n} & \cdots & b_{1p} \cdot a_{m1} + \cdots + b_{np} . a_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1p} & \cdots & b_{np} \end{bmatrix} . \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix}^{t} . \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}^{t} = B^t . A^t$$

2.4 Inversa de uma matriz

Seja A uma matriz quadrada de ordem n. Uma matriz B de ordem n é denominada a *inversa da matriz A* se $A.B = B.A = I_n$.

Exemplo:

Sejam as matrizes $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$, temos:

$$A.B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} . \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1.(-3) + 2.2 & 1.2 + 2.(-1) \\ 2.(-3) + 3.2 & 2.2 + 3.(-1) \end{bmatrix} = \begin{bmatrix} -3+4 & 2-2 \\ -6+6 & 4-3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

e

$$B.A = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} . \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} (-3).1 + 2.2 & -3.2 + 2.3 \\ 2.1 + (-1).2 & 2.2 + (-1).3) \end{bmatrix} = \begin{bmatrix} -3+4 & -6+6 \\ 2-2 & 4-3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Portanto $A.B = B.A = I_2$, o que implica que a matriz B é a inversa da matriz A.

Teorema 2.1: Se A é uma matriz invertível, então a sua inversa é única.

Demonstração: Seja A uma matriz quadrada de ordem n. Suponhamos que as matrizes B e B', ambas de ordem n, sejam matrizes inversas da matriz A.

Utilizando o produto pela matriz identidade, a definição de matriz inversa e a propriedade associativa da multiplicação de matrizes, temos:

$$B' = B'.I_n = B'.(A.B) = (B'.A).B = I_n.B = B$$

Devido à unicidade da inversa de uma matriz A, representaremo-na por A^{-1}.

Teorema 2.2: Se A é uma matriz invertível, então a sua inversa A^{-1} também é invertível e $(A^{-1})^{-1} = A$.

Demonstração: Seja A uma matriz quadrada de ordem n, invertível, então existe uma matriz quadrada A^{-1} de ordem n tal que $A.A^{-1} = I_n$.

Utilizando o produto pela matriz identidade, a definição de matriz inversa e a propriedade associativa da multiplicação de matrizes, temos:

$$(A^{-1})^{-1} = (A^{-1})^{-1}.I_n = (A^{-1})^{-1}.(A^{-1}.A) = [(A^{-1})^{-1}.A^{-1}].A = I_n.A = A$$

Portanto A^{-1} é invertível e sua inversa é A.

Teorema 2.3: Sejam A e B matrizes quadradas de ordem n e invertíveis, então $A.B$ também é invertível e $(A.B)^{-1} = B^{-1}.A^{-1}$.

Demonstração: Se A e B são matrizes quadradas de ordem n e invertíveis então existem A^{-1} e B^{-1} quadradas de ordem n, tais que $A.A^{-1} = I_n$ e $B.B^{-1} = I_n$

Assim, temos:

$$(A.B).(B^{-1}.A^{-1}) = A.(B.B^{-1}).A^{-1} = A.I_n.A^{-1} = A.A^{-1} = I_n$$

e

$$(B^{-1}.A^{-1}).(A.B) = B^{-1}.(A^{-1}.A).B = B^{-1}.I_n.B = B^{-1}.B = I_n$$

Portanto, $(A.B).(B^{-1}.A^{-1}) = (B^{-1}.A^{-1}).(A.B) = I_n$, o que implica que $A.B$ é invertível e sua inversa é $B^{-1}.A^{-1}$.

Nem todas as matrizes possuem inversa. As condições para que uma matriz seja invertível serão abordados mais a frente.

2.5 Transformações elementares de matrizes

Seja A uma matriz pertencente a $\mathcal{M}(m,n)$. Para cada $i \in \{1,2,\dots,m\}$, representaremos por L_i a $i-$ésima linha da matriz A.

Definimos as transformações elementares nas linhas da matriz A, da seguinte forma:

I) Permutação entre as linhas L_i e L_j e representamos por $L_i \leftrightarrow L_j$

II) Multiplicação de uma linha L_i por um escalar real $k \neq 0$ e representamos por $L_i \rightarrow k.L_i$

III) Substituição de uma linha, digamos L_i, pela adição da linha L_i com o produto $k.L_j$ de um escalar k, não nulo, pelos elementos da linha L_j, com $i \neq j$ e representamos por $L_i \rightarrow L_i + k.L_j$

Vejamos um exemplo da aplicação de algumas transformações elementares nas linhas de uma matriz $A = \begin{bmatrix} -1 & 2 \\ 3 & -2 \\ 0 & 4 \end{bmatrix}$:

$$\begin{bmatrix} -1 & 2 \\ 3 & -2 \\ 0 & 4 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{bmatrix} 0 & 4 \\ 3 & -2 \\ -1 & 2 \end{bmatrix} \; L_3 \xrightarrow{} -2.L_3 \; \begin{bmatrix} 0 & 4 \\ 3 & -2 \\ 2 & -4 \end{bmatrix} \; \xrightarrow{L_2 \to L_2 + \frac{1}{2}.L_1} \; \begin{bmatrix} 0 & 4 \\ 3 & 0 \\ 2 & -4 \end{bmatrix}$$

Dizemos que as matrizes $\begin{bmatrix} -1 & 2 \\ 3 & -2 \\ 0 & 4 \end{bmatrix}$, $\begin{bmatrix} 0 & 4 \\ 3 & -2 \\ -1 & 2 \end{bmatrix}$, $\begin{bmatrix} 0 & 4 \\ 3 & -2 \\ 2 & -4 \end{bmatrix}$ e $\begin{bmatrix} 0 & 4 \\ 3 & 0 \\ 2 & -4 \end{bmatrix}$ são matrizes equivalentes por linhas.

Definição: Duas matrizes A e B são equivalentes por linhas se B puder ser obtida da matriz A através de um número finito de transformações elementares sobre as linhas de A ou se A puder ser obtida de B através de um número finito de transformações elementares sobre as linhas de B.

2.5.1 Matriz elementar

Denominamos matriz elementar a toda matriz de ordem n obtida através da aplicação de uma transformação elementar sobre a matriz I_n.

Exemplo:

A matriz $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ é uma matriz elementar, pois é obtida através da transformação elementar e correspondente a permuta $L_2 \leftrightarrow L_3$ em I_3 portanto, $e(I_3) = E$.

Teorema 2.4: Seja e uma transformação elementar e E uma matriz elementar quadrada de ordem n tal que $e(I_n) = E$. Se A é uma matriz quadrada de ordem n, então $e(A) = E.A$

Demonstração: Utilizaremos na demonstração apenas a transformação elementar *permutação entre as linhas* L_i *e* L_j , sendo que para as outras transformações as demonstrações são de maneira análoga.

Seja $I_n = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \to L_i \\ \to L_j \\ \\ \\ \end{matrix}$

Seja e a transformação elementar que permuta as linhas L_i *e* L_j. Assim,

$e(I_n) = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} = E$. Seja A a matriz de ordem n a seguir:

$A = \begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ii} & a_{ij} & \cdots & a_{in} \\ a_{j1} & \cdots & a_{ji} & a_{jj} & \cdots & a_{jn} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{nj} & \cdots & a_{nn} \end{bmatrix} \begin{matrix} \\ \\ \rightarrow L_i \\ \rightarrow L_j \\ \\ \\ \end{matrix}$ Fazendo $e(A)$, temos:

$$e(A) = \begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{j1} & \cdots & a_{ji} & a_{jj} & \cdots & a_{jn} \\ a_{i1} & \cdots & a_{ii} & a_{ij} & \cdots & a_{in} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{nj} & \cdots & a_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} 1.a_{11} + \cdots + 0.a_{j1} + 0.a_{i1} + \cdots + 0.a_{n1} & \cdots & 1.a_{1i} + \cdots + 0.a_{ji} + 0.a_{ii} + \cdots + 0.a_{ni} \\ \vdots & \cdots & \vdots \\ 0.a_{11} + \cdots + 1.a_{j1} + 0.a_{i1} + \cdots + 0.a_{n1} & \cdots & 0.a_{1i} + \cdots + 1.a_{ji} + 0.a_{ii} + \cdots + 0.a_{ni} \\ 0.a_{11} + \cdots + 0.a_{j1} + 1.a_{i1} + \cdots + 0.a_{n1} & \cdots & 0.a_{1i} + \cdots + 0.a_{ji} + 1.a_{ii} + \cdots + 0.a_{ni} \\ \vdots & \cdots & \vdots \\ 0.a_{11} + \cdots + 0.a_{j1} + 0.a_{i1} + \cdots + 1.a_{n1} & \cdots & 0.a_{1i} + \cdots + 0.a_{ji} + 0.a_{ii} + \cdots + 1.a_{ni} \end{bmatrix} \cdots$$

$$\cdots \begin{bmatrix} 1.a_{1j} + \cdots + 0.a_{jj} + 0.a_{ij} + \cdots + 0.a_{nj} & \cdots & 1.a_{1n} + \cdots + 0.a_{jn} + 0.a_{in} + \cdots + 0.a_{nn} \\ \vdots & \cdots & \vdots \\ 0.a_{1j} + \cdots + 1.a_{jj} + 0.a_{ij} + \cdots + 0.a_{nj} & \cdots & 0.a_{1n} + \cdots + 1.a_{jn} + 0.a_{in} + \cdots + 0.a_{nn} \\ 0.a_{1j} + \cdots + 0.a_{jj} + 1.a_{ij} + \cdots + 0.a_{nj} & \cdots & 0.a_{1n} + \cdots + 0.a_{jn} + 1.a_{in} + \cdots + 0.a_{nn} \\ \vdots & \cdots & \vdots \\ 0.a_{1j} + \cdots + 0.a_{jj} + 0.a_{ij} + \cdots + 1.a_{nj} & \cdots & 0.a_{1n} + \cdots + 0.a_{jn} + 0.a_{in} + \cdots + 1.a_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ii} & a_{ij} & \cdots & a_{in} \\ a_{j1} & \cdots & a_{ji} & a_{jj} & \cdots & a_{jn} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{nj} & \cdots & a_{nn} \end{bmatrix} = E.A$$

Teorema 2.5: Se A e B são matrizes quadradas de ordem n, então a matriz A é equivalente por linhas à matriz B se, e somente se, existem matrizes elementares $E_1, E_2, E_3, \ldots, E_k$ quadradas de ordem n tais que $E_k . E_{k-1} . \cdots . E_2 . E_1 . A = B$.

Demonstração: Por definição, para que uma matriz A de ordem n seja equivalente por linhas a uma matriz B de mesma ordem, devem existir transformações elementares $e_1, e_2, e_3, \ldots, e_k$ tal que $e_k\left(\ldots\left(e_2\left(e_1(A)\right)\right)\right) = B$. Pelo teorema 2.4, $e_k\left(\ldots\left(e_2\left(e_1(A)\right)\right)\right) = e_k\left(\ldots\left(e_2(E_1.A)\right)\right) =$

$= e_k\big(\ldots(E_2.E_1.A)\big) = E_k.\cdots.E_2.E_1.A = B$ com cada $E_i = e_i(I_n)$, para todo $i \in \{1, 2, \ldots, k\}$.

Teorema 2.6: Toda matriz elementar é invertível e sua inversa também é uma matriz elementar.

Demonstração: Consideremos a transformação elementar e que transforma I_n na matriz elementar E, ou seja, $e(I_n) = E$. Consideremos e^{-1} a transformação elementar inversa de e, ou seja, se e for a permutação das linhas L_i e L_j da matriz I_n, então e^{-1} será a transformação permutação das linhas L_i e L_j da matriz E; se e for a multiplicação de uma linha L_i da matriz I_n por um escalar $k \neq 0$, então e^{-1} será a multiplicação da linha L_i da matriz E pelo escalar $\frac{1}{k}$ e se e for a substituição de uma linha L_i da matriz I_n pela adição da linha L_i com o produto de um escalar $k \neq 0$ por uma linha L_j, então e^{-1} será a substituição da linha L_i da matriz E pela adição da linha L_i da matriz E com o produto do escalar $-k$ pela linha L_j da matriz E.

Assim, fica evidente que $e^{-1}(E) = I_n$. Se aplicarmos a transformação e^{-1} em I_n, temos uma matriz elementar F, ou seja, $e^{-1}(I_n) = F$ e, pelo teorema 2.4, teremos $F.E = I_n$, ou seja, F é a matriz inversa da matriz E, portanto $F = E^{-1}$, concluindo então que se E é uma matriz elementar, então é invertível. Como $F = E^{-1}$ é obtida através de transformações elementares na matriz I_n, então $F = E^{-1}$ é também uma matriz elementar.

2.5.2 Matriz escalonada

Definição: Uma matriz A de ordem $m \times n$ é apresentada na forma escalonada se:

I) O primeiro elemento não nulo em cada linha da matriz A é igual a 1;

II) Cada coluna da matriz A que contém o primeiro elemento não nulo de alguma linha, possui todos os outros elementos iguais a zero;

III) Todas as linhas nulas se encontram abaixo de todas as linhas não nulas;

IV) Se as linhas não nulas da matriz A forem $L_1, L_2, L_3, \ldots, L_k$, sendo a_{1j} o primeiro elemento não nulo da linha L_1, então os elementos não nulos das linhas $L_2, L_3, \ldots, L_k$ ocuparão, respectivamente, as posições $a_{2j'}, a_{3j''}, \ldots, a_{kj'\ldots'}$ com $j < j' < j'' < \cdots < j'^{\ldots'}$.

Exemplo: Seja A a matriz de ordem 3×4 a seguir:

$A = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{bmatrix}$. A matriz apresentada se encontra na forma escalonada, pois satisfaz as condições I, II, III e IV da definição acima.

Teorema 2.7: Toda matriz é equivalente a uma matriz na forma escalonada.

Demonstração: Seja A uma matriz quadrada de ordem $m \times n$, se a primeira linha for nula então a condição (I) é satisfeita nessa linha. Se por acaso a primeira linha possuir algum elemento diferente de zero, por exemplo a_{1j}, então através da transformação elementar de multiplicar por escalar, multiplicamos a primeira linha por $\frac{1}{a_{1j}}$, satisfazendo com isso a condição (I). Para cada linha a partir da segunda, somemos $-a_{ij}, i \neq 1$ vezes a primeira linha com a i-ésima linha, assim, obtemos uma matriz cujo primeiro elemento não nulo da primeira linha é 1 e ocorre na j-ésima coluna, ademais, todos os outros elementos da j-ésima coluna são iguais a zero. Considerando a segunda linha da matriz A, se a mesma for nula, não há nada o que fazer, caso exista algum elemento diferente de zero, procedemos de forma similar ao realizado na primeira linha. Como o número de linhas da matriz é limitado, no caso m, repetindo o processo acima descrito, ao chegarmos à m-ésima linha, teremos satisfeito as condições (I) e (II). As condições (III) e (IV) poderão ser satisfeitas de maneira bastante simples através de permutações entre as linhas da matriz. Desse modo, obtemos uma matriz B na forma escalonada, equivalente por linhas à matriz A.

Teorema 2.8: Uma matriz A, quadrada de ordem n que possui uma linha nula não é invertível.

Demonstração:

Suponha $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \rightarrow L_i$

Se A for invertível, então deve existir uma matriz B, quadrada de ordem n tal que $A.B = I_n$.

Suponhamos $B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ij} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nn} \end{bmatrix}$ notemos que o produto $A.B$ terá a í-ésima linha nula, pois a i-ésima linha será determinada por:

$L_i = [0.b_{11} + \cdots + 0.b_{n1} \quad 0.b_{12} + \cdots + 0.b_{n2} \quad \cdots \quad 0.b_{1j} + \cdots + 0.b_{nj} \quad \cdots \quad 0.b_{1n} + \cdots + 0.b_{nn}]$.

E, portanto, $L_i = [0 \quad 0 \quad \cdots \quad 0 \quad \cdots \quad 0]$, fazendo com que $A.B \neq I_n$, para todo B.

Logo, A não é invertível se possuir uma linha nula.

Os teoremas vistos até agora nos dão embasamento para obter dois resultados muito importantes acerca de matrizes:

Teorema 2.9: Uma matriz A de ordem n é invertível se, e somente se for equivalente por linhas à matriz identidade.

Demonstração:

$(\Rightarrow)$ Suponhamos que A é uma matriz invertível de ordem n. Pelo teorema 2.8, A não possui linhas nulas, além disso, pelo teorema 2.7, A é equivalente por linhas a uma matriz na forma escalonada. Portanto A é equivalente por linhas a I_n.

$(\Leftarrow)$ Seja A uma matriz quadrada de ordem n, equivalente por linhas a matriz I_n. Pelo teorema 2.5, existem $E_1, E_2, E_3, \dots, E_k$ de modo que $E_k . E_{k-1} . \cdots . E_2 . E_1 . A = I_n$. Pelo teorema 2.6, temos que $E_1, E_2, E_3, \dots, E_k$ são todas invertíveis, por serem matrizes elementares, então existem $E_1^{-1}, E_2^{-1}, E_3^{-1}, \dots, E_k^{-1}$, de modo que $E_i^{-1} . E_i = I_n$ para todo $i \in \{1, 2, 3, \dots, k\}$. Assim, multiplicando à esquerda ambos os membros da igualdade $E_1, E_2, E_3, \dots, E_k . A = I_n$ por $E_k^{-1} . \dots . E_3^{-1} . E_2^{-1} . E_1^{-1}$, temos: $(E_k^{-1} . \dots . E_3^{-1} . E_2^{-1} . E_1^{-1}) . E_1, E_2, E_3, \dots, E_k . A = (E_k^{-1} . \dots . E_3^{-1} . E_2^{-1} . E_1^{-1}) . I_n$

E, utilizando a propriedade associativa do produto de matrizes, temos: $A = E_k^{-1} . \dots . E_3^{-1} . E_2^{-1} . E_1^{-1}$ e, pelo teorema 2.3, o produto de matrizes invertíveis é invertível, portanto A é uma matriz invertível.

Teorema 2.10: Se A é uma matriz invertível de ordem n e uma sequência de transformações elementares sobre as linhas de A reduz A à matriz I_n, então esta mesma sequência de transformações elementares aplicadas às linhas de I_n produzirá a matriz A^{-1}.

Demonstração: Se A é invertível, então pelo teorema 2.9, A é equivalente por linhas a matriz I_n e, pelo teorema 2.5, existem $E_1, E_2, E_3, \dots, E_k$ de modo que $E_k . E_{k-1} . \cdots . E_2 . E_1 . A = I_n$. Como por hipótese A é invertível, então existe a matriz A^{-1}. Multiplicando à direita a igualdade $E_k . E_{k-1} . \cdots . E_2 . E_1 . A = I_n$ por A^{-1}, temos $E_k . E_{k-1} . \cdots . E_2 . E_1 . A . A^{-1} = I_n . A^{-1}$, de onde obtemos $E_k . E_{k-1} . \cdots . E_2 . E_1 . I_n = A^{-1}$.

Os teoremas 2.9 e 2.10 constituem um instrumento muito importante e eficiente na determinação da invertibilidade de uma matriz e o cálculo da matriz inversa, popularmente conhecido como *método de Gauss-Jordan*.

Vejamos um exemplo:

Seja $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 2 \\ 1 & 1 & 0 \end{bmatrix}$. Apliquemos simultaneamente as transformações elementares nas linhas da matriz A e da matriz I_3 de modo a reduzir a matriz A à matriz I_3:

$$[A|I_3] = \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & -3 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] L_3 \xrightarrow{} L_1 - L_3 \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & -3 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{array}\right] L_1 \xrightarrow{} L_1 - 2L_3$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & -3 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{array}\right] L_2 \xrightarrow{} L_2 + 3L_1 \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & 2 & -3 & 1 & 6 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{array}\right] L_2 \xrightarrow{} \frac{1}{2}L_1$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & -\frac{3}{2} & \frac{1}{2} & 3 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{array}\right] L_2 \rightleftarrows L_3 \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & -1 & 0 & 2 \\ 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -\frac{3}{2} & \frac{1}{2} & 3 \end{array}\right] L_1 \rightleftarrows L_2$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & -\frac{3}{2} & \frac{1}{2} & 3 \end{array}\right] = [I_3|A^{-1}]$$

Portanto $A^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 2 \\ -\frac{3}{2} & \frac{1}{2} & 3 \end{bmatrix}$.

Isto pode ser facilmente verificado fazendo $A.A^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 2 \\ 1 & 1 & 0 \end{bmatrix}.\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 2 \\ -\frac{3}{2} & \frac{1}{2} & 3 \end{bmatrix} =$

$$= \begin{bmatrix} 2.1 + 1.(-1) + 0.\left(-\frac{3}{2}\right) & 2.0 + 1.0 + 0.\frac{1}{2} & 2.(-1) + 1.2 + 0.3 \\ 0.1 + (-3).(-1) + 2.\left(-\frac{3}{2}\right) & 0.0 + (-3).0 + 2.\frac{1}{2} & 0.(-1) + (-3).2 + 2.3 \\ 1.1 + 1.(-1) + 0.\left(-\frac{3}{2}\right) & 1.0 + 1.0 + 0.\frac{1}{2} & 1.(-1) + 1.2 + 0.3 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 - 1 + 0 & 0 + 0 + 0 & -2 + 2 + 0 \\ 0 + 3 - 3 & 0 + 0 + 1 & 0 - 6 + 6 \\ 1 - 1 + 0 & 0 + 0 + 0 & -1 + 2 + 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$. De maneira análoga, temos

$$A^{-1}.A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 2 \\ -\frac{3}{2} & \frac{1}{2} & 3 \end{bmatrix}.\begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 2 \\ 1 & 1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1.2 + 0.0 + (-1).1 & 1.1 + 0.(-3) + (-1).1 & 1.0 + 0.2 + (-1).0 \\ -1.2 + 0.0 + 2.1 & -1.1 + 0.(-3) + 2.1 & -1.0 + 0.2 + 2.0 \\ -\frac{3}{2}.2 + \frac{1}{2}.0 + 3.1 & -\frac{3}{2}.1 + \frac{1}{2}.(-3) + 3.1 & -\frac{3}{2}.0 + \frac{1}{2}.2 + 3.0 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 + 0 - 1 & 1 + 0 - 1 & 0 + 0 + 0 \\ -2 + 0 + 2 & -1 + 0 + 2 & 0 + 0 + 0 \\ -3 + 0 + 3 & -\frac{3}{2} - \frac{3}{2} + 3 & 0 + 1 + 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

3 DETERMINANTES

Consideremos $n \geq 1$ pertencente ao conjunto dos números naturais. Seja $X_n = \{1, 2, 3, \dots, n\}$. Enunciamos que toda função bijetiva $f: X_n \to X_n$ é uma permutação do conjunto X_n.

Vamos representar uma permutação f de X_n em X_n por

$$f = \begin{bmatrix} 1 & 2 & 3 & \cdots & i & \cdots & j & \cdots & n \\ f(1) & f(2) & f(3) & \cdots & f(i) & \cdots & f(j) & \cdots & f(n) \end{bmatrix}$$

Exemplos:

a) Quando $n = 1$, temos $X_1 = \{1\}$ e temos uma possível bijeção de $X_1 \to X_1$, a saber, $f_{identidade} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

b) Quando $n = 2$, temos $X_2 = \{1, 2\}$ e temos duas possíveis bijeções de $X_2 \to X_2$, a saber, $f_{identidade} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ e $f = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

c) Quando $n = 3$, temos $X_3 = \{1, 2, 3\}$ e temos $3! = 6$ possíveis bijeções de $X_3 \to X_3$, a saber, $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$ e $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

Denominaremos K ao conjunto formado pelos pares ordenados (i, j), com $1 \leq i \leq j \leq n$, nos quais $f(i) > f(j)$ e $n(K)$ ao número de elementos de K. Denominaremos ainda por $sng(f)$ ao sinal da permutação, da seguinte maneira:

$sng(f) = 1, se\ n(K)$ é par

$sng(f) = -1, se\ n(K)$ é $impar$.

Exemplos:

a) Consideremos $f = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$, os pares (i, j), com $1 \leq i \leq j \leq n$ são $(1,2), (1,3)$ e $(2,3)$, notemos que $f(1) = 1 < f(2) = 3; f(1) = 1 < f(3) = 2$ e $f(2) = 3 > f(3) = 2$, portanto $K = \{(2,3)\}$, o que implica que $n(K) = 1$, que é impar, portanto $sng(f) = -1$.

b) Consideremos $f = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$, os pares (i, j), com $1 \leq i \leq j \leq n$ são $(1,2), (1,3)$ e $(2,3)$, notemos que $f(1) = 3 > f(2) = 1; f(1) = 3 > f(3) = 2$ e $f(2) = 1 <$

$f(3) = 2$, portanto $K = \{(1,2),(1,3)\}$, o que implica que $n(K) = 2$, que é par, portanto $sng(f) = 1$.

De acordo com o sinal, classificaremos uma permutação como par, se $sng(f) = 1$, ou ímpar, se $sng(f) = -1$.

Consideremos $A = [a_{ij}]_n$ uma matriz real. Consideremos também o produto $sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$, com f sendo uma permutação do conjunto X_n. Notemos que nesse produto aparecem, como fatores, somente um elemento de cada linha da matriz A, pois os índices correspondentes às linhas variam de 1 até n, sem repetição; e aparece também, somente um elemento de cada coluna da matriz A, uma vez que os índices correspondentes às colunas não se repetem pois f é bijetiva. Notemos ainda que temos $n!$ possíveis permutações em X_n, portanto $n!$ produtos $sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$.

Definiremos o *determinante* da matriz A, como sendo a soma das $n!$ parcelas $sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$, ou ainda:

$$\det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$$

Notemos que, considerando $\mathcal{M}_n$ como sendo o conjunto de todas as matrizes reais quadradas de ordem n, temos que $\det(A)$ é uma função de $\mathcal{M}_n$ em $\mathbb{R}$, que a cada matriz $A \in \mathcal{M}_n$, faz corresponder um escalar real k tal que $\det(A) = k$. Tal função não é bijetiva, pois embora seja sobrejetiva não é injetiva.

Se $A = [a_{11}]$, temos $\det(A) = a_{11}$.

Se $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, temos as seguintes permutações:

$f_1 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$, o que implica que $K_1 = \emptyset$, ou seja, $n(K_1) = 0$, que é par, portanto $sng(f_1) = 1$.

$f_2 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, o que implica que $K_2 = \{(1,2)\}$, ou seja, $n(K_2) = 1$, que é ímpar, portanto $sng(f_2) = -1$.

Logo, $\det(A) = a_{11}.a_{22} - a_{12}.a_{21}$

Se $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, temos as seguintes permutações:

$f_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, o que implica que $K_1 = \emptyset$, ou seja, $n(K_1) = 0$, que é par, portanto $sng(f_1) = 1$.

$f_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$, o que implica que $K_2 = \{(1,3),(2,3)\}$, ou seja, $n(K_2) = 2$, que é par, portanto $sng(f_2) = 1$.

$f_3 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$, o que implica que $K_3 = \{(1,2),(1,3)\}$, ou seja, $n(K_3) = 2$, que é par, portanto $sng(f_3) = 1$.

$f_4 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$, o que implica que $K_4 = \{(2,3)\}$, ou seja, $n(K_4) = 1$, que é ímpar, portanto $sng(f_4) = -1$.

$f_5 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$, o que implica que $K_5 = \{(1,2),(1,3),(2,3)\}$, ou seja, $n(K_5) = 3$, que é ímpar, portanto $sng(f_5) = -1$.

$f_6 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$, o que implica que $K_6 = \{(1,2)\}$, ou seja, $n(K_6) = 1$, que é ímpar, portanto $sng(f_4) = -1$.

Logo:

$$\det(A) = a_{11}.a_{22}.a_{33} + a_{12}.a_{23}.a_{31} + a_{13}.a_{21}.a_{32} - a_{11}.a_{23}.a_{32} - a_{13}.a_{22}.a_{31} - a_{12}.a_{21}.a_{33}$$

Vejamos alguns exemplos:

a) Seja $A = \begin{bmatrix} 2 & 5 \\ -3 & -4 \end{bmatrix}$, uma matriz de ordem 2.

Por definição, $\det(A) = 2.(-4) - 5.(-3)$, portanto, $\det(A) = 7$

b) Seja $B = \begin{bmatrix} -2 & 1 & 0 \\ -1 & 2 & 3 \\ 0 & -2 & 5 \end{bmatrix}$, uma matriz de ordem 3.

Por definição,

$\det(A) = -2.2.5 + 1.3.0 + 0.(-1).(-2) - (-2).3.(-2) - 0.2.0 - 1.(-1).5$

$\det(A) = -20 + 0 + 0 - 12 - 0 + 5$

$\det(A) = -27$

3.1 PROPRIEADES DOS DETERMINANTES

Representando cada linha de uma matriz real A, quadrada de ordem n, por $A_1, A_2, A_3, \dots, A_n$, em que $A_i = (a_{i1}, a_{i2}, a_{13}, \dots, a_{in})$, para todo $i \in \{1, 2, 3, \dots, n\}$, podemos, com finalidade de facilitar a notação, representar a matriz A na seguinte configuração:

$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix}$, cujo determinante representaremos por $det(A) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{vmatrix}$

A função determinante goza das seguintes propriedades:

I) Dada a matriz $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{'} + k.A_i^{''} \\ \vdots \\ A_n \end{bmatrix}$, então $\begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{'} + k.A_i^{''} \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{'} \\ \vdots \\ A_n \end{vmatrix} + k. \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{''} \\ \vdots \\ A_n \end{vmatrix}$, ou seja, a função $det(A)$ é linear em cada uma das linhas separadamente da matriz A.

II) Dada matriz real A quadrada de ordem n, e um escalar real k, temos $det(k.A) = k^n.det(A)$.

III) Considerando I_n com sendo a matriz identidade de ordem n, temos $det(I_n) = 1$.

IV) Se de uma matriz A quadrada de ordem n for obtida uma matriz B através de uma transformação elementar do tipo $L_i \leftrightarrow L_j$, com $i \neq j$ (permutação entra duas linhas), então $det(B) = -det(A)$.

V) Se $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$ e $A_i = A_j$, com $i \neq j$, então $det(A) = 0$.

VI) Se $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$ e $A_i = k.A_j$, com $i \neq j$ e $k \neq 0$, então $det(A) = 0$.

VII) Dada uma matriz real A, quadrada de ordem n, temos $det(A) = det(A^t)$.

VIII) Dadas duas matrizes reais A e B, quadradas de ordem n, temos $det(A.B) = det(A).det(B)$.

IX) Uma matriz A é invertível se, e somente se, $det(A) \neq 0$.

X) Dada uma matriz real A, quadrada de ordem n, cuja uma linha, digamos L_i é a combinação linear de duas outras linhas, então $det(A) = 0$.

XI) Se uma matriz B, quadrada de ordem n é obtida a partir de uma matriz A, também quadrada de ordem n, na qual somamos uma linha, com um múltiplo de outra, deixando as demais linhas inalteradas, então $det(B) = det(A)$.

As propriedades dos determinantes são de fundamental importância para obtenção de resultados mais rápidos.

Demonstrações:

I) Como vimos, por definição a função determinante de uma matriz real A, quadrada de ordem n é $det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$, com $n!$ parcelas e que em cada parcela aparece somente um elemento de cada linha da matriz A. Sendo assim, se os elementos de uma das linhas da matriz, digamos a i-ésima linha, forem $A_i = (a_{i1}^{'} + k.a_{i1}^{''}, a_{i2}^{'} + k.a_{i2}^{''}, a_{i3}^{'} + k.a_{i3}^{''}, \ldots, a_{in}^{'} + k.a_{in}^{''})$, ao calcularmos o determinante da matriz A, em cada uma das $n!$ parcelas do somatório, aparecerá um fator do tipo $(a_{if(i)}^{'} + k.a_{if(i)}^{''})$, o que fará com que $det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.\left(a_{if(i)}^{'} + k.a_{if(i)}^{''}\right).\cdots.a_{nf(n)}$ e, pelas propriedades operacionais dos somatórios, temos:

$$det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.\left(a_{if(i)}^{'} + k.a_{if(i)}^{''}\right).\cdots.a_{nf(n)} =$$

$$= \sum_f \{sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{if(i)}^{'}.\cdots.a_{nf(n)} +$$

$$+ sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.k.a_{if(i)}^{''}.\cdots.a_{nf(n)}\} =$$

$$= \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{if(i)}^{'}.\cdots.a_{nf(n)} +$$

$$k.\sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{if(i)}^{''}.\cdots.a_{nf(n)}$$

Portanto, $\begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{'} + k.A_i^{''} \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{'} \\ \vdots \\ A_n \end{vmatrix} + k.\begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i^{''} \\ \vdots \\ A_n \end{vmatrix}$

Mostrando com isso que a função determinante é linear em cada uma das linhas de uma matriz A, separadamente.

II) É uma consequência da propriedade I, pois sendo uma matriz real A de ordem n e um escalar real k, temos $k.A = \begin{bmatrix} k.A_1 \\ k.A_2 \\ k.A_3 \\ \vdots \\ k.A_i \\ \vdots \\ k.A_n \end{bmatrix}$ e o determinante de A, por definição será

$$det(k.A) = \sum_f sng(f).k.a_{1f(1)}.k.a_{2f(2)}.k.a_{3f(3)}.\cdots.k.a_{nf(n)} =$$

$$= \sum_f sng(f).\underbrace{k.k.k.\cdots.k}_{n\ fatores}.a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)} =$$

$$= k^n.\sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)} = k^n.det(A)$$

III) Como sabemos, dada uma matriz quadrada de ordem n, temos $det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$, com $n!$ parcelas e que em cada parcela aparece somente um elemento de cada linha da matriz A. Portanto, no cálculo do determinante da matriz identidade de ordem n, somente uma das parcelas do somatório , a saber, $a_{11}.a_{22}.a_{22}.\cdots.a_{nn}$ é não nula e, por tratar-se do produto dos elementos da diagonal principal, que são todos iguais a 1, além de essa parcela ser obtida através da permutação $f_{identidade}$ e $sng(f_{identidade})$=1, temos $det(I_n) = 1$.

IV) Ao permutarmos duas linhas de uma matriz A, obtendo com isso uma matriz B, cada uma das parcelas do somatório da função determinante, $det(B)$, terá ainda os mesmos elementos das parcelas da função $det(A)$, porém com ordens de índices diferentes, o que acarretará a mudança de $sng(f)$ em cada uma das parcelas, implicando com isso que $det(B) = -det(A)$.

V) Imaginemos uma matriz quadrada A, de ordem n, com duas linhas iguais, digamos L_i e L_j. Se obtivermos através de uma operação elementar do tipo $L_i \leftrightarrow L_j$, com $i \neq j$, uma matriz B a partir da matriz A, então $A = B$, pois permutamos duas linhas iguais. Isso acarreta que $det(B) = det(A)$, mas, pela propriedade IV, vimos que ao permutarmos duas linhas de uma matriz A, obtemos uma matriz B tal que $det(B) = -det(A)$. Das duas igualdades, obtermos que $det(A) = 0$.

VI) Pela propriedade I, temos que $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ k.A_i \\ \vdots \\ A_n \end{bmatrix}$, com k um escalar real, então $det(A) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ k.A_i \\ \vdots \\ A_n \end{vmatrix} = k. \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{vmatrix}$. Suponhamos que na matriz A, exista uma linha, por exemplo a j-ésima, tal que $L_j = k.L_i$. Teremos então, $det(A) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ k.A_i \\ \vdots \\ A_n \end{vmatrix} = k. \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{vmatrix}$ e, pela propriedade V, temos $\begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{vmatrix} = 0$, portanto, $det(A) = k.0 = 0$.

VII) Seja $A = \left[a_{ij}\right]_n$ uma matriz real e A^t a sua transposta. Se f é uma permutação de n elementos, então, $a^t{}_{if(i)} = a_{f(i)i}$ para todo $i, j \in \{1,2,\dots,n\}$. Sabemos, por definição que $det(A^t) = \sum_f sng(f).a_{f(1)1}.a_{f(2)2}.a_{f(3)3}.\cdots.a_{f(n)n}$ e, como f é bijetiva, existe f^{-1} de modo que quando $i = f^{-1}(j)$, temos $a_{f(i)i} = a_{jf^{-1}(j)}$. Portanto, temos $a_{f(1)1}.a_{f(2)2}.\cdots.a_{f(n)n} = a_{1f^{-1}(1)}.a_{2f^{-1}(2)}.\cdots.a_{nf^{-1}(n)}$ e, como $fof^{-1} = f_{identidade}$ e $sng\left(f_{identidade}\right)$=+1, então f e f^{-1} possuem o mesmo sinal, ou seja, $sng(f) = sng\left(f^{-1}\right)$. Notemos ainda que f percorre todas as permutações de grau n e f^{-1} também percorre, pois é a inversa de f. Sendo assim, $det(A^t) = \sum_f sng(f).a_{f(1)1}.a_{f(2)2}.a_{f(3)3}.\cdots.a_{f(n)n} =$ $\sum_f sng\left(f^{-1}\right).a_{1f^{-1}(1)}.a_{2f^{-1}(2)}.\cdots.a_{nf^{-1}(n)} = det(A)$.

VIII) Sejam $A = [a_{ij}]_n$, $B = [b_{ij}]_n$ e $C = [c_{ij}]_n$, tal que $C = A.B$, temos por definição do produto de matrizes que $c_{ij} = \sum_{k=1}^{n} a_{ik}.b_{kj}$, para todo $i, j \in \{1, 2, 3, \dots, n\}$.

Então,

$$\det(C) = \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} = \begin{vmatrix} \sum a_{1k_1}.b_{k_1 1} & \sum a_{1k_2}.b_{k_2 2} & \cdots & \sum a_{1k_n}.b_{k_n n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum a_{nk_1}.b_{k_1 1} & \sum a_{nk_2}.b_{k_2 2} & \cdots & \sum a_{nk_n}.b_{k_n n} \end{vmatrix} =$$

Utilizando a propriedade I, que trata da linearidade da função determinante em cada uma das linhas de uma matriz, a propriedade VII, que garante a linearidade também nas colunas e, por sabermos que $det(A) = det(A^t)$, temos

$$\det(C) = \sum_{k_1}. \sum_{k_2}. \cdots . \sum_{k_n}. \begin{vmatrix} a_{1k_1}.b_{k_1 1} & a_{1k_2}.b_{k_2 2} & \cdots & a_{1k_n}.b_{k_n n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n}.b_{k_n n} & a_{nk_2}.b_{k_2 2} & \cdots & a_{nk_n}.b_{k_n n} \end{vmatrix}$$

Utilizando novamente a linearidade nas n colunas (Propriedade I e Propriedade VII), temos:

$$\det(C) = \sum_{(k_1, k_2, \cdots, k_n)} b_{k_1 1}.b_{k_2 2}.\cdots.b_{k_n n} . \begin{vmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n} & a_{nk_2} & \cdots & a_{nk_n} \end{vmatrix}$$

Eliminemos as parcelas em que $k_i = k_j$ quando $i \neq j$, pois, caso contrário, teremos $\begin{vmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n} & a_{nk_2} & \cdots & a_{nk_n} \end{vmatrix} = 0.$

$$\det(C) = \sum_{\substack{(k_1, k_2, \cdots, k_n) \\ k_i \neq k_j}} b_{k_1 1}.b_{k_2 2}.\cdots.b_{k_n n} . \begin{vmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n} & a_{nk_2} & \cdots & a_{nk_n} \end{vmatrix}$$

Com a eliminação das colunas iguais, a matriz $\begin{bmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n} & a_{nk_2} & \cdots & a_{nk_n} \end{bmatrix}$ tem as mesmas colunas da matriz A, porém permutadas através de um determinado f. Assim, a matriz $\begin{bmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k_n} & a_{nk_2} & \cdots & a_{nk_n} \end{bmatrix}$ tem determinante igual ao produto de $sng(f)$ por $det(A)$, o que implica que

$\det(C) = \sum_{(f)} b_{k_1 1}.b_{k_2 2}.\cdots.b_{k_n n} . sng(f).det(A)$

$det(C) = det(A).\sum_{(f)} sng(f).b_{k_1 1}.b_{k_2 2}.\cdots.b_{k_n n}$

Como uma permutação e sua inversa tem mesmo sinal, então

$\sum_{(f)} sng(f).b_{k_1 1}.b_{k_2 2}.\cdots.b_{k_n n} = \sum_{(f)} sng(f).b_{1k_1}.b_{2k_2}.\cdots.b_{nk_n}$, o que implica que $det(C) = det(A).\sum_{(f)} sng(f).b_{1k_1}.b_{2k_2}.\cdots.b_{nk_n} = det(A).det(B)$.

IX) ($\Longrightarrow$) Se $A = \left[a_{ij}\right]_n$ é uma matriz real invertível, então existe A^{-1} real tal que $A^{-1}.A = A.A^{-1} = I_n$. Fazendo $det(A^{-1}.A) = det(A.A^{-1}) = det(I_n)$, pela propriedade VIII, temos $det(A.A^{-1}) = det(A).det(A^{-1}) = det(A^{-1}).det(A) = det(A.A^{-1})$. Por essas duas igualdades, temos que $det(A).det(A^{-1}) = det(A^{-1}).det(A) = det(I_n)$ mas, pela propriedade III, temos que $det(I_n) = 1$. Portanto, $det(A).det(A^{-1}) = det(A^{-1}).det(A) = 1$ o que nos mostra que $det(A) \neq 0$.

($\Longleftarrow$) Se A é uma matriz quadrada de ordem n tal que $det(A) \neq 0$, então todas as linhas de A são não nulas. Pelo teorema 2.7, toda matriz é equivalente a uma matriz na forma escalonada, portanto, existe uma matriz B equivalente por linhas a matriz A com todas as linhas não nulas. Logo, $B = I_n$ e, pelo teorema 2.9 temos que A é invertível.

X) Suponhamos que $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix}$ e que $A_i = \alpha.A_j + \beta.A_k$, com $\alpha, \beta \in \mathbb{R}$. Então,

$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ \alpha.A_j + \beta.A_k \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix}$ e $det(A) = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ \alpha.A_j + \beta.A_k \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{vmatrix}$, mas pela propriedade I, temos

que $det(A) = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ \alpha.A_j + \beta.A_k \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{vmatrix} = \alpha. \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{vmatrix} + \beta. \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_k \\ \vdots \\ A_j \\ \vdots \\ A_k \\ \vdots \\ A_n \end{vmatrix}$ e, pela propriedade V, temos que $det(A) = 0$.

XI) Seja $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$. Suponhamos que uma matriz B é obtida através da soma da $i-$ésima linha da matriz A com um múltiplo da $j-$ésima linha da matriz A, permanecendo as demais linhas inalteradas. Então, $B = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i + k.A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$. Assim, temos $det(B) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i + k.A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix}$ e, pela propriedade I, $det(B) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix} + k. \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix}$. Temos ainda, pela propriedade V, $det(B) = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix} + k.0 = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix} = det(A)$.

3.1.1 Alguns comentários

A propriedade VII é de extrema importância aos determinantes, pois ela permite-nos assumir todas as outras propriedades vistas até o momento com linhas de matrizes para as colunas das matrizes.

A Propriedade IX tem fundamental importância no estudo dos determinantes. Ela estabelece um critério de invertibilidade de uma matriz, ou seja, para sabermos se uma matriz é invertível, basta verificarmos se o seu determinante é diferente de zero. Além disso, essa propriedade permite-nos, juntamente com outros conceitos, determinar a inversa de uma matriz, caso ela exista, como veremos mais adiante.

Em geral, o cálculo de determinantes através da função $det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$ é um tanto quanto trabalhoso, uma vez que para $n \geq 4$, o cálculo do determinante por essa maneira torna-se inviável, pois um conjunto com 4 elementos, já possui $4! = 24$ possíveis bijeções e portanto uma soma com 24 parcelas no determinante. Com 5 elementos, já seriam possíveis $5! = 120$ bijeções e portanto uma soma com 120 parcelas no determinante. Para isso existem outras técnicas para o cálculo dos determinantes, que abordaremos a seguir:

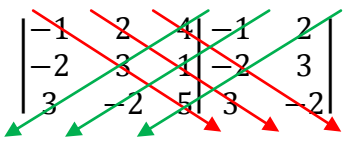
3.2 MÉTODOS PARA O CÁLCULO DE DETERMINANTES

3.2.1 Regra de Sarrus para o cálculo do determinante de uma matriz de ordem 3

Exemplo:

Seja $A = \begin{bmatrix} -1 & 2 & 4 \\ -2 & 3 & 1 \\ 3 & -2 & 5 \end{bmatrix}$ uma matriz de ordem 3.

A regra de Sarrus consiste em acrescentar, geralmente à direita do determinante, as duas primeiras colunas da matriz, obtendo a seguinte configuração:



Às diagonais traçadas em vermelho denominaremos diagonais principais e as diagonais traçadas em verde são as diagonais secundárias.

O determinante da matriz é a soma dos produtos dos elementos das diagonais principais com os simétricos dos produtos dos elementos das diagonais secundárias:

$\det(A) = -1.3.5 + 2.1.3 + 4.(-2).(-2) - 4.3.3 - (-1).1.(-2) - 2.(-2).5$

$\det(A) = -15 + 6 + 16 - 36 - 2 + 20$

$\det(A) = -11$

Notemos que a regra de Sarrus nada mais é do que a aplicação implícita da função $det(A) = \sum_f sng(f).a_{1f(1)}.a_{2f(2)}.a_{3f(3)}.\cdots.a_{nf(n)}$.

3.2.2 Regra de Laplace para o cálculo do determinante

A regra de Laplace é amplamente utilizada para o cálculo de determinantes de matrizes de ordem n, com $n \geq 4$, pois através dessa regra, de forma recorrente, diminuímos a ordem dos determinantes a serem calculados a cada interação.

3.2.2.1 Menor complementar

Seja $A = [a_{ij}]_n$ uma matriz. Considerando um elemento a_{ij} da matriz A, denominamos o *menor complementar do elemento* a_{ij} e representamos por D_{ij}, como sendo o determinante que obtemos ao suprimir na matriz A a $i - ésima$ linha e a $j - ésima$ coluna.

Exemplo:

$A = \begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1j} & a_{1k} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{h1} & \cdots & a_{hi} & a_{hj} & a_{hk} & \cdots & a_{hn} \\ a_{i1} & \cdots & a_{ii} & a_{ij} & a_{ik} & \cdots & a_{in} \\ a_{j1} & \cdots & a_{ji} & a_{jj} & a_{jk} & \cdots & a_{jn} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{nj} & a_{nk} & \cdots & a_{nn} \end{bmatrix}$. Considerando o elemento a_{ij}, temos que

$D_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1i} & a_{1k} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{h1} & \cdots & a_{hi} & a_{hk} & \cdots & a_{hn} \\ a_{j1} & \cdots & a_{ji} & a_{jk} & \cdots & a_{jn} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{nk} & \cdots & a_{nn} \end{vmatrix}$ é o seu menor complementar.

3.2.2.2 Complementar algébrico do elemento a_{ij} ou cofator de a_{ij}

O complementar algébrico ou cofator do elemento a_{ij} será representado por A_{ij} e, por definição, $A_{ij} = (-1)^{i+j} . D_{ij}$.

O determinante de uma matriz A, pela regra de Laplace é - considerando uma linha qualquer (ou coluna), por exemplo a $i - ésima$ - a soma do produto de cada elemento da linha (ou coluna) pelo seu respectivo cofator, ou seja:

$$det(A) = a_{i1} . A_{i1} + a_{i2} . A_{i2} + \cdots + a_{in} . A_{in} = \sum_{k=1}^{n} a_{ik} . A_{ik}$$

Exemplo: Seja $A = \begin{bmatrix} -1 & 2 & 4 \\ -2 & 3 & 1 \\ 3 & -2 & 5 \end{bmatrix}$ uma matriz de ordem 3. O determinante de A, segundo a regra de Laplace é dado por:

$$det(A) = -1 . (-1)^{1+1} . \begin{vmatrix} 3 & 1 \\ -2 & 5 \end{vmatrix} + 2 . (-1)^{1+2} . \begin{vmatrix} -2 & 1 \\ 3 & 5 \end{vmatrix} + 4 . (-1)^{1+3} . \begin{vmatrix} -2 & 3 \\ 3 & -2 \end{vmatrix}$$

$$det(A) = -1 . (-1)^2 . (15 + 2) + 2 . (-1)^3 . (-10 - 3) + 4 . (-1)^4 . (4 - 9)$$

$$det(A) = -1.1.17 + 2 . (-1) . (-13) + 4.1 . (-5)$$

$$det(A) = -17 + 26 - 20$$

$$det(A) = -11$$

Notemos que, para a realização do cálculo, foi escolhida a 1ª linha da matriz A, porém, poderíamos ter escolhido qualquer uma das outras linhas para o calculo do determinante e, de acordo com a propriedade VII, poderíamos também ter utilizado qualquer uma das colunas em vez das linhas.

A regra de Laplace também é uma aplicação implícita da função $det(A) = \sum_f sng(f) . a_{1f(1)} . a_{2f(2)} . a_{3f(3)} . \cdots . a_{nf(n)}$.

3.2.3 O método da eliminação de Gauss

O método da eliminação de Gauss fundamenta-se no teorema a seguir:

Teorema 3.1: Se $A = \left[a_{ij}\right]_n$ uma matriz triangular inferior (respectivamente superior), então temos $det(A) = a_{11} . a_{22} . a_{33} . \cdots . a_{nn}$, ou seja, o determinante de uma matriz triangular é o produto dos elementos da sua diagonal principal.

Demonstração: Demonstraremos esse fato utilizando indução sobre n em uma matriz triangular inferior, lembrando que para as matrizes triangulares superiores a demonstração é análoga.

Verifiquemos para $n = 2$:

Seja $A = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$. Por definição, $det(A) = a_{11}.a_{22} - a_{12}.a_{21}$, porém, em uma matriz triangular inferior, temos $a_{ij} = 0$, sempre que $i < j$, assim, temos $det(A) = a_{11}.a_{22} - 0.a_{21}$, o que implica que $det(A) = a_{11}.a_{22}$ que é o produto dos elementos da diagonal principal. Portanto, para $n = 2$ a afirmação é verdadeira.

Por hipótese de indução, seja $A = \left[a_{ij}\right]_{(n-1)}$ uma matriz triangular inferior tal que $det(A) = 0$.

Calculemos o determinante de uma matriz $A = \left[a_{ij}\right]_n$, triangular inferior:

Utilizando a regra de Laplace aplicada à 1ª linha da matriz A, temos:

$det(A) = \sum_{k=1}^{n} a_{1k}.A_{1k} = a_{11}.A_{11} + a_{12}.A_{12} + \cdots + a_{1n}.A_{1n}$, mas como $a_{ij} = 0$, sempre que $i < j$, então $det(A) = a_{11}.A_{11} + 0.A_{12} + 0.A_{13} + \cdots + 0.A_{1n} = a_{11}.A_{11}$ temos ainda que $A_{11} = (-1)^{1+1}.D_{11} = (-1)^2.D_{11} = D_{11}$, o que faz com que $det(A) = a_{11}.D_{11}$. Mas D_{11} é o determinante da matriz obtida ao suprimirmos a 1ª linha e a 1ª coluna da matriz A, portanto uma matriz quadrada de ordem $n - 1$, cuja diagonal principal são os elementos $a_{22}, a_{33}, a_{44}, \cdots, a_{nn}$ e, por hipótese de indução, $D_{11} = a_{22}.a_{33}.a_{44}.\cdots.a_{nn}$. Assim, $det(A) = a_{11}.D_{11} = a_{11}.(a_{22}.a_{33}.a_{44}.\cdots.a_{nn}) = a_{11}.a_{22}.a_{33}.a_{44}.\cdots.a_{nn}$ que é o produto dos elementos da diagonal principal da matriz A, como queríamos demonstrar.

Dada uma matriz $A = \left[a_{ij}\right]_n$, o método da eliminação de Gauss consiste em aplicar as propriedades dos determinantes com a finalidade de se obter uma matriz $B = \left[b_{ij}\right]_n$ que seja triangular, pois como vimos, calcular o determinante de uma matriz triangular é tarefa bastante simples.

Vejamos um exemplo da aplicação do método da eliminação de Gauss no cálculo do determinante da matriz $A = \begin{bmatrix} -1 & 2 & 4 \\ -2 & 3 & 1 \\ 3 & -2 & 5 \end{bmatrix}$:

$$det(A) = \begin{vmatrix} -1 & 2 & 4 \\ -2 & 3 & 1 \\ 3 & -2 & 5 \end{vmatrix} \underbrace{=}_{(Prop.\ I)} -2.\begin{vmatrix} -1 & 2 & 4 \\ 1 & -\frac{3}{2} & -\frac{1}{2} \\ 3 & -2 & 5 \end{vmatrix} \underbrace{=}_{(Prop.\ I)} -2.3.\begin{vmatrix} -1 & 2 & 4 \\ 1 & -\frac{3}{2} & -\frac{1}{2} \\ 1 & -\frac{2}{3} & \frac{5}{3} \end{vmatrix} \underbrace{=}_{(Prop.\ XI)}$$

Façamos $L_2 \to L_1 + L_2$ e $L_3 \to L_1 + L_3$.

$$\underset{(Prop.\ XI)}{=} -6.\begin{vmatrix} -1 & 2 & 4 \\ 0 & \frac{1}{2} & \frac{7}{2} \\ 1 & \frac{4}{3} & \frac{17}{3} \end{vmatrix} \underset{(Prop.\ I)}{=} -6.\frac{1}{2}.\left(-\frac{4}{3}\right).\begin{vmatrix} -1 & 2 & 4 \\ 0 & 1 & 7 \\ 0 & -1 & -\frac{17}{4} \end{vmatrix} \underset{(Prop.\ XI)}{=}$$

Façamos $L_3 \to L_2 + L_3$.

$$\underset{(Prop.\ XI)}{=} 4.\begin{vmatrix} -1 & 2 & 4 \\ 0 & 1 & 7 \\ 0 & 0 & \frac{11}{4} \end{vmatrix} = 4.(-1).1.\frac{11}{4} = -11$$

3.3 Determinantes e matriz inversa

Uma das notáveis aplicações para os determinantes é a determinação da inversa de uma matriz.

Vimos anteriormente que uma condição necessária e suficiente para que uma matriz $A = [a_{ij}]_n$ possua inversa é o fato de $det(A) \neq 0$. Veremos agora como determinar a inversa de uma matriz a partir do seu determinante.

Dada uma matriz $A = [a_{ij}]_n$, representaremos por A' a *matriz dos cofatores de A*.

Assim, se $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, então $A' = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$.

À matriz $\bar{A} = (A')^t = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{bmatrix}$, com $B_{ij} = A_{ji}, i, j \in \{1, 2, \dots, n\}$, denominamos *matriz adjunta de A*.

Teorema 3.2: Se $A = [a_{ij}]_n$ é uma matriz e I_n a matriz identidade de ordem n, então $A.\bar{A} = \bar{A}.A = det(A).I_n$

Demonstração: seja $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, cuja matriz adjunta é $\bar{A} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{bmatrix}$. Fazendo o produto $A.\bar{A}$, temos:

$$A.\bar{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} . \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}, \quad \text{onde,}$$

pela definição do produto de matrizes, $c_{ij} = \sum_{k=1}^{n} a_{ik}.B_{kj}$, mas como vimos, $B_{ij} = A_{ji}$, então, $c_{ij} = \sum_{k=1}^{n} a_{ik}.A_{jk}$. Consideremos os dois casos a seguir:

1º) $i = j$ implica que $c_{ii} = \sum_{k=1}^{n} a_{ik}.A_{ik} = det(A)$

2º) $i \neq j$ implica que $c_{ij} = \sum_{k=1}^{n} a_{ik}.A_{jk}$. Analisemos essa situação.

Considerando $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, da qual obtemos uma matriz

$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i1} & a_{i2} & \cdots & a_{in} \rightarrow \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ pela substituição da $j-$ésima pela $i-$ésima linha. j-ésima linha

Pela propriedade V, temos que $det(A') = 0$, pois A' possui duas linhas iguais. E, pela definição de determinante, aplicada à $j-$ésima linha, temos $det(A') = a_{i1}.A_{j1} + a_{i2}.A_{j2} + \cdots + a_{in}.A_{jn} = \sum_{k=1}^{n} a_{ik}.A_{jk}$, o que implica que $c_{ij} = \sum_{k=1}^{n} a_{ik}.A_{jk} = 0$, quando $i \neq j$.

Assim,

$$A.\bar{A} = \begin{bmatrix} det(A) & 0 & \cdots & 0 \\ 0 & det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & det(A) \end{bmatrix} = det(A).\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = det(A).I_n$$

De maneira análoga se demonstra que $\bar{A}.A = det(A).I_n$.

Teorema 3.3: Se A é uma matriz invertível, então $A^{-1} = \frac{1}{det\,(A)}.\bar{A}$.

Demonstração: Se A é uma matriz invertível, então $det(A) \neq 0$. Do teorema 3.2, temos que $det(A).I_n = \bar{A}.A$, multiplicando à direita ambos os membros da igualdade por A^{-1}, temos: $[det(A).I_n].A^{-1} = [\bar{A}.A].A^{-1}$. Aplicando a propriedade associativa do produto de matrizes,

temos $det(A).[I_n.A^{-1}] = \bar{A}.[A.A^{-1}]$ o que implica que $det(A).A^{-1} = \bar{A}.I_n$. Dividindo ambos os membros da igualdade por $det(A)$, temos $A^{-1} = \frac{1}{det\,(A)}.\bar{A}$.

Exemplo: Sendo $A = \begin{bmatrix} -1 & 2 & 4 \\ -2 & 3 & 1 \\ 3 & -2 & 5 \end{bmatrix}$, determinemos a inversa A^{-1} da matriz A:

Temos que $det(A) = -11$, o que, pela propriedade IX garante sua invertibilidade. Determinemos a matriz dos cofatores de A:

$A' = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$, onde:

$A_{11} = (-1)^{1+1}.\begin{vmatrix} 3 & 1 \\ -2 & 5 \end{vmatrix} = (-1)^2.[3.5 - 1.(-2)] = 1.(15 + 2) = 17$

$A_{12} = (-1)^{1+2}.\begin{vmatrix} -2 & 1 \\ 3 & 5 \end{vmatrix} = (-1)^3.[-2.5 - 1.3] = -1.(-10 - 3) = 13$

$A_{13} = (-1)^{1+3}.\begin{vmatrix} -2 & 3 \\ 3 & -2 \end{vmatrix} = (-1)^4.[-2.(-2) - 3.3] = 1.(4 - 9) = -5$

$A_{21} = (-1)^{2+1}.\begin{vmatrix} 2 & 4 \\ -2 & 5 \end{vmatrix} = (-1)^3.[2.5 - 4.(-2)] = -1.(10 + 8) = -18$

$A_{22} = (-1)^{2+2}.\begin{vmatrix} -1 & 4 \\ 3 & 5 \end{vmatrix} = (-1)^4.[-1.5 - 4.3] = 1.(-5 - 12) = -17$

$A_{23} = (-1)^{2+3}.\begin{vmatrix} -1 & 2 \\ 3 & -2 \end{vmatrix} = (-1)^5.[-1.(-2) - 2.3] = -1.(2 - 6) = 4$

$A_{31} = (-1)^{3+1}.\begin{vmatrix} 2 & 4 \\ 3 & 1 \end{vmatrix} = (-1)^4.[2.1 - 4.3] = 1.(2 - 12) = -10$

$A_{32} = (-1)^{3+2}.\begin{vmatrix} -1 & 4 \\ -2 & 1 \end{vmatrix} = (-1)^5.[-1.1 - 4.(-2)] = -1.(-1 + 8) = -7$

$A_{33} = (-1)^{3+3}.\begin{vmatrix} -1 & 2 \\ -2 & 3 \end{vmatrix} = (-1)^6.[-1.3 - 2.(-2)] = 1.(-3 + 4) = 1$

Assim, $A' = \begin{bmatrix} 17 & 13 & -5 \\ -18 & -17 & 4 \\ -10 & -7 & 1 \end{bmatrix}$, e a matriz adjunta de A é $\bar{A} = \begin{bmatrix} 17 & -18 & -10 \\ 13 & -17 & -7 \\ -5 & 4 & 1 \end{bmatrix}$.

Pelo teorema 3.3, $A^{-1} = \frac{1}{det\,(A)}.\bar{A}$, portanto,

$$A^{-1} = \frac{1}{-11}.\begin{bmatrix} 17 & -18 & -10 \\ 13 & -17 & -7 \\ -5 & 4 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{17}{11} & \frac{18}{11} & \frac{10}{11} \\ -\frac{13}{11} & \frac{17}{11} & \frac{7}{11} \\ \frac{5}{11} & -\frac{4}{11} & -\frac{1}{11} \end{bmatrix}$$

4 ALGUMAS NOÇÕES SOBRE POLINÔMIOS

Consideremos um conjunto[1] A, não vazio, com as operações de adição e multiplicação tais que para todo $a, b, c \in A$, temos $a + b \in A$ e $a.b \in A$.

Consideremos que em relação à adição no conjunto A, temos as seguintes propriedades:

I) $a + (b + c) = (a + b) + c$;

II) $a + b = b + a$;

III) $\exists o \in A; a + o = o + a = a$, onde o representa o elemento neutro aditivo;

IV) $\exists a' \in A; a + a' = a' + a = o$, no caso, representamos a' por $-a$

Com relação à multiplicação no conjunto A, temos as seguintes propriedades:

I) $a.(b.c) = (a.b).c$

II) $a.b = b.a$

III) $\exists 1 \in A; a.1 = 1.a = a$

IV) $\forall a, b \in A, a.b = o \implies a = o \; ou \; b = o$

Temos ainda no conjunto A a distributividade da multiplicação em relação à adição, ou seja, $a.(b + c) = a.b + a.c$.

Consideremos um símbolo $x \notin A$ ao qual denominaremos de *indeterminada sobre A* considerando $x^0 = 1$ e $x^1 = x$.

Para todo $n \in \mathbb{N} \cup \{0\}$, definimos um polinômio $p(x)$ com coeficientes no conjunto A como sendo a expressão formal $p(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n$, em que para $0 \leq k \leq n$ os elementos $a_k \in A$ são denominados os *coeficientes* do polinômio $p(x)$.

Denominaremos *monômios de grau k* do polinômio $p(x)$ às parcelas $a_k x^k$, com $k \neq 0$ e ao coeficiente a_0 de $p(x)$, denominaremos *termo constante*. Se $p(x) = a_0$, dizemos que $p(x)$ é um polinômio constante. Se $p(x) = 0$, então $p(x)$ é o polinômio nulo.

Em um polinômio $p(x)$, não nulo, existem $n, i \in \mathbb{N} \cup \{0\}$ tal que $a_n \neq 0$ e $a_i = 0 \;\; \forall i > n$. Neste caso, dizemos que n é o grau do polinômio $p(x)$ e representamos esse fato por $gr(p(x)) = n$. Ao coeficiente a_n, do termo de maior grau, denominamos *coeficiente líder* de $p(x)$. Caso o coeficiente líder de um polinômio $p(x)$ seja igual a 1 então dizemos que

[1] Conjuntos com as operações de adição e multiplicação que gozam das propriedades do conjunto A são classificados como *domínios de integridade* como veremos mais a frente.

$p(x)$ é um *polinômio mônico*. Não definimos grau para o polinômio nulo $o(x) = 0 + 0x + \cdots + 0x^{n-1} + 0x^n$.

Representaremos por $A[x]$ o conjunto de todos os polinômios com os coeficientes no conjunto A. Assim, $\mathbb{Z}[x]$ é o conjunto de todos os polinômios com coeficientes inteiros, assim como $\mathbb{Q}[x], \mathbb{R}[x]$ e $\mathbb{C}[x]$ são, respectivamente, os conjuntos de todos os polinômios com coeficientes racionais, reais e complexos. Notemos que os conjuntos $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$ gozam das mesmas propriedades do conjunto A.

4.1 IGUALDADE DE POLINÔMIOS

Dados dois polinômios $p(x) = \sum_{k=0}^{n} a_k x^k$ e $q(x) = \sum_{k=0}^{n} b_k x^k$ com coeficientes em A, dizemos que $p(x) = q(x)$ se $a_k = b_k$, para todo $k \in \{0, 1, 2, \ldots, n\}$.

4.2 ADIÇÃO DE POLINÔMIOS

Dados dois polinômios $p(x) = \sum_{k=0}^{n} a_k x^k$ e $q(x) = \sum_{k=0}^{m} b_k x^k$ com coeficientes em A, definimos a soma de $p(x)$ com $q(x)$, considerando $m = n$ ao reescrever $p(x)$ e $q(x)$ com as mesmas potências de x, como $p(x) + q(x) = \sum_{k=0}^{n}(a_k + b_k)x^k = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n$.

Exemplo: Dados os polinômios $p(x) = -2 + 6x - 5x^2 + 2x^4$ e $q(x) = 8 + 6x^2 + 5x^3 - 7x^4 + 3x^5$ em $\mathbb{Z}[x]$, temos:

$$p(x) + q(x) = (-2 + 8) + (6 + 0)x + (-5 + 6)x^2(0 + 5)x^3 + (2 - 7)x^4 + (0 + 3)x^5$$

$$p(x) + q(x) = 6 + 6x + x^2 + 5x^3 - 5x^4 + 3x^5$$

Uma forma prática para resolver a adição de $p(x)$ com $q(x)$ é:

$$\begin{array}{rccccccc} & -2 & +6x & -5x^2 & +0x^3 & +2x^4 & +0x^5 \\ (+) & +8 & +0x & +6x^2 & +5x^3 & -7x^4 & +3x^5 \\ \hline & +6 & +6x & +x^2 & +5x^3 & -5x^4 & +3x^5 \end{array}$$

Portanto, $p(x) + q(x) = 6 + 6x + x^2 + 5x^3 - 5x^4 + 3x^5$.

Propriedades da adição de polinômios

Para quaisquer $p(x), q(x)$ e $h(x)$, pertencentes a $A[x]$ a adição de polinômios goza das seguintes propriedades:

I) $gr\big(p(x)+\ q(x)\big) \leq max\{gr\big(p(x)\big), gr\big(\ q(x)\big)\}$

II) $[p(x)+q(x)]+h(x)=p(x)+[q(x)+h(x)]$ (associatividade)

III) $p(x)+q(x)=q(x)+p(x)$ (comutatividade)

IV) $p(x)+o(x)=p(x)$, onde $o(x)$ representa o polinômio nulo. (elemento neutro aditivo)

V) $p(x)+\big(-p(x)\big)=0$ (existência do polinômio simétrico ou inverso aditivo)

Demonstrações:

Consideremos os polinômios $p(x), q(x)$ e $h(x)$, não nulos, pertencentes a $A[x]$, tais que $p(x)=\sum_{k=0}^{n} a_k x^k$, $q(x)=\sum_{k=0}^{m} b_k x^k$ e $h(x)=\sum_{k=0}^{r} c_k x^k$. Para as demonstrações a partir da propriedade II, consideremos, sem perda de generalidade, que $n=m=r$ (basta lembrar que um polinômio de grau $m<n$ pode ser considerado como um polinômio no qual os coeficientes a partir do $m-ésimo$, exclusive, são todos iguais a zero):

I) 1º. Consideremos $n>m$:

Seja c_n o coeficiente do $n-ésimo$ termo da soma de $p(x)$ com $q(x)$. Como a partir do $m-ésimo$ termo (exclusive) de $q(x)$ os coeficientes são todos nulos, pois $gr\big(q(x)\big)=m$, temos $c_n=a_n+b_n=a_n+0=a_n\neq 0$ e, $c_i=a_i+b_i=0+0=0$ $\forall i>n$, pois $gr\big(p(x)\big)=n$, portanto:

$gr\big(p(x)+q(x)\big)=n=max\{gr\big(p(x)\big), gr\big(\ q(x)\big)\}$.

2º. Consideremos $n=m$:

Temos $c_i=a_i+b_i=0+0=0\ \forall i>n$, pois $gr\big(p(x)\big)=gr\big(q(x)\big)=n$, mas, caso $b_n=-a_n$, então $c_n=a_n+b_n=a_n-a_n=0$, implicando com isso que $gr\big(p(x)+q(x)\big)\leq n=max\{gr\big(p(x)\big), gr\big(\ q(x)\big)\}$.

Portanto, fica demonstrado que $gr\big(p(x),q(x)\big)\leq max\{gr\big(p(x)\big), gr\big(\ q(x)\big)\}$ sempre que $p(x)$ e $q(x)$ forem polinômios não nulos.

II) Por definição, $[p(x)+q(x)]+h(x)=[\sum_{k=0}^{n} a_k x^k+\sum_{k=0}^{n} b_k x^k]+\sum_{k=0}^{n} c_k x^k=$

$=\sum_{k=0}^{n}(a_k x^k+b_k x^k)+\sum_{k=0}^{n} c_k x^k=\sum_{k=0}^{n}(a_k+b_k)x^k+\sum_{k=0}^{n} c_k x^k=$

$=\sum_{k=0}^{n}[(a_k+b_k)x^k+c_k x^k]=\sum_{k=0}^{n}[(a_k+b_k)+c_k]x^k=\sum_{k=0}^{n}[a_k+(b_k+c_k)]x^k=$

$=\sum_{k=0}^{n}[a_k x^k+(b_k+c_k)x^k]=\sum_{k=0}^{n} a_k x^k+\sum_{k=0}^{n}(b_k+c_k)x^k=$

$=\sum_{k=0}^{n} a_k x^k+[\sum_{k=0}^{n} b_k x^k+\sum_{k=0}^{n} c_k x^k]=p(x)+[q(x)+h(x)]$.

III) $p(x)+q(x)=\sum_{k=0}^{n} a_k x^k+\sum_{k=0}^{n} b_k x^k=\sum_{k=0}^{n}(a_k x^k+b_k x^k)=\sum_{k=0}^{n}(a_k+b_k)x^k=$

$=\sum_{k=0}^{n}(b_k+a_k)x^k=\sum_{k=0}^{n}(b_k x^k+a_k x^k)=\sum_{k=0}^{n} b_k x^k+\sum_{k=0}^{n} a_k x^k=q(x)+p(x)$.

IV) Como $0 \in A$, então o polinômio nulo $o(x) \in A[x]$ e podemos representá-lo por $o(x) = \sum_{k=0}^{n} 0x^k$. Sendo assim, $p(x) + o(x) = \sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} 0x^k =$
$= \sum_{k=0}^{n}(a_k x^k + 0x^k) = \sum_{k=0}^{n}(a_k + 0)\, x^k = \sum_{k=0}^{n} a_k\, x^k = p(x)$, ou seja, o polinômio nulo é o elemento neutro da adição.

V) $p(x) + q(x) = o(x) \Leftrightarrow \sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} b_k x^k = \sum_{k=0}^{n} 0x^k \Leftrightarrow$
$\Leftrightarrow \sum_{k=0}^{n}(a_k x^k + b_k x^k) = \sum_{k=0}^{n} 0x^k \Leftrightarrow \sum_{k=0}^{n}(a_k + b_k)x^k = \sum_{k=0}^{n} 0x^k \Leftrightarrow$
$\Leftrightarrow\ a_k + b_k = 0 \Leftrightarrow\ b_k = -a_k$
Assim, $q(x) = \sum_{k=0}^{n} b_k x^k = \sum_{k=0}^{n}(-a_k)x^k = -\sum_{k=0}^{n} a_k x^k = -p(x)$ é o polinômio simétrico (ou o inverso aditivo) de $p(x)$.

4.3 MULTIPLICAÇÃO DE POLINÔMIOS

Dados dois polinômios $p(x)$ e $q(x)$, pertencentes a $A[x]$, tais que $p(x) = \sum_{k=0}^{m} a_k x^k$ e $q(x) = \sum_{k=0}^{n} b_k x^k$, definimos a multiplicação de $p(x)$ por $q(x)$ como sendo o polinômio $h(x) \in A[x]$; $h(x) = \sum_{k=0}^{m+n} c_k x^k$, onde $c_k = \sum_{i=0}^{k} a_i b_{k-i}$, ou seja, multiplicar $p(x)$ por $q(x)$ consiste em multiplicar cada monômio $a_i x^i$ de $p(x)$, $i \in \{0, 1, 2, \dots, m\}$, por cada termo $b_j x^j$ de $q(x)$, $j \in \{0, 1, 2, \dots, n\}$, obtendo $a_i x^i . b_j x^j = a_i b_j x^{i+j}$ e somando os resultados ao final.

Exemplo: Dados os polinômios $p(x) = -2 + 6x - 5x^2 + 2x^4$ e $q(x) = 8 + 6x^2 + 5x^3 - 7x^4 + 3x^5$ em $\mathbb{Z}[x]$, obter o produto $h(x) = p(x).q(x)$.

Solução:

$h(x) = (-2 + 6x - 5x^2 + 2x^4).(8 + 6x^2 + 5x^3 - 7x^4 + 3x^5) =$

$h(x) = -2.8 + (-2).6x^2 + (-2).5x^3 + (-2).(-7x^4) + (-2).3x^5 + 6x.8 + 6x.6x^2 +$

$+6x.5x^3 + 6x.(-7x^4) + 6x.3x^5 + (-5x^2).8 + (-5x^2).6x^2 + (-5x^2).5x^3 +$

$+(-5x^2).(-7x^4) + (-5x^2).3x^5 + 2x^4.8 + 2x^4.6x^2 + 2x^4.5x^3 + 2x^4.(-7x^4) + 2x^4.3x^5 =$

$= -16 - 12x^2 - 10x^3 + 14x^4 - 6x^5 + 48x + 36x^3 + 30x^4 - 42x^5 + 18x^6 - 40x^2 -$

$-30x^4 - 25x^5 + 35x^6 - 15x^7 + 16x^4 + 12x^6 + 10x^7 - 14x^8 + 6x^9 =$

$= -16 + 48x + (-12 - 40)x^2 + (-10 + 36)x^3 + (14 + 30 - 30 + 16)x^4 +$

$+(-6 - 42 - 25)x^5 + (18 + 35 + 12)x^6 + (-15 + 10)x^7 - 14x^8 + 6x^9 =$

$= -16 + 48x - 52x^2 + 26x^3 + 30x^4 - 73x^5 + 65x^6 - 5x^7 - 14x^8 + 6x^9$

Uma forma prática para determinar $h(x) = p(x).q(x)$ é:

$$
\begin{array}{lccccccccc}
 & & & 8 & +6x^2 & +5x^3 & -7x^4 & +3x^5 & & \\
(\cdot) & & & & -2 & +6x & -5x^2 & 2x^4 & & \\
\hline
 & -16 & & -12x^2 & -10x^3 & +14x^4 & -6x^5 & & & \\
 & & +48x & & +36x^3 & +30x^4 & -42x^5 & +18x^6 & & \\
(+) & & & -40x^2 & & -30x^4 & -25x^5 & +35x^6 & -15x^7 & \\
 & & & & & +16x^4 & & +12x^6 & +10x^7 & -14x^8 & +6x^9 \\
\hline
 & -16 & +48x & -52x^2 & +26x^3 & +30x^4 & -73x^5 & +65x^6 & -5x^7 & -14x^8 & +6x^9
\end{array}
$$

Propriedades da multiplicação de polinômios

Para quaisquer $p(x), q(x)$ e $h(x)$ pertencentes a $A[x]$ a multiplicação de polinômios goza das seguintes propriedades:

I) $gr\big(p(x)\cdot q(x)\big) = gr\big(p(x)\big) + gr\big(q(x)\big)$

II) $[p(x)\cdot q(x)]\cdot h(x) = p(x)\cdot[q(x)\cdot h(x)]$ (associatividade)

III) $p(x)\cdot q(x) = q(x)\cdot p(x)$ (comutatividade)

IV) $p(x)\cdot 1 = p(x)$, onde 1 representa o polinômio constante 1. (elemento neutro multiplicativo)

V) $p(x)\cdot[q(x)+h(x)] = p(x)\cdot q(x) + p(x)\cdot h(x)$ (distributividade da multiplicação em relação à adição)

Demonstrações:

Consideremos os polinômios $p(x), q(x)$ e $h(x)$, não nulos, pertencentes a $A[x]$, tais que $p(x) = \sum_{k=0}^{n} a_k x^k$, $q(x) = \sum_{k=0}^{m} b_k x^k$ e $h(x) = \sum_{k=0}^{r} c_k x^k$.

I) $gr\big(p(x)\big) = n$ e $gr\big(q(x)\big) = m$. Seja $c_k = \sum_{i=0}^{k} a_i b_{k-i}$ um coeficiente qualquer de $p(x)\cdot q(x)$. Assim, temos $c_{m+n} = \sum_{i=0}^{m+n} a_i b_{m+n-i} \neq 0$ e $c_k = 0\ \forall k > m+n$, portanto, $gr\big(p(x).q(x)\big) = m+n = gr\big(p(x)\big) + gr\big(q(x)\big)$.

II) Primeiramente, notemos que, decorrente da propriedade I, $gr(p(x)\cdot[q(x)\cdot h(x)]) = gr\big([p(x)\cdot q(x)]\cdot h(x)\big)$, pois $gr\big(q(x)\cdot h(x)\big) = m+r$, o que implica que $gr\Big(p(x).\big(q(x)\cdot h(x)\big)\Big) = m+n+r = t$. Em contra partida, temos $gr\big(p(x)\cdot q(x)\big) = m+n$, o que implica que $gr\big([p(x)\cdot q(x)]\cdot h(x)\big) = m+n+r = t$.

Utilizando a definição de multiplicação de polinômios, façamos $p(x) = \sum_{k=0}^{n} a_k x^k$, $q(x) = \sum_{i=0}^{m} b_i x^i$, $h(x) = \sum_{j=0}^{r} c_j x^j$, $q(x)\cdot h(x) = \sum_{l=0}^{m+r} d_l x^l$, $p(x)\cdot[q(x)\cdot h(x)] = \sum_{t=0}^{m+n+r} e_t x^t$, $p(x)\cdot q(x) = \sum_{s=0}^{m+n} f_s x^s$ e $[p(x)\cdot q(x)]\cdot h(x) =$

$\sum_{t=0}^{m+n+r} g_t x^t$. Notemos ainda que um termo de um produto, por exemplo, e_t, é determinado por $e_t = \sum_{k=0}^{t} a_k d_{t-k}$, mas, fazendo $l = t - k$, o que implica $k + l = t$, podemos reescrever esse termo da seguinte forma: $e_t = \sum_{k+l=t} a_k d_l$. De maneira análoga, escrevemos $d_l = \sum_{i+j=l} b_i c_j$, $f_s = \sum_{k+i=s} a_k b_i$ e $g_t = \sum_{s+j=t} f_s c_j$. Assim, temos $p(x) \cdot [q(x) \cdot h(x)] = \sum_{t=0}^{m+n+r} e_t x^t$, mas

$e_t = \sum_{k+l=t} a_k d_l = \sum_{k+l=t} a_k \cdot \left(\sum_{i+j=l} b_i c_j\right) = \sum_{k+i+j=t} a_k \left(b_i c_j\right) =$

$= \sum_{k+i+j=t} (a_k b_i) c_j = \sum_{s+j=t} (\sum_{k+i=s} a_k b_i) \cdot c_j = \sum_{s+j=t} f_s c_j = g_t$

Portanto, $p(x) \cdot [q(x) \cdot h(x)] = \sum_{t=0}^{m+n+r} e_t x^t = \sum_{t=0}^{m+n+r} g_t x^t = [p(x) \cdot q(x)] \cdot h(x)$

Mostrando assim, a associatividade da multiplicação de polinômios.

III) Temos que $p(x) = \sum_{k=0}^{n} a_k x^k$ e $q(x) = \sum_{k=0}^{m} b_k x^k$.

Consideremos $p(x) \cdot q(x) = \sum_{k=0}^{m+n} c_k x^k$, onde $c_k = \sum_{i=0}^{k} a_i b_{k-i}$ e $q(x) \cdot p(x) = \sum_{k=0}^{m+n} d_k x^k$, onde $d_k = \sum_{i=0}^{k} b_i a_{k-i}$. Desenvolvendo c_k temos:

$c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + a_2 b_{k-2} + \cdots + a_{k-2} b_2 + a_{k-1} b_1 + a_k b_0 =$ [2]

$b_0 a_k + b_1 a_{k-1} + b_2 a_{k-2} + \cdots + b_{k-2} a_2 + b_{k-1} a_1 + b_k a_0 = \sum_{i=0}^{k} b_i a_{k-i} = d_k$, o que implica que $\sum_{k=0}^{m+n} c_k x^k = \sum_{k=0}^{m+n} d_k x^k$, mostrando com isso que $p(x) \cdot q(x) = q(x) \cdot p(x)$. Portanto o produto de polinômios goza da propriedade comutativa.

IV) Consideremos o polinômio constante $u(x) = 1$. Como $1 \in A$, então $u(x) \in A[X]$. Temos $p(x) \cdot u(x) = \sum_{k=0}^{n+0} c_k x^k$, com $c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 . 0 + a_1 . 0 + a_2 . 0 + \cdots + a_{k-2} . 0 + a_{k-1} . 0 + a_k . 1 = a_k$, portanto, $p(x) \cdot u(x) = \sum_{k=0}^{n+0} c_k x^k = \sum_{k=0}^{n} a_k x^k = p(x)$. Logo, $u(x) = 1$ é o elemento neutro multiplicativo em $A[x]$.

V) Podemos considerar $m = r$, ao reescrever $q(x)$ e $h(x)$ com as mesmas potências de x. Pela definição de multiplicação de polinômios, temos $p(x) \cdot q(x) = \sum_{k=0}^{n+m} e_k x^k$, com $e_k = \sum_{i=0}^{k} a_i \cdot b_{k-1}$ e $p(x) \cdot h(x) = \sum_{k=0}^{n+m} f_k x^k$, com $f_k = \sum_{i=0}^{k} a_i \cdot c_{k-1}$.

Pela definição da adição de polinômios, temos $q(x) + h(x) = \sum_{k=0}^{m} (b_k + c_k) x^k$. Utilizando novamente a definição de multiplicação de polinômios, temos:

$p(x) \cdot [q(x) + h(x)] = \sum_{k=0}^{m+n} d_k x^k$, onde $d_k = \sum_{i=0}^{k} a_i \cdot (b_{k-1} + c_{k-1})$. Utilizando a propriedade distributiva da multiplicação em relação à adição em A, temos:

$d_k = \sum_{i=0}^{k} (a_i \cdot b_{k-1} + a_i \cdot c_{k-1}) = \underbrace{\sum_{i=0}^{k} a_i \cdot b_{k-1}}_{e_k} + \underbrace{\sum_{i=0}^{k} a_i \cdot c_{k-1}}_{f_k}$, portanto,

$p(x) \cdot [q(x) + h(x)] = \sum_{k=0}^{m+n} d_k x^k = \sum_{k=0}^{m+n} (e_k + f_k) x^k = \sum_{k=0}^{m+n} (e_k x^k + f_k x^k) =$

$= \sum_{k=0}^{m+n} e_k x^k + \sum_{k=0}^{m+n} f_k x^k = p(x) \cdot q(x) + p(x) \cdot h(x)$.

[2] Utilizando a comutatividade da adição e da multiplicação em A.

4.4 DIVISÃO EUCLIDIANA DE POLINÔMIOS

Dados $p(x)$ e $g(x)$ pertencentes a $A[x]$, com $g(x) \neq 0$, se existir um polinômio $q(x) \in A[x]$, tal que $p(x) = g(x) \cdot q(x)$, então, dizemos que $p(x)$ é múltiplo de $g(x)$, ou ainda que $g(x)$ divide $p(x)$.

Exemplo:

$g(x) = 5 + 2x + x^2 \in \mathbb{Z}[x]$ divide o polinômio $p(x) = 10 - 11x + x^2 - x^3 + x^4 \in \mathbb{Z}[x]$, pois existe $q(x) = 2 - 3x + x^2 \in \mathbb{Z}[x]$ tal que $p(x) = 10 - 11x + x^2 - x^3 + x^4 = (5 + 2x + x^2) \cdot (2 - 3x + x^2) = g(x) \cdot q(x)$.

Qualquer que seja $p(x) \in A[x]$, se $p(x) \neq 0$, então $p(x)$ divide 0, onde 0 representa o polinômio nulo.

Teorema 4.1: Considerando $p(x)$ e $g(x)$ polinômios não nulos do conjunto $A[x]$, se $g(x)$ tem coeficiente líder invertível e divide $p(x)$, então $gr\big(g(x)\big) \leq gr\big(p(x)\big)$.

Demonstração: Por hipótese $g(x)$ divide $p(x)$ e são ambos não nulos. Isto significa que existe $q(x) \in A[x]$, não nulo, tal que $p(x) = g(x) \cdot q(x)$, mas, pela propriedade I da multiplicação de polinômios, temos que:

$$gr\big(g(x)\big) \leq gr\big(g(x)\big) + gr\big(q(x)\big) = gr\big(g(x) \cdot q(x)\big) = gr\big(p(x)\big).$$

Notemos que em $\mathbb{Z}$ os únicos elementos invertíveis são o 1 e o -1 enquanto que em $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$, todo elemento não nulo é invertível.

Teorema 4.2 (divisão euclidiana): Consideremos o conjunto A, com suas propriedades e sejam $p(x)$ e $g(x)$ polinômios de $A[x]$, com $g(x)$ não nulo e com coeficiente líder invertível no conjunto A. Então, existem $q(x)$ e $r(x)$, unicamente determinados, pertencentes a $A[x]$, tal que $p(x) = q(x) \cdot g(x) + r(x)$, com $r(x) = 0$ ou $gr\big(r(x)\big) < gr\big(g(x)\big)$.

Demonstração: Consideraremos $g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_mx^m$, com b_m invertível em A, ou seja, admitindo que existe $b_m^{-1} \in A$, tal que $b \cdot b_m^{-1} = 1$. Dividiremos a demonstração em duas partes, a primeira trata de provar a existência de $q(x)$ e $r(x)$ enquanto a segunda provará a unicidade de $q(x)$ e $r(x)$.

1ª parte: Considerando $p(x) = 0$, então $q(x) = r(x) = 0 \in A[x]$ e $p(x) = 0 = 0 \cdot g(x) + 0 =$

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

Considerando $p(x) \in A$, com $p(x) \neq 0$ e $gr\big(p(x)\big) = n$. Se $n < m$, basta tomar $q(x) = 0$ e $r(x) = p(x)$ que teremos $p(x) = 0 \cdot g(x) + p(x) = q(x) \cdot g(x) + r(x)$.

Considerando $n \geq m$, escrevendo $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, com $a_n \neq 0$, demonstramos por indução sobre $n = gr\big(p(x)\big)$:

Se $n = 0$, temos $gr\big(p(x)\big) = 0$. Como $n \geq m$, então $0 \geq m$, o que implica que $m = 0$ e portanto, $gr\big(g(x)\big) = 0$. Logo, $p(x) = a_0 \neq 0$ e $g(x) = b_0$. Como $b_0^{-1} \in A$, podemos escrever $p(x) = a_0 = a_0 \cdot 1 + 0 = a_0b_0^{-1}b_0 + 0 = a_0b_0^{-1}g(x) + 0 = q(x) \cdot g(x) + r(x)$, com $q(x) = a_0b_0^{-1}$ e $r(x) = 0$.

Suponhamos que o resultado seja válido para polinômios com grau menor do que n. Consideremos o polinômio $p_1(x) = p(x) - a_nb_m^{-1}x^{n-m}g(x)$, notemos que $gr\big(p_1(x)\big) < gr\big(p(x)\big)$ e, por hipótese de indução, existem $q_1(x)$ e $r_1(x)$ pertencentes a $A[x]$, tais que $p_1(x) = q_1(x) \cdot g(x) + r_1(x)$, com $r_1(x) = 0$ ou $gr\big(r_1(x)\big) < gr\big(g(x)\big)$. Assim, $p_1(x) = p(x) - a_nb_m^{-1}x^{n-m}g(x)$, implicando que:

$p(x) = p_1(x) + a_nb_m^{-1}x^{n-m}g(x) = [q_1(x) \cdot g(x) + r_1(x)] + a_nb_m^{-1}x^{n-m}g(x) =$

$= [q_1(x) + a_nb_m^{-1}x^{n-m}]g(x) + r_1(x) = q(x).g(x) + r(x)$, considerando $r(x) = r_1(x)$ e $q(x) = q_1(x) + a_nb_m^{-1}x^{n-m}$.

2ª parte: Sejam $q_1(x)$, $q_2(x)$, $r_1(x)$ e $r_2(x)$, com $q_1(x) \neq q_2(x)$ e $r_1(x) \neq r_2(x)$, tais que $p(x) = q_1(x) \cdot g(x) + r_1(x)$ e $p(x) = q_2(x) \cdot g(x) + r_2(x)$, com $r_1(x) = 0$ ou $gr\big(r_1(x)\big) < gr\big(g(x)\big)$ e $r_2(x) = 0$ ou $gr\big(r_2(x)\big) < gr\big(g(x)\big)$. Temos então,

$q_1(x) \cdot g(x) + r_1(x) = q_2(x) \cdot g(x) + r_2(x)$, o que implica que $q_1(x) \cdot g(x) - q_2(x) \cdot g(x) = r_2(x) - r_1(x)$ ou ainda $[q_1(x) - q_2(x)] \cdot g(x) = r_2(x) - r_1(x)$, ou seja, $g(x)$ divide $r_2(x) - r_1(x)$, mas, por hipótese, $q_1(x) \neq q_2(x)$, o que implica que $q_1(x) - q_2(x) \neq 0$ e $r_2(x) - r_1(x) \neq 0$ e, pelo teorema 4.1, o fato de $r_1(x) = 0$ ou $gr\big(r_1(x)\big) < gr\big(g(x)\big)$ e $r_2(x) = 0$ ou $gr\big(r_2(x)\big) < gr\big(g(x)\big)$, implica que $gr\big(g(x)\big) \leq gr\big(r_2(x) - r_1(x)\big) < gr\big(g(x)\big)$, o que é um absurdo, portanto, a hipótese é falsa o que implica que $q_1(x) = q_2(x)$ e $r_1(x) = r_2(x)$.

Sejam $p(x)$, $g(x)$, $q(x)$ e $r(x)$, polinômios pertencentes a $A[x]$, tais que $p(x) = q(x) \cdot g(x) + r(x)$, denominaremos $p(x)$ de dividendo, $g(x)$ de divisor, $q(x)$ de quociente e $r(x)$ de resto. Uma maneira prática para determinar o quociente e o resto da divisão euclidiana de um polinômio $p(x)$ por um polinômio $g(x)$ com coeficiente líder invertível é a utilização do algoritmo a seguir:

$$\begin{array}{l|l} p(x) & g(x) \\ \vdots & \hline q(x) \\ \hline r(x) & \end{array}$$

Exemplos:

a) Determinar o quociente e o resto da divisão euclidiana do polinômio $p(x) = 10 - 11x + x^2 - x^3 + x^4$ pelo polinômio $g(x) = 5 + 2x + x^2$, ambos pertencentes a $\mathbb{Z}[x]$.

Solução: Observemos que $gr\big(g(x)\big) < gr(p(x))$, além do coeficiente líder de $g(x)$ ser invertível em $\mathbb{Z}$.

$$\begin{array}{rrrrr|lll} x^4 & -x^3 & +x^2 & -11x & +10 & x^2 & x^2 & +5 \\ -x^4 & -2x^3 & -5x^2 & & & x^3 & -3x & +2 \\ \hline 0 & -3x^3 & -4x^2 & -11x & +10 & & & \\ & +3x^3 & +6x^2 & +15x & & & & \\ \hline & 0 & +2x^2 & +4x & +10 & & & \\ & & -2x^2 & -4x & -10 & & & \\ \hline & & 0 & 0 & 0 & & & \end{array}$$

Notemos que $q(x) = x^3 - 3x + 2$ e o fato de $r(x) = 0$ implica que $g(x)$ divide $p(x)$.

b) Determinar o quociente e o resto da divisão euclidiana do polinômio $p(x) = -1 + 4x - 2x^2 + 5x^3$ pelo polinômio $g(x) = -2 + 3x + x^2$, ambos pertencentes a $\mathbb{Z}[x]$.

Solução: Observemos que $gr\big(g(x)\big) < gr(p(x))$, além do coeficiente líder de $g(x)$ ser invertível em $\mathbb{Z}$.

$$\begin{array}{rrrr|lll} 5x^3 & -2x^2 & +4x & -1 & x^2 & +3x & -2 \\ -5x^3 & -15x^2 & +10x & & 5x & -17 & \\ \hline 0 & -17x^2 & +14x & -1 & & & \\ & +17x^2 & +41x & -34 & & & \\ \hline & 0 & +55x & -35 & & & \end{array}$$

Assim, $q(x) = -17 + 5x$ e $r(x) = -35 + 55x$.

Suponhamos agora que o conjunto dos coeficientes de um polinômio seja K, com todas as propriedades definidas anteriormente para o conjunto A, e mais a seguinte propriedade:

$\forall a \in K, a \neq 0, \exists b \in K; a \cdot b = 1$. O elemento b é denominado inverso multiplicativo de a e, por ser único, representamos como $b = a^{-1}$. Assim, temos um conjunto em, que todos os elementos não nulos são invertíveis[3].

Representamos por $K[x]$ o conjunto de todos os polinômios com coeficientes em K.

Considerando o conjunto A, e um polinômio $p(x) \in A[x]$, dizemos que α é uma raiz de $p(x)$, se $p(\alpha) = 0$.

Considerando o conjunto K e $p(x) \in K[x]$, formulamos o seguinte teorema:

Teorema 4.3: Considerando K o conjunto com as propriedades descritas anteriormente, seja $p(x)$ um polinômio pertencente a $K[x]$ e $\alpha \in K$, dizemos que α é uma raiz de $p(x)$ se, e somente se, $(x - \alpha)$ divide $p(x)$.

Demonstração: Pelo algoritmo da divisão em $K[x]$ (idêntico a $A[x]$), temos que existem $q(x)$ e $r(x)$ pertencentes a $K[x]$ tais que $p(x) = (x - \alpha).q(x) + r(x)$, com $r(x) = 0$ ou $gr\big(r(x)\big) = 0$. Assim, $r(x) = r \in K$. Logo, α é raiz de $p(x)$ se, e somente se, $0 = p(\alpha) = (\alpha - \alpha).q(x) + r = r$, ou seja, se, e somente se, $r = 0$, quem em outras palavras significa dizer que $(x - \alpha)$ divide $p(x)$.

O resultado acima se deve ao matemático francês Jean le Rond d'Alembert e, por isso, é popularmente conhecido como *teorema de d'Alembert*.

Exemplo: Consideremos o polinômio $p(x) = 3 - 7x - x^2 + x^3$. O valor $\alpha = 3$ é uma raiz de $p(x)$, pois $(x - 3)$ divide $p(x)$:

$$
\begin{array}{rrrr|l}
x^3 & -x^2 & -7x & +3 & x \quad -3 \\
\cline{5-5}
-x^3 & +3x^2 & & & x^2 \quad +2x \quad -1 \\
\cline{1-2}
0 & +2x^2 & -7x & & \\
 & -2x^2 & +6x & & \\
\cline{2-3}
 & 0 & -x & +3 & \\
 & & x & -3 & \\
\cline{3-4}
 & & 0 & 0 &
\end{array}
$$

Notemos ainda que $p(3) = 3 - 7 \cdot 3 - 3^2 + 3^3 = 3 - 21 - 9 + 27 = 0$.

Teorema 4.4: Um polinômio de grau n com coeficientes em um conjunto K com as propriedades definidas anteriormente, possui, no máximo, n raízes distintas nesse conjunto.

Demonstração: Sejam $\alpha_1,\ \alpha_2,\ \alpha_3, \dots,\ \alpha_m$, m raízes distintas em K de um polinômio $p(x)$ de grau n. Pelo teorema 4.3, temos que existe um polinômio $q_1(x) \in K[x]$ tal que $p(x) =$

[3] Um conjunto com as propriedades do conjunto K recebe o nome de corpo, como veremos mais a frente.

$(x-\alpha_1)\cdot q_1(x)$. Como α_2 é raiz de $p(x)$, então $p(\alpha_2)=(\alpha_2-\alpha_1)\cdot q_1(\alpha_2)$. Como $\alpha_1 \neq \alpha_2$, temos que $q_1(\alpha_2)=0$, ou seja, α_2 é raiz de $q_1(x)$, logo, existe $q_2(x)\in K[x]$ tal que $p(x)=(x-\alpha_1)\cdot(x-\alpha_2)\cdot q_2(x)$. Seguimos esse procedimento até obtermos $q_m(x)\in K[x]$ tal que $P(x)=(x-\alpha_1)\cdot(x-\alpha_2)\cdot(x-\alpha_3)\cdot\ldots\cdot(x-\alpha_m)\cdot q_m(x)$.

Do exposto acima, temos que $n=gr\big(p(x)\big)=m+gr(q_m(x))$, ou seja, $m\leq n$.

4.5 INTERPOLAÇÃO

Consideremos $K[x]$ o conjunto dos polinômios com coeficientes no conjunto K com as propriedades definidas anteriormente. Sejam $a_1,a_2,a_3,\ldots,a_n$ elementos de K, distintos dois a dois. Consideremos ainda $b_1,b_2,b_3,\ldots,b_n$ elementos quaisquer de K. Nosso objetivo é determinar um polinômio $p(x)\in K[x]$ de grau menor ou igual a $n-1$, tal que $p(a_i)=b_i$, $\forall i\in\{1,2,3,\ldots,n\}$.

Teorema 4.5: Dados os elementos $a_1,a_2,a_3,\ldots,a_n$ dois a dois distintos e $b_1,b_2,b_3,\ldots,b_n$ pertencentes a K, o Polinômio $p(x)=b_1p_1(x)+b_2p_2(x)+b_3p_3(x)+\cdots+b_np_n(x)$, onde $p_j(x)=\frac{(x-a_1)\cdot(x-a_2)\cdot\ldots\cdot(x-a_{j-1})\cdot(x-a_{j+1})\cdot\ldots\cdot(x-a_{n-1})\cdot(x-a_n)}{(a_j-a_1)\cdot(a_j-a_2)\cdot\ldots\cdot(a_j-a_{j-1})\cdot(a_j-a_{j+1})\cdot\ldots\cdot(a_j-a_{n-1})\cdot(a_j-a_n)}$, é o único polinômio em $K[x]$ tal que $p(a_i)=b_i$, $\forall i\in\{1,2,3,\ldots,n\}$.

Demonstração: Escrevendo o polinômio $p(x)$, temos:

$$p(x)=c_{n-1}x^{n-1}+\cdots+c_1x+c_0$$

Tal polinômio pode ser obtido resolvendo o seguinte sistema de equações:

$$\begin{cases} c_{n-1}{a_1}^{n-1}+\cdots+c_1a_1+c_0=b_1 \\ \vdots \\ c_{n-1}{a_2}^{n-1}+\cdots+c_1a_2+c_0=b_2 \\ c_{n-1}{a_n}^{n-1}+\cdots+c_1a_n+c_0=b_n \end{cases}$$

Notemos que o sistema acima possui n equações e n incógnitas $c_1,c_2,c_3,\ldots,c_n$ e, a medida que n assume valores maiores, maior é a dificuldade em buscar a sua solução. Notemos, porém, que o fato do sistema possuir n equações e n incógnitas, implica que ele admite pelo menos uma solução. Mais ainda, afirmamos que a solução é única, pois se considerarmos outro polinômio $q(x)\in K[x]$ tal que $q(a_i)=b_i$, $\forall i\in\{1,2,3,\ldots,n\}$, então o polinômio $p(x)-q(x)$, com grau menor ou igual a $n-1$, teria $a_1,a_2,a_3,\ldots,a_n$ como raízes, o que, em virtude do teorema 4.4, é possível somente se $p(x)-q(x)=0$, o que implica que $p(x)=q(x)$, logo, o polinômio $p(x)$ é único.

Para $j \in \{1, 2, 3, \dots, n\}$, definamos os polinômios de grau $n-1$,

$$p_j(x) = \frac{(x-a_1)\cdot(x-a_2)\cdot\ldots\cdot(x-a_{j-1})\cdot(x-a_{j+1})\cdot\ldots\cdot(x-a_{n-1})\cdot(x-a_n)}{(a_j-a_1)\cdot(a_j-a_2)\cdot\ldots\cdot(a_j-a_{j-1})\cdot(a_j-a_{j+1})\cdot\ldots\cdot(a_j-a_{n-1})\cdot(a_j-a_n)}$$

Temos $p_j(a_i) = \begin{cases} 0, se\ i \neq j \\ 1, se\ i = j \end{cases}$. Como cada polinômio $p_j(x)$ tem grau $n-1$, então a soma $b_1p_1(x) + b_2p_2(x) + b_3p_3(x) + \cdots + b_np_n(x)$ tem grau menor ou igual a $n-1$, além disso, satisfaz as condições $p(a_i) = b_i,\ \forall i \in \{1, 2, 3, \dots, n\}$, logo, assumimos $p(x) = b_1p_1(x) + b_2p_2(x) + b_3p_3(x) + \cdots + b_np_n(x)$.

O polinômio acima é chamado de *polinômio de interpolação* e o processo descrito para obtê-lo é denominado *interpolação de Lagrange.*

Exemplo: Determinemos o polinômio $p(x) \in K[x]$ tal que $p(1) = 3$, $p(3) = 2$, $p(4) = 1$ e $p(6) = 4$.

Solução:

Temos: $\alpha_1 = 1$, $\alpha_2 = 3$, $\alpha_3 = 4$ e $\alpha_4 = 6$, assim:

$$p_1(x) = \frac{(x-3)\cdot(x-4)\cdot(x-6)}{(1-3)\cdot(1-4)\cdot(1-6)} = \frac{-x^3+13x^2-54x+72}{30}$$

$$p_2(x) = \frac{(x-1)\cdot(x-4)\cdot(x-6)}{(3-1)\cdot(3-4)\cdot(3-6)} = \frac{x^3-11x^2+34x-24}{6}$$

$$p_3(x) = \frac{(x-1)\cdot(x-3)\cdot(x-6)}{(4-1)\cdot(4-3)\cdot(4-6)} = \frac{-x^3+10x^2-27x+18}{8}$$

$$p_4(x) = \frac{(x-1)\cdot(x-3)\cdot(x-4)}{(6-1)\cdot(6-3)\cdot(6-4)} = \frac{x^3-8x^2+19x-12}{30}$$

Como $b_1 = 3$, $b_2 = 2$, $b_3 = 1$ e $b_4 = 4$, então:

$$p(x) = 3\cdot\left(\frac{-x^3+13x^2-54x+72}{30}\right) + 2\cdot\left(\frac{x^3-11x^2+34x-24}{6}\right) + 1$$
$$\cdot\left(\frac{-x^3+10x^2-27x+18}{8}\right) + 4\cdot\left(\frac{x^3-8x^2+19x-12}{30}\right)$$

$$p(x) = \frac{-12x^3+156x^2-648x+864+40x^3-440x^2+1360x-960-15x^3}{120} +$$
$$+\frac{+150x^2405x+270+16x^3-128x^2+304x-192}{120}$$

$$p(x) = \frac{29x^3-262x^2+611x-18}{120}$$

Portanto, $p(x) = \frac{29}{120}x^3 - \frac{131}{60}x^2 + \frac{611}{120}x - \frac{3}{20}$.

5 ESTRUTURAS ALGÉBRICAS ELEMENTARES

Abordaremos a seguir as principais estruturas algébricas elementares, destacando suas características e as propriedades que as definem.

5.1 LEI DE COMPOSIÇÃO INTERNA

Consideremos um conjunto não vazio B. Uma função f de $B \times B$ em B, que a cada par $(x, y) \in B \times B$ faz corresponder o elemento $x \oplus y \in B$, é denominada uma *lei de composição interna em* B. Assim, dizemos que B é um conjunto munido da operação $\oplus$.

Exemplo: Consideremos o conjunto $\mathbb{N}$ dos números naturais e seja f de $\mathbb{N} \times \mathbb{N}$ em $\mathbb{N}$, a função que a cada par $(x, y) \in \mathbb{N} \times \mathbb{N}$, faz corresponder ao elemento $x + y$ que também é um número natural. Assim, o conjunto $\mathbb{N}$ dos números naturais é munido da operação + (*adição*), ou ainda a *adição* é uma lei de composição interna em $\mathbb{N}$.

5.2 GRUPOS

Consideremos um conjunto G munido da operação $\oplus$. Dizemos que G é um grupo em relação à lei de composição interna $\oplus$, se, para todo $x, y, z \in G$, em relação a $\oplus$ são observadas as seguintes propriedades:

I) $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ (associatividade)

II) $\exists e \in G$ tal que $x \oplus e = e \oplus x = x$ (elemento neutro)

III) $\exists x' \in G$ tal que $x \oplus x' = x' \oplus x = e$ (todo elemento de G possui simétrico aditivo)

Quando a lei de composição interna em G for a adição $(+)$, dizemos que G é um grupo aditivo e quando a lei de composição interna em G for a multiplicação $(\cdot)$, dizemos que G é um grupo multiplicativo.

Representaremos por $(G, \oplus)$ um grupo com a lei de composição interna $\oplus$.

Se além das três propriedades mencionadas acima, dados $x, y \in G$, ocorrer que $x \oplus y = y \oplus x$, então dizemos que G é um *grupo comutativo* ou *grupo abeliano*[4].

[4] Homenagem ao matemático norueguês Niels Henrik Abel (1802-1829).

Exemplos:

1) Grupo aditivo das matrizes reais quadradas de ordem n, ou seja, $(\mathcal{M}_n, +)$.

Vimos no estudo das matrizes que a adição goza das propriedades associativa, elemento neutro (matriz nula) e simétrico aditivo (matriz oposta), portanto, o conjunto $\mathcal{M}_n$ das matrizes quadradas de ordem n, munido da operação usual de adição é um grupo. Além disso, vimos que além das três propriedades mencionadas acima, a adição de matrizes é comutativa, portanto, $(\mathcal{M}_n, +)$ é um grupo abeliano.

2) O conjunto $A[x]$ de todos os polinômios com coeficientes em A (considere A com as propriedades descritas no capítulo sobre polinômios), visto anteriormente, é um grupo aditivo para a operação usual de adição, pois vimos que a adição de polinômios em $A[x]$ é associativa, existe elemento neutro (polinômio nulo) e todo polinômio em $A[x]$ é simetrizável. Portanto, $(A[x], +)$ é um grupo. Além disso, $(A[x], +)$ é grupo abeliano, pois a adição em $A[x]$ é comutativa conforme demonstrado anteriormente.

3) O conjunto dos números inteiros para a operação usual de adição é um grupo abeliano, pois dados $x, y, z \in \mathbb{Z}$, temos $x + (y + z) = (x + y) + z$, $0 \in \mathbb{Z}$ e $0 + x = x + 0 = x$ além de $-x \in \mathbb{Z}$ e $-x + x = x + (-x) = 0$. Notemos ainda que em $\mathbb{Z}$, $x + y = y + x$, portanto, $(\mathbb{Z}, +)$ é um grupo abeliano.

4) O conjunto $\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$, dos restos das divisões euclidianas de um numero inteiro qualquer por 3 é um grupo abeliano em relação a operação de adição definida por $\bar{x} + \bar{y} = \overline{x + y}$, $\forall \bar{x}, \bar{y} \in \mathbb{Z}_3$. Observemos a tábua de operação em $\mathbb{Z}_3$:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

Notemos que quaisquer que sejam $x, y, z \in \mathbb{Z}_3$, temos $x + (y + z) = (x + y) + z$, $x + y = y + x$, $\bar{0}$ é o elemento neutro da adição, pois $x + \bar{0} = \bar{0} + x = x$ além de $\bar{1}$ e $\bar{2}$ serem os simétricos aditivos, respectivamente, de $\bar{2}$ e $\bar{1}$. Portanto, $(\mathbb{Z}_3, +)$ é um grupo abeliano.

5) O conjunto $G = \{-1,1\}$, munido da lei de composição interna multiplicação usual é um grupo abeliano, vejamos a tábua da operação:

$\cdot$	-1	1
-1	1	-1
1	-1	1

Notemos que para todo $x, y, z \in G$, temos $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, $x.y = y.x$, $x.1 = 1.x = x$, além de -1 e 1 serem os simétricos multiplicativos (inversos), respectivamente de -1 e 1.

Portanto, $(G, \cdot)$ é um grupo abeliano.

6) O Conjunto $\mathbb{N} = \{1, 2, 3, \dots\}$ com a operação usual de adição não é um grupo, pois, embora a adição de números naturais seja associativa, esse conjunto não possui elemento neutro aditivo e seus elementos não são simetrizáveis na adição.

De maneira análoga, o conjunto $\mathbb{N} = \{1, 2, 3, \dots\}$ com a operação de multiplicação usual não é um grupo, pois, embora a multiplicação de números naturais seja associativa, e o número 1 seja o elemento neutro multiplicativo, o único elemento simetrizável (invertível) na multiplicação é o número 1.

7) O conjunto $\mathbb{Z}$ dos números inteiros com a multiplicação usual não é um grupo, pois embora a multiplicação de números inteiros seja associativa e o número 1 seja o elemento neutro multiplicativo, somente os elementos 1 e -1 são invertíveis.

8) O conjunto $\mathcal{M}_n$ das matrizes quadradas de ordem n, com a operação de multiplicação usual de matrizes não é um grupo, pois embora a multiplicação de matrizes seja associativa e a matriz I_n (matriz identidade de ordem n) seja o elemento neutro multiplicativo, nem todas as matrizes quadradas de ordem n são invertíveis.

9) O conjunto $\mathbb{Z}_4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$, dos restos das divisões euclidianas de um número inteiro qualquer por 4 não é um grupo em relação à operação usual de multiplicação, pois embora a multiplicação usual seja associativa em $\mathbb{Z}_4$ e o número 1 seja o elemento neutro multiplicativo, o elemento 2 não é invertível.

5.2.1 Subgrupos

Considerando $(G, \oplus)$ um grupo e H um subconjunto não vazio de G, dizemos que H é um *subgrupo* de G se dados $x, y \in H$ temos $x \oplus y \in H$ e além disso, $(H, \oplus)$ é também um grupo.

Exemplo:

Considerando o Grupo aditivo dos números reais $(\mathbb{R}, +)$, temos que $(\mathbb{Z}, +)$ é um subgrupo de $\mathbb{R}$, pois $\mathbb{Z} \subset \mathbb{R}$, dados $x, y \in \mathbb{Z}$ temos $x + y \in \mathbb{Z}$ além de a adição nos inteiros ser associativa, possuir elemento neutro e, para todo $x \in \mathbb{Z}$, temos $-x \in \mathbb{Z}$ tal que $-x + x = x + (-x) = 0$, isto é, todo elemento de $\mathbb{Z}$ possui simétrico.

Notemos que se $(H, \oplus)$ é um subgrupo $(G, \oplus)$ e, sendo e_h e e os elementos neutros de H e G, respectivamente, então é fácil verificar que $e_h = e$, pois temos que $e_h \oplus x = x = e \oplus x$ operando a direita da dos membros da igualdade com o elemento x' que é o simétrico de x,

temos $e_h \oplus (x \oplus x') = e \oplus (x \oplus x')$ o que implica que $e_h \oplus e = e \oplus e$, concluindo com isso que $e_h = e$.

O teorema a seguir constitui uma ferramenta fácil para verificar se um subconjunto não vazio $H \subset G$ é um subgrupo de G em relação a uma lei de composição interna $\oplus$ de G:

Teorema 5.1: Seja $(G, \oplus)$ um grupo. Um conjunto não vazio $H \subset G$ é um subgrupo de G se, e somente se, $\forall x, y \in H$, temos $x \oplus y' \in H$. Onde y' representa o simétrico de y.

Demonstração:

($\Longrightarrow$) Se $H \subset G$ é um subgrupo de G então, $\forall x, y \in H$, vale que $y' \in H$ e, sendo $\oplus$ uma operação definida em H, então $x \oplus y' \in H$.

($\Longleftarrow$) Suponhamos que $\forall x, y \in H$, $x \oplus y' \in H$. Tomando $y = x$, temos que $x \oplus x' = e \in H$. Por hipótese, e pelo fato de $e \in H$, temos que $e \oplus y' = y' \in H$. Com isso garantimos a existência do elemento neutro da operação $\oplus$ em H, além de mostrar que todos os elementos de H são simetrizáveis em relação a essa operação. Dados $x, y \in H$, em virtude do que foi visto, temos que $x \oplus (y')' = x \oplus y \in H$, garantindo que H é fechado para a operação $\oplus$ que é lei de composição interna de G. Além disso, por herança, a igualdade $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ é válida em H. Portanto, H é um subgrupo de G.

5.3 ANÉIS

Consideremos um conjunto A não vazio munido das leis de composição internas $+$ (adição) e $\cdot$ (multiplicação).

Dizemos que A é um anel se, em relação à adição em A, for um grupo abeliano, ou seja, $\forall x, y, z \in A$:

I) $x + (y + z) = (x + y) + z$ (associatividade)

II) $\exists e \in G$ tal que $x + e = e + x = x$ (elemento neutro)

III) $\exists x' \in G$ tal que $x + x' = x' + x = e$ (todo elemento de G é simetrizável)

IV) $x + y = y + x$ (comutatividade)

e, se em relação a multiplicação, temos $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ (associatividade). Além disso, a multiplicação é distributiva em relação à adição, ou seja, $x \cdot (y + z) = x \cdot y + x \cdot z$.

Nas condições expostas acima, dizemos que A é um anel e representamos isso por $(A, +, \cdot)$.

Se para todo $x, y \in A$, tivermos $x \cdot y = y \cdot x$, dizemos que A é um anel comutativo. Além disso, se existir $u \in A$ tal que para todo $x \in A$ $u \cdot x = x \cdot u = x$, então A é um anel com unidade.

Em um anel A, comutativo com unidade, onde para todo $x, y \in A$, se $x.y = 0$ implicar que $x = 0$ ou $y = 0$, então dizemos que A é um *anel de integridade* ou um *domínio de integridade*. Decorre dessa observação, uma propriedade dos domínios de integridade, que conhecemos como *lei do anulamento do produto* e enunciaremos a seguir:

Sejam $x, y, z \in A$ e $(A, +, \cdot)$ é um domínio de integridade, se $z \neq 0$ e $x \cdot z = y \cdot z$, então $x = y$.

Demonstração:

Se $x, y, z \in A$ e $(A, +, \cdot)$ é um domínio de integridade, então cada elemento de A possui simétrico aditivo, além de a multiplicação ser distributiva em relação à adição e A não possuir divisores próprios de zero. Sendo assim, somando $-y \cdot z$ a ambos os membros da igualdade $x \cdot z = y \cdot z$, temos $x \cdot z - y.z = y \cdot z - y.z$, o que implica que $(x - y) \cdot z = 0$. Como $z \neq 0$, então $x - y = 0$. Somando y a ambos os membros da igualdade $x - y = 0$, temos $x - y + y = 0 + y$, o que implica que $x = y$.

Exemplos:

1) O Conjunto $\mathcal{M}_2$ das matrizes reais, quadradas de ordem 2, munido das operações usuais de adição e multiplicação é um anel, pois em relação a adição é um grupo abeliano. Em relação à multiplicação, temos a propriedade associativa e a multiplicação é distributiva em relação à adição. Portanto, $(\mathcal{M}_2, +, \cdot)$ é um anel. Além disso, a matriz identidade I_2 é o elemento neutro multiplicativo, portanto, $(\mathcal{M}_2, +, \cdot)$ é anel com unidade. Notemos, porém, que $(\mathcal{M}_2, +, \cdot)$ não é comutativo e apresenta divisores próprios de zero, vejamos: $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ e $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Logo, $(\mathcal{M}_2, +, \cdot)$ não é domínio de integridade.

2) O conjunto $(\mathbb{Z}, +, \cdot)$ é um domínio de integridade, pois $\mathbb{Z}$ é um grupo abeliano em relação à adição, como já vimos e, a multiplicação em $\mathbb{Z}$ é associativa, é distributiva em relação à adição, é comutativa e o número 1 é a unidade. Além disso, dados $x, y \in \mathbb{Z}$, se $x.y = 0$, então $x = 0$ ou $y = 0$.

3) O conjunto A dos coeficientes dos polinômios do conjunto $A[x]$, com as propriedades apresentadas no capítulo 4 é um domínio de integridade.

4) O conjunto $\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$, dos restos das divisões euclidianas de um numero inteiro qualquer por 3 é um domínio de integridade, observe as tábuas de operações em $\mathbb{Z}_3$:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Notemos que em relação à adição $\mathbb{Z}_3$ é um grupo abeliano e que a multiplicação é associativa, comutativa e possui a unidade. Além disso, a multiplicação é distributiva em relação à adição e $\mathbb{Z}_3$ não possui divisores próprios de zero.

5) O conjunto $\mathbb{Z}_4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$, dos restos das divisões euclidianas de um número inteiro qualquer por 4 é um anel comutativo com unidade, porém não é um domínio de integridade, pois em $\mathbb{Z}_4$ temos $2.2 = 4 = 0$, ou seja, $\mathbb{Z}_4$ possui divisores próprios de zero.

6) O conjunto $A[x]$ dos polinômios com coeficientes no domínio de integridade A, é também um domínio de integridade. Como vimos em 4.2 e 4.3, em relação à adição, $A[x]$ possui as propriedades associativa, comutativa, elemento neutro aditivo (polinômio nulo) e cada polinômio em $A[x]$ possui um polinômio simétrico. A multiplicação de polinômio é associativa, é comutativa, existe o polinômio constante $p(x) = 1$ que é a unidade multiplicativa e, além disso, se $p(x).q(x) = 0$ então $p(x)$ ou $q(x)$ é o polinômio nulo.

Observemos que o conjunto dos restos das divisões de um inteiro qualquer por m, o conjunto $\mathbb{Z}_m = \{\bar{1}, \bar{2}, \bar{3}, \dots, \overline{m-1}\}$ é um anel comutativo com unidade. A seguir demonstraremos um importante teorema sobre anéis:

Teorema 5.2: O anel comutativo com unidade $\mathbb{Z}_m = \{\bar{1}, \bar{2}, \bar{3}, \dots, \overline{m-1}\}$ é um domínio de integridade se, e somente se m é primo.

Demonstração:

($\Rightarrow$) Se m não for primo, então existem x e y pertencentes a $\mathbb{Z}$ de tal forma que tal que $x \cdot y = m$, com $1 < x < y < m$, o que implica que $\bar{x}, \bar{y} \in \mathbb{Z}_m$ e $\bar{x} \cdot \bar{y} = \bar{m} = \bar{0}$, ou seja, $\mathbb{Z}_m$ possui divisores próprios de zero e, portanto, não é um domínio de integridade.

($\Leftarrow$) Suponhamos que existam $\bar{x}, \bar{y} \in \mathbb{Z}_m$, de modo que $\bar{x} \cdot \bar{y} = \overline{x \cdot y} = \bar{0}$, então $x \cdot y = k \cdot m$, $k \in \mathbb{Z}$. Decorre desse fato que $m|x \cdot y$. Como m é um número primo, então $m|x$ ou $m|y$, o que implica que $x = a \cdot m$ ou $y = b \cdot m$, $a, b \in \mathbb{Z}$, portanto $\bar{x} = \bar{0}$ ou $\bar{y} = \bar{0}$, logo, $\mathbb{Z}_m$ é um domínio de integridade.

5.3.1 Subanéis

Considerando $(A, +, \cdot)$ um anel, dizemos que um subconjunto não vazio $B \subset A$ é um subanel de A, se B é fechado para as duas leis de composição interna de A, ou seja, dados $x, y \in B$, temos $x + y \in B$ e $x \cdot y \in B$ e, além disso, B for um anel em relação às operações + e $\cdot$.

Exemplo:

Considerando o anel $(\mathbb{Z}, +, \cdot)$ dos inteiros, o subconjunto $2\mathbb{Z}$ dos números inteiros pares é um subanel de $\mathbb{Z}$. Verifiquemos:

Para todo $x, y \in 2\mathbb{Z}$, temos $x = 2a$ e $y = 2b$, $a, b \in \mathbb{Z}$. Sendo assim, $x + y = 2a + 2b = 2(a + b) \in 2\mathbb{Z}$ e $x.y = 2a.2b = 4ab = 2(2ab) \in 2\mathbb{Z}$, portanto, a adição e a multiplicação de $\mathbb{Z}$ são fechadas em $2\mathbb{Z}$. Além disso, para todo $x, y, z \in 2\mathbb{Z}$, temos $x = 2a$, $y = 2b$ e $z = 2c$, com $a, b, c \in \mathbb{Z}$, assim, $x + (y + z) = 2a + (2b + 2c) = 2a + 2(b + c) = 2[a + (b + c)] = 2[(a + b) + c] = 2(a + b) + 2c = (2a + 2b) + 2c = (x + y) + z$; $0 = 2.0 \in 2\mathbb{Z}$ e é o elemento neutro da adição pois $x + 0 = 2a + 2.0 = 2(a + 0) = 2a = x$; para todo $x \in 2\mathbb{Z}$, fazendo $x + x' = 0$, temos $2a + x' = 0$, como $2a \in \mathbb{Z}$ e $\mathbb{Z}$ é uma anel, então $-2a$ é o simétrico de $2a$ em $\mathbb{Z}$. Somando $-2a$ em ambos os membros da igualdade $2a + x' = 0$ temos $-2a + 2a + x' = -2a + 0$, o que implica que $x' = -2a = -x \in 2\mathbb{Z}$; temos ainda que $x + y = 2a + 2b = 2(a + b) = 2(b + a) = 2b + 2a = y + x$. Mostrando com isso que $(2\mathbb{Z}, +)$ é um grupo abeliano.

Com relação à multiplicação, temos $x.(y.z) = 2a.(2b.2c) = (2a.2b).2c = (x.y).z$. Temos ainda $x.(y + z) = 2a.(2b + 2c) = 2a.2.b + 2a.2c = x.y + x.z$. Portanto, $(2\mathbb{Z}, +, \cdot)$ é um anel e, portanto, um subanel de $\mathbb{Z}$.

Notemos ainda que para todo $x, y \in \mathbb{Z}$, temos $x + y = 2a + 2b = 2(a + b) = 2(b + a) = 2b + 2a = y + x$, o que implica que $(2\mathbb{Z}, +, \cdot)$ é um anel comutativo, porém, enquanto $\mathbb{Z}$ é um domínio de integridade, $2\mathbb{Z}$ não o é, pois não é um anel com unidade.

O teorema a seguir constitui uma ferramenta útil e fácil para verificar se um subconjunto não vazio B é um subanel de A:

Teorema 5.3: Seja $(A, +, \cdot)$ um anel. Um conjunto não vazio $B \subset A$ é um subanel de A se, e somente se, $\forall x, y \in B$, temos $x - y \in B$ e $x.y \in B$.

Demonstração:

($\Longrightarrow$) Consideremos B um subanel de A, então, por hipótese, $(B,+)$ é um subgrupo abeliano de $(A,+)$ e, portanto, para todo $x,y \in B$, temos $x-y \in B$. Considerando ainda a hipótese, temos que $x.y \in B$, pois em um anel a multiplicação é fechada. Portanto, se B um subanel de A e $x,y \in B$, então $x-y \in B$ e $x.y \in B$.

($\Longleftarrow$) Consideremos, por hipótese, que dados $x,y \in B$, temos $x-y \in B$ e $x.y \in B$.

Como $x-y \in B$, então $(B,+)$ é um subgrupo de $(A,+)$ e, como $(A,+)$ é abeliano, então $(B,+)$ também é abeliano. Como $B \subset A$, então, para todo $x,y,z \in B$, temos que $x,y,z \in A$ e, portanto, $x.(y.z) = (x.y).z \in A$ pois A é um anel. Mas, por hipótese, a multiplicação de A é fechada em B, portanto, $x.(y.z) = (x.y).z \in B$, logo, a multiplicação é associativa em B. Além disso, pelo mesmo motivo, temos $x.(y+z) = x.y + x.z \in B$, portanto, $(B,+,\cdot)$ é um subanel de $(A,+,\cdot)$.

5.4 IDEAIS

Seja A um anel. Um subconjunto não vazio $I \subset A$ é denominado um *ideal à esquerda* de A se I é um subanel de A e $\forall x,y \in I$ e $a \in A$, temos $x-y \in I$ e $a \cdot x \in I$.

De maneira análoga, dizemos que um subconjunto não vazio $I \subset A$ é um *ideal à direita* de A se I é um subanel de A e $\forall x,y \in I$ e $a \in A$, temos $x-y \in I$ e $x \cdot a \in I$.

Quando A é um anel comutativo os ideais à esquerda e à direita coincidem e dizemos então que I é um ideal de A.

Dado um anel A, os subconjuntos $\{0\}$ e A são denominados *ideais triviais* ou *ideais próprios* de A.

Vejamos:

1) $\forall x,y \in \{0\}$ temos $x-y = 0-0 = 0 \in \{0\}$, além disso, dados $a \in A$ e $x \in \{0\}$, temos que $a \cdot x = a \cdot 0 = 0 \in \{0\}$ e $x \cdot a = x \cdot 0 = 0 \in \{0\}$ que nos mostra que $\{0\}$ é ideal à esquerda e à direita de A. Portanto, $\{0\}$ é um ideal de A.

2) Como A é um anel e $A \subset A$, então A é um subanel de A. É evidente que dados $a \in A$ e $x \in A$, temos que $a \cdot x \in A$ e $x \cdot a \in A$ que nos mostra que A é ideal à esquerda e à direita de A. Portanto, A é um ideal de A.

Seja A um anel comutativo. Sejam $a_1, a_2, a_3, \dots, a_n \in A; n \geq 1$. O conjunto $\langle a_1, a_2, \dots, a_n \rangle \subset A$, definido como:

$\langle a_1, a_2, \dots, a_n \rangle = \{x_1 \cdot a_1 + x_2 \cdot a_2 + \cdots + x_n \cdot a_n;\ x_i \in A, 1 \leq i \leq n\}$ é um ideal em A.

Vejamos:

$0 = 0 \cdot a_1 + 0 \cdot a_2$+...+$0 \cdot a_n$, portanto, $0 \in \langle a_1, a_2, \dots, a_n \rangle$.

Dados $m, n \in \langle a_1, a_2, \dots, a_n \rangle$, então existem $x_i, y_i \in A, 1 \leq i \leq n$ tais que $m = x_1 \cdot a_1 + x_2 \cdot a_2 + \cdots + x_n \cdot a_n$ e $n = y_1 \cdot a_1 + y_2 \cdot a_2 + \cdots + y_n \cdot a_n$, logo, teremos $m - n = (x_1 - y_1) \cdot a_1 + (x_2 - y_2) \cdot a_2 + \cdots + (x_n - y_n) \cdot a_n$. Como $x_i - y_i \in A$, então $m - n \in \langle a_1, a_2, \dots, a_n \rangle$. Seja $\alpha \in A$ e $m \in \langle a_1, a_2, \dots, a_n \rangle$, então existem $x_i \in A, 1 \leq i \leq n$ tal que $m = x_1 \cdot a_1 + x_2 \cdot a_2 + \cdots + x_n \cdot a_n$. Assim, temos $\alpha \cdot m = \alpha \cdot x_1 \cdot a_1 + \alpha \cdot x_2 \cdot a_2 + \cdots + \alpha \cdot x_n \cdot a_n = x_1 \cdot (\alpha \cdot a_1) + x_2 \cdot (\alpha \cdot a_2) + \cdots + x_n \cdot (\alpha \cdot a_n) = x_1 \cdot b_1 + x_2 \cdot b_2 + \cdots + x_n \cdot b_n$. Como $b_i = \alpha \cdot a_i \in A$, então $x_1 \cdot b_1 + x_2 \cdot b_2 + \cdots + x_n \cdot b_n \in \langle a_1, a_2, \dots, a_n \rangle$, portanto $\alpha \cdot m \in \langle a_1, a_2, \dots, a_n \rangle$, o que mostra que $\langle a_1, a_2, \dots, a_n \rangle$ é um ideal em A.

O ideal $\langle a_1, a_2, \dots, a_n \rangle$ obtido acima é denominado *ideal gerado por* $a_1, a_2, \dots, a_n$. Um ideal gerado por um só elemento a do anel A, representado $\langle a \rangle$ é denominado *ideal principal* gerado por a.

Exemplos:

1) Consideremos o anel das matrizes reais quadradas de ordem 2 com as leis de composição interna adição e a multiplicação usuais, ou seja, $(\mathcal{M}, +, .)$. Consideremos os conjuntos $I \subset \mathcal{M}$ tal que $I = \left\{A = [a_{ij}]_2;\ a_{ij} = 0, se\ j \neq 1\ e\ a_{i1} \neq 0\ para\ algum\ i\right\}$ e $I' \subset \mathcal{M}$ tal que $I' = \left\{B = [b_{ij}]_2;\ b_{ij} = 0, se\ i \neq 1\ e\ b_{1j} \neq 0\ para\ algum\ j\right\}$. Mostraremos que I é um ideal à esquerda de $\mathcal{M}$ e I' é um ideal à direita de $\mathcal{M}$.

Seja $M \in \mathcal{M}$ uma matriz qualquer tal que $M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$. Consideremos ainda as matrizes $A = \begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix}$ e $A' = \begin{bmatrix} a'_{11} & 0 \\ a'_{21} & 0 \end{bmatrix}$ pertencentes ao conjunto I. Fazendo $A - A'$, temos

$$\begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix} - \begin{bmatrix} a'_{11} & 0 \\ a'_{21} & 0 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix} + \begin{bmatrix} -a'_{11} & 0 \\ -a'_{21} & 0 \end{bmatrix} = \begin{bmatrix} a_{11} - a'_{11} & 0 \\ a_{21} - a'_{21} & 0 \end{bmatrix} \in I$$

Fazendo $M.A$ temos $\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} . \begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix} = \begin{bmatrix} m_{11}.a_{11} + m_{12}.a_{21} & m_{11}.0 + m_{12}.0 \\ m_{21}.a_{11} + m_{22}.a_{21} & m_{21}.0 + m_{22}.0 \end{bmatrix} =$

$= \begin{bmatrix} m_{11}.a_{11} + m_{12}.a_{21} & 0 \\ m_{21}.a_{11} + m_{22}.a_{21} & 0 \end{bmatrix} \in I$. Notemos porém que $\begin{bmatrix} a_{11} & 0 \\ a_{21} & 0 \end{bmatrix} . \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} =$

$= \begin{bmatrix} a_{11}.m_{11} + 0.m_{21} & a_{11}.m_{12} + 0.m_{22} \\ a_{21}.m_{11} + 0.m_{21} & a_{21}.m_{12} + 0.m_{22} \end{bmatrix} = \begin{bmatrix} a_{11}.m_{11} & a_{11}.m_{12} \\ a_{21}.m_{11} & a_{21}.m_{12} \end{bmatrix} \notin I$. Deduzimos com isso que I é um ideal à esquerda de $\mathcal{M}$.

De maneira análoga, considerando $M \in \mathcal{M}$ uma matriz qualquer tal que $M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$. Considerando ainda as matrizes $B = \begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix}$ e $B' = \begin{bmatrix} b'_{11} & b'_{12} \\ 0 & 0 \end{bmatrix}$ pertencentes ao conjunto I'. Fazendo $B - B'$, temos $\begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} b'_{11} & b'_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -b'_{11} & -b'_{12} \\ 0 & 0 \end{bmatrix} =$

$= \begin{bmatrix} b_{11}-b'_{11} & b_{12}-b'_{12} \\ 0 & 0 \end{bmatrix} \in I'$.

Fazendo $B.M$, temos:

$\begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix} . \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} b_{11}.m_{11} + b_{12}.m_{21} & b_{11}.m_{12} + b_{12}.m_{22} \\ 0.m_{11} + 0.m_{21} & 0.m_{12} + 0.m_{22} \end{bmatrix} =$

$= \begin{bmatrix} b_{11}.m_{11} + b_{12}.m_{21} & b_{11}.m_{12} + b_{12}.m_{22} \\ 0 & 0 \end{bmatrix} \in I'$. Em contrapartida, fazendo $M.B$, temos

$\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} . \begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} m_{11}.b_{11} + m_{12}.0 & m_{11}.b_{12} + m_{12}.0 \\ m_{21}.b_{11} + m_{22}.0 & m_{21}.b_{12} + m_{22}.0 \end{bmatrix} =$

$= \begin{bmatrix} m_{11}.b_{11} & m_{11}.b_{12} \\ m_{21}.b_{11} & m_{21}.b_{12} \end{bmatrix} \notin I'$. Deduzimos com isso que I' é um ideal à direita de $\mathcal{M}$.

Notemos porém que I e I' não são ideais de $\mathcal{M}$.

2) No anel $(\mathbb{Z}, +, \cdot)$ dos inteiros com adição e multiplicação usuais, qualquer subconjunto $n\mathbb{Z} = \{n.x; x \in \mathbb{Z}\}$ e n um inteiro fixo é um ideal de $\mathbb{Z}$, pois dados $y_1, y_2 \in n\mathbb{Z}$, temos $y_1 = n.x_1$ e $y_2 = n.x_2$, com $x, x_2 \in \mathbb{Z}$. Assim, $y_1 - y_2 = n.x_1 - n.x_2 = n.(x_1 - x_2) \in n\mathbb{Z}$. Além disso, dados $a \in \mathbb{Z}$ e $y \in n\mathbb{Z}$, temos $y = n.x$, com $x \in \mathbb{Z}$, portanto, $a.y = a.(n.x) = (a.n).x = (n.a).x = n.(a.x) \in n\mathbb{Z}$, o que mostra que $n\mathbb{Z}$ é um ideal à esquerda de $\mathbb{Z}$. Não é necessário verificar se $n\mathbb{Z}$ é um ideal à direita de $\mathbb{Z}$, uma vez que $\mathbb{Z}$ é um anel comutativo. Assim, $n\mathbb{Z}$ é um ideal de $\mathbb{Z}$ para todo $n \in \mathbb{Z}$. Além disso, como o ideal $n\mathbb{Z}$ é gerado por n, então $n\mathbb{Z} = \langle n \rangle$, ou seja, $n\mathbb{Z}$ é um ideal principal.

5.5 CORPOS

Consideremos um anel K, comutativo e com unidade. K é denominado um *corpo*, se para todo $x \in K$, $x \neq 0$, existe $y \in K$ tal que $x \cdot y = 1$, ou seja, todo elemento não nulo de K admite simétrico multiplicativo. Note que utilizamos 0 e 1 como os elementos neutro da adição e multiplicação respectivamente no corpo K, não devendo ser confundidos com os números 0 e 1.

Ao elemento $y \in K$ tal que $x \cdot y = 1$, que é o simétrico multiplicativo de x, denominaremos *inverso* de x e o representaremos por x^{-1}.

Exemplos:

1) O anel comutativo com unidade dos números reais, ou seja, $(\mathbb{R}, +, \cdot)$ é um corpo, pois para todo $x \in \mathbb{R}, x \neq 0$, existe $x^{-1} = \frac{1}{x} \in \mathbb{R}$ tal que $x \cdot x^{-1} = x \cdot \frac{1}{x} = 1$.

2) O anel comutativo com unidade dos números racionais, ou seja, $(\mathbb{Q}, +, \cdot)$ é um corpo, pois para todo $x \in \mathbb{Q}, x \neq 0$, existe $x^{-1} = \frac{1}{x} \in \mathbb{Q}$ tal que $x \cdot x^{-1} = x \cdot \frac{1}{x} = 1$

3) O anel comutativo com unidade dos números complexos, ou seja, $(\mathbb{C}, +, \cdot)$ é um corpo, pois para todo $z \in \mathbb{C}, z = x + yi, z \neq 0$, existe $z^{-1} = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2} i \in \mathbb{C}$ tal que $z \cdot z^{-1} = (x + yi).\left(\frac{x}{x^2+y^2} - \frac{y}{x^2+y^2}\right) = \frac{x^2}{x^2+y^2} - \frac{xyi}{x^2+y^2} + \frac{xyi}{x^2+y^2} + \frac{y^2}{x^2+y^2} = \frac{x^2+y^2}{x^2+y^2} = 1$.

4) O anel comutativo com unidade dos números inteiros, ou seja, $(\mathbb{Z}, +, \cdot)$ não é um corpo, pois somente os elementos $\{-1, 1\}$ possuem inversos.

5) O anel comutativo com unidade do conjunto dos restos das divisões de um inteiro por 3, com as operações de adição e multiplicação usuais, ou seja, $(\mathbb{Z}_3, +, \cdot)$, é um corpo, pois

como vimos anteriormente $\mathbb{Z}_3$ é um anel comutativo com unidade e, observando a tábua da multiplicação em $\mathbb{Z}_3$:

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

vemos que $\bar{1} \cdot \bar{1} = \overline{1.1} = \bar{1}$ e $\bar{2} \cdot \bar{2} = \overline{2.2} = \bar{4} = \bar{1}$, ou seja, todo elemento não nulo de $\mathbb{Z}_3$ possui inverso.

6) O anel comutativo com unidade do conjunto dos restos das divisões de um inteiro por 4, com as operações de adição e multiplicação usuais, ou seja, $(\mathbb{Z}_4, +, \cdot)$, não é um corpo. Vimos anteriormente que $\mathbb{Z}_4$ é um anel comutativo com unidade, mas, observando a tábua da multiplicação em $\mathbb{Z}_4$:

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

vemos que dos elementos não nulos $\bar{1}, \bar{2}$ e $\bar{3}$, que o $\bar{1}$ e o $\bar{3}$ possuem inversos, enquanto que o $\bar{2}$ não possui inverso, logo, $(\mathbb{Z}_4, +, \cdot)$ não é um corpo.

7) O conjunto $F = \{0,1\}$ cujas tábuas de operações de adição e multiplicação apresentamos a seguir, é um corpo[5]:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Teorema 5.4: Todo corpo é um domínio de integridade.

Demonstração: Suponhamos que $(K, +, \cdot)$ seja um corpo. Por simplicidade diremos apenas "o corpo K", ficando subentendidas suas leis de composição internas. A hipótese nos garante que K é um anel comutativo com unidade e que todos os elementos não nulos de K possuem inversos. Sendo assim, consideremos $x, y \in K$, com, por exemplo, $x \neq 0$, tal que $x \cdot y = 0$, então, existe $x^{-1} \in K$ tal que $x \cdot x^{-1} = 1$. Multiplicando à esquerda ambos os membros da igualdade $x \cdot y = 0$ por x^{-1}, temos $x^{-1} \cdot (x.y) = x^{-1} \cdot 0$ o que implica que $(x^{-1} \cdot x).y = 0$, implicando que $y = 0$ e, portanto, K não possui divisores próprios de zero, ou seja, $(K, +, \cdot)$ é um domínio de integridade.

A recíproca do teorema acima é falsa, pois, por exemplo, $\mathbb{Z}$ é um domínio de integridade mas não é um corpo.

Teorema 5.5: Todo domínio de integridade finito é um corpo.

Demonstração: Consideremos $(A, +, \cdot)$ tal que $A = \{x_1, x_2, x_3, \cdots, x_n\}$ é um anel de integridade com n elementos. Seja x um elemento não nulo de A, assim, $xA = \{x.x_1, x.x_2, x.x_3, \cdots, x.x_n\}$. Como A é um anel de integridade, então $x.x_i = x.x_j$ implica que $x_i = x_j$, ademais, a multiplicação é fechada em A, portanto, para cada $x_k \in A$, existe $x.x_i \in xA$ tal que $x_k = x.x_i$, portanto, $xA = \{x.x_1, x.x_2, x.x_3, \cdots, x.x_n\} = \{x_1, x_2, x_3, \cdots, x_n\} = A$. Como $1 \in A$, então para todo $x \in A$ existe um índice i para o qual temos $x.x_i = 1$ mostrando com isso que qualquer elemento de A possui inverso. Como conseqüência, $(A, +, \cdot)$ é um corpo.

Um corpo com quantidade finita de elementos é denominado *corpo finito*. A exemplo temos $(\mathbb{Z}_3, +, \cdot)$ com as operações "usuais" e $F = \{0,1\}$ com as operações de adição e multiplicação descritas no exemplo 7.

[5] Conhecido como Corpo de Galois, em homenagem ao matemático Évariste Galois, 1811-1832.

Decorre dos teoremas 5.2 e 5.5 que o Conjunto $\mathbb{Z}_m$, quando m é primo, é um corpo finito.

5.6 ESPAÇOS VETORIAIS

Dado um corpo K, um conjunto não vazio V é denominado um *espaço vetorial sobre* K ou um *K*-espaço vetorial quando:

1º) Dados $u, v \in V$, $u + v \in V$, ou seja, existe a adição em V. Considerando $u, v, w \in V$, na adição verificam-se os seguintes axiomas:

I) $u + v = v + u$

II) $u + (v + w) = (u + v) + w$

III) $\exists o \in V$ tal que $o + u = u + o = u$, com o representando o vetor nulo.

IV) $\forall u, \exists(-u) \in V$ tal que $-u + u = u + (-u) = o$

2º) Está definida uma multiplicação por escalares do corpo K em V, ou seja, dados $\alpha \in K$ e $u \in V$, temos $\alpha.u \in V$. Considerando $\alpha, \beta \in K$ e $u, v \in V$, na multiplicação por escalar verificam-se os seguintes axiomas:

I) $\alpha.(\beta.u) = (\alpha.\beta).u$

II) $(\alpha + \beta).u = \alpha.u + \beta.u$

III) $\alpha.(u + v) = \alpha.u + \alpha.v$

IV) $1.u = u$

Exemplos:

1) O conjunto dos números reais $\mathbb{R}$ é um espaço vetorial sobre o corpo dos números racionais $\mathbb{Q}$, pois dados $x, y, z \in \mathbb{R}$ e $\alpha, \beta \in \mathbb{Q}$, temos $x + y \in \mathbb{R}$; $x + y = y + x$; $x + (y + z) = (x + y) + z$; $0 \in \mathbb{R}$ e é o elemento neutro da adição; $\exists(-x) \in \mathbb{R}$ tal que $-x + x = 0 = x + (-x)$; $\alpha.x \in \mathbb{R}$; $\alpha.(\beta.x) = (\alpha.\beta).x$; $(\alpha + \beta).x = \alpha.x + \beta.x$; $\alpha.(x + y) = \alpha.x + \alpha.y$ e $1.x = x$.

2) O conjunto dos números racionais $\mathbb{Q}$ não é um espaço vetorial sobre o corpo dos números reais $\mathbb{R}$, pois por exemplo, considerando o escalar $\alpha = \sqrt{2}$, temos, para $x \in \mathbb{Q}$, $\alpha.x = \sqrt{2}.x \notin \mathbb{Q}$.

3) O conjunto dos pares ordenados de $\mathbb{R}^2$ sobre o corpo dos números reais $\mathbb{R}$, com as operações de adição e multiplicação por escalar definidas por $(x, y), (x', y') \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$,

$(x,y)+(x',y')=(x+x',y+y')$ e $\alpha.(x,y)=(\alpha.x,\alpha.y)$, constitui um espaço vetorial. Verifiquemos: Dados $(x,y),(x',y'),(x'',y'')\in\mathbb{R}^2$ e $\alpha,\beta\in\mathbb{R}$, temos:

a) $(x,y)+(x',y')=(x+x',y+y')=(x'+x,y'+y)=(x',y')+(x,y)$

b) $(x,y)+[(x',y')+(x'',y'')]=(x,y)+(x'+x'',y'+y'')=$

$=\big(x+(x'+x''),y+(y'+y'')\big)=\big((x+x')+x'',(y+y')+y''\big)=$

$=(x+x',y+y')+(x'',y'')=[(x,y)+(x',y')]+(x'',y'')$

c) $(0,0)\in\mathbb{R}^2$ e $(x,y)+(0,0)=(x+0,y+0)=(x,y)=(0+x,0+y)=(0,0)+(x,y)$

d) $(-x,-y)\in\mathbb{R}^2$ e $(-x,-y)+(x,y)=(-x+x,-y+y)=(0,0)=(x-x,y-y)=$

$=\big(x+(-x),y+(-y)\big)=(x,y)+(-x,y)$

e) $\alpha.[\beta.(x,y)]=\alpha.(\beta.x,\beta.y)=\big(\alpha.(\beta.x),\alpha.(\beta.y)\big)=\big((\alpha.\beta).x,(\alpha.\beta).y\big)=(\alpha.\beta).(x,y)$

f) $(\alpha+\beta).(x,y)=\big((\alpha+\beta).x,(\alpha+\beta).y\big)=(\alpha.x+\beta.x,\alpha.y+\beta.y)=$

$=(\alpha.x,\alpha.y)+(\beta.x,\beta.y)=\alpha.(x,y)+\beta.(x,y)$

g) $\alpha.[(x,y)+(x',y')]=\alpha.(x+x',y+y')=\big(\alpha.(x+x'),\alpha.(y+y')\big)=$

$=(\alpha x+\alpha x',\alpha y+\alpha y')=(\alpha.x,\alpha.y)+(\alpha.x',\alpha.y')=\alpha.(x,y)+\alpha(x',y')$

h) $1.(x,y)=(1.x,1.y)=(x,y)$

4) O conjunto $\mathcal{M}_2$ das matrizes reais quadradas de ordem 2, com as operações usuais de adição e multiplicação por escalar, é um espaço vetorial sobre o corpo dos números reais $\mathbb{R}$. Verifiquemos:

No capítulo 2, referente às operações com matrizes, já vimos que na adição de matrizes verificam-se os quatro primeiros axiomas, portanto, iremos verificar a validade somente dos axiomas relativos à multiplicação por escalar:

Sejam $A,B\in\mathcal{M}_2$, tais que $A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$ e $B=\begin{bmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{bmatrix}$ e $\alpha,\beta\in\mathbb{R}$. Temos:

a) $\alpha.(\beta.A)=\alpha.\Big(\beta.\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\Big)=\alpha.\begin{bmatrix}\beta.a_{11}&\beta.a_{12}\\\beta.a_{21}&\beta.a_{22}\end{bmatrix}=\begin{bmatrix}\alpha.(\beta.a_{11})&\alpha.(\beta.a_{12})\\\alpha.(\beta.a_{21})&\alpha.(\beta.a_{22})\end{bmatrix}=$

$=\begin{bmatrix}(\alpha.\beta).a_{11}&(\alpha.\beta).a_{12}\\(\alpha.\beta).a_{21}&(\alpha.\beta).a_{22}\end{bmatrix}=(\alpha.\beta).\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=(\alpha.\beta).A$

b) $(\alpha+\beta).A=(\alpha+\beta).\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=\begin{bmatrix}(\alpha+\beta).a_{11}&(\alpha+\beta).a_{12}\\(\alpha+\beta).a_{21}&(\alpha+\beta).a_{22}\end{bmatrix}=$

$=\begin{bmatrix}\alpha.a_{11}+\beta.a_{11}&\alpha.a_{12}+\beta.a_{12}\\\alpha.a_{21}+\beta.a_{21}&\alpha.a_{22}+\beta.a_{22}\end{bmatrix}=\begin{bmatrix}\alpha.a_{11}&\alpha.a_{12}\\\alpha.a_{21}&\alpha.a_{22}\end{bmatrix}+\begin{bmatrix}\beta.a_{11}&\beta.a_{12}\\\beta.a_{21}&\beta.a_{22}\end{bmatrix}=$

$=\alpha.\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}+\beta.\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=\alpha.A+\beta.A$

c) $\alpha.(A+B)=\alpha.\Big(\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}+\begin{bmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{bmatrix}\Big)=\alpha.\begin{bmatrix}a_{11}+b_{11}&a_{12}+b_{12}\\a_{21}+b_{21}&a_{22}+b_{22}\end{bmatrix}=$

$$= \begin{bmatrix} \alpha.(a_{11}+b_{11}) & \alpha.(a_{12}+b_{12}) \\ \alpha.(a_{21}+b_{21}) & \alpha.(a_{22}+b_{22}) \end{bmatrix} = \begin{bmatrix} \alpha.a_{11}+\alpha.b_{11} & \alpha.a_{12}+\alpha.b_{12} \\ \alpha.a_{21}+\alpha.b_{21} & \alpha.a_{22}+\alpha.b_{22} \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha.a_{11} & \alpha.a_{12} \\ \alpha.a_{21} & \alpha.a_{22} \end{bmatrix} + \begin{bmatrix} \alpha.b_{11} & \alpha.b_{12} \\ \alpha.b_{21} & \alpha.b_{22} \end{bmatrix} = \alpha.\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & \alpha. \end{bmatrix} + \alpha.\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \alpha.A + \alpha.B$$

d) $1.A = 1.\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1.a_{11} & 1.a_{12} \\ 1.a_{21} & 1.a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = A$

5.6.1 Algumas propriedades de um espaço vetorial

Para evitarmos possível confusão em relação às notações, representaremos por 0 o elemento neutro do corpo K, enquanto que o vetor nulo de V representaremos por o.

Sendo V um espaço vetorial sobre um corpo K, dados $x, y \in V$ e $\alpha, \beta \in K$, temos:

I) $\alpha.o = o$

II) $0.x = o$

III) $\alpha.x = o \Longrightarrow \alpha = 0$ ou $x = o$

IV) $(-\alpha).x = \alpha.(-x) = -(\alpha.x)$

V) $(\alpha - \beta).x = \alpha.x - \beta.x$

VI) $\alpha.(x - y) = \alpha.x - \alpha.y$

VII) Dados, $\alpha, \beta_1, \beta_2, \dots, \beta_n \in K$ e $x_1, x_2, \dots, x_n \in V$, então $\alpha.(\sum_{i=1}^{n} \beta_i \cdot x_i) = \sum_{i=1}^{n} (\alpha \cdot \beta_i).x_i$

Demonstrações:

I) $\alpha.o + \alpha.o = \alpha.(o + o) = \alpha.o$. Somando $-(\alpha.o)$ a ambos os membros da igualdade, temos: $\alpha.o + \alpha.o - (\alpha.o) = \alpha.o - (\alpha.o)$, o que implica que $\alpha.o = o$.

II) $0.x + 0.x = (0 + 0).x = 0.x$. Somando $-(0.x)$ a ambos os membros da igualdade, temos: $0.x + 0.x - (0.x) = 0.x - (0.x)$, o que implica que $0.x = o$.

III) Suponhamos $\alpha \neq 0$, então $\exists \alpha^{-1} \in K$ tal que $\alpha.\alpha^{-1} = 1$. Multiplicando a igualdade $\alpha.x = o$ por α^{-1}, temos $\alpha^{-1}.(\alpha.x) = o$. Levando em consideração o axioma (I) da multiplicação por escalar e a propriedade (I), temos $(\alpha^{-1}.\alpha).x = o$, mas, $\alpha^{-1}.\alpha = 1$, portanto $1.x = o$ e, pela propriedade (IV) da multiplicação por escalar, temos $x = o$. Caso consideremos $x \neq o$, como $\alpha.x = o$, considerando a propriedade (II), temos $0.x = o$, por transitividade, temos $\alpha.x = 0.x$, o que implica que $\alpha = 0$.

IV) Utilizando o axioma (II) da multiplicação por escalar e a propriedade (II), temos que $\alpha.x + (-\alpha).x = \big(\alpha + (-\alpha)\big).x = 0.x = o$. Sabemos também que $\alpha.x + (-\alpha.x) = o$, o que implica que $\alpha.x + (-\alpha).x = \alpha.x + (-\alpha.x)$ e, somando $-\alpha.x$ a ambos os membros da igualdade, temos $-\alpha.x + \alpha.x + (-\alpha).x = -\alpha.x + \alpha.x + (-\alpha.x)$, o

que resulta que $(-\alpha).x = -\alpha.x$. Por outro lado, a propriedade (I) nos garante que $\alpha.o = o$, temos também, pelo axioma (IV) da adição, que $x + (-x) = o$, assim, temos $\alpha.(x + (-x)) = o$ e, pelo axioma (III) da multiplicação, $\alpha.x + \alpha.(-x) = o$. Somando $-\alpha.x$ a essa última igualdade, temos $-\alpha.x + \alpha.x + \alpha.(-x) = -\alpha.x + o$, o que implica que $\alpha.(-x) = -\alpha.x$. Logo, $(-\alpha).x = \alpha.(-x) = -\alpha.x$.

V) $(\alpha - \beta).x = \big(\alpha + (-\beta)\big).x$. Pelo axioma (II) da multiplicação por escalar, temos $\big(\alpha + (-\beta)\big).x = \alpha.x + (-\beta).x$ e, pela propriedade (IV), $\alpha.x + (-\beta).x = \alpha.x - \beta.x$. Portanto, $(\alpha - \beta).x = \alpha.x - \beta.x$

VI) $\alpha.(x - y) = \alpha.(x + (-y))$. Pelo axioma (III) da multiplicação por escalar, temos que $\alpha.\big(x + (-y)\big) = \alpha.x + \alpha.(-y)$ mas, pela propriedade (IV), $\alpha.(-y) = -\alpha.y$, assim, $\alpha.x + \alpha.(-y) = \alpha.x - \alpha.y$. Portanto, $\alpha.(x - y) = \alpha.x - \alpha.y$.

VII) Utilizando indução sobre n, temos, para $n = 1$, $\alpha.(\sum_{i=1}^{1} \beta_i \cdot x_i) = \alpha.(\beta_1 \cdot x_1)$. Utilizando o axioma (II) da multiplicação por escalar, temos $\alpha.(\beta_1 \cdot x_1) = (\alpha.\beta_1).x_1 = \sum_{i=1}^{1}(\alpha.\beta_i).x_i$. Portanto, a igualdade é válida para $n = 1$.

Suponhamos que a igualdade seja válida para $n = k$, ou seja, $\alpha.\big(\sum_{i=1}^{k} \beta_i \cdot x_i\big) = \sum_{i=1}^{k}(\alpha \cdot \beta_i).x_i$ e verifiquemos se essa hipótese implica na validez da igualdade para $n = k + 1$: $\alpha.\big(\sum_{i=1}^{k+1} \beta_i \cdot x_i\big) = \alpha.[\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_k \cdot x_k + \beta_{k+1} \cdot x_{k+1}] =$

$= \alpha.[(\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_k \cdot x_k) + \beta_{k+1} \cdot x_{k+1}]$. Pelo axioma (III) da multiplicação por escalar, temos $\alpha.(\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_k \cdot x_k) + \alpha.(\beta_{k+1} \cdot x_{k+1})$, mas $\alpha.(\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_k \cdot x_k) = \alpha.\big(\sum_{i=1}^{k} \beta_i \cdot x_i\big)$, assim, $\alpha.(\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_k \cdot x_k) + \alpha.(\beta_{k+1} \cdot x_{k+1}) = \alpha.\big(\sum_{i=1}^{k} \beta_i \cdot x_i\big) + \alpha.(\beta_{k+1} \cdot x_{k+1})$. Por hipótese de indução e utilizando o axioma (I) da multiplicação por escalar, temos

$\alpha.\big(\sum_{i=1}^{k} \beta_i \cdot x_i\big) + \alpha.(\beta_{k+1} \cdot x_{k+1}) = \big[\sum_{i=1}^{k}(\alpha \cdot \beta_i).x_i\big] + (\alpha.\beta_{k+1}).x_{k+1} =$

$= \sum_{i=1}^{k+1}(\alpha \cdot \beta_i).x_i$

5.6.2 Subespaços vetoriais

Seja V um espaço vetorial sobre um corpo K. Dizemos que um conjunto não vazio $W \subset V$ é um subespaço vetorial de V se W é um espaço vetorial em relação à adição e à multiplicação por escalar em V.

Em outras palavras, dizer que W é um subespaço vetorial de V, é afirmar que para as operações de Adição e de multiplicação por escalar do espaço V, são verificados os 8 axiomas que definem um espaço vetorial em W.

Exemplo:

O conjunto $W \subset \mathbb{R}^2$ tal que $W = \{(x, y) \in \mathbb{R}^2; x = y\}$ é um subespaço vetorial de $\mathbb{R}^2$, pois, dados $(x, x), (y, y), (z, z) \in W$ e $\alpha, \beta \in \mathbb{R}$, temos:

$(x, x) + (y, y) = (x + y, x + y) \in W$ e $\alpha.(x, x) = (\alpha.x, \alpha.x) \in W$. Além disso, temos:

I) $(x, x) + (y, y) = (x + y, x + y) = (y + x, y + x) = (y, y) + (x, x)$

II) $(x, x) + [(y, y) + (z, z)] = (x, x) + (y + z, y + z) = (x + (y + z), x + (y + z)) = \big((x + y) + z, (x + y) + z\big) = (x + y, x + y) + (z, z) = [(x, x) + (y, y)] + (z, z)$

III) $(0,0) \in W$, pois $0 = 0$ e $(x, x) + (0,0) = (x + 0, x + 0) = (x, x)$ e $(0,0) + (x, x) = (0 + x, 0 + x) = (x, x)$

IV) $(-x, -x) \in W$, pois $-x = -x$ e $(x, x) + (-x, -x) = \big(x + (-x), x + (-x)\big) = (0,0)$ e $(-x, -x) + (x, x) = (-x + x, -x + x) = (0,0)$

V) $\alpha.\big(\beta.(x, x)\big) = \alpha.(\beta.x, \beta.x) = \big(\alpha.(\beta.x), \alpha.(\beta.x)\big) = \big((\alpha.\beta).x, (\alpha.\beta).x\big) = (\alpha.\beta).(x, x)$

VI) $(\alpha + \beta).(x, x) = \big((\alpha + \beta).x, (\alpha + \beta).x\big) = (\alpha.x + \beta.x, \alpha.x + \beta.x) = (\alpha.x, \alpha.x) + (\beta.x, \beta.x) = \alpha.(x, x) + \beta.(x, x)$

VII) $\alpha.[(x, x) + (y, y)] = \alpha.(x + y, x + y) = \big(\alpha.(x + y), \alpha.(x + y)\big) = (\alpha.x + \alpha.y, \alpha.x + \alpha.y) = (\alpha.x, \alpha.x) + (\alpha.y, \alpha.y) = \alpha.(x, x) + \alpha.(y, y)$

VIII) $1.(x, x) = (1.x, 1.x) = (x, x)$

Portanto, $W = \{(x, y) \in \mathbb{R}^2; x = y\}$ é subespaço vetorial de $\mathbb{R}^2$.

Teorema 5.6: Seja V um espaço vetorial sobre um corpo K. Um subconjunto não vazio $W \subset V$ é um subespaço vetorial de V se, e somente se, dados $x, y \in W$ e $\alpha, \beta \in K$, temos $\alpha.x + \beta.y \in W$.

Demonstração:

($\Longrightarrow$) Seja $W \subset V$ um subespaço vetorial de V. Por definição de subespaço, dados $x, y \in W$ e $\alpha, \beta \in K$, temos $\alpha.x \in W$ e $\beta.y \in W$. Como a soma de vetores é fechada em W, então $\alpha.x + \beta.y \in W$.

($\Longleftarrow$) Consideremos $W \subset V$, $W \neq \emptyset$ tal que para todo $x, y \in W$ e $\alpha, \beta \in K$, temos $\alpha.x + \beta.y \in W$. Para $\alpha = \beta = 1$ temos $\alpha.x + \beta.y = 1.x + 1.y = x + y \in W$ e para $\beta = 0$, temos

$\alpha.x + \beta.y = \alpha.x + 0.y = \alpha.x \in W$, ou seja, as operações de adição e multiplicação por escalar em V são fechadas em W. Como para todo $x, y, z \in V$, temos que $x + y = y + x$ e $x + (y + z) = (x + y) + z$ e como W é fechado para a operação de adição de V, então, se $x, y, z \in W$, as igualdades $x + y = y + x$ e $x + (y + z) = (x + y) + z$ são válidas em W, garantindo com isso a comutatividade e a associatividade da adição em W. Como a multiplicação por escalar é fechada em W, ou seja, $\forall \alpha \in K$ e $x \in W$, temos $\alpha.x \in W$, então, tomando $\alpha = 1$, temos $\alpha.x = 1.x = x \in W$ e, tomando $\alpha = -1$, temos $\alpha.x = -1.x = -x \in W$. Como por hipótese $\alpha.x + \beta.y \in W$ para todo $x, y \in V$ e $\alpha, \beta \in K$, tomando $\beta = -\alpha$ e $y = x$, temos $\alpha.x + \beta.y = \alpha.x + (-\alpha).x = \alpha.x - \alpha.x = o \in W$. Como para todo $\alpha, \beta \in K$ e $x, y \in V$ temos $\alpha.(\beta.x) = (\alpha.\beta).x$, $(\alpha + \beta).x = \alpha.x + \beta.x$ e $\alpha.(x + y) = \alpha.x + \alpha.y$ e, como a multiplicação por escalar é fechada em W, então essas igualdades são válidas em W, mostrando com isso que W é um espaço vetorial. Por hipótese, temos $W \subset V$, o que nos mostra que W é um subespaço vetorial de V, completando a demonstração.

Retomando o exemplo anterior, consideremos o conjunto $W \subset \mathbb{R}^2$ tal que $W = \{(x, y) \in \mathbb{R}^2; x = y\}$, para mostrar que W é um subespaço vetorial de $\mathbb{R}^2$, de acordo com o teorema 5.6, basta mostrar que dados $(x, x), (y, y) \in W$ e $\alpha, \beta \in K$, temos $\alpha.(x, x) + \beta.(y, y) \in W$, o que é fácil de comprovar:
$\alpha.(x, x) + \beta.(y, y) = (\alpha.x, \alpha.x) + (\beta.y, \beta.y) = (\alpha.x + \beta.y, \alpha.x + \beta.y) \in W$ por terem coordenadas iguais. Portanto, $W = \{(x, y) \in \mathbb{R}^2; x = y\}$ é um subespaço vetorial de $\mathbb{R}^2$.

5.6.3 Base e Dimensão

5.6.3.1 Independência Linear

Considerando V um espaço vetorial sobre um corpo K. Dizemos que um conjunto $I \subset V$, $I = \{x_1, x_2, x_3, \dots, x_n\}$ é *linearmente independente*, quando para todo $\alpha \in K$, a igualdade $\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2 + \alpha_3 \cdot x_3 + \cdots + \alpha_n \cdot x_n = o$ for verdadeira somente se $\alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$. Caso exista algum $i \in \{1, 2, 3, \dots, n\}$, para o qual se tenha $\alpha_i \neq 0$, então dizemos que I é *linearmente dependente*. A uma igualdade do tipo $\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2 + \alpha_3 \cdot x_3 + \cdots + \alpha_n \cdot x_n = x$ denominamos *combinação linear*.

5.6.3.2 Conjunto de geradores

Trataremos nesse tópico somente de espaços vetoriais finitamente gerados.

Dados um espaço vetorial V sobre um corpo K, dizemos que um conjunto finito $G \subset V$, $G = \{u_1, u_2, u_3, \dots, u_n\}$ é um gerador do espaço vetorial V, se para todo $x \in V$, existirem $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in K$ tal que, $x = \alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \alpha_3 \cdot u_3 + \cdots + \alpha_n \cdot u_n$. Assim, dizemos que V é gerado por G ou que G gera V ou ainda que $V = [G]$.

5.6.3.3 Base e dimensão de um espaço vetorial

Seja $B \subset V$ um conjunto finito, onde V é um espaço vetorial sobre um corpo K. Se $[B] = V$ e B é um conjunto de vetores linearmente independentes, então, dizemos que B é uma *base* do espaço vetorial V.

Exemplo:

Vimos anteriormente que o conjunto dos pares ordenados de coordenadas reais $\mathbb{R}^2$ é um espaço vetorial sobre o corpo $\mathbb{R}$ para as operações convencionais de soma e multiplicação por escalar. Consideremos o conjunto $B = \{(1,0), (0,1)\}$ contido em $\mathbb{R}^2$. Observemos que qualquer que seja o vetor $v = (x, y) \in \mathbb{R}^2$, vale que $v = (x, y) = (x, 0) + (0, y) = x \cdot (1,0) + y \cdot (0,1)$, assim, vemos que $[B] = \mathbb{R}^2$. Além disso, para $\alpha, \beta \in \mathbb{R}$, temos $\alpha \cdot (1,0) + \beta \cdot (0,1) = (0,0) \Leftrightarrow (\alpha, 0) + (0, \beta) = (0,0) \Leftrightarrow (\alpha, \beta) = (0,0) \Leftrightarrow \alpha = 0$ e $\beta = 0$, o que acarreta que B é linearmente independente e portanto é uma base de $\mathbb{R}^2$.

Teorema 5.7: Seja V um espaço vetorial sobre um corpo K. Se $u_1, u_2, u_3, \dots, u_n$ são vetores de V e $[u_1, u_2, u_3, \dots, u_n] = V$, então existe $B \subset \{u_1, u_2, u_3, \dots, u_n\}$ tal que B é uma base de V.

Demonstração: Se $u_1, u_2, u_3, \dots, u_n$ são linearmente independentes, então $B = \{u_1, u_2, u_3, \dots, u_n\}$ é uma base de V e não há nada o que demonstrar. Caso $u_1, u_2, u_3, \dots, u_n$ sejam vetores linearmente dependentes, então, na combinação linear $\alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \alpha_3 \cdot u_3 + \cdots + \alpha_n \cdot u_n = o$, $\alpha_i \in K, 1 \leq i \leq n$ existe pelo menos um coeficiente não nulo. Suponhamos, sem perda de generalidade, que $\alpha_1 \neq 0$, então "dividimos" a equação $\alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \alpha_3 \cdot u_3 + \cdots + \alpha_n \cdot u_n = o$ por α_1, obtendo $u_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) \cdot u_2 + \left(-\frac{\alpha_3}{\alpha_1}\right) \cdot u_3 + \left(-\frac{\alpha_4}{\alpha_1}\right) \cdot u_4 + \cdots + \left(-\frac{\alpha_n}{\alpha_1}\right) \cdot u_n$, fazendo $\beta_j = -\frac{\alpha_{j+1}}{\alpha_1}$, $1 \leq j \leq n-1$, temos

$u_1 = \beta_1 \cdot u_2 + \beta_2 \cdot u_3 + \beta_3 \cdot u_4 + \cdots + \beta_{n-1} \cdot u_n$, ou seja, u_1 é combinação linear dos $n-1$ vetores $u_2, u_3, u_4, \ldots, u_n$, ou seja, $u_2, u_3, u_4, \ldots, u_n$ ainda geram V. Caso $u_2, u_3, u_4, \ldots, u_n$ forem linearmente independentes, então temos $B = \{u_2, u_3, u_4, \ldots, u_n\}$ uma base de V. Caso $u_2, u_3, u_4, \ldots, u_n$ forem linearmente dependentes, repetimos o processo anterior e encontramos um vetor dentre os vetores $u_2, u_3, u_4, \ldots, u_n$ que é escrito como combinação linear dos outros $n-2$ vetores e, portanto, os $n-2$ vetores ainda geram V. Após uma quantidade finita de repetições do processo descrito anteriormente teremos, dentre os $u_i's$, um conjunto com vetores linearmente independentes, constituindo assim, uma base B do espaço V.

Teorema 5.8: Seja V uma espaço vetorial sobre um corpo K. Se $u_1, u_2, u_3, \ldots, u_n$ são vetores de V e $[u_1, u_2, u_3, \ldots, u_n\ ] = V$, então qualquer conjunto com mais de n vetores é linearmente dependente, ou seja, qualquer conjunto linearmente independente de vetores de V tem no máximo n vetores.

Demonstração: Como $[u_1, u_2, u_3, \ldots, u_n\ ] = V$, pelo teorema 5.7, podemos extrair uma base dentre os vetores $u_1, u_2, u_3, \ldots, u_n$. Suponhamos que os vetores $x_1, x_2, x_3, \ldots, x_t$, com $t \leq n$ formam essa base. Consideremos agora os vetores $v_1, v_2, v_3, \ldots, v_m$ de V, com $m > n$. Existem então escalares $\alpha_{ij} \in K$, com $1 \leq i \leq m$ e $1 \leq j \leq n$ tais que:

$$\text{(I)} \qquad \begin{aligned} v_1 &= \alpha_{11} \cdot x_1 + \alpha_{12} \cdot x_2 + \cdots \alpha_{1t} \cdot x_t \\ v_2 &= \alpha_{21} \cdot x_1 + \alpha_{22} \cdot x_2 + \cdots \alpha_{2t} \cdot x_t \\ &\vdots \qquad\qquad\qquad \vdots \\ v_m &= \alpha_{m1} \cdot x_1 + \alpha_{m2} \cdot x_2 + \cdots \alpha_{mt} \cdot x_t \end{aligned}$$

Consideremos agora uma combinação linear nula dos vetores $v_1, v_2, v_3, \ldots, v_m$:

$$\text{(II)} \qquad \beta_1 \cdot v_1 + \beta_2 \cdot v_2 + \cdots + \beta_m \cdot v_m = o$$

Substituindo (I) em (II), temos:

$$\beta_1 \cdot (\alpha_{11} \cdot x_1 + \alpha_{12} \cdot x_2 + \cdots \alpha_{1t} \cdot x_t) + \beta_2 \cdot (\alpha_{21} \cdot x_1 + \alpha_{22} \cdot x_2 + \cdots \alpha_{2t} \cdot x_t) + \cdots$$
$$\ldots + \beta_m \cdot (\alpha_{m1} \cdot x_1 + \alpha_{m2} \cdot x_2 + \cdots \alpha_{mt} \cdot x_t) = o$$

Reagrupando temos:

$$(\alpha_{11} \cdot \beta_1 + \alpha_{21} \cdot \beta_2 + \cdots + \alpha_{m1} \cdot \beta_m).x_1 + (\alpha_{12} \cdot \beta_1 + \alpha_{22} \cdot \beta_2 + \cdots + \alpha_{m2} \cdot \beta_m).x_2 + \cdots$$
$$\ldots + (\alpha_{1t} \cdot \beta_1 + \alpha_{2t} \cdot \beta_2 + \cdots + \alpha_{mt} \cdot \beta_m).x_t = o,$$ mas, por hipótese, $x_1, x_2, x_3, \ldots, x_t$ é uma base e portanto, são vetores linearmente independentes, o que implica que temos:

$$\begin{cases} \alpha_{11} \cdot \beta_1 + \alpha_{21} \cdot \beta_2 + \cdots + \alpha_{m1} \cdot \beta_m = 0 \\ \alpha_{12} \cdot \beta_1 + \alpha_{22} \cdot \beta_2 + \cdots + \alpha_{m2} \cdot \beta_m = 0 \\ \qquad\qquad\qquad \vdots \\ \alpha_{1t} \cdot \beta_1 + \alpha_{2t} \cdot \beta_2 + \cdots + \alpha_{mt} \cdot \beta_m = 0 \end{cases}$$

Que é um sistema linear homogêneo com t equações e m incógnitas $\beta_1,\ \beta_2, \dots, \beta_m$. Como $t \leq n < m$, então esse sistema admite uma solução não trivial, ou seja, existe algum $\beta_i \neq 0$, com $1 < i < m$, acarretando com isso que os vetores $v_1, v_2, v_3, \dots, v_m$ são linearmente independentes.

Teorema 5.9: Se V um *K*-espaço vetorial finitamente gerado, então duas bases quaisquer de V tem o mesmo número de vetores.

Demonstração: Sejam $B_1 = \{u_1, u_2, u_3, \dots, u_n\}$ e $B_2 = \{v_1, v_2, v_3, \dots, v_m\}$ duas bases do espaço vetorial V. Então, por definição, $u_1, u_2, u_3, \dots, u_n$ são linearmente independentes e geram V e $v_1, v_2, v_3, \dots, v_m$ são linearmente independentes e geram V. Como $u_1, u_2, u_3, \dots, u_n$ geram V e $v_1, v_2, v_3, \dots, v_m$ são linearmente independentes, então, pelo teorema 5.8, temos $m \leq n$. Em contrapartida, como $v_1, v_2, v_3, \dots, v_m$ geram V e $u_1, u_2, u_3, \dots, u_n$ são linearmente independentes, então, pelo teorema 5.8, temos $n \leq m$. Essas duas desigualdades são possíveis somente se $m = n$ e, portanto, B_1 e B_2 possuem o mesmo número de vetores.

O teorema 5.9 permite-nos apresentar a seguinte definição:

Seja V um espaço vetorial finitamente gerado, denominamos *dimensão de* V e representamos por $\dim V$ o número de elementos de qualquer uma de suas bases. Neste caso, dizemos que V é um espaço de dimensão finita.

Exemplo: Vimos anteriormente que $B = \{(1,0), (0,1)\}$ é uma base de $\mathbb{R}^2$. Pelo teorema 5.9, qualquer outra base de $\mathbb{R}^2$ possuirá também dois vetores, o que, de acordo com o que acabamos de definir, faz com que $\dim \mathbb{R}^2 = 2$.

O *K*-espaço vetorial $E = \{o\}$ (contendo apenas o vetor nulo) tem dimensão zero, ou seja, $\dim E = 0$, pois para todo $\alpha \in K$ tem-se $\alpha \cdot o = o$, o que implica que $E = [\emptyset]$.

Teorema 5.10: Sendo V um *K*-espaço vetorial de dimensão finita, qualquer conjunto de vetores linearmente independentes de V pode ser completado de modo a se obter uma base de V.

Demonstração: Consideremos $\dim V = n$ e sejam $v_1, v_2, v_3, \dots, v_t$ pertencentes a V e linearmente independentes. Pelo teorema 5.8, temos $t \leq n$. Se $[v_1, v_2, v_3, \dots, v_t\ ] = V$, então $v_1, v_2, v_3, \dots, v_t$ formam uma base e não a nada o que demonstrar. Caso $[v_1, v_2, v_3, \dots, v_t\ ] \neq V$, então existe $v_{t+1} \in V$ tal que $v_{t+1} \notin [v_1, v_2, v_3, \dots, v_t]$, o que implica que v_{t+1} não é escrito como combinação linear dos vetores $v_1, v_2, v_3, \dots, v_t$, logo $v_1, v_2, v_3, \dots, v_t, v_{t+1}$ são linearmente independentes e caso $[v_1, v_2, v_3, \dots, v_t, v_{t+1}] = V$, então $v_1, v_2, v_3, \dots, v_t, v_{t+1}$

formam uma base de V, caso contrário, repetimos o processo por no máximo $n - t$ vezes e obtemos assim uma base para V.

5.6.4 Noções sobre transformação linear

Sejam V e W dois K-espaços vetoriais. Uma função $T: V \to W$ é denominada uma *transformação linear de V em W* quando para todo $v_1, v_2 \in V$ e $\alpha \in K$, temos $T(v_1 + v_2) = T(v_1) + T(v_2)$ e $T(\alpha . v_1) = \alpha . T(v_1)$.

Exemplo:

Consideremos a função $T: \mathbb{R}^3 \to \mathbb{R}^2$, definida por $T(x, y, z) = (3x, 2x - y + 5z)$. Notemos que dados dois vetores $v_1 = (x_1, y_1, z_1)$ e $v_2 = (x_2, y_2, z_2)$ e $\alpha \in \mathbb{R}$, temos:

$$T(v_1 + v_2) = T\big((x_1, y_1, z_1) + (x_2, y_2, z_2)\big) = T(x_1 + x_2, y_1 + y_2, z_1 + z_2) =$$
$$= \big(3 \cdot (x_1 + x_2), 2 \cdot (x_1 + x_2) - (y_1 + y_2) + 5 \cdot (z_1 + z_2)\big) =$$
$$= (3 \cdot x_1 + 3 \cdot x_2, 2 \cdot x_1 + 2 \cdot x_2 - y_1 - y_2 + 5 \cdot z_1 + 5. z_2) =$$
$$= ((3 \cdot x_1) + (3 \cdot x_2), (2 \cdot x_1 - y_1 + 5 \cdot z_1) + (2 \cdot x_2 - y_2 + 5. z_2)) =$$
$$= (3 \cdot x_1, 2 \cdot x_1 - y_1 + 5 \cdot z_1) + (3 \cdot x_2, 2 \cdot x_2 - y_2 + 5. z_2) =$$
$$= T(x_1, y_1, z_1) + T(x_2, y_2, z_2) = T(v_1) + T(v_2)$$

E $T(\alpha . v_1) = T\big(\alpha . (x_1, y_1, z_1)\big) = T(\alpha . x_1, \alpha . y_1, \alpha . z_1) =$

$$= (3\alpha . x_1, 2\alpha . x_1 - \alpha . y_1 + 5\, \alpha . z_1) = (\alpha . (3x_1), \alpha . (2x_1 - y_1 + 5z_1)) =$$
$$= \alpha . (3x_1, 2x_1 - y_1 + 5z_1) = \alpha . T(x_1, y_1, z_1) = \alpha . T(v_1)$$

Portanto, T é uma transformação linear de $\mathbb{R}^3$ em $\mathbb{R}^2$.

5.6.4.1 Núcleo e Imagem de uma Transformação Linear

Sejam V e W dois K-espaços vetoriais e $T: V \to W$ uma transformação linear de V em W. Denominamos *núcleo da transformação linear T* e representamos por $Ker\,(T)$ ao seguinte subconjunto de V:

$$Ker\,(T) = \{v \in V; T(v) = 0\}$$

Exemplo: Considerando a transformação linear $T: \mathbb{R}^3 \to \mathbb{R}^2$, definida por $T(x, y, z) = (3x, 2x - y + 5z)$, para todo $(x, y, z) \in Ker(T)$, temos $T(x, y, z) = (0,0)$, o que implica que $(3x, 2x - y + 5z) = (0,0)$, ou seja, $x = 0$ e $y = 5z$, portanto, $Ker\,(T) = \{(0, 5z, z); z \in \mathbb{R}\}$.

Teorema 5.11: Sejam V e W dois *K*-espaços vetoriais e $T: V \to W$ uma transformação linear, então $Ker\ (T)$ é um subespaço vetorial de V.

Demonstração: Dados $v_1, v_2 \in Ker\ (T)$, por definição de núcleo, temos que $T(v_1) = 0$ e $T(v_2) = 0$. Sejam $\alpha, \beta \in K$, para que $Ker\ (T)$ seja um subespaço vetorial de V devemos mostrar que $\alpha v_1 + \beta v_2 \in Ker\ (T)$, ou seja, devemos ter $T(\alpha v_1 + \beta v_2) = 0$, o que é fácil de comprovar, pois, como T é linear, temos $T(\alpha v_1 + \beta v_2) = T(\alpha v_1) + T(\beta v_2) = \alpha T(v_1) + \beta T(v_2) = \alpha.0 + \beta.0 = 0 = 0 = 0$. Logo, $Ker\ (T)$ é um subespaço vetorial de V.

Teorema 5.12: Sejam V e W dois *K*-espaços vetoriais e $T: V \to W$ uma transformação linear, então T é injetiva se, e somente se, $Ker\ (T) = \{0\}$.

Demonstração:

($\Longrightarrow$) Suponhamos que T seja injetiva e seja $v \in Ker\ (T)$, então temos $T(v) = 0$. Mas, no teorema 5.11 vimos que $Ker\ (T)$ é um subespaço vetorial de V, então $0 \in Ker(T)$, o que implica que $T(0) = 0$ e, portanto, $T(v) = T(0)$, mas, por hipótese T é injetiva, o que acarreta que $v = 0$ e, portanto, $Ker\ (T) = \{0\}$.

($\Longleftarrow$) Suponhamos $v_1, v_2 \in V$ e $Ker\ (T) = \{0\}$. Se $T(v_1) = T(v_2)$ então, subtraindo $T(v_2)$ de ambos os membros da igualdade, temos $T(v_1) - T(v_2) = 0$ e, como T é linear, $T(v_1 - v_2) = 0$, ou seja, $v_1 - v_2 \in Ker\ (T)$, ou seja, $v_1 - v_2 = 0$, o que implica que $v_1 = v_2$ ou seja, T é injetiva.

Dados V e W dois *K*-espaços vetoriais e $T: V \to W$ uma transformação linear de V em W. Denominamos *imagem da transformação T* e representamos por $Im(T)$ ao conjunto:

$$Im(T) = \{T(v); v \in V\}$$

A imagem de uma transformação linear $T: V \to W$ é um subespaço vetorial de W, pois dados $w_1, w_2 \in Im(T)$ e $\alpha, \beta \in K$, existem $v_1, v_2 \in V$ tais que $w_1 = T(v_1)$ e $w_2 = T(v_2)$, assim:

$$\alpha \cdot w_1 + \beta \cdot w_2 = \alpha \cdot T(v_1) + \beta \cdot T(v_2) = T(\alpha \cdot v_1) + T(\beta \cdot v_2) = T(\alpha \cdot v_1 + \beta \cdot v_2) \in Im(T)$$

Se $Im(T) = W$, então T é sobrejetiva.

Teorema 5.13: Sejam V e W dois *K*-espaços vetoriais de dimensão finita e $T: V \to W$ uma transformação linear, então, $\dim V = \dim Ker(T) + \dim Im(T)$.

Demonstração: Seja $B_1 = \{v_1, v_2, v_3, \ldots, v_t\}$ uma base de $Ker(T)$. De acordo com o teorema 5.10, essa base pode ser completada de modo a se obter uma base $B_2 =$

$\{v_1, v_2, v_3, \dots, v_t, u_1, u_2, u_3, \dots, u_m\}$ do espaço vetorial V. Devemos demonstrar então que $T(u_1), T(u_2), T(u_3), \dots, T(u_m)$ é uma base de $Im(T)$:

Qualquer que seja $w \in Im(T)$, existe $v \in V$ tal que $T(v) = w$. Como $v \in V$, então $\alpha_i, \beta_j \in K$, $1 \leq i \leq t$ e $1 \leq j \leq m$ tal que $v = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_t \cdot v_t + \beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \cdots + \beta_m \cdot u_m$, mas, $w = T(v)$, então, $w = T(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_t \cdot v_t + \beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \cdots + \beta_m \cdot u_m)$ e, $w = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_t T(v_t) + \beta_1 T(u_1) + \beta_2 T(u_2) + \cdots + \beta_m T(u_m)$. Como $v_1, v_2, v_3, \dots, v_t$ pertencem a $Ker(T)$, então $T(v_1) = T(v_2) = T(v_3) = \cdots = T(v_t) = 0$, logo $w = \beta_1 T(u_1) + \beta_2 T(u_2) + \cdots + \beta_m T(u_m)$, ou seja $[T(u_1), T(u_2), T(u_3), \dots, T(u_m)] = Im(T)$. Por outro lado, considerando a combinação linear $\beta_1 T(u_1) + \beta_2 T(u_2) + \beta_3 T(u_3) + \cdots + \beta_m T(u_m) = 0$ como T é uma transformação linear, então temos $T(\beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \beta_3 \cdot u_3 + \cdots + \beta_m \cdot u_m) = 0$, o que implica que $\beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \beta_3 \cdot u_3 + \cdots + \beta_m \cdot u_m \in Ker(T)$ e, portanto, pode ser escrito como combinação linear dos vetores da $v_1, v_2, v_3, \dots, v_t$ que constituem uma base de $Ker(T)$, ou seja, existem escalares $\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_t$ tais que $\beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \beta_3 \cdot u_3 + \cdots + \beta_m \cdot u_m = \gamma_1 . v_1 + \gamma_2 . v_2 + \gamma_3 . v_3 + \cdots + \gamma_t . v_t$, ou seja, $\beta_1 \cdot u_1 + \beta_2 \cdot u_2 + \beta_3 \cdot u_3 + \cdots + \beta_m \cdot u_m - \gamma_1 . v_1 - \gamma_2 . v_2 - \gamma_3 . v_3 - \cdots - \gamma_t . v_t = 0$, porém $v_1, v_2, v_3, \dots, v_t, u_1, u_2, u_3, \dots, u_m$ é uma base de V, logo $\beta_1 = \beta_2 = \beta_3 = \cdots = \beta_m = \gamma_1 = \gamma_2 = \gamma_3 = \cdots \gamma_t = 0$. Como $\beta_1 = \beta_2 = \beta_3 = \cdots = \beta_m = 0$, então $T(u_1), T(u_2), T(u_3), \dots, T(u_m)$ são linearmente independentes, e portanto formam uma base para $Im(T)$. Como uma base de V tem $t + m$ elementos e, portanto $\dim V = t + m$, sendo $\dim Ker(T) = t$, acabamos de verificar que $\dim Im(T) = m$. Logo, $\dim V = \dim Ker(T) + \dim Im(T)$.

Teorema 5.14: Sejam V e W dois K-espaços vetoriais de dimensão finita e $T: V \to W$ uma transformação linear injetiva. Se $\dim V = \dim W$ então T transforma uma base qualquer de V em uma base de W.

Demonstração: Consideremos que $\dim V = \dim W = n$ e $B = \{v_1, v_2, v_3, \dots, v_n\}$ seja uma das bases de V. Mostraremos que o conjunto $S \subset W$ tal que $S = \{T(v_1), T(v_2), T(v_3), \dots, T(v_n)\}$ é linearmente independente e tem uma quantidade n de vetores, e, portanto, é uma das bases de W: Como T é injetiva, então $T(v_1) \neq T(v_2) \neq T(v_3) \neq \cdots \neq T(v_n)$, ou seja, S possui exatamente n vetores. Sejam $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$ escalares do corpo K, tais que $\alpha_1 \cdot T(v_1) + \alpha_2 \cdot T(v_2) + \alpha_3 \cdot T(v_3) + \cdots + \alpha_n \cdot T(v_n) = 0$, como T é linear, então $T(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 + \cdots + \alpha_n \cdot v_n) = 0$, o que implica que $\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 + \cdots + \alpha_n \cdot v_n = 0$, mas $v_1, v_2, v_3, \dots, v_n$ são linearmente

independentes, portanto, $\alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, o que implica que $S = \{T(v_1), T(v_2), T(v_3), \ldots, T(v_n)\}$ é um conjunto de n vetores linearmente independentes, e portanto, uma base de W.

5.6.5 Noções sobre produto interno

Definição: Considerando V um espaço vetorial finitamente gerado sobre um corpo K, denominamos *produto interno* sobre V à função que transforma cada par de vetores $(u, v) \in V \times V$ em um escalar $a \in K$, o qual representaremos por $\langle u, v\rangle$, com as seguintes propriedades:

$\forall u, v, w \in V$ e $\alpha \in K$, temos:

I) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$

II) $\langle \alpha \cdot u, v\rangle = \alpha . \langle u, v\rangle$

III) $\langle u, v\rangle = \langle v, u\rangle$

IV) $\langle u, u\rangle > 0_k$, para todo $u \neq 0_v$, onde 0_k representa o elemento neutro da adição no corpo K e 0_v representa o vetor nulo no espaço V.

Exemplo: Consideremos o espaço vetorial $\mathbb{R}^3$ sobre o corpo $\mathbb{R}$. Dados $u, v, w \in \mathbb{R}^3$, e $\alpha \in \mathbb{R}$, a operação $\langle u, v\rangle = \langle (x_1, y_1, z_1) \cdot (x_2, y_2, z_2)\rangle = x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2$ é um produto interno sobre V. Verifiquemos:

I) $\langle u + v, w\rangle = \langle [(x_1, y_1, z_1) + (x_2, y_2, z_2)], (x_3, y_3, z_3)\rangle =$

$= \langle (x_1 + x_2, y_1 + y_2, z_1 + z_2), (x_3, y_3, z_3)\rangle =$

$= (x_1 + x_2) . x_3 + (y_1 + y_2) \cdot y_3 + (z_1 + z_2) \cdot z_3 =$

$= x_1 \cdot x_3 + x_2 \cdot x_3 + y_1 \cdot y_3 + y_2 \cdot y_3 + z_1 \cdot z_3 + z_2 \cdot z_3 =$

$= (x_1 \cdot x_3 + y_1 \cdot y_3 + z_1 \cdot z_3) + (x_2 \cdot x_3 + y_2 \cdot y_3 + z_2 \cdot z_3) =$

$= \langle (x_1, y_1, z_1), (x_3, y_3, z_3)\rangle + \langle (x_2, y_2, z_2), (x_3, y_3, z_3)\rangle = \langle u, w\rangle + \langle v, w\rangle$

II) $\langle \alpha \cdot u, v\rangle = \langle \alpha \cdot (x_1, y_1, z_1), (x_2, y_2, z_2)\rangle = \langle (\alpha \cdot x_1, \alpha \cdot y_1, \alpha \cdot z_1), (x_2, y_2, z_2)\rangle =$

$= (\alpha \cdot x_1) \cdot x_2 + (\alpha \cdot y_1) \cdot y_2 + (\alpha \cdot z_1) \cdot z_2 =$

$= \alpha \cdot (x_1 \cdot x_2) + \alpha \cdot (y_1 \cdot y_2) + \alpha \cdot (z_1 \cdot z_2) = \alpha \cdot (x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2) =$

$= \alpha \cdot \langle (x_1, y_1, z_1), (x_2, y_2, z_2)\rangle = \alpha \cdot \langle u, v\rangle$

III) $\langle u, v\rangle = \langle (x_1, y_1, z_1), (x_2, y_2, z_2)\rangle = x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2 =$

$= x_2 \cdot x_1 + y_2 \cdot y_1 + z_2 \cdot z_1 = \langle (x_2, y_2, z_2), (x_1, y_1, z_1)\rangle = \langle v, u\rangle$

IV) Se $u \neq (0,0,0)$, então, temos $x_1 \neq 0$ ou $y_1 \neq 0$ ou $z_1 \neq 0$, assim, temos $\langle u, u\rangle = \langle (x_1, y_1, z_1), (x_1, y_1, z_1)\rangle = x_1 \cdot x_1 + y_1 \cdot y_1 + z_1 \cdot z_1 = x_1^2 + y_1^2 + z_1^2 > 0$.

6 CÓDIGOS CORRETORES DE ERROS

O avanço rápido da tecnologia está presente no nosso cotidiano. Dispomos hoje de grande facilidade em armazenar dados ou nos comunicar de maneira prática e rápida utilizando aparelhos eletrônicos como celulares, *tablets*, microcomputadores, além dos meios de comunicação convencionais tais como televisão, rádio etc., que geram ao mesmo tempo eficiência, conforto e lazer para nós usuários. No entanto passam despercebidos à maioria dos usuários todo um contexto matemático utilizado e necessário para o funcionamento desses aparelhos. Embora não utilizemos matemática de maneira direta ao, por exemplo, enviarmos uma mensagem via celular, de maneira indireta isso só é possível por meio da utilização indireta da matemática.

Um aspecto importante no envio ou armazenamento de informações consiste na incerteza em saber se a informação por nós enviada através de um dispositivo eletrônico de comunicação será recebida tal qual enviamos ou se um dado hoje armazenado será acessado amanhã com o mesmo grau de fidedignidade. Informações enviadas ou armazenadas serão passíveis de erros? Caso haja um erro na transmissão de uma informação ou no armazenamento da mesma, serão possíveis as detecções e correções? Pensando nessas questões, abordaremos a seguir a teoria dos Códigos Corretores de Erros.

A teoria dos códigos foi criada pelo matemático americano Claude Elwood Shannon, no laboratório Bell, e foi apresentada de um trabalho publicado no ano de 1948. Nas décadas de 50 e 60 vários matemáticos que se interessaram pelo assunto, contribuíram de forma considerável com o desenvolvimento dessa teoria. A partir da década de 70, profissionais engenheiros passaram a ter interesse pela teoria em virtude das pesquisas espaciais, telecomunicações e o uso difundido de computadores. Nos dias atuais, qualquer aparelho que seja utilizado para a transmissão ou armazenamento de dados, faz uso dessa teoria, portanto, a teoria dos códigos está a cada dia mais presente em nossas vidas.

6.1 O QUE É UM CÓDIGO?

Podemos citar como exemplo de um código o idioma que usamos. Consideremos um alfabeto formado por 38 caracteres, sendo 37 deles os caracteres a seguir:

$\{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, á, à, â, ã, é, ê, í, ó, ô, õ, ú\}$ e mais um caractere correspondendo ao espaço entre palavras. Denominaremos esse alfabeto de P. Consideremos que a maior palavra desse alfabeto seja "pneumoultramicroscopicossilicovulcanoconiótico"[6]. Percebe-se que P possui 38 elementos (caracteres) e sua maior palavra possui 46 caracteres. Podemos fazer com que cada palavra de P possua exatamente o mesmo número de caracteres da sua maior palavra, ou seja, por meio do acréscimo de espaços no fim de cada palavra de P, fazemos com que todas elas possuam exatamente 46 caracteres. Definimos assim um código como sendo um conjunto $C \subset P^{46}$ de todas as palavras existentes no nosso idioma. Notemos, porém que o código C não é eficiente para detectar e corrigir erros, por exemplo, se transmitíssemos a palavra *"telefone"* e ocorresse um erro na transmissão, de modo que a palavra recebida fosse *"belefone"*, o código C detectaria que houve um erro, pois *belefone* não pertence ao conjunto C. Uma vez detectado o erro, seria fácil corrigi-lo, pois a palavra pertencente a C que mais se aproxima de *belefone* é *telefone*. Sendo assim, saberíamos que a palavra transmitida, na realidade havia sido *telefone*. Em contrapartida, se a palavra transmitida fosse *"bola"* e por ventura ocorresse um erro na transmissão, de modo que a palavra recebida fosse *"wola"*, o erro seria detectado, pois *wola* não é uma palavra pertencente a C, porém, a correção seria impossível, uma vez que em C existem várias palavras que igualmente se aproximam de *wola*, por exemplo *bola*, *cola*, *mola*, *sola* e *gola*. Em uma terceira hipótese, se a palavra transmitida fosse *"caneca"* e ocorresse um erro na transmissão, de modo que a palavra recebida fosse *"canela"*, o código C nem detectaria o erro, pois a palavra *canela* também pertence a C.

Para exemplificar os princípios da teoria dos códigos, analisemos o seguinte caso:

Suponhamos que o braço mecânico de base fixa da figura 2, através de comandos digitais, possibilite quatro movimentos básicos: *para cima*, *para baixo*, *para a direita* e *para a esquerda*:

Figura 2: Braço mecânico

Fonte: <http://thing-better.blogspot.com.br/2013/04/o-que-e-robotica_6379.html>

[6] Doença pulmonar causada pela inalação de cinzas de origem vulcânica.

Aos comandos acima denominaremos *"fonte"*.

Os circuitos digitais (ou circuitos lógicos) baseiam seu funcionamento na lógica binária, ou seja, cada informação deve ser expressa utilizando-se de dois dígitos, a saber, 0 e 1. Como temos dois dígitos disponíveis para expressar os comandos e dispomos de quatro comandos básicos para o braço mecânico, considerando o conjunto $F = \{0,1\}$, podemos codificar os quatro comandos como elementos de $F^2 = F \times F = \{(0,0), (0,1), \}(1,0), (1,1)\}$. Por simplicidade de notação, consideraremos cada par $(a, b) \in F^2$ simplesmente como ab e a cada um dos quatro comandos 00, 01, 10 e 11 denominaremos *"código da fonte"*. Por exemplo:

Fonte	**Código da fonte**
Para a esquerda:	00
Para a direita:	01
Para cima:	10
Para baixo:	11

Imaginemos agora que os comandos (mensagens) sejam transmitidos ao braço mecânico via sinais de rádio frequência, através de um controle remoto por exemplo. Suponhamos que seja dado ao braço mecânico o comando *"para a esquerda"*, o que será convertido para o código de fonte 00 e enviado ao braço mecânico, indicando para que ele se mova para a esquerda. Suponhamos ainda que a transmissão do sinal, por alguma interferência externa, sofra um erro e chegue até o braço mecânico como 10, o que acarreta que o braço em vez de mover-se para a esquerda, fosse movido para cima. Observemos que o circuito digital do braço mecânico seria incapaz de detectar o erro, pois 10 é um comando existente em seu banco de dados.

Diante de uma situação como a descrita acima, o que fazemos é inserir redundâncias, através do acréscimo de dígitos nos códigos da fonte, de modo que se possa detectar e corrigir possíveis erros de transmissão, dando origem a um novo código ao qual denominamos de *"código de canal"*:

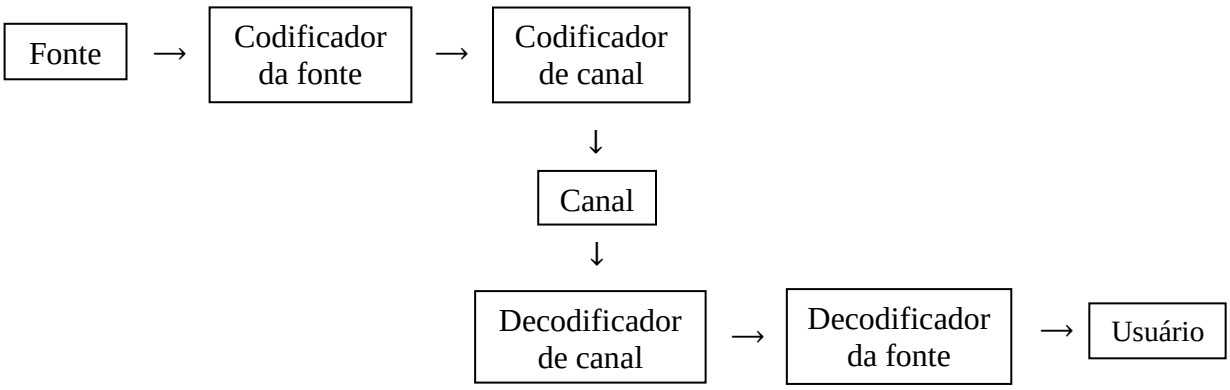
Fonte	**Código da fonte**	**Código de canal**
Para a esquerda:	00	00000
Para a direita:	01	01011
Para cima:	10	10110
Para baixo:	11	11101

Nesta nova codificação, as duas primeiras posições representam o código da fonte enquanto as três últimas posições são as redundâncias inseridas.

Façamos a seguir uma nova análise:

Suponhamos que seja dado ao braço mecânico o comando *"para a direita"*, o que será convertido para o código da fonte 01 e em seguida acrescido de redundâncias será convertido para o código de canal 01011 e enviado ao braço mecânico, indicando para que ele se mova para a direita. Suponhamos ainda que a transmissão do sinal, por alguma interferência externa, sofra um erro e chegue até o braço mecânico como 01010. Notemos que esse comando não existe em banco de dados e isso acarretaria a identificação de um erro pelo circuito digital do braço mecânico. Consultando seu banco de dados, o comando que mais se aproxima de 01010 é 01011 e, portanto, o circuito digital faria a correção, interpretando o comando recebido como 01011 e movendo o braço mecânico para a direita.

Vejamos em diagrama de blocos a seguir, todas as etapas desde o comando dado até a chegada da mensagem transmitida:



O estudo da teoria dos códigos apresentado nesse trabalho, objetivará a transformação de códigos da fonte em códigos de canal, as detecções e correções de possíveis erros ocorridos durante o processo de transmissão e a decodificação de códigos de canal em códigos da fonte. Consideraremos nesse estudo apenas canais simétricos, canais estes que possuem as seguintes características:

- Todos os caracteres transmitidos tem a mesma probabilidade (ínfima) de serem recebidos errados;
- Se um caractere é recebido errado, a probabilidade de ele ser qualquer um dos outros caracteres disponíveis é a mesma.

6.2 MÉTRICA DE HAMMING

Primeiramente entendamos o que é uma métrica.

Uma métrica é uma generalização do conceito geométrico de distância. Dizemos que, dado um conjunto T, uma métrica em T é uma função $d: T \times T \rightarrow \mathbb{R}$ que a cada $(x, y) \in T \times T$ faz corresponder o elemento $d(x, y) \in \mathbb{R}$, denominado a distância de x a y, tal que para todo $x, y, z \in T$, temos:

- $d(x, y) \geq 0$, valendo a igualdade quando $x = y$
- $d(x, y) = d(y, x)$ (simetria)
- $d(x, z) \leq d(x, y) + d(y, z)$ (desigualdade triangular)

Dado um conjunto A, finito, ao qual denominaremos de alfabeto. Representaremos o número de elementos de A por $|A| = q$. Definimos um código corretor de erros como sendo um subconjunto próprio qualquer de A^n, com $n \in \mathbb{N}$. Dados $u, v \in A^n$, denominamos de *"distância de Hamming"* ao valor $d(u, v) = |\{i, u_i \neq v_i, 1 \leq i \leq n\}|$.

Vejamos um exemplo:

Sendo $A = \{0,1\}$, para $n = 4$, temos $|A^4| = 16$ e $\{0000, 0001, 1010, 1011, 1111\} \subset A^4$. Assim:

$d(1010,1011) = 1$

$d(0001,1011) = 2$

$d(0001,1111) = 3$

$d(0000, 1111) = 4$

Consideremos, de maneira geral, os elementos $u, v, w \in A^n$, tais que $u = u_1 u_2 u_3 \ldots u_n$, $v = v_1 v_2 v_3 \ldots v_n$ e $w = w_1 w_2 w_3 \ldots w_n$. Como $u_i, v_i \in \{0,1\}$ para todo $i \in \{1, 2, \ldots, n\}$, se tivermos $u_i = v_i$, então $d(u, v) = 0$, caso existam k índices i para os quais $u_i \neq v_i$ então, por definição, $d(u, v) = k > 0$, logo, deduzimos que $d(u, v) \geq 0$. Por outro lado, se tivermos $u_i = v_i$ para todo $i \in \{1, 2, \ldots, n\}$, então $u = v$, o que implica que $d(u, v) = 0$ e $d(v, u) = 0$, acarretando que $d(u, v) = d(v, u)$. Caso existam k índices i para os quais $u_i \neq v_i$ então, por definição, $d(u, v) = k$ e $d(v, u) = k$, implicando que $d(u, v) = d(v, u)$.

Para cada índice i, a contribuição para a distância $d(u, v)$, das $i - ésimas$ coordenadas de u e v é igual a 0 ou 1, respectivamente se $u_i = v_i$ ou $u_i \neq v_i$. De maneira análoga a contribuição para a distância $d(v, w)$, das $i - ésimas$ coordenadas de v e w é igual a 0 ou 1, respectivamente se $v_i = w_i$ ou $v_i \neq w_i$ e a contribuição para a distância $d(u, w)$, das $i - ésimas$ coordenadas de u e w é igual a 0 ou 1, respectivamente se $u_i = w_i$

ou $u_i \neq w_i$. Considerando que a contribuição para a distância $d(u,w)$, das $i-$ésimas coordenadas de u e w seja 0, ou seja, $u_i = w_i$, então temos $d(u,w) \leq d(u,v) + d(v,w)$, pois a contribuição das $i-$ésimas coordenadas de u_i e v_i e v_i e w_i em $d(u,v) + d(v,w)$ é igual a $0, 1$ ou 2. Caso consideremos $u_i \neq w_i$, então não se tem $u_i = v_i$ e $v_i = w_i$, pois seria contrário a hipótese, assim, temos que a contribuição das $i-$ésimas coordenadas de u_i e v_i e v_i e w_i em $d(u,v) + d(v,w)$ é maior ou igual a 1, que, por hipótese, é a contribuição das $i-$ésimas coordenadas de u_i e w_i em $d(u,w)$. Portanto, temos sempre $d(u,w) \leq d(u,v) + d(v,w)$.

Concluímos com isso que a distância de Hamming entre os elementos de A^n cumpre as três condições necessárias para classificá-la como uma métrica, portanto, a partir desse momento a denominaremos de *métrica de Hamming*.

6.2.1 Disco e esfera de centro c e raio r

Consideremos um elemento $c \in A^n$ e $r \in \mathbb{R}$, tal que $r \geq 0$.

Dizemos que um *disco* de centro c e raio r é um conjunto $D(c,r) = \{u \in A^n; d(u,c) \leq r\}$. De maneira análoga, definimos uma *esfera* de centro c e raio r como um conjunto $S(c,r) = \{u \in A^n; d(u,c) = r\}$.

Discos e esferas são conjuntos finitos como veremos a seguir:

Sendo $|A| = q$ o número de elementos do alfabeto A e $u \in A$ uma palavra desse alfabeto, em cada coordenada de u temos $q-1$ elementos de A^n distintos, que podem variar nas i coordenadas de u, obtendo com isso $(q-1)^i$. Como u tem tamanho n e as i entradas distintas podem percorrer qualquer coordenada de u, temos então a combinação $\binom{n}{i}$. Se $|S(c,i)|$ representa o número de elementos da esfera S de centro c e raio i, então $|S(c,i)| = \binom{n}{i}.(q-1)^i$, o que nos mostra que $S(c,i)$ possui um número finito de elementos. Notemos ainda que $S(c,i) \cap S(c,j) = \emptyset$ quando $i \neq j$ e que $\bigcup_{i=0}^{r} S(c,i) = D(c,r)$, portanto, o número de elementos do conjunto $D(c,r)$, representado por $|D(c,r)|$, também é finito, pois $|D(c,r)| = |\bigcup_{i=0}^{r} S(c,i)| = \sum_{i=0}^{r} |S(c,i)| = \sum_{i=0}^{r} \binom{n}{i}.(q-1)^i$.

Vejamos um exemplo:

Consideremos o alfabeto $A = \{0,1\}$, portanto $|A| = q = 2$. Para $n = 3$, temos $A^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Consideremos $c = 010 \in A^3$ e $r = 2$. Temos

então o conjunto $S(c,r) = S(010,2) = \{001, 100, 111\}$ e $D(c,r) = D(010,2) = \{000, 001, 010, 011, 100, 110, 111\}$. Notemos que $|S(010,2)| = \binom{3}{2}.(2-1)^2 = 3.1 = 3$ e que $|D(010,2)| = \sum_{i=0}^{3}\binom{3}{i}.(2-1)^i = \binom{3}{0}.1^0 + \binom{3}{1}.1^1 + \binom{3}{2}.1^2 = 1+3+3 = 7$.

6.2.2 Distância mínima de um código

Dado um código C, definimos sua distância mínima como sendo um número d, tal que $d = min\{d(u,v); u,v \in C\ e\ u \neq v\}$.

No exemplo do braço mecânico, tínhamos o alfabeto $F = \{0,1\}$, do qual obtivemos o código de fonte $F^2 = \{00, 01, 10, 11\}$ e, através do acréscimo de redundâncias, obtivemos o código $C \subset F^5$ tal que $C = \{u_1, u_2, u_3, u_4\} = \{00000, 01011, 10110, 11101\}$. Notemos que $d(u_1,u_2) = 3$, $d(u_1,u_3) = 3$, $d(u_1,u_4) = 4$, $d(u_2,u_3) = 4$, $d(u_2,u_4) = 3$ e $d(u_3,u_4) = 3$, portanto, $d = min\{d(u,v); u,v \in C\ e\ u \neq v\} = min\{3,4\} = 3$.

Percebe-se que para determinarmos d no exemplo dado, foram necessários os cálculos de seis distâncias e que à medida que C possua um número maior de palavras, mais cálculos de distâncias serão necessários para a determinação de d. De maneira geral, para o cálculo de d são necessários os cálculos de $\binom{|C|}{2}$ distâncias, onde $|C|$ representa o número de elementos do conjunto C. Porém, o cálculo de $\binom{|C|}{2}$ demanda um custo computacional exagerado, o que inviabiliza esse método e, para tanto, veremos mais adiante outras maneiras para se encontrar d com um esforço computacional minimizado.

Considerando C um código de distância mínima d, definimos $\kappa = \left[\frac{d-1}{2}\right]$, onde $\left[\frac{d-1}{2}\right]$ representa a parte inteira do número real $\frac{d-1}{2}$.

Teorema 6.1: Considerando C um código de distância mínima d. Se $c,c' \in C$ e $c \neq c'$, então $D(c,\kappa) \cap D(c',\kappa) = \emptyset$.

Demonstração: Suponhamos que $D(c,\kappa) \cap D(c',\kappa) \neq \emptyset$, ou seja, existe $u \in D(c,\kappa) \cap D(c',\kappa)$, então temos que $d(u,c) \leq \kappa$ e $d(u,c') \leq \kappa$, mas pela métrica de Hamming, temos $d(u,c) = d(c,u)$ e $d(c,c') \leq d(c,u) + d(u,c')$, o que implica que $d(c,c') \leq \kappa + \kappa = 2\kappa \leq d-1$, o que contradiz a hipótese, pois d é a distância mínima, ou seja, $d(c,c') \geq d$. Portanto se $c,c' \in C$ e $c \neq c'$, então $D(c,\kappa) \cap D(c',\kappa) = \emptyset$.

6.2.3 Número de detecções e número de correções de erros

A distância mínima d de um código C tem grande relevância nos processos de detecção e correção de erros.

Teorema 6.2: Considere um código C com distância mínima d. C pode detectar até $d-1$ erros.

Demonstração: Sendo d a distância mínima de um código C, sabemos que dada uma palavra $c \in C$, qualquer outra palavra c' do código C está a uma distância no mínimo igual a d da palavra c. Isso significa que podemos introduzir em uma palavra qualquer de C até $d-1$ erros sem encontrar outra palavra de C, tornando possível a detecção do erro.

Teorema 6.3: Consideremos um código C com distância mínima d. O código C pode corrigir até $\kappa = \left[\frac{d-1}{2}\right]$ erros.

Demonstração: Suponhamos que uma palavra $c \in C$ sofra t erros, com $t \leq \kappa$, ao ser transmitida, de modo que r seja a palavra recebida. Temos então $d(r,c) = t \leq \kappa$ e, pelo teorema 6.1, a distância de r a qualquer outra palavra de C é maior do que κ, assim, a palavra c é univocamente determinada a partir da palavra r.

Exemplo:

Considerando o código C dos comandos do braço mecânico, como vimos que $d = 3$, então $d - 1 = 3 - 1 = 2$ e $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{3-1}{2}\right] = 1$, portanto, no código C é possível detectar até 2 erros e corrigir 1 erro.

6.2.4 Códigos perfeitos

Um código $C \subset A^n$, com distância mínima d e $\kappa = \left[\frac{d-1}{2}\right]$ é denominado *código perfeito* se $\bigcup_{c \in C} D(c,k) = A^n$.

Observemos que o código C do braço mecânico não é perfeito, pois $C \subset F^5$, $C = \{u_1, u_2, u_3, u_4\} = \{00000, 01011, 10110, 11101\}$ e $\kappa = 1$, mas considerando a palavra $p = 11010 \in F^5$, vemos que $d(u_1, p) = d(u_4, p) = 3$ e $d(u_2, p) = d(u_3, p) = 2$. Como

$k = 1$, significa que $p \notin D(u_1, 1) \cup D(u_2, 1) \cup D(u_3, 1) \cup D(u_4, 1)$, o que implica que $\bigcup_{u_i \in C} D(u_i, 1) \neq F^n$, ou seja, $C \subset F^5$ não é código perfeito.

6.2.5 Equivalência de códigos

Primeiramente falemos sobre isometrias.

Dados dois espaços métricos (conjuntos munidos de uma métrica) X e Y e dados dois elementos $x_1, x_2 \in X$, de modo que a distância entre x_1 e x_2 no espaço X seja $d_X(x_1, x_2)$. Uma função $f: X \to Y$ que a cada $x \in X$ faça corresponder a $f(x) \in Y$ é denominada uma *isometria* se em relação a distância no espaço métrico Y, for válida a igualdade $d_Y\big(f(x_1), f(x_2)\big) = d_X(x_1, x_2)$ para todo $x_1, x_2 \in X$, ou seja, f é uma transformação que preserva a distância.

Considerando um conjunto A, ao qual denominamos alfabeto e um número natural n, o conjunto A^n de todas as palavras de tamanho n é um espaço métrico, pois nele temos definida a métrica de Hamming. Sendo assim, uma função $f: A^n \to A^n$ é uma isometria de A^n se preservar distâncias de Hamming, ou seja, $d\big(f(x), f(y)\big) = d(x, y)$ para todo $x, y \in A^n$.

Considerando isometrias para a métrica de Hamming:

Teorema 6.4: Se $f: A^n \to A^n$ é uma isometria, então f é uma bijeção.

Demonstração: Consideremos $f: A^n \to A^n$ uma isometria. Suponhamos que dados $x, y \in A^n$, tenhamos $f(x) = f(y)$, o que implica que $d\big(f(x), f(y)\big) = 0$. Mas, por hipótese, $f: A^n \to A^n$ é uma isometria, então $d(x, y) = d\big(f(x), f(y)\big)$ o que implica que $d(x, y) = 0$ e, portanto, $x = y$, mostrando que f é injetiva. Como A^n é um conjunto finito e toda bijeção de um conjunto finito nele próprio é uma sobrejeção, temos com isso que f é sobrejetiva e portanto bijetiva.

Teorema 6.5: A função identidade $I_{A^n}: A^n \to A^n$ é uma isometria.

Demonstração: Temos que para todo $x, y \in A^n$, $I_{A^n}(x) = x$ e $I_{A^n}(y) = y$, o que implica $d\big(I_{A^n}(x), I_{A^n}(y)\big) = d(x, y)$, mostrando que I_{A^n} é uma isometria.

Teorema 6.6: Se f é uma isometria de A^n, então f^{-1} também o é.

Demonstração: Se f é uma isometria, então pelo teorema 6.4, f é bijetiva, o que garante a existência de f^{-1}. Como por hipótese f é uma isometria, então $d\big(f^{-1}(x), f^{-1}(y)\big) = d\big(f\big(f^{-1}(x)\big), f\big(f^{-1}(y)\big)\big) = d(x, y)$, mostrando com isso, que f^{-1} é uma isometria.

Teorema 6.7: Se f_1 e f_2 são isometrias de A^n, então $f_1 \circ f_2$ é uma isometria de A^n.

Demonstração: Se f_1 e f_2 são isometrias de A^n, então $d\left(f_1\big(f_2(x)\big), f_1\big(f_2(y)\big)\right) = d\big(f_2(x), f_2(y)\big) = d(x, y)$, mostrando com isso que $f_1 \circ f_2$ é uma isometria de A^n.

Dados dois códigos C_1 e C_2 contidos em A^n, dizemos que C_1 e C_2 são códigos equivalentes quando existe uma isometria f de A^n tal que $f(C_1) = C_2$.

Os parâmetros fundamentais de um código $C \subset A^n$ são o seu comprimento n, o seu número de elementos $|C| = M$ e a sua distância mínima d. Representamos os parâmetros de um código $C \subset A^n$ pela terna $[n, M, d]$.

Teorema 6.8: Dois códigos equivalentes C_1 e C_2 de A^n possuem os mesmos parâmetros.

Demonstração: Suponhamos que os parâmetros do código C_1 são $[n, M, d]$. Como C_2 é código de A^n, então todas as suas palavras tem comprimento n. Como C_1 e C_2 são equivalentes, então existe uma isometria f de A^n tal que $f(C_1) = C_2$ e, pelo teorema 6.4 f é bijetiva, logo $|C_2| = |C_1| = M$. Por fim, sejam $x, y \in C_1$ tais que $d(x, y) = d$, temos então, $d(x, y) = d\big(f(x), f(y)\big) = d$, mostrando que a mínima distância em C_2 também é d. Assim, os parâmetros do código C_2 são $[n, M, d]$.

6.3 CÓDIGOS LINEARES

Consideremos um corpo finito K com um número q de elementos, ao qual denominaremos alfabeto.

Seja $n \in \mathbb{N}$, temos que K^n é um k-espaço vetorial de dimensão n.

Um código $C \subset K^n$ é classificado como um código linear quando C for um subespaço vetorial de K^n.

O exemplo do braço mecânico utilizado anteriormente é um código linear, pois o conjunto $C = \{00000, 01011, 10110, 11101\}$, contido em F^5 é um subespaço vetorial de F^5, verifiquemos:

O corpo de escalares $F = \{0,1\}$ contém dois elementos e, qualquer que seja o vetor $u \in C$, temos $0.u = 00000$ e $1.u = u$. Notemos também que $00000 + 01011 = 01011$, $00000 + 10110 = 10110$, $00000 + 11101 = 11101$, $01011 + 10110 = 11101$, $01011 + 11101 = 10110$ e $10110 + 11101 = 01011$. Desse modo, para todo $\alpha, \beta \in F$ e $u, v \in C$, temos $\alpha u + \beta v \in C$, mostrando com isso que C é subespaço vetorial de F^5.

Como um código linear é um subespaço de um *K*-espaço vetorial de dimensão finita, então todo código linear é, também, um *K*-espaço vetorial de dimensão finita. Sendo k o número de elementos de uma das bases de C (dimensão de C) e, sendo $u_1, u_2, u_3, \dots, u_k$ uma dessas bases , então qualquer que seja $u \in C$, u se escreve de maneira única como $u = \alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \alpha_3 \cdot u_3 + \cdots + \alpha_k \cdot u_k$, $\forall \alpha_i \in K$ e, portanto, o número de elementos do código C é $M = |C| = q^k$ ou seja, $\dim C = k = \log_q q^k = \log_q M$.

No exemplo do braço mecânico, temos $C = \{00000, 01011, 10110, 11101\}$ e $F = \{0,1\}$, o que implica que $M = 4$ e $q = 2$, portanto $\dim C = \log_2 4 = 2$, ou seja, qualquer base de C possui dois vetores.

6.3.1 Peso de um código

Considerando d a métrica de Hamming, dado um vetor u do *K*-espaço vetorial K^n, definimos o peso de u como sendo o número inteiro $\omega(u) = d(u, 0)$ e, o peso de um código C é definido como sendo a distância de Hamming mínima não nula dos vetores de C ao vetor nulo. Em outras palavras:

$\omega(C) = min\{\omega(u); u \in C \backslash \{0\}\}$.

Teorema 6.9: Considerando um código linear $C \subset K^n$, com distância mínima d, temos $\forall u, v \in K^n$, $d(u, v) = \omega(u - v)$ e $d = \omega(C)$.

Demonstração: Dados $u, v \in K^n$, pela definição de distância, temos $d(u, v) = d(u - v, 0) = \omega(u - v)$ e, se $u, v \in C$ e $u \neq v$, e a distância mínima do código C é $d = d(u, v)$, então existe $w \in C \backslash \{0\}$ tal que $w = u - v$ e $d = d(u, v) = \omega(u - v) = \omega(w) = \omega(C)$.

6.3.2 Matriz geradora de um código

Consideramos um corpo finito K com q elementos e um código linear $C \subset K^n$. À terna (n, k, d) *denominamos parâmetros do código linear* C. O parâmetro n representa o número de coordenadas de cada vetor (palavra) do código C; o parâmetro k representa a dimensão do Código (espaço vetorial) C sobre o corpo K e o parâmetro d representa a distância mínima do código C, que é igual ao peso $\omega(C)$ do código C.

Consideremos $B = \{u_1, u_2, u_3, \dots, u_k\}$ uma base ordenada de C, onde cada vetor $u_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{in})$, com $1 \leq i \leq k$ e uma matriz $G = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}$, ou seja, $G = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kn} \end{bmatrix}$. A matriz G é denominada *matriz geradora de C associada à base B* e não é a única matriz geradora de C, pois, para cada base diferente de C , obtemos uma geradora diferente. Notemos que uma matriz geradora de um código C pode ser obtida de outra matriz geradora através de transformações elementares sobre matrizes, vistas em 2.5.

Consideremos agora uma transformação linear $T: K^k \to K^n$ de modo que dado $x \in K^k$, tenhamos $T(x) = x \cdot G$. Como $x \in K^k$, então possui k coordenadas $x_1, x_2, x_3, \dots, x_k$ e, portanto, $T(x) = x_1 \cdot v_1 + x_2 \cdot v_2 + x_3 \cdot v_3 + \cdots + x_k \cdot v_k$, o que implica que $T(K^k) = C$, assim, temos K^k o código da fonte, C é o código de canal e T é a codificação, que leva o código da fonte ao código de canal.

Para obter uma matriz geradora de um código de dimensão k, contido em um espaço K^n, basta tomar uma matriz com k linhas linearmente independentes e n colunas. Por exemplo, considerando o corpo galoisiano $F = \{0,1\}$ e uma matriz $G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$. Consideremos as palavras do código da fonte como sendo vetores de F^3. Uma palavra x do código da fonte é codificada em código de canal através da transformação $T: F^3 \to F^5$, tal que $T(x) = x \cdot G$. Suponhamos $x = 110$, assim, temos:

$$T(110) = [1 \quad 1 \quad 0].\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} =$$

$$= [1.1 + 1.1 + 0.0 \quad 1.0 + 1.1 + 0.1 \quad 1.0 + 1.0 + 0.1 \quad 1.1 + 1.0 + 0.1 \quad 1.0 + 1.1 + 0.0] =$$

$= [0 \quad 1 \quad 0 \quad 1 \quad 1]$, portanto, a palavra 110 do código da fonte é codificada como 01011 no código do canal.

Notemos nesse exemplo que $q = 2$, pois adotamos o corpo finito (galoisiano) $F = \{0,1\}$ e $\mathrm{k} = \dim C = 3$, logo, o número de elementos de C é $M = 2^3 = 8$. O código C é, portanto, o seguinte conjunto:

$C = \{00000, 10101, 11010, 11111, 01111, 01010, 00101, 10000\}$, que facilmente pode ser verificado que foi obtido através dos vetores 10010, 11001 e 01110, que constituem uma base de C.

Caso deseje-se decodificar as palavras do código de canal C de modo a obter as palavras do código de fonte, basta tomar os vetores $x = x_1x_2x_3$ de F^3 e resolver a equação $x.G = y$, onde $y = y_1y_2y_3y_4y_5$é uma palavra do código de canal C. Porém, esse procedimento consiste em resolver a equação matricial:

$[x_1 \quad x_2 \quad x_3] \cdot \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5]$, que gera o sistema $\begin{cases} x_1 + x_2 = y_1 \\ x_2 + x_3 = y_2 \\ x_3 = y_3 \\ x_1 + x_3 = y_4 \\ x_2 = y_5 \end{cases}$ que, em geral exige um alto custo computacional e, portanto, é inviável. Mas efetuando as operações elementares (vistas em 2.5) sobre as linhas da matriz G, obtemos uma matriz G' com a seguinte forma:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{L_2 \rightarrow L_1 + L_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{L_3 \rightarrow L_2 + L_3} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = G'.$$

A matriz G' é equivalente por linhas à matriz G e, portanto, suas linhas são linearmente independentes e em consequência disso formam outra base de C. Assim, podemos obter as palavras do código de fonte resolvendo a equação matricial $x.G' = y$, que equivale a $[x_1 \quad x_2 \quad x_3] \cdot \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5]$, ou ainda a $[x_1 \quad x_2 \quad x_3 \quad x_1 + x_2 \quad x_2 + x_3] = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5]$, ou seja as palavras do código da fonte são obtidas considerando as três primeiras coordenadas das palavras do código de canal. Assim, as palavras do código de fonte são $\{000, 101, 110, 111, 011, 010, 001, 100\}$.

Dizemos que uma matriz geradora de um código C se encontra na forma padrão, se for apresentada na forma $[I_k|A]$, com I_k sendo a matriz identidade de ordem k e A uma matriz cuja ordem é $k \times (n - k)$. No exemplo anterior a matriz G não estava na forma padrão,

porém ao serem efetuadas operações elementares sobre as linhas de G, foi obtida uma matriz G', equivalente por linhas à matriz G e apresentada na forma padrão.

Outra maneira de se obter uma matriz G' na forma padrão, equivalente a matriz G, é resolver o produto $M^{-1} \cdot G$, onde M é a matriz quadrada de ordem k obtida pelo bloco das k primeiras colunas da matriz G. Pelo teorema 3.3 do capítulo 3, temos que $M^{-1} = \frac{1}{\det(M)} \cdot \overline{M}$, quando $\det(M) \neq 0$. No exemplo anterior, $M = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ e pelo teorema 3.1, $\det(M) = 1$. Calculemos a seguir a matriz dos cofatores de M:

$$M' = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$$

$M_{11} = (-1)^{1+1} \cdot \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 \cdot 1 = 1$ $\quad$ $M_{12} = (-1)^{1+2} \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1 \cdot 1 = -1 = 1$

$M_{13} = (-1)^{1+3} \cdot \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \cdot 1 = 1$ $\quad$ $M_{21} = (-1)^{2+1} \cdot \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} = -1 \cdot 0 = 0$

$M_{22} = (-1)^{2+2} \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \cdot 1 = -1 = 1$ $\quad$ $M_{23} = (-1)^{2+3} \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1 \cdot 1 = -1 = 1$

$M_{31} = (-1)^{3+1} \cdot \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} = 1 \cdot 0 = 0$ $\quad$ $M_{32} = (-1)^{3+2} \cdot \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} = -1 \cdot 0 = 0$

$M_{33} = (-1)^{3+3} \cdot \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 \cdot 1 = 1$

Logo, $M' = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Como $\overline{M} = (M')^t$, temos $\overline{M} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Assim, temos:

$$M^{-1} = \frac{1}{\det(M)} \cdot \overline{M} = \frac{1}{1} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Como $G' = M^{-1} \cdot G$, então $G' = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$

Notemos que $G' = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = [I_3|A]$, com $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$. Cabe salientar que nem sempre é possível obter uma matriz G na forma padrão, de um código C, apenas realizando operações elementares sobre linhas. Veja o exemplo:

Consideremos um código $C \subset F^5$ cuja matriz geradora é $G = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. É fácil verificar que nenhuma operação elementar sobre as linhas de G fará com que a matriz se apresente na forma padrão, uma vez que todos os elementos da primeira coluna de G são nulos. Porém, aplicando as operações de permutação entre duas colunas da matriz G e multiplicação de uma coluna de G por um escalar não nulo, podemos obter uma matriz G',

geradora na forma padrão, de um código $C' \subset F^5$ que é equivalente ao código C. Por exemplo:

$$G = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \overrightarrow{c_1 \to c_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \overrightarrow{c_2 \to c_3} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \overrightarrow{c_3 \to c_5}$$

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

Teorema 6.10: Sendo C um código, existe um código C', equivalente a C, cuja matriz geradora se apresenta na forma padrão.

Demonstração: Seja C um código cuja matriz geradora é $G = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix}$.

Utilizando as operações elementares sobre as linhas de um matriz (vistas no capítulo 2) e operações sobre as colunas de G, temos:

Como as linhas de G constituem uma base de C, então são linearmente independentes e, portanto nenhuma linha é nula. Consideremos, sem perda de generalidade, que $x_{11} \neq 0$. Como x_{11} é elemento de um corpo, então possui um inverso multiplicativo x_{11}^{-1} tal que $x_{11} \cdot x_{11}^{-1} = 1$. Multiplicando a primeira linha de G por x_{11}^{-1}, obtermos: $\begin{bmatrix} 1 & y_{12} & \cdots & y_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix}$.

Substituindo cada linha dessa matriz, a partir da segunda, pela soma da respectiva linha com a primeira multiplicada por $-x_{21}$, ..., $-x_{k1}$, respectivamente, temos a seguinte matriz: $\begin{bmatrix} 1 & y_{12} & \cdots & y_{1n} \\ 0 & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & y_{k2} & \cdots & y_{kn} \end{bmatrix}$. A segunda linha dessa matriz possui algum elemento não nulo e, por meio de uma permutação entre colunas, é possível fazer com que esse elemento não nulo ocupe a posição segunda linha e segunda coluna. Multiplicando a segunda linha pelo inverso desse elemento não nulo e somando cada uma das linhas restantes, pela segunda linha multiplicada respectivamente por $-y_{12}$, $-y_{13}$, ..., $-y_{k2}$, temos a matriz $\begin{bmatrix} 1 & 0 & \cdots & z_{1n} \\ 0 & 1 & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{kn} \end{bmatrix}$.

Repetindo o processo descrito acima, uma quantidade de até k vezes, obtemos uma matriz $G' = [I_k|A]$, na forma padrão.

Uma matriz geradora que não se apresenta na forma padrão, gera palavras código não sistemáticas, ou seja, palavras código nas quais os dígitos das palavras do código da fonte estão misturados com os dígitos da redundância acrescentada. Enquanto que uma matriz geradora que se apresente na forma padrão, gera palavras do código sistemáticas, nas quais os k primeiros dígitos correspondem aos dígitos do código da fonte, enquanto que os $n-k$ últimos dígitos correspondem aos dígitos da redundância acrescida.

6.3.3 Códigos duais

Considerando C um código linear contido em um espaço vetorial K^n definimos o complemento ortogonal de C como sendo o conjunto $C^{\perp} = \{v \in K^n;\ \langle u, v\rangle = 0, \forall u \in C\}$.

Se temos um código linear $C \subset K^n$, então $C^{\perp}$ é um subespaço vetorial de K^n, pois dados $u, v \in C^{\perp}$, $\alpha, \beta \in K$ e $w \in C$, temos $\langle \alpha \cdot u + \beta \cdot v, w\rangle = \alpha \cdot \langle u, w\rangle + \beta \cdot \langle v, w\rangle = 0$. Além disso, se G é matriz geradora do código C e $w \in C^{\perp}$, então $G \cdot w^t = 0$, o que é facilmente verificável, uma vez que cada linha $v_1, v_2, v_3, \ldots, v_k$ de G é um vetor de uma das bases de C e, portanto, $\langle v_1, w^t\rangle = \langle v_2, w^t\rangle = \langle v_3, w^t\rangle = \cdots = \langle v_k, w^t\rangle = 0$. Como foi mostrado, $C^{\perp}$ é um subespaço vetorial de K^n, portanto, por definição, $C^{\perp}$ é também um código linear.

Definição: Um subespaço vetorial $C^{\perp} \subset K^n$ que é um complemento ortogonal do código C e é também um código linear, é denominado *código dual* de C.

Teorema 6.11: Considerando C um código linear contido em K^n, com dimensão k, cuja matriz geradora na forma padrão é $G = [I_k|A]$, temos $\dim C^{\perp} = n - k$.

Demonstração: Vimos anteriormente que $w = w_1, w_2, w_3, \ldots, w_n$ pertence a $C^{\perp}$, quando $G \cdot w^t = 0$. Como $G = [I_k|A]$ se apresenta na forma padrão, então, temos:

$$G \cdot w^t = \begin{bmatrix} 1 & 0 & \cdots & 0 & g_{(k+1)1} & g_{(k+2)1} & \cdots & g_{n1} \\ 0 & 1 & \cdots & 0 & g_{(k+1)2} & g_{(k+2)2} & \cdots & g_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & g_{(k+1)k} & g_{(k+2)k} & \cdots & g_{nk} \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \begin{bmatrix} w_1 + g_{(k+1)1} \cdot w_{k+1} + \cdots + g_{n1} \cdot w_n \\ w_2 + g_{(k+1)2} \cdot w_{k+1} + \cdots + g_{n2} \cdot w_n \\ \vdots \\ w_k + g_{(k+1)k} \cdot w_{k+1} + \cdots + g_{nk} \cdot w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} + \begin{bmatrix} g_{(k+1)1} \cdot w_{k+1} + \cdots + g_{n1} \cdot w_n \\ g_{(k+1)2} \cdot w_{k+1} + \cdots + g_{n2} \cdot w_n \\ \vdots \\ g_{(k+1)k} \cdot w_{k+1} + \cdots + g_{nk} \cdot w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} = \begin{bmatrix} -g_{(k+1)1} \cdot w_{k+1} - \cdots - g_{n1} \cdot w_n \\ -g_{(k+1)2} \cdot w_{k+1} - \cdots - g_{n2} \cdot w_n \\ \vdots \\ -g_{(k+1)k} \cdot w_{k+1} - \cdots - g_{nk} \cdot w_n \end{bmatrix} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} = - \begin{bmatrix} g_{(k+1)1} & g_{(k+2)1} & \cdots & g_{n1} \\ g_{(k+1)2} & g_{(k+2)2} & \cdots & g_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ g_{(k+1)k} & g_{(k+2)k} & \cdots & g_{nk} \end{bmatrix} \cdot \begin{bmatrix} w_{k+1} \\ w_{k+2} \\ \vdots \\ w_n \end{bmatrix}.$$

Os $n-k$ elementos $w_{k+1}, w_{k+2}, \dots, w_n$ podem ser escolhidos de forma aleatória. Logo, temos que $\dim C^{\perp} = n-k$.

Teorema 6.12: Considerando C um código linear contido em K^n, com dimensão k, cuja matriz geradora na forma padrão é $G = [I_k|A]$, temos $H = [-A^t|I_{n-k}]$ é uma matriz geradora de $C^{\perp}$.

Demonstração: Considerando $i \in \{1, 2, \dots, k\}$, temos que cada coordenada w_i de um vetor $w \in C^{\perp}$ é escrita como $w_i = -g_{(k+1)i} \cdot w_{k+1} - g_{(k+2)i} \cdot w_{k+2} - \cdots - g_{ni} \cdot w_n$, o que implica que $w = (-g_{(k+1)1} \cdot w_{k+1} - g_{(k+2)1} \cdot w_{k+2} - \cdots - g_{n1} \cdot w_n, -g_{(k+1)2} \cdot w_{k+1} - g_{(k+2)2} \cdot w_{k+2} - \cdots - g_{n2} \cdot w_n, \dots, -g_{(k+1)k} \cdot w_{k+1} - g_{(k+2)k} \cdot w_{k+2} - \cdots - g_{nk} \cdot w_n, w_{k+1}, w_{k+2}, \dots, w_n)$, o que implica que

$$\left\{ \begin{matrix} \left(-g_{(k+1)1}, -g_{(k+1)2}, \dots, -g_{(k+1)k}, 1, 0, \dots, 0\right), \left(-g_{(k+2)1}, -g_{(k+2)2}, \dots, -g_{(k+2)k}, 0, 1, \dots, 0\right), \dots \\ \dots, \left(-g_{(k+1)3}, -g_{(k+1)3}, \dots, -g_{(k+1)3}, 0, 0, 1, \dots, 0\right), \left(-g_{n1}, -g_{n2}, \dots, -g_{nk}, 0, 0, 0, \dots, 1\right) \end{matrix} \right\}$$

é uma base de $C^{\perp}$, portanto $H = \begin{bmatrix} -g_{(k+1)1} & -g_{(k+1)2} & \cdots & -g_{(k+1)k} & 1 & 0 & \cdots & 0 \\ -g_{(k+2)1} & -g_{(k+2)2} & \cdots & -g_{(k+2)k} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -g_{n1} & -g_{n2} & \cdots & -g_{nk} & 0 & 0 & \cdots & 1 \end{bmatrix}$ é uma matriz geradora de $C^{\perp}$ na forma $H = [-A^t|I_{n-k}]$.

Teorema 6.13: Considerando C um código linear de dimensão k, contido em K^n, cuja matriz geradora seja G. Uma matriz H de ordem $(n-k) \times n$, com elementos pertencentes a K, cujas linhas sejam linearmente independentes é geradora do código $C^{\perp}$ se, e somente se, $G \cdot H^t = 0$.

Demonstração: Como as linhas de H são linearmente independentes, então formam uma base de um subespaço vetorial de K^n, cuja dimensão é $n-k$, mas $\dim C^{\perp} = n-k$. O produto

$G \cdot H^t$ consiste no produto interno de dos vetores linhas de G pelos vetores colunas de H^t, mas os vetores colunas de H^t são os vetores linhas de H e, caso se tenha $G \cdot H^t = 0$, então os vetores linhas de G e os vetores linhas de H são, entre si, ortogonais, logo, todos os vetores do subespaço gerado por H estão em $C^\perp$ e, portanto, H é matriz geradora de $C^\perp$.

Teorema 6.14: Seja C um código linear contido em um espaço K^n, temos $(C^\perp)^\perp = C$.

Demonstração: Consideremos as matrizes G e H geradoras dos códigos C e $C^\perp$, respectivamente. Pelo teorema 6.13, $G \cdot H^t = 0$. Mas se $G \cdot H^t = 0$, então $(G \cdot H^t)^t = 0$ e, pela propriedade IV, apresentada em 2.3, temos que $(G \cdot H^t)^t = (H^t)^t \cdot G^t = 0$ e pela propriedade I em 2.3, temos $(H^t)^t = H$. Assim, $H \cdot G^t = 0$, o que implica que G é matriz geradora de $(C^\perp)^\perp$, mas por hipótese, G é matriz geradora de C, portanto, $(C^\perp)^\perp = C$.

Teorema 6.15: Considerando C um código linear e H a matriz geradora do código $C^\perp$, um vetor v pertence ao código C se, e somente se, $H \cdot v^t = 0$.

Demonstração: Pelo teorema 13, temos que $(C^\perp)^\perp = C$, portanto, $v \in C$ se, e somente se, $v \in (C^\perp)^\perp$. Vimos anteriormente que o produto de uma matriz geradora de um código pela matriz transposta cuja coluna é vetor pertencente ao complemento ortogonal desse código é igual ao vetor nulo, sendo assim, $v \in (C^\perp)^\perp$ se, e somente se, $H \cdot v^t = 0$.

O teorema 6.15 constitui uma ferramenta eficiente pra determinar se um dado vetor $v \in K^n$ pertence a um dado código linear $C \subset K^n$, bastando para isso, verificar se $H \cdot v^t = 0$.

À matriz H, geradora de $C^\perp$, denominamos *matriz teste de paridade* do código C e ao vetor $H \cdot v^t$, com $v \in K^n$, denominamos *síndrome* do vetor v.

Exemplo:

Considere $C \subset F^6$ um código linear sobre $F = \{0,1\}$, cuja matriz geradora é $G = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$. Dados dois vetores $u, v \in F^6$, tal que $u = (111101)$ e $v = (010101)$, desejamos verificar se u e v são vetores de C.

Observemos que a matriz G não se apresenta na forma padrão, porém, por meio de operações elementares sobre as linhas de G é possível obter uma matriz G' que se apresente na forma padrão:

$$\begin{bmatrix}1&0&1&0&1&1\\1&1&0&0&0&1\\0&1&0&1&1&0\end{bmatrix}\xrightarrow{L_2 \to L_2 + L_3}\begin{bmatrix}1&0&1&0&1&1\\1&0&0&1&1&1\\0&1&0&1&1&0\end{bmatrix}\xrightarrow{L_1 \to L_1 + L_2}$$

$$\begin{bmatrix}0&0&1&1&0&0\\1&0&0&1&1&1\\0&1&0&1&1&0\end{bmatrix}\xrightarrow{L_1 \leftrightarrow L_2}\begin{bmatrix}1&0&0&1&1&1\\0&0&1&1&0&0\\0&1&0&1&1&0\end{bmatrix}\xrightarrow{L_1 \leftrightarrow L_2}\begin{bmatrix}1&0&0&1&1&1\\0&1&0&1&1&0\\0&0&1&1&0&0\end{bmatrix} = G'.$$

Assim, $G' = \begin{bmatrix}1&0&0&1&1&1\\0&1&0&1&1&0\\0&0&1&1&0&0\end{bmatrix}$ é uma matriz geradora do código C, que se apresenta na forma padrão. Notemos que $G' = [I_3|A]$, o que implica que $A = \begin{bmatrix}1&1&1\\1&1&0\\1&0&0\end{bmatrix}$ e $-A^t = \begin{bmatrix}1&1&1\\1&1&0\\0&0&1\end{bmatrix}$. Pelo teorema 6.12, a matriz teste de paridade H, do código C é $H = [-A^t|I_3]$, portanto, $H = \begin{bmatrix}1&1&1&1&0&0\\1&1&0&0&1&0\\1&0&0&0&0&1\end{bmatrix}$.

A síndrome de u é $H \cdot u^t = \begin{bmatrix}1&1&1&1&0&0\\1&1&0&0&1&0\\1&0&0&0&0&1\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\\1\\0\\1\end{bmatrix} = \begin{bmatrix}1+1+1+1+0+0\\1+1+0+0+0+0\\1+0+0+0+0+1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$, 0

que implica que $u \in C$.

A síndrome de v é $H \cdot v^t = \begin{bmatrix}1&1&1&1&0&0\\1&1&0&0&1&0\\1&0&0&0&0&1\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\\1\\0\\1\end{bmatrix} = \begin{bmatrix}0+1+0+1+0+0\\0+1+0+0+0+0\\0+0+0+0+0+1\end{bmatrix} = \begin{bmatrix}0\\1\\1\end{bmatrix}$, o

que implica que $v \notin C$.

Teorema 6.16: Consideremos H uma matriz teste de paridade de um código linear C sobre um corpo K. Então o peso $\omega(C)$ do código C é maior ou igual p se, e somente se, quaisquer $p-1$ colunas da matriz H são linearmente independentes. Valendo a igualdade se, e somente se, quaisquer $p-1$ colunas de H forem linearmente independentes e existirem p colunas de H linearmente dependentes.

Demonstração: Dividiremos a demonstração em duas partes, sendo que a segunda será encarregada de demonstrar a igualdade:

1ª parte: $(\Leftarrow)$ Suponhamos que cada $(p-1)$-uplas de colunas da matriz H sejam linearmente independentes e que $\omega(v) \leq p-1$. Seja $v = v_1 v_2 \ldots v_n$ uma palavra não nula de

C. Sabemos que $H.v^t = 0$, o que implica que $H.v^t = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \cdots & h_{(n-k)n} \end{bmatrix}$.

$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, o que gera o sistema $\begin{cases} h_{11} \cdot v_1 + h_{12} \cdot v_2 + \cdots + h_{1n} \cdot v_n = 0 \\ h_{21} \cdot v_1 + h_{22} \cdot v_2 + \cdots + h_{2n} \cdot v_n = 0 \\ \vdots \\ h_{(n-k)1} \cdot v_1 + h_{(n-k)2} \cdot v_2 + \cdots + h_{(n-k)1n} \cdot v_n = 0 \end{cases}$.

Somando as equações, e reagrupando, temos:

$(h_{11} + h_{21} + \cdots + h_{(n-k)1}) \cdot v_1 + \cdots + (h_{1n} + h_{2n} + \cdots + h_{(n-k)n}) \cdot v_n = 0$. Como $\omega(v)$ representa o número de coordenadas não nulas de v, teríamos então uma combinação linear nula com no máximo $p-1$ colunas da matriz H, contradizendo a hipótese inicial de que $\omega(v) \leq p-1$. Assim, $\omega(v) > p-1$, o que implica que $\omega(v) \geq p$ e, portanto, $\omega(C) \geq p$.

($\Rightarrow$) Em contrapartida, se considerarmos $\omega(C) \geq p$ e suponhamos que existam $p-1$ colunas linearmente dependentes na matriz H, então existem, por exemplo, $v_1 v_2 \ldots v_{p-1} \in K$, nem todos nulos, tal que $(h_{11} + h_{21} + \cdots + h_{(n-k)1}) \cdot v_1 + (h_{12} + h_{22} + \cdots + h_{(n-k)2}) \cdot v_2 + \cdots$ $\ldots + (h_{1(p-1)} + h_{2(p-1)} + \cdots + h_{(n-k)(p-1)}) \cdot v_{p-1} = 0$, o que implica que $v = v_1 v_2 \ldots 0 \ldots 0 \ldots v_{p-1} \ldots 0$ pertence ao código C, implicando com isso, que $\omega(v) \leq p - 1 < p$, e, portanto, $\omega(C) < p$, contradizendo a hipótese. Logo, H possui $p-1$ colunas linearmente independentes.

2ª parte: Para demonstrar a igualdade, suponhamos $\omega(C) = p$, temos que todo conjunto de $p-1$ colunas de H é linearmente independente. Se existissem p colunas linearmente independentes em H, então, pelo que foi visto anteriormente, teríamos $\omega(C) \geq p+1$, logo, em H existem p colunas linearmente dependentes. Por outro lado, se na matriz H existem $p-1$ colunas linearmente independentes e p colunas linearmente dependentes então temos $\omega(C) \geq p$, Mas se $\omega(C) > p$, por exemplo $\omega(C) \geq p+1$, pelo visto anteriormente, teríamos em H que todo conjunto com p colunas seria linearmente independente, contradizendo a hipótese, logo $\omega(C) = p$.

Teorema 6.17: Os parâmetros (n, d, k) de um código linear C satisfazem a desigualdade $d \leq n - k + 1$.

Demonstração: Seja C um código linear sobre um corpo K, tal que $C \subset k^k$. Uma matriz teste de paridade H do código linear C tem ordem $(n-k) \times n$ ou seja, possui $n-k$ linhas linearmente independentes, o que implica que H tem colunas em K^{n-k}. Pelo teorema 6.16, quaisquer $d-1$ colunas de H são linearmente independentes e como K^{n-k} possui no máximo

$n-k$ vetores linearmente independentes, então $d-1 \leq n-k$, o que implica que $d \leq n-k+1$.

À desigualdade acima denominamos *cota de Singleton*.

Um código C no qual valha a igualdade $d = n-k+1$ é denominado de *MDS*, que representa as iniciais das palavras *Maximum Distance Separable*.

6.3.4 Decodificação

O processo de decodificação consiste em ao ser recebida uma palavra através do canal de comunicação, o decodificador de canal se incumbe da detecção e correção da palavra recebida se, por acaso, por alguma interferência, tenha sofrido algum erro, para depois enviá-la ao decodificador de fonte e por fim chegar ao usuário. Para que o processo de decodificação seja eficiente, deve possuir um custo computacional baixo tornando viável sua utilização.

A seguir apresentaremos o processo de decodificação.

Consideremos o vetor c como sendo uma palavra transmitida e o vetor r a palavra recebida com erro. Definimos o vetor erro e como a diferença entre a palavra recebida e a palavra transmitida:

$$e = r - c$$

Quando $e = 0$ significa que a palavra recebida é igual a palavra transmitida e, neste caso, não houve erro na transmissão. Caso $e \neq 0$, entendemos que houve erro na transmissão. Notemos, ainda, que o peso do vetor e define o número de erros ocorridos na transmissão, ou seja, $\omega(e) = p$ implica em p erros na palavra recebida.

Vejamos um exemplo:

Suponha que de um código C sobre o corpo galoisiano F={0,1}, seja transmitida uma a palavra (0101100) e por alguma interferência no canal de transmissão, a palavra recebida seja (1001010). Temos então, $c = 0101100$ e $r = 1001010$, logo, $e = 1001010 - 0101100$, ou seja, $e = 1100110$. Como $\omega(e) = \omega(1100110) = 4$, vemos que ocorreram 4 erros na transmissão.

Considerando H a matriz teste de paridade de um código C, considerando c um vetor (palavra) de C, sabemos que a síndrome de c é nula, ou seja, $H \cdot c^t = 0$. Portanto, a síndrome do vetor erro e é dada por:

$$H \cdot e^t = H \cdot (r-c)^t = H.(r^t - c^t) = H \cdot r^t - H \cdot c^t = H \cdot r^t - 0 = H \cdot r^t$$

Portanto, a síndrome do erro é igual a síndrome da palavra recebida. De uma outra forma, considerando $e = (\alpha_1, \alpha_2, \dots, \alpha_n)$, temos:

$$H \cdot r^t = H.e^t = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \cdots & h_{(n-k)n} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} =$$

$$= \begin{bmatrix} h_{11} \cdot \alpha_1 + h_{12} \cdot \alpha_2 + \cdots + h_{1n} \cdot \alpha_n \\ h_{21} \cdot \alpha_1 + h_{22} \cdot \alpha_2 + \cdots + h_{2n} \cdot \alpha_n \\ \vdots \\ h_{(n-k)1} \cdot \alpha_1 + h_{(n-k)2} \cdot \alpha_2 + \cdots + h_{(n-k)1n} \cdot \alpha_n \end{bmatrix} =$$

$$= \alpha_1 \cdot \begin{bmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{(n-k)1} \end{bmatrix} + \alpha_2 \cdot \begin{bmatrix} h_{12} \\ h_{22} \\ \vdots \\ h_{(n-k)2} \end{bmatrix} + \cdots + \alpha_n \cdot \begin{bmatrix} h_{1n} \\ h_{2n} \\ \vdots \\ h_{(n-k)1n} \end{bmatrix} =$$

$= \alpha_1 \cdot h^1 + \alpha_2 \cdot h^2 + \cdots + \alpha_n \cdot h^n = \sum_{i=1}^{n} \alpha_i \cdot h^i$, onde h^i representa a $i-$ésima coluna da matriz H.

Teorema 6.18: Considerando C um código linear contido em K^n, capaz de corrigir até κ erros. Se uma palavra recebida r pertence ao espaço K^n e a palavra transmitida c pertence ao código C são tais que $d(c,r) \leq \kappa$, então existe um único vetor e tal que $\omega(e) \leq \kappa$, cuja síndrome é igual a síndrome de r, ou seja, $H \cdot e^t = H \cdot r^t$, tal que $c = r - e$.

Demonstração: Para provar a existência, vejamos que pelo enunciado do teorema, temos $d(c,r) \leq \kappa$ e, por tratar-se de uma métrica, sabemos que $d(c,r) = d(r,c)$ e pelo teorema 6.9, $d(r,c) = d(r-c) = \omega(r-c)$, logo, $\omega(r-c) \leq \kappa$ implica que $\omega(e) \leq \kappa$, mostrando a existência de e.

Para provar a unicidade, suponhamos H seja a matriz teste de paridade de um código C em K^n e que existam $e = (\alpha_1, \alpha_2, \dots, \alpha_n)$ e $e' = (\beta_1, \beta_2, \dots, \beta_n)$ tais que $\omega(e) \leq \kappa$, $\omega(e') \leq \kappa$ e $H \cdot e^t = H \cdot e'^t = H \cdot r$, com r sendo uma palavra recebida. Temos então:

$H \cdot e^t = H \cdot e'^t \Rightarrow \alpha_1 \cdot h^1 + \alpha_2 \cdot h^2 + \cdots + \alpha_n \cdot h^n = \beta_1 \cdot h^1 + \beta_2 \cdot h^2 + \cdots + \beta_n \cdot h^n$, onde h^i representa a $i-$ésima coluna de H. Daí, temos:

$(\alpha_1 - \beta_1) \cdot h^1 + (\alpha_2 - \beta_2) \cdot h^2 + \cdots + (\alpha_n - \beta_n) \cdot h^n = 0$ e, pelo teorema 6.16, quaisquer $d-1$ colunas de H são linearmente independentes, portanto, temos $\alpha_i = \beta_i \ \forall i$, logo, $e = e'$.

Para a determinação do vetor e, quando $\omega(e) \leq 1$, ou seja, quando ocorreu no máximo um erro entre a palavra transmitida c e a palavra recebida r, considerando um código C com $d \geq 3$, temos:

I) Se $H.e^t = 0$, então $\omega(e) = 0$, o que implica que $r \in C$ e não ocorreu erro, portanto, tomamos $c = r$.

II) Se $H.e^t \neq 0$, então $\omega(e) = 1$ e temos um coordenada não nula no vetor e, por exemplo a $i-$ésima, ou seja, $e = (0, 0, \dots, \alpha_i, \dots, 0)$. Como $H \cdot e^t = H \cdot r^t = \sum_{i=1}^{n} \alpha_i \cdot h^i$ e, no caso, e possui coordenadas nulas, com exceção da $i-$ésima, então $H \cdot e^t = H \cdot r^t = \alpha_i \cdot h^i$, onde h^i é a $i-$ésima coluna da matriz H.

Exemplo:

Suponhamos o código do braço mecânico visto anteriormente e que uma palavra recebida pelo circuito do braço seja $r = (10101)$.

O código C do braço mecânico está contido em F^5 e o código de canal está contido em F^2, então, tomando quaisquer dois vetores linearmente independentes de C, por exemplo (10110) e (01011), constituímos uma base de C e por conseqüência uma matriz geradora do código C:

$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$. Note que G apresenta-se na forma padrão, ou seja, $G = [I_2|A]$, com $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Como devemos ter $H = [-A^t|I_3]$, então $H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$ é a matriz teste de paridade do código C.

Calculando a síndrome de r, temos:

$H \cdot r^t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+0+1+0+0 \\ 1+0+0+0+0 \\ 0+0+0+0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. Vemos que $H \cdot r^t = 1 \cdot h^2$.

Como $H \cdot e^t = H \cdot r^t$, então $H \cdot e^t = 1 \cdot h^2$, o que implica que $e = (01000)$ e, por consequência, $c = r - e = (10101) - (01000) = (11101)$.

6.3.4.1 Classe lateral

Consideremos um código corretor de erros C contido em K^n, com matriz teste de paridade H, com distância mínima d e capacidade de correção $\kappa = \left[\frac{d-1}{2}\right]$. Como vimos, $H \cdot e^t = H \cdot r^t$ e se $\omega(e) \leq \kappa$, então e é univocamente determinado por r.

Considerando um vetor v do espaço K^n, definimos o conjunto $v + C$, denominado *classe lateral de v segundo* C, da seguinte forma:

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

Teorema 6.19: Dados dois vetores $u, v \in K^n$, $H \cdot u^t = H \cdot v^t$ se, e somente se, $u \in v + C$.

Demonstração: $H \cdot u^t = H \cdot v^t \Leftrightarrow H \cdot u^t - H \cdot v^t = 0 \Leftrightarrow H \cdot (u^t - v^t) = 0 \Leftrightarrow$
$\Leftrightarrow H \cdot (u - v)^t = 0 \Leftrightarrow u - v \in C \Leftrightarrow u \in v + C$.

O conjunto $v + C$ goza das seguintes propriedades:

I) $v + C = v^{'} + c \Leftrightarrow v - v' \in C$

II) $(v + C) \cap (v^{'} + C) \neq \emptyset \Longrightarrow v + C = v^{'} + c$

III) $\bigcup_{v \in K^n} (v + C) = K^n$

IV) $|(v + C)| = |C| = q^k$

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

Demonstrações:

I) ($\Rightarrow$) Se $v + C = v^{'} + C$, então existem $c_1, c_2 \in C$ tais que $v + c_1 = v^{'} + c_2$, o que implica que $v - v^{'} = c_2 - c_1$, mas C é um subespaço vetorial de K^n, portanto, $c_2 - c_1 \in C$, o que implica que $v - v^{'} \in C$.

($\Leftarrow$) Suponhamos que $v - v^{'} \in C$, isso implica que $c_1, c_2 \in C$ tais que $c_1 + (v - v^{'}) = c_2$, ou seja, $v + c_1 = v^{'} + c_2$. Notemos que $v + c_1 \in v + C$ e $v^{'} + c_2 \in v^{'} + C$, portanto, $v + C = v^{'} + C$.

II) Se $(v + C) \cap (v^{'} + C)$, então existe $u \in (v + C) \cap (v^{'} + C)$, o que implica que $u \in (v + C)$ e $u \in (v' + C)$. Então, existem $c_1, c_2 \in C$ tais que $u = v + c_1$ e $u = v' + c_2$ e por conseqüência $v + c_1 = v' + c_2$. Da igualdade anterior temos $v - v^{'} = c_2 - c_1$. Como $c_2 - c_1 \in C$, então $v - v^{'} \in C$ e, pela propriedade I, temos que $v + C = v^{'} + c$.

III) K^n é um espaço vetorial sobre o corpo K, logo, $0 \in K^n$. Como C é um subespaço vetorial de K^n, então $0 \in C$. Para todo $v \in K^n$, v pode ser escrito como $v + 0$, o que implica que v pertence a uma classe lateral $v + C$, portanto, temos $\bigcup_{v \in K^n} (v + C) = K^n$.

IV) Sabemos que $|C| = q^k = M$, ou seja, $C = \{c_1, c_2, c_3, \dots, c_M\}$. Seja v um vetor do conjunto $v + C$. Por definição, $v + C = \{v + c, c \in C\}$, assim, $v + C = \{v +$

$c_1, v + c_2, v + c_3, \dots, v + c_M\}$, ou seja, $v + C = \{c'_1, c'_2, c'_3, \dots, c'_M\}$, ou seja, $|v + C| = M = q^k = |C|$.

V) Notemos que $C = \{c_1, c_2, c_3, \dots, c_m\} = \{0 + c_1, 0 + c_2, 0 + c_3, \dots, 0 + c_m\} = 0 + C$. Se $v + C = C$, então $v + C = 0 + C$ e pela propriedade I, temos que $v + C = 0 + C \Leftrightarrow v - 0 \in C \Leftrightarrow v \in C$.

Pela propriedade II, temos que classes laterais diferentes segundo C são disjuntas. Sabemos que $|K| = q$, o que implica que $|K^n| = q^n$. Pela propriedade III, $\bigcup_{v \in K^n}(v + C) = K^n$, o que implica que $|\bigcup_{v \in K^n}(v + C)| = |K^n| = q^n$ e pela propriedade IV, temos que $|(v + C)| = |C| = q^k$. Assim, o número de classes laterais segundo C é dado por $\frac{|\bigcup_{v \in K^n}(v+C)|}{|(v+C)|} = \frac{q^n}{q^k} = q^{n-k}$.

Exemplo:

Considerando o código linear C utilizado no exemplo do braço mecânico, vimos que a matriz geradora de C é $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$, o que implica que $C = \{00000, 10110, 01011, 11101\}$. Dados $v_1, v_2, v_3 \in F^5$, tais que $v_1 = (00000)$, $v_2 = (01000)$ e $v_3 = (01110)$.As classes laterais de v_1, v_2, v_3 segundo C são:

$$00000 + C = \{00000, 10110, 01011, 11101\}$$
$$01000 + C = \{01000, 11110, 00011, 10101\}$$
$$01110 + C = \{01110, 11000, 00101, 10011\}$$

O teorema 6.19 garante uma correspondência biunívoca entre classes laterais e síndromes, de modo que todos os vetores de uma classe lateral possuam síndromes iguais e vetores de classes laterais diferentes possuem síndromes diferentes.

Seja x um vetor pertencente a uma classe lateral de v segundo C. Se $\omega(x) = min\{\omega(v_i);\ v_i \in v + C\}$, então dizemos que x é o *líder* de $v + C$.

No exemplo anterior, temos que 00000 é o líder de $00000 + C$, 01000 é o líder de $01000 + C$ e 11000 e 00101 são os líderes de $01110 + C$. Notemos que o líder de uma classe não necessariamente é único.

Teorema 6.20: Considerando $C \subset K^n$ um código com distância mínima d. Se $v \in K^n$ é um vetor tal que $\omega(v) \leq \left[\frac{d-1}{2}\right] = \kappa$, então v é o único elemento líder em sua classe lateral.

Demonstração: Sejam $v_1, v_2 \in K^n$ tais que $\omega(v_1) \leq \left[\frac{d-1}{2}\right]$ e $\omega(v_2) \leq \left[\frac{d-1}{2}\right]$. Se $v_1 - v_2 \in C$, então $\omega(v_1 - v_2) \leq \omega(v_1) + \omega(v_2) \leq \left[\frac{d-1}{2}\right] + \left[\frac{d-1}{2}\right] \leq d - 1$, portanto, $v_1 - v_2 = 0$, o que implica que $v_1 = v_2$.

O teorema 6.19 constitui uma ferramenta importante para a determinação dos líderes de classes de peso menor ou igual a $\left[\frac{d-1}{2}\right]$. Para isso, basta tomar os vetores $v_i \in K^n$, para os quais se tenha $\omega(v_i) \leq \left[\frac{d-1}{2}\right]$. Cada um dos v_i é líder de uma e somente uma classe.

Exemplo:

Vimos que a matriz teste de paridade do código $C \subset F^5$, do braço mecânico é $H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$. Sabemos que nesse código, a distância mínima $d = 3$, pois vemos facilmente que quaisquer duas colunas de H são linearmente independentes enquanto que três colunas de H são linearmente dependentes (teorema 6.16), o que implica que $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{3-1}{2}\right] = 1$, ou seja, C tem capacidade de correção de 1 erro. Os vetores $v_i \in F^5$ tais que $\omega(v_i) \leq 1$ são $00000, 00001, 00010, 00100, 01000, 10000$. Os líderes v_i e suas respectivas síndromes $H \cdot v_i^t$ são apresentados na tabela a seguir:

Líder	Síndrome
00000	000
00001	001
00010	010
00100	100
01000	011
10000	110

Suponhamos que duas palavras (comandos) c_1 e c_2 sejam transmitidas ao braço mecânico e, devido a algum ruído, as palavras (comandos) recebidas sejam $r_1 = (11110)$ e $r_2 = (11010)$. Temos que $H \cdot r_1^t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1+0+1+0+0 \\ 1+1+0+1+0 \\ 0+1+0+0+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} =$ $[0 \quad 1 \quad 1]^t$, comparando com a tabela, temos $e = (01000)$. Como $c = r - e$, temos $c_1 = (11110) - (01000) = (10110)$. O comando transmitido foi *para cima*. Por outro lado,

$H \cdot r_2^t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1+0+0+0+0 \\ 1+1+0+1+0 \\ 0+1+0+0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = [1 \quad 1 \quad 1]^t$. Vemos que a síndrome (111) não é encontrada na tabela, portanto em r_2 ocorreu mais de um erro e o código C não é capaz de corrigir.

Outro exemplo:

Suponha que desejemos transmitir a mensagem **PROFMAT BRASIL** através de um código linear sobre o corpo galoisiano $F = \{0,1\}$.

Abaixo mostraremos uma lista de procedimentos necessários até a obtenção do código de canal necessário à transmissão da mensagem:

1) Fonte: (espaço), A, B, C, D, E, F, G, H, I, J, L, M, N, O, P, Q, R, S, T, U, V, X e Z, com 24 caracteres.
2) Código da fonte: notemos que o código da fonte deve possuir no mínimo 24 palavras código, portanto, adotaremos $k = 5$, o que implica que o código de fonte está contido em F^5. Utilizemos as seguintes informações:

$espaço = 00000$	$E = 00001$	$J = 01100$	$P = 00011$	$U = 11001$
$A = 10000$	$F = 11000$	$L = 01010$	$Q = 11100$	$V = 01110$
$B = 01000$	$G = 10100$	$M = 01001$	$R = 10110$	$X = 00111$
$C = 00100$	$H = 10010$	$N = 00110$	$S = 10101$	$Z = 11110$
$D = 00010$	$I = 10001$	$O = 00101$	$T = 11010$	

Observemos que ao transmitir a fonte P, utilizando o código de fonte 00011, se ocorrer um erro, por exemplo, na quinta coordenada, o código recebido será 00010, que equivale a fonte D e por consequência o erro não seria detectado.

3) Código de canal: por meio do acréscimo de redundâncias, o código de fonte é convertido em código de canal. Suponhamos que o código de canal tenha comprimento $n = 9$. Temos então que o código C é um subespaço vetorial do espaço F^9. Ou seja, C é obtido através de uma transformação linear $T: F^5 \to F^9$ e, pelo que foi visto, $\dim C = k = 5$. Tomemos quaisquer cinco vetores linearmente independentes de F^9 para obtermos uma base de C e, consequentemente, uma matriz G geradora do código C:

$\{(110010000), (100100010), (001001001), (000010110), (000101010)\}$ é uma base de C, pois os cinco vetores desse conjunto são linearmente independentes, logo, $G = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$ é uma matriz geradora do código C. Por praticidade, determinaremos $G' = [I_5|A]$ equivalente por linhas à matriz G, apresentada na forma padrão:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} \overrightarrow{L_1 \rightarrow L_1 + L_4} \\ \overrightarrow{L_2 \rightarrow L_2 + L_5} \end{matrix} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \overrightarrow{L_1 \rightarrow L_1 + L_2}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} \overrightarrow{L_1 \leftrightarrow L_2} \\ \overrightarrow{L_4 \leftrightarrow L_5} \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} = G'.$$

Assim temos:

Fonte	Código da fonte	(Código daFonte).G	Código do canal
espaço	00000	00000.G	000000000
A	10000	10000.G	100001000
B	01000	01000.G	010001110
C	00100	00100.G	001001001
D	00010	00010.G	000010110
E	00001	00001.G	000101010
F	11000	11000.G	110000110
G	10100	10100.G	101000001
H	10010	10010.G	100100010
I	10001	10001.G	100011110
J	01100	01100.G	011000111
L	01010	01010.G	010100100
M	01001	01001.G	010011000
N	00110	00110.G	010100100
O	00101	00101.G	001011111
P	00011	00011.G	000111100
Q	11100	11100.G	111001111
R	10110	10110.G	101101011
S	10101	10101.G	101010111
T	11010	11010.G	110101100
U	11001	11001.G	110010000
V	01110	01110.G	011101101
X	00111	00111.G	001110101
Z	11110	11110.G	111100101

Portanto, as palavras do código a serem transmitidas, na ordem em que aparecem, são:

000111100 101101011 001011111 110000110 010011000 100001000 110101100
000000000 010001110 101101011 100001000 101010111 100011110 010100100

Suponhamos que ao utilizar o código acima, a seguinte mensagem seja recebida:

010011000 000101010 101010110 110101100 101101111 100001000 000010110
101011111 000000010 000111100 101101011 001011110 110000110 100011110
101010101 101010111 100011110 101011111 010100100 100001000 010100110

a qual desejamos decodificar. Suponhamos ainda que no máximo um erro tenha sido introduzido em cada palavra transmitida.

Da matriz geradora $G' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$ do código linear C, apresentada na forma padrão, obtemos uma matriz teste de paridade $H = [-A^t | I_{n-k}]$. Como $n = 9$ e $k = 5$, então $H = [-A^t | I_4]$. Notemos que $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$, então, $-A^t = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$, o que implica que $H = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

Vemos que quaisquer duas colunas de H são linearmente independentes enquanto que três colunas de H são linearmente dependentes. Pelo teorema 6.16 temos que $\omega(C) = 3$, o que implica que $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{3-1}{2}\right] = 1$, ou seja, C tem capacidade de detecção de 2 erros e correção de 1 erro.

Os vetores $v_i \in F^9$, para os quais se tenha $\omega(v_i) \leq 1$, classificados como líderes de classe são: 000000000, 00000001, 000000010, 000000100, 000001000, 000010000, 00010000, 001000000, 010000000 e 100000000. Resolvendo os produtos $H \cdot v_i^t$, determinamos as síndromes dos líderes de classe, conforme a seguinte tabela:

Líder	Síndrome
000000000	0000
000000001	0001
000000010	0010
000000100	0100
000001000	1000

000010000	0110
000100000	1010
001000000	1001
010000000	1110
100000000	1000

Das vinte e uma palavras recebidas r_i, calcularemos suas respectivas síndromes $H \cdot r_i^t$ e os erros e_i, comparando com a tabela anterior e determinando as palavras transmitidas c_i, identificando suas respectivas fontes. O procedimento descrito acima é apresentado na tabela a seguir:

Palavra recebida (r_i)	**Síndrome ($H \cdot r_i^t$)**	**Erro (e_i) (líder)**	**Observação**	**Palavra transmitida ($c_i = r_i - e_i$)**	**Fonte**
010011000	0000	000000000	Não houve erro	010011000	M
000101010	0000	000000000	Não houve erro	000101010	E
101010110	0001	000000001	Houve um erro	101010111	S
110101100	0000	000000000	Não houve erro	110101100	T
101101111	0100	000000100	Houve um erro	101101011	R
100001000	0000	000000000	Não houve erro	100001000	A
000010110	0000	000000000	Não houve erro	000010110	D
101011111	1000	100000000	Houve um erro	001011111	O
000000010	0010	000000010	Houve um erro	000000000	espaço
000111100	0000	000000000	Não houve erro	000111100	P
101101011	0000	000000000	Não houve erro	101101011	R
001011110	0001	000000001	Houve um erro	001011111	O
110000110	0000	000000000	Não houve erro	110000110	F
100011110	0000	000000000	Não houve erro	100011110	I
101010101	0010	000000010	Houve um erro	101010111	S
101010111	0000	000000000	Não houve erro	101010111	S
100011110	0000	000000000	Não houve erro	100011110	I
101011111	1000	100000000	Houve um erro	001011111	O
010100100	0000	000000000	Não houve erro	010100100	N
100001000	0000	000000000	Não houve erro	100001000	A
010100110	0010	000000010	Houve um erro	010100100	L

Portanto, a mensagem transmitida foi **MESTRADO PROFISSIONAL**.

6.3.5 Alguns exemplos de códigos lineares

6.3.5.1 Código de repetição

As características fundamentais de um código $R(n) \subset F^n$ de repetição sobre o corpo galoisiano $F = \{0,1\}$ são sua dimensão $k = 1$, o número de palavras do código é $M = 2^k = 2^1 = 2$ e sua distância mínima $d = n$.

O código de repetição descrito acima detecta até $n-1$ erros e sua capacidade de correção é $\kappa = \left[\frac{n-1}{2}\right]$ erros, o que leva-nos a deduzir que se n for ímpar, então o código $R(n)$ corrige até $\frac{n-1}{2}$ erros enquanto que se n for par, então o código $R(n)$ corrige até $\frac{n-2}{2}$ erros.

O processo de decodificação consiste na contagem do número de "zeros" e do número de "uns" na palavra recebida, sendo que se houver um número maior de "uns", ou seja, esse número estiver entre $\frac{n+1}{2}$ e n, para n ímpar ou estiver entre $\frac{n+2}{2}$ e n para n par, então a palavra é corrigida para $(\underbrace{111\ldots1}_{n\ dígitos})$, caso contrário, a palavra é corrigida para $(\underbrace{000\ldots0}_{n\ dígitos})$.

Exemplo:

Um circuito digital comandado por controle remoto entra em funcionamento quando o comando acionado é "on", e deixa de funcionar quando o comando acionado é "off".

Temos então a fonte como sendo os comandos "on" e "off" e podemos codificar esses comandos de modo a se obter o código de fonte 1 e 0 respectivamente.

Já vimos anteriormente que a transmissão direta do código de fonte não é viável pois caso ocorra um erro não é possível sua detecção e sua correção. Em virtude disso, utilizemos, por exemplo, o código de repetição 6 representado por $R(6)$, temos então:

Fonte	Código de fonte	Código de canal
Off	0	000000
On	1	111111

Ao ser transmitido o comando "on", suponhamos que a transmissão sofra um erro e a palavra código de F^6 recebida seja 101101. O decodificador de fonte detectará o erro, pois 101101 não pertence a $R(6)$. Como o $n = 6$ é par e o número de dígitos 1 está entre $\frac{n+2}{2} = \frac{6+2}{2} = 4$ e $n = 6$, então a palavra código 101101 é corrigida para 111111.

Notemos porém que um código de repetição não é viável, seja pela demora na transmissão de palavras código com grande número de dígitos (bits), ou ainda pela ineficiência na correção, pois no exemplo acima, se a palavra recebida fosse 101100, seria impossível ao decodificador de canal decidir se a palavra correta transmitida era 000000 ou 111111.

6.3.5.2 Código de um dígito de paridade (Código de peso par)

Em alguns casos é mais interessante detectar a ocorrência de um erro do que corrigi-lo propriamente, pois podemos ter um custo elevado na construção de redundâncias que sejam capazes de corrigir esses erros, sendo mais viável a retransmissão da informação do que a correção do erro detectado.

Uma forma de detectar um erro, com um baixo custo computacional, é o acréscimo de um único dígito (bit) no código da fonte, obtendo assim, um código de canal com comprimento maior que o código da fonte, por um dígito.

O dígito acrescido ao código de fonte, de modo a se obter o código de canal é denominado *dígito de verificação de paridade* e é univocamente determinado em cada palavra do código da fonte, com finalidade de obter um número par de dígitos "uns", de modo que se houver um (único) erro de transmissão, o mesmo seja detectado, mas não corrigido.

Em um código $C \subset K^n$ de um dígito de paridade com dimensão k, temos $n - k = 1$. Como o dígito de paridade é acrescido para que seja obtido um número par de "uns", então a quantidade mínina de "uns" em um vetor não nulo de C é igual a 2, o que implica que a distância mínima desse código (ou o seu peso) é $d = 2$. Essa observação permite-nos verificar que a capacidade de detecção do código C é $d - 1 = 2 - 1 = 1$ erro e sua capacidade de correção é $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{2-1}{2}\right] = 0$, ou seja, C não é capaz de corrigir erros como mencionado anteriormente.

Exemplo:

Voltemos ao braço mecânico apresentado anteriormente, do qual temos a seguinte tabela:

Fonte	**Código da fonte**
Para a esquerda:	00
Para a direita:	01
Para cima:	10
Para baixo:	11

Diferente do que foi feito anteriormente, acrescentaremos como redundância apenas um dígito, de modo que a quantidade de “uns” seja par, obtendo com isso o código de canal apresentado a seguir:

Fonte	**Código da fonte**	**Código de canal**
Para a esquerda:	00	000
Para a direita:	01	011
Para cima:	10	101
Para baixo:	11	110

É evidente que se um comando fosse dado ao braço mecânico, por exemplo “para baixo”, cujo código de fonte e de canal são respectivamente 11 e 110 e ocorresse um erro, de modo que o decodificador de canal recebesse a palavra 100, um erro seria imediatamente detectado, pois 100 possui uma quantidade ímpar de “uns”, porém, a correção seria impossível uma vez que qualquer um dos três dígitos de100 poderia estar errado, acarretando três possibilidades para a palavra transmitida: 000, 101 e 110.

6.3.5.3 Código de Hamming

Um código C sobre o corpo galoisiano F, cuja matriz teste de paridade é H_m, de ordem $m \times n$, com colunas em $F^m \backslash \{0\}$, em qualquer ordem, é denominado *código de Hamming*.

Como a matriz H_m possui colunas em $F^m \backslash \{0\}$, então, o seu número de colunas é dado por $2^m - 1$, o que implica que cada palavra de C tem comprimento $n = 2^m - 1$. A dimensão do código C é dada por $k = n - m$, ou seja, $k = (2^m - 1) - m$, o que implica que $k = 2^m - m - 1$. A distância mínima (ou o peso) em um código de Hamming é $\omega(C) = 3$ (teorema 6.16).

As matrizes $H_3 = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$ e $H_4 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$ são matrizes teste de paridade dos códigos de Hamming correspondentes a $m = 3$ e $m = 4$ respectivamente.

Teorema 6.21: Todo código de Hamming é perfeito.

Demonstração: Por definição, um código é perfeito se $\bigcup_{c\in C} D(c,k) = F^n$. Consideremos o código $C \subset F^n$ como sendo um código de Hamming, então C possui distância mínima $d = 3$, o que implica que $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{3-1}{2}\right] = 1$. Seja c um vetor de F^n. Sabemos que o número de discos de centro c e raio r em F^n é dado por $|D(c,r)| = \sum_{i=0}^{r} \binom{n}{i} \cdot (q-1)^i$, portanto, $|D(c,1)| = \binom{n}{0} \cdot (2-1)^0 + \binom{n}{1} \cdot (2-1)^1 = 1 \cdot 1 + n \cdot 1 = 1 + n$. Assim, $|\bigcup_{c\in C} D(c,1)| = (1+n) \cdot 2^k = (1 + 2^m - 1) \cdot 2^{n-m} = 2^n$, ou seja, $|\bigcup_{c\in C} D(c,1)| = F^n$. Portanto, C é um código perfeito.

Teorema 6.22: Um código de Hamming é MDS se, e somente se, $m = 2$.

Demonstração:

($\Rightarrow$) Se um código de Hamming é MDS, então $d = n - k + 1$, mas em todo código de Hamming, $d = 3$, o que implica que $n - k + 1 = 3$, portanto, temos:

$(2^m - 1) - (2^m - m - 1) + 1 = 3$, o que implica que $m = 2$.

($\Leftarrow$) Em um código de Hamming de ordem $m = 2$, temos $n = 2^2 - 1 = 3$ e $k = 2^2 - 2 - 1 = 1$. Como em um código de Hamming temos sempre $d = 3$, então $d = 3 - 1 + 1 = n - k + 1$, portanto para $m = 2$ um código de Hamming é MDS.

6.3.5.4 Código de Reed-Solomon

Consideremos um corpo finito K e um espaço vetorial $K[X]_{k-1}$ de todos os polinômios $p(x)$ em $K[X]$ cujo grau seja menor ou igual a $k-1$, juntamente com o polinômio nulo. Uma base para o espaço vetorial $K[X]_{k-1}$ é o conjunto $B = \{1, X, X^2, X^3, \dots, X^{k-1}\}$. Portanto, $\dim K[X]_{k-1} = k$.

Consideremos $n \in \mathbb{N}$ e $\alpha_i \in K$, com $i \in \{1, 2, \dots, n\}$, tal que $\alpha_i \neq \alpha_j$ sempre que $i \neq j$. Uma função $T: K[X]_{k-1} \to K^n, k < n$, que a cada elemento $p(x) \in K[X]_{k-1}$ associa a $n-$upla $\big(p(\alpha_1), p(\alpha_2), p(\alpha_3), \dots, p(\alpha_n)\big) \in K^n$.

T é uma transformação linear, pois dados $p(x), q(x) \in K[X]_{k-1}$ e $\beta \in K$, temos:

$T\big(p(x) + q(x)\big) = \big(p(\alpha_1) + q(\alpha_1), p(\alpha_2) + q(\alpha_2), \dots, p(\alpha_n) + q(\alpha_n)\big) =$

$= \big(p(\alpha_1), p(\alpha_2), \dots, p(\alpha_n) + q(\alpha_n)\big) + \big(q(\alpha_1), q(\alpha_2), \dots, q(\alpha_n)\big) = T\big(p(x)\big) + T(q(x))$ e

$T\big(\beta \cdot p(x)\big) = \big(\beta \cdot p(\alpha_1), \beta \cdot p(\alpha_2), \dots, \beta \cdot p(\alpha_n)\big) = \beta \cdot \big(p(\alpha_1), p(\alpha_2), p(\alpha_3), \dots, p(\alpha_n)\big) =$

$= \beta \cdot T\big(p(x)\big)$.

Além disso, T é uma transformação linear injetiva, pois $T(p(x)) = 0$ implica que $\big(p(\alpha_1), p(\alpha_2), p(\alpha_3), \dots, p(\alpha_n)\big) = (0, 0, 0, \dots, 0)$, ou seja, $p(\alpha_1) = p(\alpha_2) = p(\alpha_3) = \dots = p(\alpha_n) = 0$, pois um polinômio de grau $k-1$ não pode possuir n raízes distintas. Assim, temos $Ker\ (T) = \{p(x) \in K[X]_{k-1}; p(\alpha_1) = p(\alpha_2) = \dots = p(\alpha_n) = 0\}$, o que implica que $Ker\ (T) = \{o\}$, logo, T é injetiva.

Pelo Teorema 5.13, temos que $\dim K[X]_{k-1} = \dim \ker(\mathrm{T}) + \dim Im(T)$. Como $Ker\ (T) = \{o\}$, então $\dim \ker(\mathrm{T}) = 0$, o que implica que $\dim Im(T) = \dim K[X]_{k-1} = k$, ou seja, $Im(T)$ é um subespaço vetorial de K^n, com dimensão k. Podemos adotar $K[X]_{k-1}$ com sendo o código de fonte, $Im(T) = C$ como sendo o código de canal (um código linear) e a transformação T como sendo a codificação.

Ao código descrito acima denominamos *código de Reed- Solomon*, com comprimento n e dimensão k, definido por $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$.

Notemos que dado $c \in C$, tal que $c \neq 0$, então $\exists p(x) \in K[X]_{k-1}$ tal que $c = \big(p(\alpha_1), p(\alpha_2), p(\alpha_3), \dots, p(\alpha_n)\big)$. Por definição de peso de um código, temos que $\omega(c) = |\{i \in \{1, 2, 3, \dots, n\}; p(\alpha_i) \neq 0\}|$. Como c possui n coordenadas, então $|\{i \in \{1, 2, 3, \dots, n\}; p(\alpha_i) \neq 0\}| = n - |\{i \in \{1, 2, 3, \dots, n\}; p(\alpha_i) = 0\}| \geq n - gr\big(p(x)\big)$.

Portanto, $\omega(c) \geq n - (k-1)$, o que implica que $\omega(c) \geq n - k + 1$, portanto, $d \geq n - k + 1$. Pelo teorema 6.17 vimos que parâmetros (n, d, k) de um código linear C satisfazem a desigualdade $d \leq n - k + 1$. Das duas desigualdades acima temos $d = n - k + 1$.

Como $B = \{1, X, X^2, X^3, \dots, X^{k-1}\}$ é uma base de $K[X]_{k-1}$, temos que $B' = \{T(1), T(X), T(X^2), T(X^3), \dots, T(X^{k-1})\}$ é uma base de C, logo uma matriz geradora do código C pode ser representada por:

$$G = \begin{bmatrix} T(1) \\ T(X) \\ T(X^2) \\ \vdots \\ T(X^{k-1}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{bmatrix}.$$

Exemplo: Consideremos um corpo finito $K = \mathbb{Z}_7$, $k = 3$, $n = 5$ e $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 3$, $\alpha_4 = 4$ e $\alpha_5 = 5$. Pela definição de matriz geradora, temos que

$G = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 2 & 4 \end{bmatrix}$ é uma matriz geradora do código de Reed-Solomon de comprimento 5, dimensão 3 e definida pelos elementos 1, 2, 3, 4, 5 do corpo $\mathbb{Z}_7$, com distância mínima $d = n - k + 1 = 5 - 3 + 1 = 3$

Para determinar uma matriz $G' = [I_k|A]$, na forma padrão, equivalente à matriz G, observemos que, através do polinômio de Lagrange, visto em 4.5, podemos obter os polinômios $p_1(x), p_2(x), p_3(x) \in K[X]_{k-1}$, tais que $p_1(\alpha_1) = p_2(\alpha_2) = p_3(\alpha_3) = 1$ e $p_1(\alpha_2) = p_1(\alpha_3) = p_2(\alpha_1) = p_2(\alpha_3) = p_3(\alpha_1) = p_3(\alpha_2) = 0$, da seguinte forma:

$$p_1(x) = \frac{(x-\alpha_2)\cdot(x-\alpha_3)}{(\alpha_1-\alpha_2)\cdot(\alpha_1-\alpha_3)} = \frac{(x-2)\cdot(x-3)}{(1-2)\cdot(1-3)} = \frac{x^2-5x+6}{2} = 4x^2 + x + 3$$

$$p_2(x) = \frac{(x-\alpha_1)\cdot(x-\alpha_3)}{(\alpha_2-\alpha_1)\cdot(\alpha_2-\alpha_3)} = \frac{(x-1)\cdot(x-3)}{(2-1)\cdot(2-3)} = \frac{x^2-4x+3}{-1} = 6x^2 + 4x + 4$$

$$p_3(x) = \frac{(x-\alpha_1)\cdot(x-\alpha_2)}{(\alpha_3-\alpha_1)\cdot(\alpha_3-\alpha_2)} = \frac{(x-1)\cdot(x-2)}{(3-1)\cdot(3-2)} = \frac{x^2-3x+2}{2} = 4x^2 + 2x + 1$$

Notemos que $p_1(x), p_2(x)$ e $p_3(x)$ são linearmente independentes, pois dados $\beta_1, \beta_2, \beta_3 \in K$, $\beta_1 \cdot p_1(x) + \beta_2 \cdot p_2(x) + \beta_3 \cdot p_3(x) = 0$ implica que $\beta_1 \cdot (4x^2 + x + 3) + \beta_2 \cdot (6x^2 + 4x + 4) + \beta_3 \cdot (4x^2 + 2x + 1) = 0$, o que equivale a ter o sistema de equações lineares $\begin{cases} 4\beta_1 + 6\beta_2 + 4\beta_3 = 0 \\ \beta_1 + 4\beta_2 + 2\beta_3 = 0 \\ 3\beta_1 + 4\beta_2 + \beta_3 = 0 \end{cases}$, que implica que $\beta_1 = \beta_2 = \beta_3 = 0$. Assim, $p_1(x), p_2(x)$ e $p_3(x)$ formam uma base para $K[X]_{k-1}$ e, como T é injetiva, então $T(p_1(x)), T(p_2(x))$ e $T(p_3(x))$ formam uma base do código de Reed-Solomon de comprimento 5, dimensão 3 e definida pelos elementos 1, 2, 3, 4, 5 do corpo $\mathbb{Z}_7$, com distância mínima $d = n - k + 1 = 5 - 3 + 1 = 3$. Logo, $G' = \begin{bmatrix} T(p_1) \\ T(p_2) \\ T(p_3) \end{bmatrix} = \begin{bmatrix} p_1(\alpha_1) & p_1(\alpha_2) & p_1(\alpha_3) & p_1(\alpha_4) & p_1(\alpha_5) \\ p_2(\alpha_1) & p_2(\alpha_2) & p_2(\alpha_3) & p_2(\alpha_4) & p_2(\alpha_5) \\ p_3(\alpha_1) & p_3(\alpha_2) & p_3(\alpha_3) & p_3(\alpha_4) & p_3(\alpha_5) \end{bmatrix}$, portanto, $G' = \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 4 & 6 \\ 0 & 0 & 1 & 3 & 6 \end{bmatrix}$ é a matriz geradora na forma padrão.

6.3.5.5 O código do Mariner 9 (Código de Reed-Muller de 1ª ordem)

A Mariner 9 foi uma sonda espacial lançada ao espaço em 30 de maio de 1971, com objetivo de explorar o planeta Marte. Durante seu período de atividade, a sonda Mariner 9 enviou à terra mais de 7.000 fotos desse planeta.

O código utilizado para a detecção e correção de erros dos dados enviados pela sonda Mariner 9 à terra, pertence a uma família de códigos $R(1,m)$ sobre $F=\{0,1\}$, denominados *Códigos de Reed-Muller de Primeira Ordem.*

A matriz G geradora desse código é obtida através da matriz teste de paridade de um código de Hamming de dimensão $m-n$, ou seja, a matriz H_m.

A matriz G possui ordem $(m+1)\times 2^m$ e, para construí-la, consideremos a matriz $H_m=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1(2^m-1)} \\ h_{21} & h_{22} & \cdots & h_{2(2^m-1)} \\ \vdots & \vdots & \ddots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{m(2^m-1)} \end{bmatrix}$. A matriz G possui a primeira linha com todos os elementos iguais a 1 e a coluna de ordem 2^m (última coluna) possui todos os elementos nulos, com exceção do primeiro. O bloco $\begin{bmatrix} g_{21} & g_{22} & \cdots & g_{2(2^m-1)} \\ g_{31} & g_{32} & \cdots & g_{3(2^m-1)} \\ \vdots & \vdots & \ddots & \vdots \\ g_{(m+1)1} & g_{(m+1)2} & \cdots & g_{(m+1)(2^m-1)} \end{bmatrix}$ é igual a matriz H_m, ou seja, $G=\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ g_{21} & g_{22} & \cdots & g_{2(2^m-1)} & 0 \\ g_{31} & g_{32} & \cdots & g_{3(2^m-1)} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{(m+1)1} & g_{(m+1)2} & \cdots & g_{(m+1)(2^m-1)} & 0 \end{bmatrix}=\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ h_{11} & h_{12} & \cdots & h_{1(2^m-1)} & 0 \\ h_{21} & h_{22} & \cdots & h_{2(2^m-1)} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{m(2^m-1)} & 0 \end{bmatrix}$

Notemos que a primeira linha de G é linearmente independente das demais, basta observar a última coluna dessa matriz. Temos também que todas as outras linhas de G, a partir da segunda, também são linearmente independentes, pois o bloco H_m garante esse fato. Sendo assim, a dimensão do código $R(1,m)$ é $m+1$. O comprimento do código $R(1,m)$ é evidentemente 2^m, pois G possui 2^m colunas. Determinemos agora a distância mínima do código $R(1,m)$.

Primeiramente, observemos que uma matriz cujas colunas sejam cada um dos 2^m vetores de F^m possui linhas com 2^{m-1} elementos iguais a 1 e 2^{m-1} elementos iguais a 0. Como as colunas de uma matriz H_m de um código de Hamming são todos os vetores de $F^m\backslash\{0\}$, então cada linha de H_m possui 2^{m-1} elementos iguais a 1 e $2^{m-1}-1$ dígitos iguais a 0. Na matriz G geradora do código $R(1,m)$, excetuando-se a primeira linha, todas as demais

são formadas por linhas da matriz H_m, acrescidas do dígito zero. Sendo assim, qualquer linha da matriz G, com exceção da primeira, tem a metade dos elementos iguais a 1 e a outra metade igual a zero, ou seja, cada vetor formado pela combinação linear dessas linhas possui 2^{m-1} dígitos iguais a 1 e 2^{m-1} dígitos iguais a zero, acarretando que o peso de cada um desses vetores é igual a 2^{m-1}. Resta-nos avaliar as combinações lineares com o vetor primeira linha da matriz G, que possui peso 2^m, pois possui suas 2^m coordenadas iguais a 1. Notemos que qualquer que seja $v \in R(1,m)$, 2^{m-1} coordenadas de v são iguais a 1. Consideremos um vetor $u = v + (\underbrace{111 \ldots 11}_{2^n \ dígitos})$, o que implica que a soma das 2^{m-1} coordenadas do vetor v com o vetor $(\underbrace{111 \ldots 11}_{2^n \ dígitos})$ resultará em zero, fazendo com que u possua 2^{m-1} dígitos iguais a zero. De maneira análoga, somando as 2^{m-1} coordenadas iguais a zero do vetor v, com as 2^{m-1} coordenadas do vetor $(\underbrace{111 \ldots 11}_{2^n \ dígitos})$, obtermos 2^{m-1} coordenadas iguais a 1, fazendo com que u possua 2^{m-1} coordenadas iguais a 1.

Portanto, temos que no código $R(1,m)$, a distância mínima é $d = 2^{m-1}$.

Diante do exposto, os parâmetros (n,k,d) do código $R(1,m)$ são $(2^m, m+1, 2^{m-1})$.

A capacidade de detecção e de correção de erros num código $R(1,m)$ são, respectivamente, $d-1 = 2^{m-1}-1$ e $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{2^{m-1}-1}{2}\right]$.

O código de detecção e correção de erros, utilizado nas transmissões da sonda espacial Mariner 9, corresponde a $R(1{,}5)$. Seus parâmetros (n,k,d) são $(2^5, 5+1, 2^{5-1}) = (32{,}6{,}16)$ e sua capacidade de detecção e correção de erros são, respectivamente, $d-1 = 2^{5-1}-1 = 15$ e $\kappa = \left[\frac{2^{m-1}-1}{2}\right] = \left[\frac{2^{5-1}-1}{2}\right] = 7$.

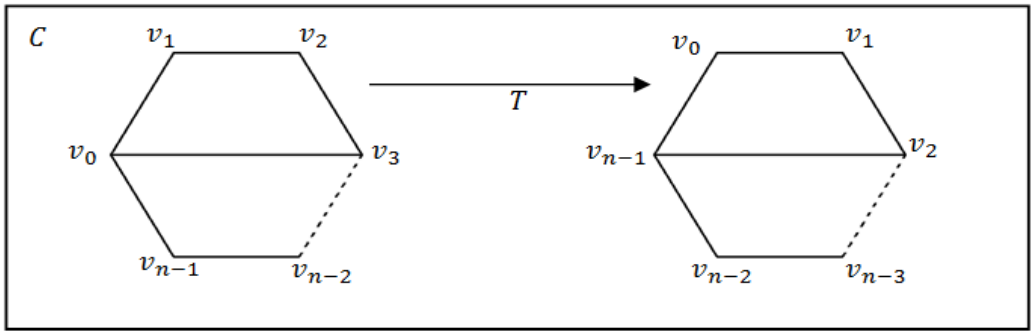
6.4 ALGUMAS NOÇÕES SOBRE CÓDIGOS CÍCLICOS

Consideremos K um corpo finito e um espaço vetorial K^n no qual um vetor v possua coordenadas $(v_0, v_1, v_2, \ldots, v_{n-1})$.

Dizemos que um código C contido no espaço vetorial K^n é um código cíclico quando C é linear e $\forall v \in C$, $v = (v_0, v_1, v_2, \ldots, v_{n-1})$, o vetor $v' = (v_{n-1}, v_0, v_1, \ldots, v_{n-2})$ também pertence a C.

Em outras palavras, C é um código cíclico quando existe uma transformação permutação cíclica $T: C \to C$ que a cada vetor $v = (v_0, v_1, v_2, \dots, v_{n-1})$ associa o vetor $T(v) = (v_{n-1}, v_0, v_1, \dots, v_{n-2})$, conforme a figura a seguir:

Figura 3: Esquema de uma permutação cíclica



Fonte: figura produzida pelo autor

Notemos que de acordo com exposto, temos $T(C) \subset C$ e que n transformações aplicadas a um vetor v retorna ao vetor v.

T é uma transformação linear, pois dados $\alpha \in K$ $u, v \in C$, com $u = (u_0, u_1, u_2, \dots, u_{n-1})$ e $v = (v_0, v_1, v_2, \dots, v_{n-1})$, temos:

$$T(u+v) = T[(u_0, u_1, u_2, \dots, u_{n-1}) + (v_0, v_1, v_2, \dots, v_{n-1})] =$$
$$= T(u_0 + v_0, u_1 + v_1, u_2 + v_2, \dots, u_{n-1} + v_{n-1}) =$$
$$= (u_{n-1} + v_{n-1}, u_0 + v_0, u_1 + v_1, \dots, u_{n-2} + v_{n-2}) =$$
$$= (u_{n-1}, u_0, u_1, \dots, u_{n-2}) + (v_{n-1}, v_0, v_1, \dots, v_{n-2}) =$$
$$= T(u_0, u_1, u_2, \dots, u_{n-1}) + T(v_0, v_1, v_2, \dots, v_{n-1}) = T(u) + T(v) \text{ e}$$
$$T(\alpha \cdot u) = T[\alpha \cdot (u_0, u_1, u_2, \dots, u_{n-1})] = T(\alpha \cdot u_0, \alpha \cdot u_1, \alpha \cdot u_2, \dots, \alpha \cdot u_{n-1}) =$$
$$= (\alpha \cdot u_{n-1}, \alpha \cdot u_0, \alpha \cdot u_1, \dots, \alpha \cdot u_{n-2}) = \alpha \cdot (u_{n-1}, u_0, u_1, \dots, u_{n-2}) =$$
$$= \alpha \cdot T(u_0, u_1, u_2, \dots, u_{n-1}) = \alpha \cdot T(u).$$

Ao vetor $(v_{n-1}, v_0, v_1, \dots, v_{n-2})$ denominamos *desvio cíclico* de $v \in C$ e representamos por $T(v)$.

$T(v)$ é uma transformação que leva o vetor $v \in C$ ao eu desvio cíclico $(v_{n-1}, v_0, v_1, \dots, v_{n-2})$. Notemos que T é uma permutação cíclica, portanto, T é bijetiva, logo, existe T^{-1} tal que $u \in C$ tal que $T^{-1}(u) \in C$.

Exemplos:

a) O código $C_1 = \{000, 110, 101, 011\}$ contido em F^3 é cíclico, pois:

$T(000) = 000 \in C_1$

$T(110) = 011 \in C_1$

$T(101) = 110 \in C_1$

$T(011) = 101 \in C_1$

b) O código $C_2 = \{00000, 10110, 01011, 11101\}$ contido em F^5, apresentado anteriormente no exemplo do braço mecânico, não é cíclico, pois:

$T(00000) = 00000 \in C_2$

$T(10110) = 01011 \in C_2$

$T(01011) = 10101 \notin C_2$

$T(11101) = 11110 \notin C_2$.

Consideremos um corpo K^n e o anel $R_n = K[x]/\langle x^n - 1\rangle$ quociente de $K[x]$ pelo ideal gerado por $x^n - 1$. Temos que $R_n = \{v_0 + v_1x + v_2x^2 + \cdots + v_{n-1}x^{n-1};\ v_i \in K\}$. Notemos que existe uma correspondência biunívoca entre os vetores de K^n e os elementos de R_n, que são polinômios (restos de divisão por $x^n - 1$):

$v = (v_0, v_1, v_2, \ldots, v_{n-1}) \leftrightarrow v(x) = v_0 + v_1x + v_2x^2 + \cdots + v_{n-1}x^{n-1}$.

Como K^n é um espaço vetorial, então R_n tem natureza de espaço vetorial.

Notemos que se $v' = (v_{n-1}, v_0, v_1, \ldots, v_{n-2})$ é uma permutação cíclica do vetor v, então $xv(x)$ deixa resto $v'(x)$ quando dividido por $x^n - 1$, pois $xv(x) = x \cdot (v_0 + v_1x + v_2x^2 + \cdots + v_{n-1}x^{n-1}) = v_0x + v_1x^2 + v_2x^3 + \cdots + v_{n-1}x^n$ e $v_{n-1} \cdot (x^n - 1) + (v_{n-1} + v_0x + v_1x^2 + \cdots + v_{n-2}x^{n-1}) = v_0x + v_1x^2 + \cdots + v_{n-1}x^n$, ou seja, $xv(x) = v_{n-1} \cdot (x^n - 1) + v'(x)$. Assim, qualquer código cíclico C contido em R_n pode ser definido como um subespaço de R_n tal que, se $v(x) \in C$, então $xv(x) \in C$.

Do exposto acima decorre que, se o código cíclico C está contido em R_n e $v(x) \in C$, então $xv(x) \in C$, $x^2v(x) \in C, x^3v(x) \in C$, etc. Observemos ainda que multiplicar por x significa realizar um deslocamento cíclico de uma posição e multiplicar por x^i significa realizar i deslocamentos cíclicos.

Teorema 6.23: Um código linear $C \subset R_n$ é um código cíclico se, e somente se, C é um ideal de R_n.

Demonstração:

($\Rightarrow$) Consideremos quem C é um código linear cíclico. Sejam $u(x) = (u_0 + u_1x + u_2x^2 + \cdots + u_{n-1}x^{n-1}) \in R_n$ e $v(x) = (v_0 + v_1x + v_2x^2 + \cdots + v_{n-1}x^{n-1}) \in C$. Vejamos que $u(x) \cdot v(x) = u_0 \cdot v(x) + u_1x \cdot v(x) + u_2x^2 \cdot v(x) + \cdots + u_{n-1}x^{n-1} \cdot v(x)$. Como C é cíclico, então, $x^i \cdot v(x) \in C$, além disso, C é linear, o que implica que a multiplicação de um

vetor $v(x) \in C$ por um escalar $u^i \in K$ e a soma de vetores em C resultam em vetores de C, logo $u(x) \cdot v(x) \in C$, o que mostra que C é um ideal de R_n.

($\Leftarrow$) Se C é um ideal de R_n, então, dado $x \in R_n$ e $v(x) \in C$, então $xv(x) \in C$. Mas, como vimos anteriormente, $xv(x)$ representa um deslocamento cíclico de uma posição, logo, C é cíclico.

R_n é um anel de ideais principais, pois R_n recebe essa propriedade de $K[x]$ e, como vimos em 5.4, todos os ideais de $K[x]$ são principais.

Qualquer que seja o ideal I de $K[x]$, tal que $I \supseteq \langle x^n - 1 \rangle$, então I é gerado por polinômios mônicos que dividem $x^n - 1$, ou seja, $I = \langle g(x) \rangle$ tal que $g(x) | x^n - 1$. Desta forma, um código cíclico C é um ideal gerado por $g(x)$:

$C = \langle g(x) \rangle;\ g(x) | x^n - 1$, daí, dizemos que $g(x)$ é um polinômio gerador de C.

Como o comprimento de cada palavra código da fonte é k e o comprimento de cada palavra do código de canal é n, então, temos $gr\big(g(x)\big) = n - k$.

6.4.1 Codificação em código cíclico

6.4.1.1 Codificação polinomial

Voltemos ao exemplo do braço mecânico, porém, nesse momento, desejamos obter um código C cíclico sobre o corpo galoisiano $F = \{0,1\}$. Vimos que o código de fonte necessário para a codificação dos quatro comandos é o conjunto $F^2 = \{00, 01, 10, 11\}$, cujos elementos representaremos por $m_0 = (0,0)$, $m_1 = (0,1)$, $m_2 = (1,0)$ e $m_3 = (1,1)$. A representação dos códigos do canal na forma polinomial é $m_0(x) = 0$, $m_1(x) = 1$, $m_2(x) = x$ e $m_3(x) = 1 + x$. Vamos obter um código C cíclico de comprimento $n = 6$ e dimensão $k = 2$. Um polinômio $g(x)$ gerador do código C é tal que $gr\big(g(x)\big) = 6 - 2 = 4$ e $g(x) | x^6 - 1$. Decompondo $x^6 - 1$, encontramos os polinômios fatores $(x - 1)$, $(x + 1)$ e $(x^4 + x^2 + 1)$, ou seja, $x^6 - 1 = (x - 1) \cdot (x + 1) \cdot (x^4 + x^2 + 1)$. Como $gr\big(g(x)\big) = 4$, então o polinômio gerador de C é $g(x) = 1 + x^2 + x^4$. Seja $c_i(x) \in C$ uma palavra do código, na forma polinomial não sistemática, assim, $c_i(x) = m_i(x) \cdot g(x)$, portanto, temos:

$$c_0(x) = m_0(x) \cdot g(x) = 0 \cdot (1 + x^2 + x^4) = 0$$
$$c_1(x) = m_1(x) \cdot g(x) = 1 \cdot (1 + x^2 + x^4) = 1 + x^2 + x^4$$
$$c_2(x) = m_2(x) \cdot g(x) = x \cdot (1 + x^2 + x^4) = x + x^3 + x^5$$

$c_3(x) = m_3(x) \cdot g(x) = (1 + x) \cdot (1 + x^2 + x^4) = 1 + x + x^2 + x^3 + x^4 + x^5$,

assim, temos $C = \{000000, 101010, 010101, 111111\}$.

Se quisermos determinar as palavras de um código cíclico C, na forma sistemática, ou seja, onde os k últimos dígitos correspondem aos dígitos da palavra do código de fonte e os $n - k$ primeiros dígitos representam a redundância acrescida, então, procedemos da seguinte forma:

Dada uma palavra m do código da fonte, à qual desejamos acrescentar redundâncias e obter uma palavra c do código de canal, escrevemos m na forma polinomial $m(x) = m_0 + m_1x + m_2x^2 + \cdots + m_{k-1}x^{k-1}$ e multiplicamos por x^{n-k}, obtendo um polinômio $p(x) = m(x) \cdot x^{n-k}$ de grau $n - 1$. Em seguida, dividimos o polinômio $p(x)$ pelo polinômio $g(x)$ gerador do código do código cíclico, obtendo um polinômio resto $\rho(x)$, ou seja, $m(x) \cdot p(x) = q(x) \cdot g(x) + \rho(x)$. Note que $q(x) \cdot g(x) = p(x) - \rho(x)$, ou seja, $p(x) - \rho(x) = c(x) \in C$. Logo, determinamos a palavra código do canal, na forma polinomial, fazendo $c(x) = -\rho(x) + m(x) \cdot x^{n-k}$. A palavra c do código cíclico é obtida dos coeficientes do polinômio $c(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1}$.

Observemos que, se o código C é sobre um corpo $K = F = \{0,1\}$, temos $-\rho(x) = \rho(x)$, o que implica que $c(x) = -\rho(x) + m(x) \cdot x^{n-k} = \rho(x) + m(x) \cdot x^{n-k}$.

Exemplo:

Vimos acima que um código cíclico C de comprimento $n = 6$, para os comandos do braço mecânico, possui dimensão $k = 2$ e é gerado pelo polinômio $g(x) = 1 + x^2 + x^4$. Apresentaremos na tabela a seguir as etapas da codificação polinomial:

$\boldsymbol{m}$	$\boldsymbol{m(x)}$	$\boldsymbol{x^{n-k} \cdot m(x)}$	$\boldsymbol{\rho(x)}$	$\boldsymbol{c(x)}$	**c**
00	0	0	0	0	000000
01	x	x^5	$x + x^3$	$x + x^3 + x^5$	010101
10	1	x^4	$1 + x^2$	$1 + x^2 + x^4$	101010
11	$1 + x$	$x^4 + x^5$	$1 + x + x^2 + x^3$	$1 + x + x^2 + x^3 + x^4 + x^5$	111111

6.4.1.2 Codificação matricial

Teorema 6.24: Seja C um código cíclico gerado por um polinômio $g(x)$ tal que $gr\big(g(x)\big) = n - k$ então os polinômios código $g(x), x \cdot g(x), x^2 \cdot g(x), \ldots, x^{k-1} \cdot g(x)$ formam uma base para C.

Demonstração: Para demonstrar esse fato, devemos provar que os polinômios $g(x), x \cdot g(x), x^2 \cdot g(x), \dots, x^{k-1} \cdot g(x)$ são linearmente independentes e que geram o Código C (que é um espaço vetorial sobre K).

Dados $\alpha_0, \alpha_1, \dots, \alpha_{k-1}$ pertencentes a K tais que, se $\alpha_0 \cdot g(x) + \alpha_1 \cdot x \cdot g(x) + \cdots + + \alpha_{k-1} \cdot x^{k-1} \cdot g(x) = 0$, então, temos $(\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{k-1} x^{k-1}) \cdot g(x) = 0$ como $gr((\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{k-1} x^{k-1}) \leq k-1$ e $gr\big(g(x)\big) = n-k$, então $gr\Big((\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{k-1} x^{k-1}) \cdot g(x)\Big) \leq n-k+k-1 = n-1 < n$, o que implica que $\alpha_0 \cdot g(x) + \alpha_1 \cdot x \cdot g(x) + \alpha_2 \cdot x^2 \cdot g(x) + \cdots + \alpha_{k-1} \cdot x^{k-1} \cdot g(x) = 0$ só se verifica quando $\alpha_0 = \alpha_1 = \alpha_2 = \cdots, = \alpha_{k-1} = 0$, portanto $g(x), x \cdot g(x), x^2 \cdot g(x), \dots, x^{k-1} \cdot g(x)$ são polinômios linearmente independentes. Além disso, pelo teorema 6.17, temos que $\forall v(x) \in C$, temos que $v(x) = u(x) \cdot g(x)$, onde $u(x) \in R_n$, $gr\big(u(x)\big) \leq k-1$ e $g(x)$ é o gerador de C. Assim:

$v(x) = u(x) \cdot g(x) = (u_0 + u_1 x + u_2 x^2 + \cdots + u_{k-1} x^{k-1}) \cdot g(x) =$

$= u_0 \cdot g(x) + u_1 x \cdot g(x) + u_2 x^2 \cdot g(x) + \cdots + u_{k-1} x^{k-1} \cdot g(x) =$

$= u_0 \cdot \big(g(x)\big) + u_1 \cdot \big(x \cdot g(x)\big) + u_2 \cdot \big(x^2 \cdot g(x)\big) + \cdots + u_{k-1} \cdot \Big(x^{k-1} \cdot g(x)\Big)$, ou seja, qualquer que seja $v(x) \in C$, temos que $v(x)$ é escrito como combinação linear de $g(x), x \cdot g(x), x^2 \cdot g(x), \dots, x^{k-1} \cdot g(x)$. Portanto, $g(x), x \cdot g(x), x^2 \cdot g(x), \dots, x^{k-1} \cdot g(x)$ é uma base de C.

O teorema 6.24 é de extrema importância para a representação matricial de um código cíclico, pois, se um código C tem comprimento n, dimensão k, é gerado pelo polinômio $g(x)$, com $gr\big(g(x)\big) = n-k$, então, como $g(x), x \cdot g(x), x^2 \cdot g(x), \dots, x^{k-1} \cdot g(x)$ é uma base de C, a matriz G de ordem $k \times n$, geradora desse código é representada como:

$$G = \begin{bmatrix} g(x) \\ xg(x) \\ x^2 g(x) \\ \vdots \\ x^{k-1} g(x) \end{bmatrix} = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k} & 0 & \cdots & 0 & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k-1} & g_{n-k} & 0 & \dots & 0 \\ 0 & 0 & g_0 & \cdots & g_{n-k-2} & g_{n-k-1} & g_{n-k} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & g_0 & g_1 & \cdots & g_{n-k} \end{bmatrix},$$ onde g_i são coeficientes do polinômio gerador.

Exemplo:

Seja C um código cíclico de comprimento $n = 7$ e dimensão $k = 4$, sobre o corpo galoisiano $F = \{0,1\}$. Um polinômio $g(x)$, gerador de C possui grau $gr(g(x) = n-k = 7-4 = 3$ e divide o polinômio $x^7 - 1$. Decompondo $x^7 - 1$,

encontramos os fatores $(x-1)$, (x^3+x^2+1) e (x^3+x+1), ou seja, $(x^7-1)=(x-1)\cdot(1+x^2+x^3)\cdot(1+x+x^3)$.

Isso pode ser facilmente verificado efetuando o produto $(x-1)\cdot(1+x^2+x^3)\cdot(1+x+x^3)=(-1+x-x^2+x^4)\cdot(1+x+x^3)=(-1-x-x^3+x+x^2+x^4-x^2-x^3-x^5+x^4+x^5+x^7)=(-1-2x^3+2x^4+x^7)$, não devemos esquecer que os coeficientes desses polinômios são elementos de $F=\{0,1\}$, portanto, a aritmética utilizada deve ser compatível com esse corpo, assim, $(-1-2x^3+2x^4+x^7)=(-1-0x^3+0x^4+x^7)=(x^7-1)$. Dessa decomposição, vemos que existem dois polinômios de grau 3 que dividem x^7-1, a saber $1+x^2+x^3$ e $1+x+x^3$, sendo que qualquer um desses polinômios é um gerador do código C.

Um polinômio gerador de um código cíclico C é da forma $g(x)=g_0+g_1x+g_2x^2+\cdots+g_{k-1}x^{n-k}$. Como o comprimento do código C acima é $n=7$ e sua dimensão é $k=4$, então polinômio gerador de C é da forma $g(x)=g_0+g_1x+g_2x^2+g_3x^3$. Comparando com um dos polinômios geradores acima, por exemplo, o polinômio $1+x^2+x^3$, temos:

$g_0+g_1x+g_2x^2+g_3x^3=1+x^2+x^3=1+0x+x^2+x^3$, o que implica que $g_0=1,\ g_1=0,\ g_2=1$ e $g_3=1$. Assim, uma matriz G de ordem $k\times n$, geradora do código C é:

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

Note que a matriz G não se apresenta na forma padrão, portanto, gera palavras código na forma não sistemática, mas, por meio de transformações elementares sobre as linhas de G, podemos determinar uma matriz $G'=[R|I_k]$ equivalente a G e apresentada na forma padrão. A partir de G' geramos palavras código na forma sistemática.

Voltando ao exemplo anterior, do código cíclico C referente aos comandos do braço mecânico, um polinômio gerador de um código cíclico C é da forma $g(x)=g_0+g_1x+g_2x^2+\cdots+g_{k-1}x^{n-k}$. Como o comprimento do código C acima é $n=6$ e sua dimensão é $k=2$, então polinômio gerador de C é da forma $g(x)=g_0+g_1x+g_2x^2+g_3x^3+g_4x^4$. Comparando com o polinômio gerador $g(x)=1+x^2+x^4$, temos:

$g_0+g_1x+g_2x^2+g_3x^3+g_4x^4=1+x^2+x^4=1+0x+x^2+0x^3+x^4$, o que implica que $g_0=1,\ g_1=0,\ g_2=1,\ g_3=0$ e $g_4=1$. Assim, uma matriz G de ordem $k\times n=2\times 6$, geradora do código C é:

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

Note que a matriz G já se encontra na forma padrão $G = [R|I_2]$.

Para obter os elementos (palavras) do código C é suficiente realizar o produto dos vetores do código de canal pela matriz G:

$$[0 \quad 0] \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$[0 \quad 1] \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1]$$

$$[1 \quad 0] \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = [1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0]$$

$$[1 \quad 1] \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1]$$

Observe que as palavras código foram obtidas na forma sistemática em virtude de G ser apresentada na forma padrão.

6.4.2 Código dual de um código cíclico

Teorema 6.25: Da do um código cíclico $C \subset K^n$, um código $C^{\perp}$ dual do código C, é também um código cíclico.

Demonstração: Se C é um código cíclico, então, por definição, C é linear. Como o código dual de um código linear é também um código linear, então o código dual $C^{\perp}$ é um código linear. Verifiquemos então que $C^{\perp}$ é cíclico. Para todo $c \in C$, temos que $T^{-1}(c) \in C$, dado $v \in C^{\perp}$, temos que $v \cdot T^{-1}(c) = 0$, mas $v \cdot T^{-1}(c) = \sum_{i=0}^{n-1} v_i \cdot c_{i+1} = T(v) \cdot c$, portanto, o fato de $v \in C^{\perp}$, implica que $T(v) \in C^{\perp}$, portanto, $C^{\perp}$ é um código cíclico.

Se C é um código cíclico de comprimento n, com polinômio gerador $g(x)$, cujo grau é $gr\big(g(x)\big) = n - k$, então existe $h(x) \in K[x]$ tal que $h(x) \cdot g(x) = x^n - 1$. Ao polinômio $h(x)$ denominamos *polinômio de paridade* do código C. Como $g(x)$ e $x^n - 1$ são polinômios mônicos, então $h(x)$ é também um polinômio Mônico. Notemos ainda que se $h(x) \cdot g(x) = x^n - 1$, então temos $gr\big(h(x) \cdot g(x)\big) = gr(x^n - 1)$ e, portanto, $gr\big(h(x)\big) + gr\big(g(x)\big) = gr(x^n - 1)$, ou seja, $gr\big(h(x)\big) + n - k = n$, o que implica que $gr\big(h(x)\big) = k$.

Em geral, o polinômio $h(x)$ de paridade de um código C não é o gerador do código dual $C^{\perp}$.

Teorema 6.26: $c(x) \in C$ se, e somente se, $c(x) \cdot h(x)$ deixa resto zero quando dividido por $x^n - 1$.

Demonstração: Dizer que $c(x) \cdot h(x)$ deixa resto zero quando dividido por $x^n - 1$ equivale a dizer que existe $z(x) \in K[x]$ tal que $c(x) \cdot h(x) = z(x) \cdot (x^n - 1)$, o que equivale a $c(x) = z(x) \cdot g(x)$ ou ainda que $c(x) \in \langle g(x) \rangle = C$.

Teorema 6.27: Seja C um código cíclico de comprimento n e dimensão k, cujo polinômio de paridade é $h(x) = h_0 + h_1 x + h_2 x^2 + \cdots + h_k x^k$, então a matriz $H_{(n-k)\times n}$, definida por

$$H = \begin{bmatrix} h_k & h_{k-1} & \cdots & h_0 & 0 & \cdots & \cdots & 0 \\ 0 & h_k & h_{k-1} & \cdots & h_0 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & h_k & h_{k-1} & \cdots & h_0 & 0 \\ 0 & \cdots & \cdots & 0 & h_k & h_{k-1} & \cdots & h_0 \end{bmatrix}$$ é a matriz teste de paridade para o código C.

Demonstração: $h_k = 1$, pois $h(x)$ é um polinômio mônico e as linhas de H são linearmente independentes. Dado $c \in C$ tal que $c(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$, pelo teorema 6.26 $c(x) \in C$ se, e somente se, $c(x) \cdot h(x)$ deixa resto zero quando dividido por $x^n - 1$, assim, desenvolvendo $c(x) \cdot h(x)$, temos:

$$c(x) \cdot h(x) = c_0 h_0 + (c_0 h_1 + c_1 h_0) + \cdots + \left(\sum_{i+j=k} c_i h_j \right) x^k + \cdots + \left(\sum_{i+j=n-1} c_i h_j \right) x^{n-1}$$

$$+ \cdots + \left(\sum_{i+j=n} c_i h_j \right) x^n + \cdots + c_{n-1} h_k x^{n+k-1}$$

O desenvolvimento acima deixa resto zero quando dividido por $x^n - 1$, portanto, os termos de graus n até $n + k - 1$ transformam-se em termos de grau 0 a $k - 1$, respectivamente, portanto, os termos de graus k a $n - 1$ deixam resto zero quando divididos por $x^n - 1$. Assim, $c(x) \in C$ se, e somente se, é solução do seguinte sistema de equações:

$$\begin{cases} c_0 h_k + c_1 h_{k-1} + \cdots + c_k h_0 = 0 \\ c_1 h_k + \cdots + c_k h_1 + c_{k+1} h_0 = 0 \\ \vdots \\ c_{n-k-1} h_k + \cdots + c_{n-1} h_0 = 0 \end{cases}$$

Representando o sistema acima em notação matricial, temos $H \cdot c^t$, o que implica que H é a matriz teste de paridade do código C.

Teorema 6.28: Seja C um código cíclico de comprimento n e dimensão k, cujo polinômio de paridade é $h(x) = h_0 + h_1x + h_2x^2 + \cdots + h_kx^k$, então o polinômio $h^*(x) = x^k h(t^{-1}) \in K[x]$ é um gerador do código dual $C^{\perp}$. (O polinômio $h^*(x)$ é denominado polinômio recíproco de $h(x)$)

Demonstração: Temos $h_k = 1$ pois h é um polinômio Mônico, portanto, resta-nos provar que $h^*(x)|x^n - 1$. Como $h(x) \cdot g(x) = x^n - 1$, então, $h(x^{-1}) \cdot g(x^{-1}) = x^{-n} - 1$. Multiplicando a igualdade por $-x^n$, temos $-x^n \cdot h(x^{-1}) \cdot g(x^{-1}) = -x^n \cdot (x^{-n} - 1)$, que equivale a $x^k \cdot h(x^{-1}) \cdot \left(-x^{n-k} \cdot g(x^{-1})\right) = x^n - 1$, ou seja, $h^*(x) \cdot \left(-x^{n-k} \cdot g(x^{-1})\right) = x^n - 1$. Logo, $h^*(x)$ divide $x^n - 1$ e, portanto, é um polinômio gerador do código dual $C^{\perp}$ do código C.

Exemplo: Considerando código C de comprimento 7 e dimensão 4 sobre $F\{0,1\}$, visto anteriormente, cujo polinômio gerador é $g(x) = 1 + x^2 + x^4$. Da igualdade $h(x) \cdot g(x) = x^7 - 1$, temos:

$$
\begin{array}{llllll|lll}
x^7 & -1 & & & & & x^3 & +x^2 & +1 \\
\cline{7-9}
-x^7 & -x^6 & -x^4 & & & & x^4 & -x^3 & +x^2 & +1 \\
\cline{1-4}
 & -x^6 & -x^4 & -1 & & & & & & \\
 & x^6 & +x^5 & +x^3 & & & & & & \\
\cline{2-5}
 & & +x^5 & -x^4 & +x^3 & -1 & & & & \\
 & & -x^5 & -x^4 & -x^2 & & & & & \\
\cline{3-6}
 & & & & +x^3 & -x^2 & -1 & & & \\
 & & & & -x^3 & -x^2 & -1 & & & \\
\cline{5-7}
 & & & & 0 & 0 & 0 & & &
\end{array}
$$

Note que estamos operando em $F = \{0,1\}$, portando, $x^4 - x^3 + x^2 + 1 = x^4 + x^3 + x^2 + 1$. Assim, temos $h(x) = 1 + x^2 + x^3 + x^4$ como polinômio de paridade de C. Isso implica que a matriz teste de paridade para C é:

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

Como $h(x) = 1 + x^2 + x^3 + x^4 = 1 \cdot x^0 + 0 \cdot x + 1 \cdot x^2 + 1 \cdot x^3 + 1 \cdot x^4$, então o polinômio recíproco de $h(x)$ é:

$h^*(x) = 1 \cdot (x^0)^{-1} + 0 \cdot (x^1)^{-1} + 1 \cdot (x^2)^{-1} + 1.(x^3)^{-1} + 1 \cdot (x^4)^{-1}$

$h^*(x) = 1 \cdot x^4 + 0 \cdot x^3 + 1 \cdot x^2 + 1.x^1 + 1 \cdot x^0$

$h^*(x) = 1 + x + x^2 + x^4$

Portanto, $h^*(x) = 1 + x + x^2 + x^4$ é o polinômio gerador do código dual $C^{\perp}$ de C.

6.4.3 Decodificação em código cíclico

Seja C um código cíclico, $c \subset C$ a mensagem enviada e r a mensagem recebida. Na forma polinomial, temos $c(x)$ e $r(x)$.

Notemos inicialmente que se G é uma matriz geradora do código C, então, por meio de transformações elementares sobre as linhas de G obtemos uma matriz G' na forma padrão, ou seja, $G^{'} = [R|I_k]$, da qual se obtém uma matriz teste de paridade $H = [I_{n-k}| - R^T]$. Representemos por $-\rho_i(x)$ o polinômio correspondente à $i-$ésima linha da matriz R. Assim, a $i-$ésima linha da matriz G' é representada pelo polinômio $-\rho_i(x) + x^{n-k+i}$, com $gr\big(-\rho_i(x)\big) \leq n-k-1$, pois a matriz R possui $n-k$ colunas.

Sabemos que cada linha da matriz G' é uma palavra do código cíclico C. Como C é gerado por $g(x)$, então, cada linha de G' é um múltiplo de $g(x)$, ou seja, existe $q_i(x)$, tal que $-\rho_i(x) + x^{n-k+i} = q_i(x) \cdot g(x)$ ou ainda que $x^{n-k+i} = q_i(x) \cdot g(x) + \rho_i(x), \quad \forall i \in \{1, 2, \dots, k-1\}$, ou seja, $\rho_i(x)$ é o resto da divisão de x^{n-k+i} por $g(x)$.

Como um código cíclico é linear, então, podemos utilizar a decodificação por síndrome para códigos lineares, ou seja, se $r(x) \subset K^n$, é uma palavra recebida, onde K é um corpo finito, então devemos determinar uma síndrome $s(r)$.

Teorema 6.29: A síndrome $s(r(x))$ de uma palavra r recebida é o resto da divisão de $r(x)$ pelo polinômio gerador $g(x)$ do código C.

Demonstração: Considerando que C é um código linear com matriz geradora na forma padrão $G^{'} = [R|I_k]$, então existe uma matriz teste de paridade de C na forma padrão, representada por $H = [I_{n-k}| - R^T]$. Notemos que as colunas de $-R^T$ são os vetores representados por $\rho_0(x), \rho_1(x), \dots, \rho_{k-1}(x)$. Como $r = (r_0, r_1, \dots, r_{n-1})$, então a representação polinomial de r é $r(x) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-1}x^{n-1}$. Por definição, a síndrome de r é $s(r) = H \cdot r^t$ e, em representação polinomial, utilizando a matriz $H = [I_{n-k}| - R^T]$, temos:

$$s\big(r(x)\big) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-k-1}x^{n-k-1} + r_{n-k}\rho_0(x) + r_{n-k+1}\rho_1(x) + \cdots + r_{n-1}\rho_{k-1}(x)$$

$$s\big(r(x)\big) = r(x) - \sum_{i=0}^{k-1} r_{n-k+i}(\rho_i(x) - x^{n-k+i}$$

$$s\big(r(x)\big) = r(x) - \sum_{i=0}^{k-1} r_{n-k+i}(q_i(x) \cdot g(x)) = r(x) - \left(\sum_{i=0}^{k-1} r_{n-k+i}q_i(x)\right) \cdot g(x)$$

Sendo $Q(x) = \sum_{i=0}^{k-1} r_{n-k+i} q_i(x)$, temos $s\big(r(x)\big) = r(x) - Q(x) \cdot g(x)$, ou seja, $r(x) = Q(x) \cdot g(x) + s\big(r(x)\big)$, o que implica que a síndrome $s\big(r(x)\big)$ é o resto da divisão do polinômio $r(x)$ pelo polinômio gerador $g(x)$.

Temos ainda que caso haja erros na transmissão de uma palavra código c de modo que a palavra recebida seja r, então o vetor erro e é tal que $r = c + e$ e $s(e) = s(r)$. Caso $s(r)$ seja o vetor nulo, então $s(e)$ é o vetor nulo e não $r = c$.

O peso de um elemento $v(x) \in R_n$, com $v(x) = v_0 + v_1 x + \cdots + v_{n-1} x^{n-1}$, é definido como sendo $\omega\big(v(x)\big) = \omega(v) = \omega(v_0, v_1, v_2, \dots, v_{n-1})$. Pelo teorema 6.29, se um código linear possui distância mínima d, e, se $s(r(x))$ é tal que $\omega\big(s(r(x))\big) \leq \left[\frac{d-1}{2}\right] = \kappa$, então $s(r(x))$ é um líder da classe $s(r(x)) + C$, logo, decodificamos $r(x)$, obtendo $c(x) = r(x) - s(r(x))$, onde $s(r(x))$ é o resto da divisão de $r(x)$ pelo polinômio gerador $g(x)$.

Exemplos:

1) Considerando o código cíclico C do braço mecânico. O polinômio gerador de C é $g(x) = 1 + x^2 + x^4$ e $\kappa = 1$. Seja $c = 111111$ a palavra transmitida, que na forma polinomial é $c(x) = 1 + x + x^2 + x^3 + x^4 + x^5$. Suponhamos que a palavra recebida tenha sido $r = 110111$, cuja representação polinomial é $r(x) = 1 + x + x^3 + x^4 + x^5$.

Resolvendo o quociente $\frac{r(x)}{g(x)}$, temos:

$$\begin{array}{rrrrr|l}
x^5 & +x^4 & +x^3 & +x & +1 & x^4 + x^2 + 1 \\
\cline{6-6}
-x^5 & & -x^3 & -x & & x + 1 \\
\cline{1-4}
 & +x^4 & & & +1 & \\
 & -x^4 & -x^2 & & -1 & \\
\cline{2-4}
 & & -x^2 & & &
\end{array}$$

Veja que $r(x) = -x^2 + (x + 1) \cdot (1 + x^2 + x^4)$, portanto, a síndrome da palavra recebida é $s(r(x)) = -x^2$. Como $\omega\big(s(r(x))\big) = 1$, então $c(x) = r(x) - s(r(x))$ equivale a $c(x) = (1 + x + x^3 + x^4 + x^5) - (-x^2) = 1 + x + x^2 + x^3 + x^4 + x^5$, o que implica que a palavra transmitida foi $c = 111111$.

2) Suponhamos agora que a palavra transmitida seja $c = 010101$, cuja forma polinomial equivalente é $c(x) = x + x^3 + x^5$ e que a palavra recebida seja $r = 010111$ cuja representação polinomial é $r(x) = x + x^3 + x^4 + x^5$.

Resolvendo o quociente $\frac{r(x)}{g(x)}$, temos:

$$\begin{array}{rrrr|l} x^5 & +x^4 & +x^3 & +x & x^4 \quad +x^2 \quad +1 \\ \cline{5-5} -x^5 & & -x^3 & -x & x+1 \\ \cline{1-3} & +x^4 & & & \\ & -x^4 & -x^2 & -1 & \\ \cline{2-4} & & -x^2 & -1 & \end{array}$$

Veja que $r(x) = (1 + x^2) + (x + 1) \cdot (1 + x^2 + x^4)$, portanto, a síndrome da palavra recebida é $s(r(x)) = 1 + x^2$, como $\omega\big(s(r(x))\big) = 2 > 1 = \kappa$, então não podemos ainda decodificar $r(x)$ utilizando $c(x) = r(x) - s(r(x))$. Porém, os polinômios $r(x) = x + x^3 + x^4 + x^5$ e $r'(x) = x^3 + x^4 + x^5 + x^7$ deixam mesmo resto quando divididos por $x^6 - 1$, assim, façamos o quociente $\frac{r'(x)}{g(x)}$:

$$\begin{array}{rrrrr|l} x^7 & +x^5 & +x^4 & +x^3 & x^4 \quad +x^2 \quad +1 \\ \cline{5-5} -x^7 & -x^5 & & -x^3 & x^3 \\ \cline{1-4} & & +x^4 & & \end{array}$$

Note que $gr(x^4) = gr(g(x))$, o que implica que x^4 não é o resto da divisão de $r(x)$ por $g(x)$, porém, $\omega(x^4) = 1 = \kappa$, logo, x^4 é líder da classe $r(x) + C$, portanto, $c(x) = r(x) - x^4$, que equivale a $c(x) = x + x^3 + x^4 + x^5 - x^4 = x + x^3 + x^5$, ou seja, a palavra transmitida foi $c = 010101$

Vimos que obter o líder de classe não acontece de forma direta com a aplicação do sintoma se esse tiver peso maior que κ. Notemos, porém, que um desvio cíclico $x^i r(x) \in R_n$ carrega a mesma informação que R_n, portanto, se decodificarmos $x^i r(x)$ para algum i, como consequência, decodificaremos $r(x)$.

Teorema 6.30: Seja $r(x) \in R_n$. A síndrome de um desvio cíclico $xr(x)$ de $r(x)$ é $s\big(r(x)\big) = xs\big(r(x)\big) - s_{n-k-1}g(x)$, onde s_{n-k-1} é o coeficiente do termo de grau $n-k-1$ de $s\big(r(x)\big)$.

Demonstração: Considerando $s\big(r(x)\big)$ como a síndrome de $r(x)$, temos que $r(x) = q(x) \cdot g(x) + s(r(x))$, com grau de $s\big(r(x)\big) \leq n-k-1$, para algum $q(x)$. Escrevendo $s\big(r(x)\big) = s_{n-k-1}x^{n-k-1} + s^{'}(x)$ e $g(x) = x^{n-k} + g(x)$, onde $gr\big(s(x)\big) < n-k-1$ e $gr\big(g^{'}(x)\big) < n-k$, temos:

$xr(x) = xq(x) \cdot g(x) + xs\big(r(x)\big) = x(q(x) + s_{n-k-1}) \cdot g(x) + (xs\big(r(x)\big) - s_{n-k-1}g(x))$.

Mas $xs\big(r(x)\big) - s_{n-k-1} = xs^{'}\big(r(x)\big) - s_{n-k-1}g^{'}(x)$ tem grau menor que $n-k$, logo, pelo teorema 6.29, temos que $s(xr(x) = xs\big(r(x)\big) - s_{n-k-1}g(x)$.

Com base do descrito acima, levando em consideração o exemplo anterior (2), onde $r(x) = x + x^3 + x^4 + x^5$, $g(x) = 1 + x^2 + x^4$ e $n-k-1 = 6-4-1 = 3$, vimos que a síndrome de $r(x)$ é $s\big(r(x)\big) = 1 + x^2$, portanto, as demais síndromes são:

$s_1\big(xr(x)\big) = xs\big(r(x)\big) - s_3 g(x) = x \cdot (1 + x^2) - 0 \cdot (1 + x^2 + x^4) = x + x^3$

$s_2\big(x^2 r(x)\big) = xs_1\big(xr(x)\big) - s_3 g(x) = x \cdot (x + x^3) - 1 \cdot (1 + x^2 + x^4) =$

$= x^2 + x^4 - 1 - x^2 - x^4 = 1$

$s_3\big(x^3 r(x)\big) = xs_2\big(x^2 r(x)\big) - s_3 g(x) = x \cdot 1 - 0 \cdot (1 + x^2 + x^4) = x$

$s_4\big(x^4 r(x)\big) = xs_3\big(x^3 r(x)\big) - s_3 g(x) = x \cdot x - 0 \cdot (1 + x^2 + x^4) = x^2$

$s_5\big(x^5 r(x)\big) = xs_4\big(x^4 r(x)\big) - s_3 g(x) = x \cdot x^2 - 0 \cdot (1 + x^2 + x^4) = x^3$.

Dizemos que um vetor $v = (v_0, v_1, v_2, \dots, v_{n-1})$ de k^n contém uma sequência cíclica de k zeros se existe j tal que $v_j = v_{j+1} = v_{j+2} = \cdots = v_{j+k-1} = 0$.

Exemplo: O vetor $v_1 = (1,0,0,0,1,0,1) \in F^7$ possui uma sequência cíclica de 3 zeros. O vetor $v_2 = (0,0,1,1,0,1,0,0,0,0,0) \in F^{11}$ possui uma sequência cíclica de sete zeros pois basta realizar um deslocamento cíclico de cinco unidades para a direita que obtemos o vetor $T(v_2) = (0,0,0,0,0,0,0,1,1,0,1)$. Se $v = (v_0, v_1, v_2, \dots, v_{n-1})$ contém uma sequência cíclica de k zeros, então existe $i \in \{0,1,2,\dots,n-1\}$ tal que $gr\left(x^i v(x)\right) \leq n-k-1$, para isso, basta permutar ciclicamente as coordenadas de v de modo que os k zeros ocupem as k últimas coordenadas de v, que obtemos um vetor com no máximo as $n-k-1$ primeiras coordenadas não nulas, que corresponde a um polinômio de grau no máximo $n-k-1$.

Teorema 6.31: Se um vetor erro $e \in k^n$, cujo peso é $\omega(e) \leq \left[\frac{d-1}{2}\right] = \kappa$ e possui uma sequência cíclica de k zeros, então $\omega\left(s\left(x^i e(x)\right)\right) \leq \left[\frac{d-1}{2}\right] = \kappa$, para algum $i \in \{0,1,2,\dots,n-1\}$.

Demonstração: Pelo visto anteriormente, se e possui uma sequência cíclica de k zeros, então existe $i \in \{0, 1, 2, \dots, n-1\}$ tal que $gr\left(x^i e(x)\right) \leq n-k-1$. Pelo teorema 6.30 temos que $s\left(x^i e(x)\right) = x^i s(e(x))$. Como as permutações entre as coordenadas de um vetor não alteram seu peso, então temos que se $\omega(e) \leq \left[\frac{d-1}{2}\right] = \kappa$ então $\omega\left(s\left(x^i e(x)\right)\right) \leq \left[\frac{d-1}{2}\right] = \kappa$.

Com base no exposto, uma sequência para decodificação consiste em determinar os sintomas de um vetor $r(x)$ recebido e, se $\omega(s_i(r(x))) \leq \kappa$ para algum $i \in \{0, 1, 2, \dots, n-1\}$, então assumimos que $e(x) = x^{n-1} s_i(r(x))$ é o vetor erro e decodificamos $r(x)$ como $c(x) = r(x) - x^{n-1} s_i(r(x))$. Caso contrário, o erro não é corrigível.

Para concluir, apresentamos um último exemplo referente à decodificação em códigos cíclicos:

Consideremos um código cíclico C contido em F^7, onde $F = \{0,1\}$ é um corpo galoisiano, tal que $\dim C = 4$ e um polinômio gerador de C é $g(x) = 1 + x + x^3$. A distância mínima desse código é $d = 3$, o que implica que ele tem capacidade de detecção de $d - 1 = 3 = 1 = 2$ erros e capacidade de correção $\kappa = \left[\frac{d-1}{2}\right] = \left[\frac{3-1}{2}\right] = 1$ erro. Como $\kappa = 1$, então os líderes de classe são os vetores com peso $\omega \leq 1$.

Consideremos que a palavra recebida seja $r = 1000100$ cuja representação polinomial é $r(x) = 1 + x^4$. Resolvendo o quociente $\frac{r(x)}{g(x)}$, temos:

$$\begin{array}{rrr|l} x^4 & +1 & & x^3 + x + 1 \\ \cline{4-4} -x^4 & -x^2 & -x & x \\ \cline{1-3} -x^2 & -x & +1 & \end{array}$$

O resto da divisão de $r(x)$ por $g(x)$ é $-x^2 - x + 1 = x^2 + x + 1$.

Assim, $x^4 + 1 = x \cdot (x^3 + x + 1) + (1 + x + x^2)$, o que implica que $s\left(r(x)\right) = 1 + x + x^2$. Mas $\omega(s\left(r(x)\right) = 3 > 1 = \kappa$, sendo assim, não podemos ainda decodificar $r(x)$ utilizando $c(x) = r(x) - s(r(x))$, portanto, determinaremos as síndromes de todos os desvios cíclicos de $r(x)$:

$s_1\left(xr(x)\right) = xs\left(r(x)\right) - s_2 g(x) = x \cdot (1 + x + x^2) - 1 \cdot (1 + x^2 + x^3) = x^2 - 1 = 1 + x^2$

$s_2\left(x^2 r(x)\right) = xs_1\left(xr(x)\right) - s_2 g(x) = x \cdot (1 + x^2) - 1 \cdot (1 + x^2 + x^3) =$

$= x^2 + x^3 - 1 - x - x^3 = -1 = 1$

$s_3\left(x^3 r(x)\right) = xs_2\left(x^2 r(x)\right) - s_2 g(x) = x \cdot 1 - 0 \cdot (1 + x + x^3) = x$

$s_4(x^4 r(x)) = x s_3(x^3 r(x)) - s_2 g(x) = x \cdot x - 0 \cdot (1 + x + x^3) = x^2$

$s_5(x^5 r(x)) = x s_4(x^4 r(x)) - s_2 g(x) = x \cdot x^2 - 1 \cdot (1 + x + x^3) = x^3 - 1 - x - x^3 =$

$= -1 - x = 1 + x$

$s_6(x^5 r(x)) = x s_5(x^5 r(x)) - s_2 g(x) = x \cdot (1 + x) - 0 \cdot (1 + x + x^3) = x + x^2.$

Qualquer um dos $s_i\left(x^i r(x)\right)$ tal que $\omega\left(s_i\left(x^i r(x)\right)\right) = 1$ é um líder da classe, portanto, qualquer um deles pode ser utilizado no processo de decodificação. Utilizemos, por exemplo, $s_2(x^2 r(x)) = 1$: pelo que foi visto anteriormente, $e(x) = x^{n-1} s_i(r(x))$, logo, $e(x) = x^{7-2} s_2(r(x)) = x^5 \cdot 1 = x^5$, assim, a decodificação correta de $r(x)$ é $c(x) = r(x) - x^5 = 1 + x^4 - x^5 = 1 + x^4 + x^5$. A palavra transmitida é $c = 100011$.

7 ATIVIDADES POPOSTAS

Este capítulo tem como objetivo apresentar uma coletânea de atividades propostas sobre os conteúdos apresentados neste trabalho, cujo objetivo principal é a aplicação das propriedades estudas e, por conseguinte, a familiarização dos estudantes aos conceitos aqui apresentados.

A rotina de atividades propostas segue a seguinte ordem:

1º - Atividades sobre matrizes, operações e propriedades;

2º - Atividades sobre determinantes e suas propriedades;

3º - Atividades sobre polinômios, operações e propriedades;

4º - Atividades sobre códigos corretores de erros, aplicação de matrizes, determinantes e polinômios.

7.1 MATRIZES REAIS

1) Indique explicitamente os elementos da matriz $A = (a_{ij})_{3\times4}$ tal que $a_{ij} = i^2 - 3ij + j^2$.

2) Construa a matriz $A = (a_{ij})_{4\times4}$ tal que $a_{ij} = \begin{cases} 1, se\ i = j \\ 0, se\ i \neq j \end{cases}$. O que se pode dizer a respeito dessa matriz?

3) Construa a matriz $A = (a_{ij})_{5\times5}$ tal que $a_{ij} = \begin{cases} i^2 + 2j, se\ i \leq j \\ 0, se\ i > j \end{cases}$. O que se pode dizer a respeito dessa matriz?

4) Dadas as matrizes $A = \begin{bmatrix} 5a & 4b \\ 6 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} a+4 & 6b \\ 6 & b+3 \end{bmatrix}$, quais são os valores de a e b para que se tenha $A = B$?

5) Dadas as matrizes $A = \begin{bmatrix} 2 & -3 & 4 \\ -1 & -6 & 0 \\ 3 & 5 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 & 1 \\ -2 & 5 & 1 \\ 3 & -3 & -1 \end{bmatrix}$ e $C = \begin{bmatrix} 3 & 1 & -5 \\ -3 & 2 & 4 \\ 2 & 0 & 2 \end{bmatrix}$, determine a matriz X tal que $A^2 + X = 3 \cdot B - C^t$.

6) Obtenha todas as matrizes M que comutam com $A = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}$.

7) Determine todas as matrizes A quadradas de ordem 2 tais que $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

8) Dadas as matrizes $A = (a_{ij})_{3\times4}$ tal que $a_{ij} = i^3 - 4j$, $B = (b_{ij})_{3\times4}$ tal que $b_{ij} = 8i - j^2$ e $C = (c_{ij})_{3\times4}$ tal que $c_{ij} = i^2 - j^2$, determine:

a) $A + B$
b) $A - B$
c) $-A + (-B)$
d) $B + A$
e) $A + C$
f) $(A + B) + C$
g) $A + (B + C)$
h) $3A - 5C$
i) $(B \cdot C^t)^t$
j) $B^t \cdot C$
l) $A \cdot B^t$
m) $(A \cdot B^t) \cdot C$
n) $A \cdot (B^t \cdot C)$
o) $-(C^t)^t$
p) $(A + C)^t$
q) $A^t + C^t$
r) $(A + B + C)^t$
s) $I_3 \cdot (A + B)$
t) $-6A^t + 4B^t - \frac{1}{2}C^t$
u) $\frac{3}{5}(A \cdot B^t) \cdot C + \frac{2}{3}(A \cdot C^t) \cdot B$
v) $(A \cdot A^t)^2$

9) Considerando as matrizes A, B e C do exercício anterior, é possível obter $A \cdot B, A \cdot C$ e $B \cdot C$? Justifique sua resposta.

10) Se M e N são matrizes quadradas de ordem 2 que comutam com a matriz $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, mostre que $M \cdot N = N \cdot M$.

11) Determine os valores desconhecidos nas sentenças abaixo:

a) $\begin{bmatrix} 2x + y & 7 \\ -7 & 3x - 2y \end{bmatrix} = \begin{bmatrix} 5 & w + z \\ 3w - 4z & 4 \end{bmatrix}$

b) $\begin{bmatrix} 2 & 5 & 3 \\ -1 & 3 & 0 \\ 4 & -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 9 \end{bmatrix}$

12) Considere uma matriz M de ordem n que não é equivalente por linhas à matriz identidade I_n. A matriz M é invertível? Justifique sua resposta.

13) Seja N uma matriz invertível, cuja a inversa é N^{-1}. A matriz N^{-1} é invertível? Caso seja, qual é a sua inversa? Caso não seja, justifique.

14) Sendo A e uma matriz quadrada de ordem n tal que B e C são suas matrizes inversas. Qual é a relação que existe entre B e C? B e C são matrizes quadradas? Qual a ordem de B?

15) Escreva a matriz $M = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 4 & 3 \\ 3 & 0 & 0 & 2 \\ 4 & 3 & 2 & -5 \end{bmatrix}$ na forma escalonada.

16) Dada a matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 1 \\ 2 & 1 & -1 \end{bmatrix}$, efetue operações elementares sobre suas linhas e verifique se A é invertível. Caso seja, explicite sua inversa.

17) Para as matrizes A, B e C, abaixo, utilizando operações elementares sobre as linhas, determine suas inversas:

a) $A = \begin{bmatrix} 2 & -3 \\ -4 & 1 \end{bmatrix}$

b) $B = \begin{bmatrix} 1 & -1 & 0 \\ 2 & -3 & 1 \\ -2 & 1 & -3 \end{bmatrix}$

c) $C = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 1 & 0 & 2 & 3 \\ 3 & 2 & 0 & 1 \\ 0 & 3 & 1 & 2 \end{bmatrix}$

18) Resolva as seguintes equações matriciais:

a) $\begin{bmatrix} 3 & -1 \\ -2 & 5 \end{bmatrix} \cdot X = \begin{bmatrix} 6 \\ -4 \end{bmatrix}$

b) $X \cdot \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = [-3 \quad 2]$

c) $\begin{bmatrix} 2 & 0 & -1 \\ -1 & 3 & 1 \\ 3 & -2 & 1 \end{bmatrix} \cdot X = \begin{bmatrix} -1 & 2 \\ 0 & -3 \\ 2 & 1 \end{bmatrix}$

19) É possível obter x e y de modo que a matriz $\begin{bmatrix} 2 & 3 \\ x & y \end{bmatrix}$ seja ortogonal?

20) Dada uma matriz A invertível, podemos afirmar que A^t é invertível? Justifique.

7.2 DETERMINANTES DE MATRIZES REAIS

21) Calcule o determinante de cada uma das matrizes a seguir e identifique se as mesmas são invertíveis ou não:

a) $\begin{bmatrix} 0 & 2 \\ -3 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 4 & 8 \\ 1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} \frac{1}{3} & -\frac{3}{4} \\ -\frac{2}{5} & \frac{4}{3} \end{bmatrix}$ d) $\begin{bmatrix} -\sqrt{2} & -\frac{3}{4} \\ \left(\frac{1}{2}\right)^{-4} & 3\sqrt{8} \end{bmatrix}$

22) Aplicando a regra de Sarrus, calcular o determinante de cada uma das seguintes matrizes:

a) $\begin{bmatrix} -1 & 0 & 3 \\ 4 & -2 & -5 \\ 0 & 3 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 6 & \frac{1}{2} \\ \frac{4}{5} & -3 & 0 \\ -2 & -\frac{1}{3} & 4 \end{bmatrix}$

c) $\begin{bmatrix} \sqrt{2} & -1 & 0 \\ \frac{1}{2} & \sqrt{3} & -\frac{3}{5} \\ 1 & -1 & \sqrt{6} \end{bmatrix}$ d) $\begin{bmatrix} a & -\frac{1}{2} & -b \\ -\frac{2}{7} & -1 & 5 \\ a & -2 & \frac{1}{3} \end{bmatrix}$

23) Aplique a triangulação de Gauss e calcule os seguintes determinantes:

a) $\begin{vmatrix} -1 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & -2 & 3 \end{vmatrix}$ b) $\begin{vmatrix} -2 & 1 & 1 & -1 \\ 1 & -3 & 2 & 1 \\ -1 & -2 & 1 & 3 \\ 0 & 2 & -2 & 1 \end{vmatrix}$ c) $\begin{vmatrix} 3 & 2 & 1 & 0 & -1 \\ 2 & -2 & 1 & -1 & 0 \\ -1 & 1 & 0 & 4 & -3 \\ 2 & -4 & 1 & -1 & 4 \\ 0 & 0 & -4 & 3 & 3 \end{vmatrix}$

24) Determine x para o qual se tenha $\begin{vmatrix} x & -2 & \frac{1}{2} \\ 2x & -3 & 2 \\ -1 & \frac{1}{5} & x \end{vmatrix} \neq 0$.

25) Determine a matriz dos cofatores de cada uma das matrizes abaixo:

a) $\begin{bmatrix} -3 & 2 \\ -\frac{1}{2} & \frac{2}{3} \end{bmatrix}$ b) $\begin{bmatrix} -1 & 2 & -3 \\ \frac{2}{3} & 5 & -1 \\ 4 & -3 & 1 \end{bmatrix}$

26) Calcule os determinantes abaixo utilizando a regra de Laplace:

a) $\begin{vmatrix} -2 & 3 & 0 & 1 \\ 2 & -1 & 3 & -4 \\ 3 & 1 & 0 & 1 \\ 0 & -5 & 2 & -3 \end{vmatrix}$ b) $\begin{vmatrix} -9 & 3 & 0 & -1 \\ 0 & 5 & 0 & 0 \\ -3 & 2 & 1 & 1 \\ 1 & 3 & -4 & 1 \end{vmatrix}$

27) Mostre que para todo $a, b, c \in \mathbb{R}$, sempre temos $\begin{vmatrix} -1 & a & 3 \\ 5 & b & -15 \\ -3 & c & 9 \end{vmatrix} = 0$.

28) Prove que $\begin{vmatrix} 1 & 1 & 1 \\ x^2 & y^2 & z^2 \\ x^3 & y^3 & z^3 \end{vmatrix} = \begin{vmatrix} yz & xz & xy \\ x & y & z \\ x^2 & y^2 & x^2 \end{vmatrix}$.

29) Quais as condições necessárias e suficientes para que um determinante seja nulo?

30) Aplicando as propriedades dos determinantes, mostre que $\begin{vmatrix} \cos(2x) & cos^2(x) & sen^2(x) \\ \cos(2y) & cos^2(y) & sen^2(y) \\ \cos(2z) & cos^2(z) & sen^2(z) \end{vmatrix} = 0$

31) Prove que o determinante $\begin{vmatrix} 1 & 5 & 4 \\ 1 & 2 & 6 \\ 1 & 8 & 2 \end{vmatrix}$ é múltiplo de 7, sem desenvolvê-lo.

32) Por qual motivo se tem $\begin{vmatrix} -2 & 4 & 1 & 2 \\ 1 & -2 & 2 & -1 \\ 3 & -1 & 1 & 7 \\ -1 & 3 & 0 & 3 \end{vmatrix} = 0$?

33) Se o determinante de uma matriz A é igual a zero, a matriz A é invertível? Justifique.

34) Dada a matriz $A = \begin{bmatrix} -3 & 1 & -1 \\ 2 & a & a \\ 5 & -3 & a \end{bmatrix}$, para quais valores de a não existe A^{-1}?

35) Dada a matriz $A = \begin{bmatrix} 3 & 3 & 1 \\ -3 & -2 & 1 \\ 2 & 0 & -5 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & -3 & 3 \\ 0 & -2 & 3 \\ -5 & 1 & 1 \end{bmatrix}$, qual a relação entre $\det(A)$ e $\det(B)$? Justifique.

36) Determine a adjunta das seguintes matrizes:

a) $A = \begin{bmatrix} 5 & -2 & -3 \\ 4 & -1 & 2 \\ 3 & 1 & 0 \end{bmatrix}$ b) $B = \begin{bmatrix} 0 & 7 & \frac{1}{3} \\ 4 & -4 & 0 \\ \frac{2}{5} & -1 & 2 \end{bmatrix}$ c) $C = \begin{bmatrix} -\frac{1}{7} & 6 & 7 \\ 13 & -8 & 2 \\ 1\frac{1}{3} & -5 & 3 \end{bmatrix}$

37) Utilizando determinantes, calcule a matriz inversa das seguintes matrizes:

a) $A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ b) $B = \begin{bmatrix} -\frac{2}{3} & \frac{1}{5} \\ \frac{3}{7} & -\frac{2}{7} \end{bmatrix}$

c) $C = \begin{bmatrix} 0 & 2 & 3 \\ -3 & 5 & 2 \\ 4 & -2 & 1 \end{bmatrix}$ d) $D = \begin{bmatrix} -6 & 3 & 1 \\ 2 & 5 & -1 \\ 4 & 0 & 0 \end{bmatrix}$

38) Sabendo que a matriz inversa de uma matriz A é $A^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{3}{7} & -2 \\ \frac{2}{5} & -\frac{3}{4} & \frac{2}{3} \\ -1 & 3 & 5 \end{bmatrix}$ e que $\det(A) = -\frac{5}{7}$, determine a matriz dos cofatores da matriz A.

39) Dadas duas matrizes A e B, quadradas de ordem n e invertíveis, mostre que:

a) $\det(A^t \cdot B^t) = \det(A) \cdot \det(B)$

b) $\det(A^t \cdot B^{-1}) = \frac{1}{\det(A^{-1}) \cdot \det(B^t)}$

40) Seja $\mathcal{M}_n$ o conjunto de todas as matrizes quadradas de ordem n, definimos o determinante como uma função de $\mathcal{M}_n$ em $\mathbb{R}$, que a cada matriz $M \in \mathcal{M}_n$ faz corresponder o número real $\det(M)$. Mostre que a função determinante não é bijetiva.

7.3 POLINÔMIOS EM $\mathbb{R}[X]$

41) Justifique por qual motivo a expressão $p(x) = 3x^2 - 2x + \frac{1}{3} - \frac{5}{x} + \frac{3}{x^2}$ não é um polinômio em $\mathbb{R}[x]$.

42) Determine a condição necessária e suficiente para que a expressão $\frac{a_0 + a_1x + a_2x^2 + a_3x^3}{b_0 + b_1x + b_2x^2 + b_3x^3}$ seja independente de x.

43) Determinar os valores reais de a, b e c, de modo que se tenha o polinômio $p(x) = (a^2 - 4)x^3 + (3b^2 - 5a)x^2 - (a + b + c)x + (ab - c)$ igual ao polinômio nulo.

44) Dados os polinômios $p(x) = (a+3)x^4 - (2a+b)x^2 + (3a-b+c)$ e $q(x) = (4a-2)x^2 + (3-5b)x^2 + (b-4c)$. Determine os valores de a, b e c para que se tenha $p(x) = q(x)$.

45) Dados os polinômios $p(x) = 2x^2 + 5x - 3$, $q(x) = 3x^3 - 4x^2 + 5x - 1$ e $g(x) = -6x^3 + 7x - 2$, calcular:

a) $p(x) + q(x)$ b) $p(x) + g(x)$ c) $q(x) + g(x)$ d) $p(x) + q(x) + g(x)$

46) Dados os polinômios $p(x) = 7x^4 - 5x^3 - x$, $q(x) = -3x^3 + 5x^2 - 12x + 8$ e $g(x) = 3x^4 + x^3 - 9x^2 + 3x - 4$, determine:

a) $p(x) - q(x)$ b) $p(x) - g(x) - q(x)$ c) $p(x) - q(x) + g(x)$ d) $g(x) - q(x) + p(x)$

47) Dados os polinômios $p(x) = -x^2 + 3x - 4$, $q(x) = 4x^3 - 2x + 1$ e $g(x) = x^3 - 3x^2 + 2x - 4$, determine:

a) $p(x) \cdot q(x)$ b) $p(x) \cdot g(x)$ c) $p(x) \cdot q(x) \cdot g(x)$ d) $[p(x)]^2 - [g(x)]^2$

48) Dados os polinômios $p(x) = x^4 - 4x^2 - 3x + 1$, $q(x) = x^3 - 3x^2 + 2$ e $g(x) = x^2 - 5x + 1$, determine os polinômios quociente e resto das seguintes divisões euclidianas:

a) $p(x) : q(x)$ b) $p(x) : g(x)$ c) $q(x) : g(x)$ d) $[q(x)]^2 : g(x)$ e) $[p(x)]^2 : [g(x)]^3$

49) Se $p(x)$ e $q(x)$ são dois polinômios não nulos de grau respectivamente m e n, qual o grau de $p(x) + q(x)$? E de $p(x) \cdot q(x)$?

50) Considerando os polinômios $p(x)$ e $q(x)$ do exercício anterior, se $m > n$, qual o grau do quociente $\frac{p(x)}{q(x)}$? Qual o grau máximo do polinômio $r(x)$ que é o resto dessa divisão?

51) Dividindo um polinômio $p(x)$ por $g(x) = x^2 + 5x - 3$, obtemos o quociente $q(x) = x^3 + 4x^2 - 3$ e o resto $r(x) = 3x + 4$. Determine o polinômio $p(x)$.

52) Determinar os números reais a e b de modo que o polinômio $p(x) = x^6 - 4$ seja divisível por $x^2 + ax + b$.

53) Mostrar que se $p(x)$ e $g(x)$ são polinômios divisíveis pelo polinômio $q(x)$, então o resto da divisão de $p(x)$ por $g(x)$ também é divisível por $q(x)$.

54) Dados $p(x)$, $g(x)$ e $q(x)$ polinômios de $\mathbb{R}[x]$ tais que $q(x)|p(x)$ e $q(x)|g(x)$, prove que $q(x)|[p(x)+g(x)]$, $q(x)|p(x)\cdot g(x)$ e $q(x)|[p(x)-g(x)]$.

55) Verificar se o polinômio $p(x)=x^4+5x^3-4x^2+3x+11$ é divisível por $x-2$ e por $x+1$. Justifique.

56) Determinar o quociente e o resto da divisão de $p(x)=x^m-\alpha^m$ por $g(x)=x-\alpha$.

57) Consideremos o polinômio de coeficientes reais $p(x)=a_0+a_1x+a_2x^2+\cdots+a_{30}x^{30}$. Se $a_0=a_1=a_2=\cdots=a_{30}$, mostre que $p(x)$ é divisível por $x-1$. Generalize.

58) Mostre que se $p(x)$ é divisível por $x-\alpha$, então $p(\alpha)=0$.

59) Seja $p(x)\in\mathbb{R}[x]$ um polinômio tal que $p(2)=1, p(1)=2$ e $p(3)=-1$. Utilizando a interpolação de Lagrange, determine $p(x)$.

60) Encontre o polinômio $p(x)$, de coeficientes reais, que atende as seguintes condições: $p(a)=b, p(b)=c,\ p(c)=d$ e $p(d)=a$.

7.4 CÓDIGOS CORRETORES DE ERROS

61) Considere um código C linear de comprimento $n=4$ sobre o corpo galoisiano $F\{0,1\}$, obtido da seguinte maneira: para cada $x_1x_2\in F^2$, obtemos o elemento $x_2x_1x_2x_1\in F^4$. Determinar a distância mínima de C e a capacidade de correção desse código.

62) Mostre que a distância de Hamming cumpre as condições necessárias para caracterizá-la como uma métrica.

63) Mostre que os códigos $C=\{00000, 00100, 01010\}$ e $C'=\{00000, 10000, 10101\}$ contidos em F^5 possuem os mesmos parâmetros porém, não são equivalentes.

64) Imagine que um braço mecânico que possua os movimentos: “para cima”, “para baixo”, “para a esquerda” e “para a direita”. Além disso, o mesmo possua uma base móvel que se desloca horizontalmente locomovendo o braço, com comandos: “para o norte”, “para o sul”, “para o leste” e “para o oeste”.
Construa um código de fonte sobre $F = \{0,1\}$ e um código de canal C para esses comandos.
C é um código perfeito? Justifique.

65) Obtenha uma matriz geradora para o código C do problema anterior.

66) Dado um código C definido sobre $F = \{0,1\}$, com matriz geradora $G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$, obtenha uma matriz geradora para C que se apresente na forma padrão $G' = [I_4|A]$, onde A é uma matriz de ordem $k \times (n-k)$. Mostre ainda que existe um código C' equivalente ao código C, tal que sua matriz geradora na forma padrão seja $G'' = [B|I_4]$, onde B é uma matriz de ordem $k \times (n-k)$.

67) Obtenha a matriz teste de paridade H do código C do exercício 64, apresentando-a na forma padrão.

68) Seja C um código sobre $F = \{0,1\}$, cuja matriz geradora seja $G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$.

a) Determine o comprimento, a dimensão e o número de elementos do código C;

b) Encontre uma matriz teste de paridade do código C e determine a sua distância mínima.

69) Seja C um código sobre $F = \{0,1\}$, com matriz teste de paridade $H = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$.

a) Obtenha a matriz teste de paridade H' na forma padrão.

b) C é um código perfeito? Justifique.

c) Determine todas as palavras do código C.

d) Determine os líderes de classes e as síndromes.

e) Decodifique as mensagens: $v_1 = 10101$, $v_2 = 11111$ e $v_3 = 10111$.

70) Considere C um código linear sobre $F = \{0,1\}$ tal que sua matriz geradora seja $G = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$.

a) Determine a dimensão, o comprimento e o número de elementos de C.

b) Construa uma matriz teste de paridade H do código C e determine o peso de C.

c) Dada a tabela de informações:

$espaço = 00000$	$E = 00001$	$J = 01100$	$P = 00011$	$U = 11001$
$A = 10000$	$F = 11000$	$L = 01010$	$Q = 11100$	$V = 01110$
$B = 01000$	$G = 10100$	$M = 01001$	$R = 10110$	$X = 00111$
$C = 00100$	$H = 10010$	$N = 00110$	$S = 10101$	$Z = 11110$
$D = 00010$	$I = 10001$	$O = 00101$	$T = 11010$	

Utilizando as informações acima, codifique as mensagens:

m_1: RUMO AO SUCESSO

m_2: MUNDO DE PAZ

71) Calcule uma tabela de líderes e síndromes referentes ao código C do problema anterior.

72) Supondo que no máximo ocorreu um erro por palavra, levando em consideração o código C do exercício 70, decodifique a mensagem:

1001010011	1010111010	1100100101	0111000001	0010110011
0010000000	0100110011	1011010011	1001010010	1010111010

73) Determine o polinômio $g(x)$ gerador de um código cíclico C sobre $F = \{0,1\}$, cujo comprimento seja $n = 9$ e a dimensão $k = 5$. A partir de $g(x)$ determine uma matriz G geradora do código C.

74) Determine o polinômio de paridade $h(x)$ do código C do problema anterior. A partir de $h(x)$, determine uma matriz verificação de paridade do código C.

75) Determine o polinômio recíproco $h^*(x)$ de $h(x)$ do problema 74. Determine uma matriz geradora do código dual $C^{\perp}$ de C. Multiplique essa matriz pela matriz G geradora do código C. O que acontece? Explique.

76) Construa dez palavras na forma sistemática e na forma não sistemática de um código cíclico C sobre $F = \{0,1\}$, cujo polinômio gerador é $g(x) = 1 + x^2 + x^3$.

77) Considere o código cíclico C sobre $F = \{0,1\}$, de comprimento $n = 15$ e dimensão $k = 11$, cujo polinômio gerador é $g(x) = 1 + x + x^4$.

a) Determine a matriz G geradora de C.

b) Determine o polinômio de paridade do C.

c) Determine a matriz teste de paridade H do código C.

d) Determine o polinômio gerador do código dual de C.

e) Coloque as matrizes G e H na forma padrão.

78) Supondo que $m = 10110110010$ é uma palavra do código de fonte, escreva essa palavra na forma polinomial. Utilize o código C do exercício anterior para obter um código de canal para m. Suponha que $c \in C$ é uma palavra, cuja forma polinomial é $c(x)$, tenha sido enviada e, por algum motivo a palavra recebida tenha sido $r(x) = 1 + x^3 + x^4 + x^9 + x^{10} + x^{12}$. Sabendo que ocorreu um único erro nessa transmissão, determine as síndromes de $r(x)$ e de todos os seus desvios cíclicos e decodifique $r(x)$, corrigindo o erro e determinando $c(x)$.

79) Forme grupos com seus alunos. Ofereça mensagens curtas, de no máximo três palavras da língua portuguesa. Proponha a cada grupo que crie códigos lineares de comprimentos e dimensões diferentes e codifique cada um dos caracteres das mensagens propostas. Compare os resultados. Teste a capacidade de detecção e correção de cada código e eleja o mais eficiente. Dê bastante ênfase a cada uma das operações e propriedades matriciais utilizadas.

80) Forme grupos com seus alunos. Ofereça mensagens curtas, de no máximo três palavras da língua portuguesa. Proponha a cada grupo que crie códigos cíclicos de comprimentos e dimensões diferentes e codifique cada um dos caracteres das mensagens propostas dando preferência para a forma polinomial de codificação. Compare os resultados. Teste a capacidade de detecção e correção de cada código e eleja o mais eficiente. Dê bastante ênfase a cada uma das operações e propriedades polinomiais utilizadas.

81) Forme grupos com seus alunos. Ofereça mensagens a serem decodificadas, que contenham erros, solicite que façam as correções e decodifiquem as mensagens.

CONSIDERAÇÕES FINAIS

O ensino/aprendizagem de matemática no Brasil, ao longo dos últimos anos, tem passado por várias transformações na busca pelo oferecimento de uma educação com melhor qualidade aos estudantes do ensino básico.

Índices como o IDEB mostram um discreto progresso na educação, porém, a passos lentos.

Temos observado que o distanciamento entre matemática praticada nas escolas e as experiências vivenciadas pelos alunos em ambiente extraescolar pode contribuir de forma negativa para o alcance dos objetivos da educação. A contextualização da matemática consiste em uma proposta pedagógica que busca estreitar essa lacuna, apresentando uma matemática não mais vista como uma mera disciplina escolar, mas como uma necessidade humana de interpretar o mundo, adaptar-se e interagir com ele.

Nas relações sociais, a matemática também se faz presente e, portanto, o domínio dos seus conceitos e da sua linguagem é de fundamental importância nessas relações. O sujeito Cidadão participa do meio social ao qual está inserido e para tanto, deve ter o conhecimento matemático necessário para sua atuação.

O material apresentado nesse trabalho não tem como objetivo formar especialistas em teoria dos códigos corretores de erros e sim usar essa teoria como um fator motivacional ao aluno que, ao conhecê-la, poderá perceber o quão útil é a matemática. É de extrema importância que ao ser utilizada a teoria dos códigos corretores de erros como uma estratégia para o ensino de matrizes, determinantes e polinômios, seja dada a devida atenção a cada uma das propriedades e operações utilizadas, pois, nessa teoria, frequentemente é trabalhado conceitos como matriz inversa, matriz identidade, transposição de matrizes, operações elementares sobre linhas, determinantes, produto de matrizes, operações e propriedades dos polinômios, interpolação de Lagrange, enfim, uma vastidão de operações e propriedades, que quando bem evidenciadas durante o processo de aplicação na teoria dos códigos, poderá construir um aprendizado consolidado desses conceitos.

Achamos importante ainda mencionar que o uso da teoria dos códigos corretores de erros, por si só, não consolidará a aprendizagem dos conceitos utilizados se for apresentada aos alunos meramente como mais um conteúdo a ser estudado. É importante ao profissional docente que, ao apresentar essa teoria aos seus alunos, utilize estratégias com ênfase na prática, sejam oficinas, dinâmicas ou qualquer recurso em que os alunos tenham a oportunidade de manipular esses conceitos na codificação e decodificação de mensagens, de

modo que a sua participação, não apenas como expectador, seja de fundamental importância no decorrer desse processo.

Por fim, além esse trabalho apropriar-se da teoria dos códigos corretores de erros, na tentativa de contextualizar a matemática ensinada nas séries finais do ensino médio, tem também como proposta, fomentar a discussão sobre teoria e prática, abstrato e concreto, escolar e extraescolar, com intuito de incentivar o gosto pela matemática com toda sua abstração e rigor.

REFERÊNCIAS

BAHIA, Flaviano. Um primeiro curso sobre códigos corretores de erros. ERMAC 2010: I Encontro Regional de Matemática Aplicada e Computacional, 2010. Disponível em: <http://www.ufsj.edu.br/portal2-repositorio/File/i-ermac/anais/minicursos/mc8.pdf> acesso em 13 de setembro de 2014.

BARBOSA, Tauan de S.; ASSIS, Aline M. Princípio teóricos dos códigos corretores de erros: códigos lineares e cíclicos. Disponível em: <seer.ucg.br/index.php/estudos/article/download/3364/1951> acesso em 21 de outubro de 2014.

BOLDRINI, José Luiz [et al]. **Álgebra linear**. 2. ed. São Paulo: Harper & Row do Brasil, 1980.

BRASIL, Secretaria do Ensino Médio. Parâmetros Curriculares Nacionais. Ensino Médio. Brasília: MEC/SEM, 2002.

CALLIOLI, Carlos A.; COSTA, Roberto C. F.; DOMINGUES, Hygino H. **Álgebra linear e aplicações**. 7. ed. São Paulo: Atual, 1990.

COLOMBO, Jones. Códigos cíclicos: códigos BCH. Disponível em <http://www.professores.uff.br/jcolombo/artigos/codigosCiclicosBCH.pdf> acesso em 06 de agosto de 2014.

D'Ambrosio, Ubiratan. **Educação matemática:** Da teoria à prática. 23 ed. Campinas, SP: Papirus, 2012.

DOMINGUES, Hygino H.; IEZZI, Gelson. **Álgebra moderna**. 3. ed. São Paulo: Atual, 1982.

FOGAÇA, Jennifer. **Contextualização**. Disponível em: <http://educador.brasilescola.com/trabalho-docente/contextualizacao.html>. Acesso em: 02 de outubro de 2014.

GARCIA, Arnaldo; LEQUAIN, Yves. **Elementos de álgebra**. 5. ed. Rio de Janeiro: IMPA, 2010.

GONÇALVES, Adilson. **Introdução à álgebra**. 5. ed. Rio de Janeiro: IMPA, 2009.

GONZÁLEZ, Mario Enrique Duarte. Monografia – Códigos cíclicos, anéis e corpos. Disponível em: <http://www.ime.unicamp.br/~ftorres/ENSINO/MONOGRAFIAS/mono2_mario.pdf> acesso em 28 de agosto de 2014.

HENRIQUES, Ana Cláudia C. B. **O pensamento matemático avançado e a aprendizagem da análise numérica num contexto de actividades de investigação**. 2010. Tese (Doutorado em Educação - Didáctica da Matemática). Instituto de Educação, Universidade de Lisboa, Lisboa (Portugal).

HERNÁNDEZ, Cruz Enrique Borges. Códigos cíclicos. Disponível em: <http://paginaspersonales.deusto.es/cruz.borges/Papers/05Codigos.pdf> acesso em 21 de setembro de 2014.

HOFFMAN, Kenneth; KUNZE, Ray. **Álgebra linear**. Trad. Renate Watanabe. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S. A., 1979.

HEFEZ, Abramo; FERNANDEZ, Cecília S. **Introdução à álgebra linear**. 1. ed. Rio de Janeiro: Sociedade Brasileira de Matemática, 2012.

HEFEZ, Abramo; VILLELA, Maria Lúcia T. **Códigos corretores de erros**. 2. ed. Rio de Janeiro: IMPA, 2008.

———. **Polinômios e equações algébricas**. 1. ed. Rio de Janeiro: Sociedade Brasileira de Matemática, 2012.

LIMA, Elon Lages. **Álgebra linear**. 5. ed. Rio de Janeiro: IMPA, 2001.

LIPSCHUTS, Seymour. **Álgebra linear**. Trad. Roberto Ribeiro Baldino. 2. ed. Rio de Janeiro: McGRAW-HILL do Brasil LTDA, 1978.

MENEGHESSO, Carla. Códigos corretores de erros. Disponível em: <http://www.dm.ufscar.br/dm/attachments/article/5/monografia_carla%20TCC.pdf> acesso em 21 de agosto de 2014.

MEYER, João Frederico da Costa de Azevedo (org.); CALDEIRA, Ademir Donizeti; MALHEIROS, Ana Paula dos Santos. **Modelagem em Educação Matemática.** Belo Horizonte: Autêntica Editora, 2011.

MILIES, César Polcino. Breve introdução à teoria dos códigos corretores de erros. Disponível em: < http://www.sbm.org.br/docs/coloquios/CO-1-09.pdf> acesso em 07 de outubro de 2014.

PIAGET, Jean. **Seis estudos de psicologia.** Trad. Maria Alice Magalhães D'Amorim e Paulo Sergio Lima Silva. 21. ed. Rio de Janeiro: Forense Universitária, 1995.

POZO, Juan Ignacio e ECHEVERRÍA, María Del Pui Pérez. **Aprender a resolver problemas e resolver problemas para aprender**. Porto Alegre: Artes Médicas, 1988.

SOUZA, Mário José. Códigos corretores de erros. Disponível em: <http://semanadoime.mat.ufg.br/up/34/o/min_mario.pdf> acesso em 02 de outubro de 2014.

VOLOCH, José Felipe. Códigos corretores de erros. Disponível em <http://www.impa.br/opencms/pt/biblioteca/cbm/16CBM/16_CBM_87_06.pdf> acesso em 28 de agosto de 2014.