

UNIVERSIDADE FEDERAL DE SANTA CATARINA

CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS

DEPARTAMENTO DE MATEMÁTICA

Curso de Especialização em Matemática

Formação de Professor na Modalidade a Distância

CONTROLE E ESTABILIZAÇÃO EM DIMENSÃO FINITA

VANESSA OECHSLER

FLORIANÓPOLIS - SC

março de 2011



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
Departamento de Matemática
Curso de Especialização em Matemática-Formação de Professor na modalidade a distância

"Controle e Estabilização em Dimensão Finita"

Monografia submetida à Comissão de avaliação do Curso de Especialização em Matemática-Formação do professor em cumprimento parcial para a obtenção do título de Especialista em Matemática.

APROVADA PELA COMISSÃO EXAMINADORA em 04/03/2011

Dr. Felix Pedro Quispe Gomez (CFM/UFSC - Orientador)

Dr. Marcio Rodolfo Fernandes (CFM/UFSC - Examinador)

Dr. Sergio Eduardo Michelin (CFM/UFSC - Examinador)

Dra. Neri Terezinha Both Carvalho

Coordenadora do Curso de Especialização em Matemática-Formação de Professor

Florianópolis, Santa Catarina, março de 2011.

AGRADECIMENTOS

Ao meu orientador, professor Félix Pedro Q. Gómez, pela paciência, compreensão e pelas grandes contribuições dadas ao trabalho.

À coordenação da especialização e do pólo de Blumenau por atenderam prontamente às nossas solicitações, sanando nossas dúvidas e tornando essa caminhada mais fácil.

À tutora da especialização Carla Peres Souza pela amizade e pelas dicas durante essa caminhada.

Aos colegas da especialização do pólo de Blumenau, Orêncio e Hiandra, pelos momentos de estudo e dedicação que passamos juntos, bem como pelas palavras de incentivo nos momentos difíceis.

À professora Rosinéte Gaertner, professora da FURB e grande amiga por sua paciência e conselhos durante essa jornada.

À professora Andresa Pescador, professora da UDESC, pela disponibilidade em me auxiliar em algumas dúvidas que surgiram ao longo do caminho.

À amiga Adriana Kuehn, pelo incentivo dado para que eu iniciasse essa especialização.

Aos meus pais, Carlos e Svea, por estarem sempre ao meu lado, dando apoio e incentivando para que eu seguisse em frente.

Aos meus familiares, vó Siny, vó Tereza, vô Vilídio, tia Marisa, tio Ronaldo e primo Darlan pela compreensão nos momentos em que estive ausente nesse período de estudos.

SUMÁRIO

INTRODUÇÃO.....	1
CAPÍTULO I.....	2
O PROBLEMA DE CONTROLE E ESTABILIZA-ÇÃO EM DIMENSÃO FINITA	2
CAPÍTULO II.....	13
PRELIMINARES E NOTAÇÕES	13
2.1. Equação Diferencial	13
2.2. Noções de Análise Matricial	38
2.3. Independência e Dependência Linear de Vetores.....	47
2.4. Teoria Espectral Elementar	48
CAPÍTULO III	80
CONTROLE EXATO DE SISTEMAS DE DIMENSÃO FINITA	80
3.1. Sistemas de Controle	80
3.2. O ambiente da Teoria do Controle	84
3.3. Os métodos da Teoria de Controle	84
3.4. Controle de Tempo Mínimo para um Trem sem Atrito	86
3.5. O problema de Controle de um Sistema Regido por EDOs	94
3.6. Controle de Sistemas Lineares.	96
3.7. Teorema de Existência e Unicidade.	96
3.8. Existência e Unicidade de Soluções em espaços L^2	97
3.9. Propriedade da Observabilidade.....	102
3.10. Condições de Kalman.....	107
3.11. Estabilização de Sistemas Lineares de Dimensão Finita.....	112
CONSIDERAÇÕES FINAIS	119
REFERÊNCIAS	120

LISTA DE FIGURAS

Figura 1: Processo com entrada previsível com certas restrições. (Fonte: GUZMAN, 1980, p.239).....	83
Figura 2: Trajetórias do trem (Fonte: GUZMAN, 1980, p.245).....	89
Figura 3: Gráfico do controle no tempo ótimo (Fonte: GUZMAN, 1980, p.247).....	91
Figura 4: Trajetórias da parábola (do autor).....	92
Figura 5: Trajetórias da parábola no sentido negativo (do autor).....	93
Figura 6: Solução para o problema do trem (Fonte: GUZMAN, 1980, p.248).....	94

RESUMO

O presente trabalho aborda a Teoria do Controle e sua Estabilização em um sistema de Dimensão Finita. Essa Teoria é uma área da matemática que estuda o controle de sistemas de modo que, através do controle de um objeto, pode-se influenciar o seu comportamento para se atingir um determinado objetivo. O estudo desse ramo é relativamente recente, mas a natureza faz o uso do mecanismo de controle desde a época da criação do Universo. Atualmente, campos como robótica, lógica de computador, bio-engenharia fazem o uso dessa Teoria para criar e controlar seus objetos.

INTRODUÇÃO

A Teoria do Controle é uma área da matemática que estuda o controle de sistemas. Para efetuar o controle de objetos, é preciso intervir no seu comportamento de modo que se consiga chegar ao resultado esperado. E, para que isso ocorra, os engenheiros são obrigados a construir dispositivos que incorporam diversas técnicas matemáticas. Cita-se como exemplo desse dispositivo o motor a vapor, projetado durante a Segunda Guerra Mundial e processadores sofisticados encontrados nos artigos de consumo da atualidade, como CD players, automóveis e até em robôs industriais e piloto automático de aeronaves.

O trabalho aborda essa teoria e sua estabilização em um sistema de dimensão finita. No primeiro capítulo, faz-se uma breve retrospectiva do surgimento da Teoria do Controle, desde a época da criação do Universo até os desafios para os dias atuais.

No capítulo II são explorados alguns conteúdos matemáticos que são utilizados para a resolução dos problemas da Teoria do Controle, como Equações Diferenciais Ordinárias e, dentre elas, a equação diferencial linear de 1ª ordem não-homogênea e a equação diferencial linear de 2ª ordem não-homogênea com coeficientes constantes; a Álgebra Linear e, em seus assuntos, a álgebra de matrizes, o escalonamento de matrizes, o posto de uma matriz a independência e dependência linear de vetores, o determinante e a Teoria Espectral.

Embasado nas definições e resoluções do capítulo anterior, o terceiro capítulo apresenta um problema típico da teoria do controle: um trem de massa m que se encontra em um ponto A de uma via retilínea e vai de A para B , com velocidade v . Deseja-se que o trem esteja em repouso no ponto de chegada B no menor tempo possível. Dessa forma, tem-se a seguinte pergunta: qual a estratégia a seguir, ou seja, qual deve ser a força u , de aceleração ou de frenagem que se deve aplicar em cada ponto da trajetória? Para responder à essa questão, no terceiro capítulo são apresentados os cálculos matemáticos utilizados para a resolução deste problema.

CAPÍTULO I

O PROBLEMA DE CONTROLE E ESTABILIZAÇÃO EM DIMENSÃO FINITA

A Teoria do Controle é uma área da matemática que estuda o controle de sistemas. Para efetuar o controle de um objeto, é necessário influenciar o seu comportamento de forma a atingir uma meta desejada. E, para que isso ocorra, os engenheiros são obrigados a construir dispositivos que incorporam diversas técnicas matemáticas. Pode-se citar, como exemplo de dispositivo, o motor a vapor, projetado durante a Segunda Guerra Mundial e processadores sofisticados encontrados nos artigos de consumo da atualidade, como CD players, automóveis e até em robôs industriais e piloto automático de aeronaves.

Uma breve história de controle

De acordo com Sontag [16] (1998) mecanismos de controle são amplamente distribuídos na natureza, sendo utilizados por organismos vivos para manter a temperatura do corpo, os níveis de açúcar no sangue. Ou seja, na natureza, o controle é utilizado para manter as variáveis essenciais para o funcionamento de um ser vivo em equilíbrio.

Além de mecanismos da natureza, encontram-se mecanismos de controle elaborados pelo homem. Por exemplo, os romanos utilizavam o conhecimento de controle para manter o nível de água constante nos aquedutos, através de uso de várias combinações de válvulas.

Dessa forma, a história dos mecanismos de controle, tanto naturais quanto os elaborados pelo homem, pode ser dividida em diferentes épocas da história da humanidade.

CRIAÇÃO DO UNIVERSO

Acredita-se que o universo físico tenha se originado a partir da explosão do Big Bang. Do ponto de vista da Teoria do Controle, a primeira questão a ser levantada sobre a Teoria do Big Bang é:

“Foi a criação de um comando de circuito aberto iniciando a resposta em um tempo específico, em $t = 0$?¹” MARKUS [14] (1993) (tradução nossa)

ou

“Este foi um comando de circuito fechado para desencadear uma ação em um determinado estado do universo, como um quantum-mecânico em um vácuo negativo?²” MARKUS [14] (1993) (tradução nossa)

Essas respostas são muito difíceis de ser respondidas, pois a criação do universo é um experimento muito difícil de repetir e o espaço é bastante complexo para se analisar.

ERA DO CONTROLE PRÉ-HUMANO

Nos primeiros segundos após o Big Bang, a expansão da esfera de energia se tornou menos densa e quente até a radiação começar a condensar em prótons, elétrons e partículas de matéria elementar, como átomos de hidrogênio simples. Esses átomos gravitaram em enormes bolas de gás com pressões e temperaturas centrais tão elevadas que os átomos de hidrogênio se fundiram e produziram o gás hélio em um incêndio nuclear. Surge, nesse momento, a estrela e monta-se o palco para o segundo grande experimento do controle.

Uma estrela é um sistema dinâmico instável, desde que a atração gravitacional dentro das camadas de superfície da estrela se intensifica à medida que diminui o raio. E, portanto, a gravidade atua como realimentação positiva. Mas o raio da estrela pode ser estabilizado através da pressão da radiação externa gerada pela fusão nuclear no centro.

Questiona-se: o que acontece com a estrela, quando todo o hidrogênio disponível no centro foi convertido em hélio? Por que então o equilíbrio falha e os colapsos no raio intensificam a pressão central até que o hélio começa a se fundir? Isto estabelece um novo equilíbrio entre a gravidade e a radiação. Mas o que acontece quando o hélio se esgota? Por que então os equilíbrios começam a falhar até que a estrela colapsa para o ponto onde superfusões de elementos maiores ocorrem causando uma explosão, que fragmenta a estrela e a leva a ejetar os restos de carbono, oxigênio, nitrogênio, fósforo e ferro - os elementos da vida. Dessa forma, uma estrela é uma falha no controle do sistema.

¹ “Was the creation an open-loop command initiating the response at a specified time-say, at $t = 0$?”

² “Was this a closed-loop command set to trigger action at a specified state of the universe-say, at a quantum-mechanical negative-vacuum?”

Durante a criação do universo, outro fenômeno de controle natural pôde ser observado: considere o problema de controle de estabilizar a relação entre a densidade relativa de CO₂ e O₂ atmosfera. Se a densidade de CO₂ aumentar, então a folhagem das florestas tropicais e o plâncton dos oceanos da Antártida crescem, resultando no efeito estufa. Estas plantas crescem, então, com a energia decorrente de fotossíntese, consumindo o CO₂ e liberando O₂. Por outro lado, se a densidade de O₂ aumentar, então os incêndios florestais aumentam e mais CO₂ é produzido. Por mecanismos de realimentação regulador como este, o equilíbrio atmosférico foi mantido quase estável por quase um bilhão de anos.

Equívocos populares destas interações bioquímicas levaram à crença de que a “Natureza” tem consciência ou vontade inteligente, mais do que o reconhecimento do fenômeno da teoria do sistema de realimentação.

ERA DO CONTROLE ARTESÃO E TÉCNICO

Desde os tempos mais antigos, artesãos e técnicos têm habilmente empregado métodos de controle no refino de minério, produzindo vidro, e também na regulação dos processos de irrigação em larga escala, através da utilização de rodas d'água, bombas de parafuso, e barragens.

Não está inteiramente claro que esses métodos se encaixam no quadro de controle automático, mas há muitos exemplos individuais de autenticidade. Registros ao longo dos séculos descrevem a perfeição do cálculo de flutuação, para regular o nível de água em um tanque, pelos gregos; o desenvolvimento dos relógios de água no Egito e na China; e da introdução de reguladores de temperatura do forno por C. Drebbel, em 1600 d.C.

No início da revolução industrial, com os grandes avanços de instrumentação, Christian Huygens aperfeiçoou o relógio de pêndulo, trabalhando com o problema do controle da velocidade. O desenvolvimento de relógios precisos a partir de Huygens e seu contemporâneo Hooke serviu para atender às necessidades da navegação, que precisavam de relógios precisos para determinar o tempo solar. Em 1763, John Harisson alcançou a perfeição do cronômetro, com um complexo design de compensador de temperatura e movimento. Mas, certamente, a aplicação de Watt do regulador centrífugo para motores a vapor pode ser considerado como o início da moderna engenharia de controle.

INÍCIO DA ERA DE CONTROLE CIENTÍFICO (1850 – 1900)

Em 1868, J. C. Maxwell publicou uma análise matemática do regulador mecânico de Watt. Em seu famoso escrito “Dos Governadores³” Maxwell aplicou os princípios mecânicos de Newton para explicar a ação de um governador por equações lineares diferenciais com coeficientes constantes. Ele reconheceu que uma condição necessária e suficiente para a estabilidade era que o polinômio característico tivesse raízes somente no quarto esquerdo do plano complexo. Depois de resolver a questão da estabilidade para equações do terceiro grau, ele então declarou: “Eu não tenho sido completamente capaz de determinar essas condições para equações de grau maior que três, mas eu espero que o conteúdo obtenha a atenção dos matemáticos”.

Para polinômios reais, o problema da estabilidade foi atacado por C. Hermite e totalmente resolvido no famoso escrito de A. Hurwitz, em 1894. As aplicações para a engenharia foram enfatizadas por J. Vishnegradskii, em 1876.

E, aproximadamente um século depois, ainda para equações diferenciais parciais e equações diferenciais funcionais usa-se o espaço Hardy, a teoria Paley-Wiener e matrizes sobre anéis Noetherianos.

As pesquisas em dinâmica, especialmente na teoria da estabilidade, incluindo o trabalho de Routh (1877), são bem relatados nos dois volumes da “Teoria do Som” de Baron Rayleigh (1894).

O outro grande tratado desta era nos sistemas de estabilidade dinâmica era o volume “Problema geral da estabilidade do movimento” de A. Lyapunov. Neste trabalho, o autor introduziu dois métodos para analisar a estabilidade de um sistema diferencial de vetores não-lineares. O primeiro método consistia na linearização e se relaciona com um comportamento perto do equilíbrio. O segundo método depende da função de dissipação de energia e gera um resultado global. Estes dois métodos de análise de estabilidade geraram diversas investigações no século seguinte.

ERA DO CONTROLE CLÁSSICO DA ENGENHARIA (1900 – 1940)

³ “On Governors”

O problema central da engenharia de controle na primeira metade do século XX foi a reprodução fiel e amplificação de um sinal em um cabo de telefone. Esta ampliação seguiu a invenção de Forest do tubo tríodo e foi essencial para a comunicação de longa distância, e os problemas levantados e resolvidos levaram aos métodos clássicos de engenharia de controle.

Nas primeiras décadas do século XX, O. Heaviside e C. Steinmetz foram pioneiros no caminho da algebrização de sistemas diferenciais lineares por meio do cálculo do operador e análise complexa. Embora esses métodos tenham sido conhecidos no contexto matemático desde os tempos de Lagrange e Laplace, eles foram agora avaliados no formato que podem ser utilizados pelos engenheiros elétricos.

Mas os descobrimentos mais importantes foram na década de 30, em que as pesquisas na teoria do controle se destacaram no Laboratório de Telefones Bell, onde foi desenvolvida uma teoria dos amplificadores, com base na garantia de estabilidade e de resposta adequada para circuitos elétricos. Este trabalho, desenvolvido por H. Nyquist, H.W. Bode, entre outros, constitui ainda a base para a concepção de frequência.

Enquanto esse trabalho era desenvolvido, outro trabalho, menos conhecido era elaborado. Este outro trabalho tinha grandes consequências, como explicado por Nier, professor de Física da Universidade de Minnesota. Em 1935, Nier publicou um pequeno texto sobre a operação do espectrômetro de massa, no qual ele descreve o problema:

“Em praticamente todos os métodos de análise de raios positivos de uma combinação de um campo elétrico e magnético é usado para separar os íons de diferentes massas... O fato de que a deformação sofrida por um íon depende tanto do campo magnético quanto de um campo elétrico, sugere que ao invés de tentar manter o campo magnético constante, pode provocar o campo elétrico a oscilar automaticamente junto com o campo magnético de tal maneira que a deflexão de um íon seria independente das flutuações do campo magnético.”⁴

MARKUS [14] (1993) (tradução nossa)

Esta importante descoberta de Nier fez a massa de espectrômetro um instrumento para orientar a separação de Urânio 235 em 238. Era a chave para o todo o futuro desenvolvimento na energia atômica e pesquisas em física nuclear.

⁴ In practically all methods of positive ray analysis a combination of an electric and a magnetic field is used to separate the ions of different masses... The fact that the deflection suffered by an ion depends upon both the magnetic field and the electric field suggests that instead of attempting to keep the magnetic field constant one could cause the electric field to fluctuate automatically along with the magnetic field in such a manner that the deflection of an ion would be independent of the magnetic field fluctuations”

Em resumo, a era do controle clássico era preocupada com a estabilização dos circuitos de compensação de realimentação, que foram descritas matematicamente em termos de equações diferenciais ordinárias lineares com coeficientes constantes. Algumas tentativas foram feitas para estender estes métodos para equações com coeficientes variáveis no tempo, como acontece com a descrição da função de Kochenburger, e para permitir algum atraso estocástico ruído e do sinal.

INOVAÇÕES DO CONTROLE NA ERA DA SEGUNDA GUERRA MUNDIAL

Durante a Segunda Guerra Mundial, o principal estímulo para a engenharia de controle surgiu a partir do desenvolvimento de novas tecnologias de armamento, especialmente na aeronáutica e marinha. Estas construções, de caças a jato e de aviões de bombardeio com mísseis de longo alcance, aviões, bem como radar e detecção eletrônica e navegação na defesa, exigiram inovações de alto nível da ciência e da engenharia.

Nos Estados Unidos, muitas destas pesquisas eram conduzidas no MIT, particularmente no Laboratório de Radiação, que tinha sua principal preocupação na perfeição do radar. O problema do controle era usar os dados do radar como realimentação para orientar o acompanhamento de arma de aeronaves inimigas. Dessa forma, houve um grande esforço no desenvolvimento da teoria do controle. Novas e sofisticadas teorias matemáticas envolvendo processos estocásticos e métodos estatísticos de análise de funções ganharam destaque, lideradas pelas pesquisas de Norbert Wiener.

Na União Soviética, a teoria das vibrações e oscilações não-lineares avançaram com Krylow e Bogoliubov, e seus seguidores, enquanto grupos fortes de engenheiros teóricos eram liderados por A.M. Letov, A. I. Luré e muitos outros. A teoria da previsão de Kolmogorov foi desenvolvida independentemente da de Wiener.

ERA DO CONTROLE MATEMÁTICO (1950 – 1970)

Esta parte da breve história do controle é relatada por Lawrence Markus. Lawrence Markus nasceu em 1922, na cidade de North Hibbing, no estado de Minnesota, Estados

Unidos. Anos depois, a cidade foi tragada por uma mina de ferro a céu aberto, o que levou a família de Markus a se mudar para Minneapolis e, mais tarde, para Chicago.

Lawrence, conhecido como Larry, ingressou na faculdade de Chicago aos 16 anos, tendo estudado matemática, física e astronomia.

Com o ingresso da América na Segunda Guerra Mundial, Larry foi impelido a estudar Meteorologia, tendo se tornado mestre na área. E, aos dezenove anos de idade, Markus tornou-se Instrutor de Meteorologia da Universidade de Chicago. Após um curto período de tempo com o Projeto Atômico Manhattan, Larry ofereceu-se como voluntário para a marinha dos Estados Unidos e serviu vários anos a bordo de uma fragata do tempo no Atlântico Norte.

Após a guerra, Larry foi para Harvard, onde completou o seu PHD em Matemática indo, em seguida, para Paris, onde conseguiu uma bolsa. Markus retornou, depois, à faculdade de Harvard como instrutor e, mais tarde, tornou-se instrutor de Yale, conferencista em Princeton.

Em 1957, Larry juntou-se à Escola de Matemáticos na Universidade de Minnesota onde, rapidamente tornou-se professor (em 1960) e, mais tarde, em 1980, professor regente.

Lawrence era um matemático incomum, tendo um foco na aplicação matemática na dinâmica e na teoria do controle, outro nos aspectos teóricos da geometria e topologia e foco também na álgebra abstrata. Seus cursos de graduação em estruturas algébricas e geométricas do controle dinâmico eram um escândalo ou uma fonte de inspiração, dependendo dos alunos. Com um estilo de aula divertido, ele recebeu o Prêmio de Ilustre Professor da Universidade, em 1968.

Larry é um grande pesquisador, com 130 publicações científicas, incluindo o famoso tratado “Fundamentos da Teoria do Controle Ótimo”, escrito durante um torneio na Suíça, além de outras 15 monografias e coleções editoriais.

A partir de agora, é relatada a história da era do controle matemático de acordo com Markus:

Após a guerra, Solomon Lefschetz, um famoso topologista mundial e ex-presidente do AMS, alegou que a América deveria ter uma política nacional de incentivo aos campos da matemática. Com o encorajamento do Escritório de Pesquisa Naval, Lefschetz estabeleceu um centro de pesquisa em equações diferenciais ordinárias não-lineares e dinâmicas na Universidade de Princeton, em paralelo com o Instituto Courant em NYU, especializado em equações diferenciais parciais. Lefschetz foi o principal interessado na teoria da estabilidade e das oscilações não-lineares. Depois que Lefschetz se aposentou em Princeton, ele transferiu

sua pesquisa para o RIAS (Instituto de Pesquisa da Corporação Martin) e mais tarde, para a Universidade Brown, onde se tornou o Centro Lefschetz para Sistemas Dinâmicos, mais tarde na direção de Joseph La Salle, com a cooperação de Wendell Fleming E., Harols Kushner e outros.

Larry Markus tornou-se membro do grupo em Princeton e ajudou Lefschetz na organização de um de seus seminários. Muitos outros matemáticos participaram de seus programas, como R. Bellman, L. Cesari, J. La Salle, A. Antosiewicz, R. Bass, S. Diliberto, R. Kalman, J. Moser e S. Smale.

Lefschetz nasceu em Moscou em 1884, mas começou sua carreira profissional como um engenheiro graduado na Ecole Centrale de Paris. Mais tarde, ele foi aos Estados Unidos como um engenheiro industrial. Mas em um trágico acidente engenheiro, ele perdeu seus dois braços, e posteriormente, usou membros artificiais com grande habilidade em sua vida diária e até mesmo em seus escritos e palestras. Suas descobertas em topologia e geometria algébrica eram o ponto alto da matemática do século XX.

Muitos dos matemáticos e engenheiros teóricos de Pinceton-RIAS-Brown estabeleceram seus próprios centros de pesquisa. Na Universidade de Minnesota, Larry Markus e Bruce Lee organizaram o Centro de Controle de Ciência e Sistemas Dinâmicos, cuja cooperação culminou no tratado “Fundações da Teoria do Controle Ótimo”, que foi amplamente utilizado e depois traduzido para o Russo e o Japonês.

Em 1968, Larry fez uma associação com a Universidade de Warwick, na Inglaterra, onde estabeleceu e dirigiu o Centro de Teoria do Controle.

A lista de professores e visitantes destes dois centros incorporaram a maioria dos cientistas de controle e líderes mundiais nas respectivas áreas de matemática e engenharia. As atividades de pesquisa enfatizavam as equações diferenciais ordinárias não-lineares (especialmente os aspectos de geometria e álgebra da teoria do controle). Bruce Lee tornou-se uma autoridade mundial com controle de sistemas.

Enquanto isso, na União Soviética, no Instituto Steklov, em Moscou, L. S. Pontryagin estabeleceu uma contração semelhante de jovens matemáticos com grande interesse pelas equações diferenciais não-lineares. Estes incluem seus colegas vencedores do prêmio Lênin, V. Boltyanski, R. Gamkrelidze, e E. Mischenko, que foram co-autores do famoso livro “A Teoria Matemática dos Processos Ótimos”.

Como Lefschetz, Pontyagin era famoso mundialmente como um topólogo algébrico, e como o criador da teoria da dualidade para grupos topológicos. Ele também sofreu um

acidente precoce que o deixou cego. Embora não fosse muito conhecido pela criação matemática, Pontryagin havia muito tempo perseguiu uma carreira criativa na matemática aplicada. Por exemplo, Pontryagin generalizou o critério de estabilidade Hurwitz para cobrir as equações diferenciais lineares, que tinham funções polinomiais exponenciais características. Também foi ele quem introduziu o conceito de estabilidade estrutural de sistemas dinâmicos.

Outros importantes centros de pesquisa na União Soviética se desenvolveram ao redor de Mitropolsky em Kiev, Krasovskii e Mishkis em Sverdlovsk, tão bem quanto outras instituições científicas poderosas em Leningrado e Moscou.

A pesquisa na teoria de controle também foi desenvolvida nas Universidades de Oxford, Cambridge, Warwick, Manchester e o Imperial College, em Londres. Outros institutos também se dedicaram à pesquisa do controle no resto da Europa, Japão e América do Sul.

Para Markus, a ciência do controle está expandindo em novas áreas:

- i.) Sistema de parâmetros distribuídos, baseados nos tratados revolucionários de A. G. Butkovskiy e J.- L. Lions.
- i.) Sistemas diferenciais não-lineares, envolvendo novos métodos de álgebra e grupos.
- ii.) Teoria dos sistemas lineares fundamentados nas concepções de controlabilidade, observabilidade e identificação. Aqui a interação entre o domínio do tempo e da frequência de domínio era explorada pela álgebra abstrata, com a liderança assumida por Rudy Kalman e H. Rosenbrock.

UMA HISTÓRIA DE DUAS ESTARRECEDORAS SURPRESAS

No Congresso Internacional de Matemáticos em Edimburgo, em 1958, Pontryagin apresentou uma resposta aos principais convidados: “Regulação dos Processos Ótimos”. Naquele tempo, os russos no Ocidente foram fenômenos exóticos, e a emoção da ocasião foi inflada pela fama e eminência matemática de Pontryagin - o topólogo que era renomado por suas pesquisas em cosmologia e pela teoria da dualidade dos grupos topológicos. O auditório estava lotado e Lipman Bers assumiu seu lugar para uma tradução em Inglês simultânea da palestra.

O público matemático internacional, liderado por uma concentração de topólogos, ficaram boquiabertos e atônitos como a palestra desenvolvida. Pontryagin parecia estar

falando sobre algum tipo de engenharia, fazendo-os se sentirem ignorantes e confusos. O Princípio do Máximo de Pontryagin parecia misterioso e incompreensível, exceto para aqueles poucos que eram especialistas na teoria de controle e que já estavam familiarizados com o princípio Bang-Bang, no caso linear tratado anteriormente por Bellman.

Depois de alguns dias de reflexão e murmúrios no Congresso, os matemáticos decidem que o Princípio do Máximo não se tratava de engenharia, apesar de tudo, mas que era um tema clássico no cálculo das variações. Assim, o consenso do Congresso chegou à conclusão de que a teoria de controle pode ser matematicamente respeitável, mas que era monótono e enfadonho - ainda a posição defendida por muitos dos matemáticos abstratos de vanguarda.

Em 1961, Kalman e Bucy publicaram seu notável escrito no controle ótimo dos sistemas lineares, com distúrbios de Gauss, e relativo a um custo funcional quadrático. O resultado matemático é que o controlador ótimo poderia ser sintetizado através da solução de uma equação diferencial da matriz de Riccati. Assim, um programa de computador razoavelmente elementar poderia gerar o controlador desejado ideal, que também serviu como um estabilizador de realimentação.

Não foi apenas esta análise matemática elegante, mas mais importante ainda, levou diretamente ao hardware práticas para o controle de engenharia de dispositivos de orientação. Em particular, o filtro de Kalman-Bucy era um componente essencial do sistema de orientação para a Apollo Moon-Rocket.

A comunidade de engenharia industrial ficou boquiaberta e atônita que os tais “Matemáticos de fantasia” podiam levar a produtos úteis e importantes para a engenharia. Esta teoria de controle matemático foi gradualmente aceita e considerada potencialmente útil e prática. Por muitos anos, sempre que os teóricos de controle foram desafiados a defender o seu ofício aos diretores das empresas de engenharia ou de agências de fomento governamentais, a resposta seria "filtro de Kalman".

NOVAS TENDÊNCIAS PARA A TEORIA DE CONTROLE (1970 – 1990)

Durante os últimos anos do século passado, novas tendências apareceram na teoria matemática do controle de sistemas. Alguns dos recentes ramos na árvore de controle matemático são:

- i.) Algebrização de análise do domínio da frequência, levando para a esfera da teoria do anel e da geometria algébrica. Os trabalhos de Kalman e seus colegas, particularmente Sontag, foram desenvolvimentos fundamentais.
- ii.) Equações diferenciais parciais e análise funcional para sistemas de parâmetros distribuídos, e sistemas com espaços de infinitas dimensões. Pioneiros neste campo incluem Fattorini, Lions, Russel.
- iii.) Grupos de Lie e Álgebras em um tratamento geométrico dos sistemas de controle não-linear. Veja escritos de Brockett, Hermes, Isidori e Sussmann.
- iv.) Equações diferenciais funcionais, especialmente equações diferenciais de atraso, são um campo ativo para investigações por Hale e Lee.
- v.) Domínio da frequência para análise de sistemas de dimensão infinita, levando a novos tipos de problemas na análise complexa.
- vi.) Campos tão novos e tão turbulentos sob reavaliação contínua (robótica, lógica de computador, bio-engenharia), que a sua influência na teoria de controle, ou vice-versa, é impossível avaliar no momento.

CAPÍTULO II

PRELIMINARES E NOTAÇÕES

Antes de abordar o assunto de Controle, é necessário fazer uma revisão de conceitos básicos de Álgebra Linear e Equações Diferenciais para um entendimento do próximo capítulo.

2.1. EQUAÇÃO DIFERENCIAL

As equações diferenciais são de importância capital na resolução de problemas matemáticos que envolvem movimento, crescimento, vibrações, eletricidade, magnetismo, aerodinâmica, termodinâmica, hidrodinâmica, energia nuclear, e todo tipo de fenômeno que envolva taxas de variação. Ver Swokowski [18] (1994)

Uma equação que contém as derivadas ou diferenciais de uma ou mais variáveis dependentes em relação a uma ou mais derivadas independentes é chamada de equação diferencial.

Para uma melhor análise das equações diferenciais, as classificamos por tipo, ordem e linearidade. Se uma equação diferencial contiver somente derivadas ordinárias de uma ou mais variáveis dependentes em relação a uma única variável independente, será chamada de *equação diferencial ordinária*, EDO.

Uma equação que envolve as derivadas parciais de uma ou mais variáveis dependentes de duas ou mais variáveis independentes é chamada de *equação diferencial parcial*, EDP.

A classificação pela ordem de uma EDO ou EDP, é dada pela ordem da maior derivada que aparece na equação diferencial em questão.

Definição. Podemos dizer que uma equação do tipo

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0 \quad (2.1.1)$$

onde $y = y(x)$ é a função incógnita de uma variável e suas derivadas definidas em algum intervalo $I \subseteq \mathbb{R}$, é chamada de equação diferencial ordinária de *ordem n*. Observe que F é

uma função de $n+2$ variáveis, $x, y, y', y'', \dots, y^{(n)}$. É complicado estudar (2.1.1) diretamente, por isso consideramos a hipótese de que o teorema das funções implícitas é aplicável. Logo, sempre é possível resolver uma equação diferencial dada na forma (2.1.1), de maneira única, para que a derivada mais alta $y^{(n)}$ se escreva em termos das $n+1$ variáveis restantes, a fim de obter,

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}) \quad (2.1.2)$$

onde f é uma função contínua de valores reais. A equação diferencial (2.1.2) é conhecida por *forma normal*.

Assim sendo, quando servir a nossos objetivos, utilizaremos a forma normal

$$y' = f(x, y) \quad \text{e} \quad y'' = f(x, y, y')$$

para representar equações diferenciais ordinárias gerais de primeira e segunda ordem.

Toda equação do tipo (2.1.2) pode ser reduzida ao estudo de uma equação de primeira ordem, isto é a equação (2.1.2) com $n=1$.

Exemplos.

- (1) $y' + 2y = 0$ é uma equação diferencial ordinária de primeira ordem.
- (2) $y'' + 3y' - 2y - 6e^x = 0$ é uma equação diferencial ordinária de segunda ordem.
- (3) $(y''')^2 - 2y y' + (y'') = 7$ é uma equação diferencial ordinária de terceira ordem.
- (4) $e^y \frac{d^2 y}{dx^2} + 2 \left(\frac{dy}{dx} \right)^3 = 0$ é uma equação diferencial ordinária de segunda ordem.

A solução de uma equação diferencial é uma função $y = \psi(x)$ que transforma a equação (2.1.1) numa identidade quando y e suas derivadas são substituídas por $\psi(x)$ e suas derivadas. Enquanto que o gráfico de $\psi(x)$ é chamado de *curva integral*. A solução representada implicitamente por $\phi(x, y) = 0$ é chamada de uma **integral** de (2.1.1).

Exemplo 1. Seja função exponencial $y = e^{-2x}$. Mostre que é **uma solução** da equação diferencial ordinária,

$$y' + 2y = 0.$$

Solução. Para verificar tal afirmação, basta substituir $y = f(x)$ e sua derivada na equação diferencial dada, isto é,

$$y = e^{-2x} \quad \text{e} \quad y' = -2e^{-2x}$$

Então, substituindo

$$y' + 2y = -2e^{-2x} + 2e^{-2x} = 0,$$

o que satisfaz a equação fornecida.

Da mesma forma, pode-se mostrar que

$$y = 2e^{-2x}, \quad y = -5e^{-2x}, \quad y = \frac{1}{4}e^{-2x}$$

são também soluções da equação diferencial dada.

Na verdade, qualquer função da forma $y = Ce^{-2x}$ onde $C \in \mathbb{R}$, é uma solução da equação. Esta família de soluções é chamada de **solução geral** da equação diferencial.

Definição. Uma função $y = y(x)$ definida implicitamente por $\phi(x, y, c_1, c_2, \dots, c_n) = 0$ que contém n constantes arbitrárias e é solução da equação diferencial (2.1.1), é chamada de **solução geral** da equação diferencial na região onde (2.1.1) é dada. Atribuindo valores às constantes c_1, c_2, \dots, c_n , obtemos uma **solução particular** de (2.1.1). Ver Leithold [9] (1994)

Exemplo 2. Verifique que $y = Cx^3$ é uma solução da equação diferencial

$$xy' - 3y = 0 \quad \text{para qualquer } C$$

Encontre a solução particular determinada pela condição inicial: $y = 2$ quando $x = -3$

Solução. Do enunciado do problema, tem-se que $y = Cx^3$ é a **solução geral** da equação diferencial. Para se constatar essa afirmação, é necessário substituir, na equação $xy' - 3y = 0$, a função y e sua derivada:

$$y = Cx^3 \quad \text{e} \quad y' = 3Cx^2$$

Substituindo na equação, tem-se:

$$xy' - 3y = 0$$

de onde obtemos

$$3Cx^2 - 3Cx^3 = 3Cx^3 - 3Cx^3 = 0$$

Desta forma, verificou-se que $y = Cx^3$ é a **solução geral** da equação diferencial.

Para se encontrar a **solução particular**, será utilizado o Problema do Valor Inicial: $y = 2$ quando $x = -3$. Então, a partir da solução geral,

$$y = Cx^3$$

e substituindo os valores dados, obtemos

$$2 = C(-3)^3 \quad \text{logo} \quad C = \frac{2}{27}$$

Logo, $y = \frac{2}{27}x^3$ é uma **solução particular** da equação diferencial dada.

Observação. O valor atribuído à função que elimina a constante é chamado de *condição*. As condições estão organizadas de acordo com o ponto em que são avaliadas. Assim, se a equação diferencial ordinária estiver definida no intervalo $[a, b]$ e a condição estiver dada em $x = a$ teremos uma *condição inicial*. De outra maneira, se a condição for dada num ponto $x \neq a$, será chamada de *condição de contorno*. A quantidade de condições necessárias para a total eliminação das constantes é igual à ordem da equação diferencial ordinária.

O *grau* de uma equação diferencial é o expoente da derivada ordinária de maior ordem que aparece na equação diferencial. Como ilustração, vejamos as seguintes equações diferenciais ordinárias:

- a.) $y' + e^x y = 0$, equação diferencial ordinária de primeiro grau.
- b.) $(y')^2 - 3xy + 5y = 0$, equação diferencial ordinária de segundo grau.
- c.) $y'' - 7y' + 5y = 0$, equação diferencial ordinária de primeiro grau.

EQUAÇÕES DIFERENCIAIS ORDINÁRIAS LINEARES

De acordo a classificação, temos a linearidade. Dizemos que uma EDO de ordem n (2.1.1) é linear se F for linear em $y, y', y'', \dots, y^{(n-1)}$. Isso significa que uma EDO de ordem n é linear quando (2.1.1) for escrita da seguinte forma,

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = k(x) \quad (2.1.3)$$

Na escrita da equação diferencial (2.1.3) observamos duas características importantes de uma equação diferencial linear: a variável dependente e todas as suas derivadas são do primeiro grau; cada coeficiente depende, no máximo, da variável independente.

Em princípio, estudaremos as equações diferenciais lineares de primeira ordem não-homogêneas e a equação diferencial linear de segunda ordem não-homogênea com coeficientes constantes.

Equação Diferencial Linear de Primeira Ordem

Denotaremos por \mathbb{R} o conjunto dos números reais e Ω um subconjunto aberto conexo de \mathbb{R}^{m+1} cujos elementos serão escritos na forma (y, x) onde $y \in \mathbb{R}^m$ e $x \in \mathbb{R}$ e consideraremos

$$f : \Omega \subseteq \mathbb{R}^{m+1} \rightarrow \mathbb{R}$$

uma função contínua. Diremos que $y = y(x)$ é solução da equação diferencial

$$y' = f(x, y) \quad \text{em} \quad I \subseteq \mathbb{R} \quad (2.1.4)$$

sempre que $(x, y) \in \Omega$ para $x \in I$ e $y = y(x)$ satisfaz a equação diferencial (2.1.4) para $x \in I$.

Observações. Como a função f é contínua segue-se que uma solução $y = y(x)$ de (2.1.4), se existir, será continuamente diferenciável. Em geral a equação (2.1.4) terá mais de uma solução, mas estamos interessados na solução de uma ED sujeita a determinadas restrições ou condições prescritas que serão impostas à solução incógnita $y = y(x)$ e suas derivadas.

Em algum intervalo I contendo o ponto x_0 o problema

$$\begin{aligned} y^{(n)} &= f(x, y, y', \dots, y^{(n-1)}) \\ y(x_0) &= y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1} \end{aligned}$$

onde y_0, y_1, \dots, y_{n-1} são constantes reais, é chamado de *problema de valor inicial* PVI ou *problema de Cauchy*. Os valores de $y = y(x)$ e suas $n-1$ derivadas num único ponto x_0 são chamados de *condições iniciais*.

Definição. Uma **equação diferencial linear de primeira ordem** é uma equação da forma

$$y' = f(x, y) = -P(x)y + Q(x) \quad \Leftrightarrow \quad y' + P(x)y = Q(x)$$

onde P e Q são funções contínuas em algum intervalo $I \subseteq \mathbb{R}$. Ver Swokowski [18] (1994)

Se $Q(x) = 0$, então, a equação diferencial linear de primeira ordem é homogênea. Desta forma, a equação diferencial ordinária (EDO) de primeira ordem homogênea com coeficientes constantes tem o seguinte formato:

$$y' - k y = 0$$

Como solução desta equação diferencial ordinária, tem-se:

$$y = c e^{kt} \quad \text{para } c \in \mathbb{R}$$

Analisando os seguintes casos:

1) Se $c = 0$, então:

$$y = 0 \cdot e^{ky} = 0$$

2) Se $c \in \mathbb{R} \setminus \{0\}$, então:

$$y' = \frac{d}{dt} y = \frac{d}{dt} (c e^{kt}) = c \frac{d}{dt} (e^{kt}) = c k e^{kt}$$

Mas k é uma constante. Logo, $y = c e^{kt}$ é uma família de soluções para a EDO.

Para obter apenas uma única solução do problema, deve-se trabalhar com o problema do valor inicial, (PVI)

$$y' - ky = 0, \quad y(t_0) = y_0$$

onde y_0 é um valor conhecido.

Quando $t_0 = 0$, tem-se:

$$y_0 = y(t_0) = y(0) = c e^{k \cdot 0} = c$$

Desta forma, a única solução da equação será:

$$y(t) = y_0 e^{kt}$$

Analisando a constante k :

- 1) Se $k > 0$, então a solução y é crescente de forma exponencial e k é a constante de crescimento.
- 2) Se $k < 0$, então a solução y decai exponencialmente e k é a constante de decaimento.

Se considerarmos $Q(x) \neq 0$ na equação diferencial de primeira ordem linear,

$$y' + P(x) y = Q(x)$$

então, a equação diferencial linear de primeira ordem é não-homogênea.

A equação diferencial ordinária de primeira ordem não-homogênea com coeficientes constantes é dada por

$$y' + k y = f \quad \text{com} \quad f \neq 0$$

é possível obter uma família de soluções por intermédio de uma função chamada de multiplicador. Mas desejamos uma única solução. Para isso acontecer, devemos adicionar mais uma condição à equação diferencial, chamada de condição inicial. Logo, o problema a resolver é conhecido como problema de valor inicial, PVI.

Para encontrar a solução explícita do PVI, não homogêneo,

$$y' + a y = f(t), \quad y(0) = y_0 \quad (2.1.5)$$

procedemos da seguinte maneira: Seja o multiplicador $M(t) = e^{at}$. Multiplicando $M(t)$ a ambos os membros da primeira equação de (2.1.5), tem-se:

$$M(t)[y' + ay] = M(t)f$$

Substituindo $M(t)$ por e^{at} :

$$e^{at} [y' + ay] = e^{at} f$$

Aplicando a propriedade distributiva:

$$e^{at} y' + e^{at} ay = e^{at} f \quad (2.1.1)$$

Tem-se que:

$$\frac{d}{dt}(ye^{at}) = y'e^{at} + ya e^{at} \quad (2.1.2)$$

Substituindo (2.1.2) em (2.1.1), tem-se:

$$\frac{d}{dt}(ye^{at}) = e^{at} f(t)$$

Logo, escrevendo na sua forma de diferenciais obtemos,

$$d(ye^{at}) = e^{at} f(t) dt \quad (2.1.3)$$

Aplicando a integral indefinida em ambos os lados da igualdade, tem-se:

$$\int d(ye^{at} - C) = \int e^{at} f(t) dt \quad \text{então} \quad ye^{at} = \int e^{at} f(t) dt + C$$

Isolando a função y

$$y = Ce^{-at} + e^{-at} \int e^{at} f(t) dt$$

Substituindo a variável t de dentro da integral pela variável δ ,

$$y = Ce^{-at} + e^{-at} \int e^{a\delta} f(\delta) d\delta$$

onde, após simplificações obtemos

$$y = Ce^{-at} + \int e^{-at} e^{a\delta} f(\delta) d\delta \quad \text{então} \quad y = Ce^{-at} + \int e^{a(\delta-t)} f(\delta) d\delta$$

Desta forma,

$$y = Ce^{-at} + \int e^{a(\delta-t)} f(\delta) d\delta$$

é a **família de soluções (solução geral)** da equação diferencial dada.

Para encontrar uma **solução única (solução particular)** dessa EDO, aplica-se a integral definida em (2.1.3)

$$\int_0^t d(ye^{as}) = \int_0^t e^{as} f(s) ds$$

Aplicando o Teorema Fundamental do Cálculo:

$$y(t)e^{at} - y(0)e^{a0} = \int_0^t e^{as} f(s) ds \quad \Rightarrow \quad y(t)e^{at} - y(0) = \int_0^t e^{as} f(s) ds$$

Isolando a função $y(t)$:

$$y(t) = e^{-at} y(0) + e^{-at} \int_0^t e^{as} f(s) ds$$

obtemos,

$$y(t) = e^{-at} y_0 + \int_0^t e^{a(s-t)} f(s) ds, \quad t \in [0, T] \quad (2.1.4)$$

Desta forma, (2.1.4) é a solução única do problema de valor inicial.

A seguir, consideramos o seguinte PVI não homogêneo com a e f funções contínuas,

$$y' + a(t)y = f(t), \quad y(0) = y_0 \quad (2.1.5)$$

Para encontrar a sua solução consideramos o multiplicador

$$M(t) = e^{\int a(t) dt} = e^{P(t)} \quad \text{com} \quad P(t) = \int a(t) dt$$

Multiplicando $M(t)$ a ambos os membros da primeira equação de (2.1.5), tem-se:

$$M(t)[y'(t) + a(t)y(t)] = M(t)f(t)$$

Aplicando a propriedade distributiva:

$$M(t)y'(t) + M(t)a(t)y(t) = M(t)f(t) \quad (2.1.6)$$

Por outro lado obtemos,

$$\frac{d}{dt}[y(t)M(t)] = y'(t)M(t) + M'(t)y(t) \quad (2.1.7)$$

Pela Regra da Cadeia, tem-se:

$$M' = \frac{d}{dt}M(t) = e^{P(t)} \frac{d}{dt}P(t) \quad \text{logo} \quad M' = \frac{d}{dt}M(t) = e^{P(t)}[a(t) + c]$$

Para $c = 0$,

$$\frac{d}{dt}M(t) = a(t)e^{P(t)}$$

Mas $e^{P(t)} = M(t)$. Então, substituindo na equação anterior obtemos,

$$M' = \frac{d}{dt}M(t) = a(t)M(t) \quad (2.1.8)$$

Substituindo (2.1.8) em (2.1.7), tem-se:

$$\frac{d}{dt}[y(t)M(t)] = y'(t)M(t) + a(t)M(t)y(t) \quad (2.1.9)$$

Substituindo (2.1.9) em (2.1.6), tem-se:

$$\frac{d}{dt}[y(t)M(t)] = M(t)f(t)$$

Em termos de diferenciais, temos

$$d[y(t)M(t)] = M(t)f(t)dt \quad (2.1.10)$$

Aplicando a integral indefinida em ambos os lados da igualdade, tem-se:

$$\int d[y(t)M(t) - C] = \int M(t)f(t)dt \quad \text{logo} \quad y(t)M(t) = C + \int M(t)f(t)dt$$

Mas $M(t) = e^{P(t)}$, substituindo no resultado anterior, temos

$$y(t)e^{P(t)} = C + \int e^{P(t)}f(t)dt$$

Isolando a função $y(t)$:

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int e^{P(t)}f(t)dt$$

Mas $P(t) = \int a(t)dt$, novamente substituindo obtemos

$$y(t) = Ce^{-\int a(t)dt} + e^{-\int a(t)dt} \int e^{\int a(t)dt} f(t)dt$$

Substituindo a variável t de dentro da integral pela variável δ :

$$y(t) = Ce^{-\int a(t)dt} + e^{-\int a(t)dt} \int e^{\int a(\delta)d\delta} f(\delta)d\delta$$

Desta forma,

$$y(t) = C e^{-\int a(t)dt} + e^{-\int a(t)dt} \int e^{\int a(\delta)d\delta} f(\delta) d\delta$$

é a família de soluções do problema de valor inicial não homogêneo.

Para encontrar uma solução única dessa EDO, aplica-se a integral definida em (2.1.10),

$$\int_0^t d[y(t)M(t)] = \int_0^t M(s)f(s)ds$$

Aplicando o Teorema Fundamental do Cálculo:

$$y(t)M(t) - y(0)M(0) = \int_0^t M(s)f(s)ds \quad (2.1.11)$$

Por outro lado,

$$M(t) = e^{\int_0^t a(s)ds}$$

Substituindo $t=0$ obtemos,

$$M(0) = e^{\int_0^0 a(s)ds} = e^0 = 1 \quad (2.1.12)$$

Substituindo (2.1.12) em (2.1.11), tem-se:

$$y(t)M(t) - y(0) = \int_0^t M(s)f(s)ds$$

Mas $M(t) = \exp\left(\int_0^t a(s)ds\right)$. Então:

$$y(t)e^{\int_0^t a(s)ds} = y(0) + \int_0^t e^{\int_0^s a(r)dr} f(s)ds$$

Isolando a função $y(t)$

$$y(t)e^{\int_0^t a(s)ds} = y(0) + \int_0^t e^{\int_0^s a(r)dr} f(s)ds$$

Sendo $P(t) = \int_0^t a(s)ds$. Então, substituindo no resultado anterior,

$$y(t) = e^{-P(t)}y(0) + e^{-P(t)} \int_0^t e^{P(s)} f(s)ds$$

Utilizando o valor fornecido da condição inicial temos

$$y(t) = e^{-P(t)}y_0 + \int_0^t e^{P(s)-P(t)} f(s)ds \quad (2.1.13)$$

Por outro lado,

$$P(s) - P(t) = \int_0^s a(r)dr - \int_0^t a(s)ds = -\left[\int_0^t a(s)ds - \int_0^s a(r)dr\right]$$

Seja

$$a(w) = \begin{cases} a(r) & r \in [0, s] \\ a(s) & s \in [0, t] \end{cases}$$

Por outro lado sabemos que

$$\int_0^s a(w)dw + \int_s^t a(w)dw = \int_0^t a(w)dw$$

Assim obtemos,

$$\int_s^t a(w)dw = \int_0^t a(s)ds - \int_0^s a(r)dr \quad (2.1.14)$$

Substituindo (2.1.14) em (2.1.13), tem-se:

$$y(t) = e^{-P(t)} y_0 + \int_0^t e^{-\int_s^t a(w)dw} f(s)ds \quad (2.1.15)$$

Desta forma, (2.1.15) é a solução única do PVI não homogêneo.

Método de Lagrange ou de Variação dos Parâmetros.

Consideremos o problema de encontrar $y = y(t)$, tal que

$$\begin{aligned} y' &= a(t)y + f(t), \quad \text{com } t_0 \in I \subseteq \mathbb{R} \\ y(t_0) &= y_0 \end{aligned} \quad (2.1.16)$$

onde a e f funções contínuas numa vizinhança de t_0 .

Em primeiro lugar, se

$$f(t) \equiv 0 \quad \text{e} \quad y_0 \equiv 0$$

o problema é trivial, e a sua solução é $y(t) \equiv 0$.

Se a seguinte possibilidade, acontecer

$$f(t) \equiv 0 \quad \text{e} \quad y_0 \neq 0$$

o problema é de variáveis separáveis e logo a solução é obtida com relativa facilidade,

$$y(t) = y_0 \exp\left(\int_{t_0}^t a(s)ds\right)$$

Se a função $f(t) \neq 0$ podemos conjecturar que talvez substituindo na expressão anterior o valor de y_0 por uma função apropriada $z(t)$ diferenciável numa vizinhança de t_0 podemos obter uma solução de (2.1.16).

Portanto, ensaiaremos com a expressão,

$$y(t) = z(t) \exp\left(\int_{t_0}^t a(s) ds\right)$$

tratando de obter $z(t)$ tal que $y = y(t)$ verifique o PVI, (2.1.16). Isto é,

$$\begin{aligned} y'(t) &= z'(t) \exp\left(\int_{t_0}^t a(s) ds\right) + a(t) z(t) \exp\left(\int_{t_0}^t a(s) ds\right) \\ &= a(t) z(t) \exp\left(\int_{t_0}^t a(s) ds\right) + f(t) \end{aligned}$$

Assim, resulta o seguinte

$$z'(t) \exp\left(\int_{t_0}^t a(s) ds\right) = f(t)$$

Por outro lado temos que a condição inicial,

$$y_0 = y(t_0) = z(t_0)$$

Assim sendo,

$$z(t) = \int_{t_0}^t f(s) \exp\left(-\int_{t_0}^s a(r) dr\right) ds + y_0$$

Portanto a seguinte expressão final,

$$y(t) = y_0 \exp\left(\int_{t_0}^t a(s) ds\right) + \int_{t_0}^t f(s) \exp\left(\int_s^t a(r) dr\right) ds$$

é a solução do PVI em questão.

Equação Diferencial Ordinária Linear de Segunda Ordem

As equações diferenciais ordinárias de segunda ordem lineares, escritas na sua forma normal são dadas pela expressão geral,

$$y'' = f(x, y, y')$$

Neste estudo, será abordada apenas a equação diferencial linear de segunda ordem com coeficientes constantes, que tem a seguinte forma:

$$y'' = f(x, y, y') = -cy - by' + k(x) \Rightarrow y'' + by' + cy = k(x),$$

onde b e c são constantes.

Por solução da equação diferencial entendemos uma função que satisfaça a equação diferencial.

Exemplo 1. Seja a equação diferencial ordinária e uma função,

$$y'' = 3y, \quad y = \varphi(x) = e^{3x}$$

Verifique que a função $\varphi(x)$ é solução da EDO em questão.

Solução. Derivando duas vezes a função fornecida obtemos,

$$y' = 3e^{3x}, \quad y'' = 9e^{3x}.$$

Finalmente, substituindo as derivadas anteriores na equação diferencial dada, obtemos a identidade desejada.

Exemplo 2. Seja a equação diferencial ordinária de segunda ordem

$$y'' = 0, \quad x \in [a, b]$$

Solução. Por intermédio do teorema fundamental do cálculo, aplicado duas vezes obtemos a função solução

$$y = \varphi(x) = c_1x + c_2$$

Neste exemplo, a **solução geral** depende de duas constantes arbitrárias. Muitas vezes devemos resolver problemas com duas condições. No caso de serem fornecidas as condições $y(a)$ e $y'(a)$, teremos um problema do valor inicial. Caso se tenha uma condição em $x = a$ e outra condição em $x = b$ teremos um problema de contorno.

A partir de agora, será estudada a resolução das equações diferenciais de segunda ordem com coeficientes constantes. Para resolver a equação diferencial de segunda ordem homogênea,

$$y''(t) + by'(t) + cy(t) = 0$$

com b e c constantes, é necessário apresentar alguns teoremas e definições.

Teorema 1. Se as funções y_1 e y_2 forem soluções definidas em algum intervalo $I \subseteq \mathbb{R}$ da equação diferencial homogênea linear de segunda ordem,

$$y'' + ay' + by = 0$$

então a nova função,

$$Y = c_1y_1(x) + c_2y_2(x), \quad c_1, c_2 \in \mathbb{R}$$

é também solução da equação diferencial em algum intervalo.

Devemos exigir alguma garantia na forma de condições sobre a solução Y para que esta seja chamada de solução geral da equação diferencial homogênea. Em outras palavras, se W for uma outra solução da equação diferencial homogênea, então podemos selecionar as constantes c_1 e c_2 de maneira que $Y(x) = W(x)$.

Teorema 2. Se as funções y_1 e y_2 forem soluções definidas em algum intervalo $I \subseteq \mathbb{R}$ da equação diferencial homogênea linear de segunda ordem,

$$y'' + ay' + by = 0 \quad \text{em} \quad [\alpha, \beta]$$

com a seguinte expressão,

$$y_1(x)y_2'(x) - y_1'(x)y_2(x) \neq 0 \quad \text{em} \quad [\alpha, \beta]$$

Então, a nova função dada por

$$Y = c_1y_1(x) + c_2y_2(x)$$

é a *solução geral* da equação diferencial homogênea.

Definição. A expressão $y_1(t)y_2'(t) - y_1'(t)y_2(t)$ descrita no teorema anterior ou representada por um determinante,

$$\begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix}$$

é chamada de **wronskiano** das funções y_1 e y_2 e é denotado pelo seguinte símbolo,

$$W(t) = W[y_1, y_2](t)$$

Desta forma, será, a partir de agora, estudado um método, extraído de Leithold [9] (1994), para a resolução da equação homogênea

$$ay''(t) + by'(t) + cy(t) = 0.$$

Se forem encontradas duas soluções cujo wronskiano seja diferente de zero, então, tem-se a **solução geral** da equação. Ao se analisar a equação, percebe-se que a soma de $ay''(t)$ com $by'(t)$ e $cy(t)$ deve ser nula. Para isso, as funções $y''(t)$, $y'(t)$ e $y(t)$ devem ter o mesmo aspecto, ou seja, não se pode pensar em $y(t)$ como uma função polinomial, pois suas derivadas sucessivas reduziriam o grau do polinômio e a soma destes resultados jamais seria zero. Pode-se tentar deduzir que $y(t) = e^{\lambda t}$, onde λ é um parâmetro a ser determinado.

Desta forma, tornar-se-ia o primeiro membro da equação como uma combinação linear da função exponencial $e^{\lambda t}$:

$$y(t) = e^{\lambda t}, \quad y'(t) = \lambda e^{\lambda t} \quad \text{e} \quad y''(t) = \lambda^2 e^{\lambda t}$$

Então, substituindo as funções anteriores em,

$$ay''(t) + by'(t) + cy(t) = 0$$

obtemos

$$a\lambda^2 e^{\lambda t} + b\lambda e^{\lambda t} + ce^{\lambda t} = 0 \quad \text{logo} \quad (a\lambda^2 + b\lambda + c)e^{\lambda t} = 0$$

Como a função exponencial é diferente de zero, $e^{\lambda t}$ é solução da equação homogênea se, e somente se, λ é solução da equação $a\lambda^2 + b\lambda + c = 0$, denominada de *equação característica* da equação dada.

Encontrando a solução da equação $a\lambda^2 + b\lambda + c = 0$ tem-se:

$$\lambda_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{e} \quad \lambda_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Analisando $b^2 - 4ac$

Caso 1: Se a expressão $b^2 - 4ac > 0$. A equação possui duas raízes reais distintas e, portanto, duas soluções $e^{\lambda_1 t}$ e $e^{\lambda_2 t}$.

Calculando o wronskiano:

$$W[e^{\lambda_1 t}, e^{\lambda_2 t}] = \begin{vmatrix} e^{\lambda_1 t} & e^{\lambda_2 t} \\ \lambda_1 e^{\lambda_1 t} & \lambda_2 e^{\lambda_2 t} \end{vmatrix} = (\lambda_2 - \lambda_1)e^{(\lambda_1 + \lambda_2)t}$$

Como $W[e^{\lambda_1 t}, e^{\lambda_2 t}] \neq 0$, a solução geral da equação é dada por

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

onde c_1 e c_2 são constantes arbitrárias.

Exemplo 1. Dada a equação diferencial

$$y''(t) - y'(t) - 2y(t) = 0$$

Determine a solução geral desta equação.

Solução. Seja $y(t) = e^{\lambda t}$. Então, as derivadas sucessivas são dadas por

$$y'(t) = \lambda e^{\lambda t} \quad \text{e} \quad y''(t) = \lambda^2 e^{\lambda t}.$$

Substituindo na equação dada, tem-se:

$$(\lambda^2 - \lambda - 2)e^{\lambda t} = 0$$

Calculando a equação característica $\lambda^2 - \lambda - 2 = 0$, tem-se:

$$\lambda_1 = \frac{-(-1) + \sqrt{(-1)^2 - 4(1)(-2)}}{2(1)} \quad \text{e} \quad \lambda_2 = \frac{-(-1) - \sqrt{(-1)^2 - 4(1)(-2)}}{2(1)}$$

$$\lambda_1 = \frac{1 + \sqrt{9}}{2} = 2 \quad \text{e} \quad \lambda_2 = \frac{1 - \sqrt{9}}{2} = -1$$

Então:

$$y_1(t) = c_1 e^{2t} \quad \text{também} \quad y_2(t) = c_2 e^{-t}$$

Além disso,

$$W[y_1(t), y_2(t)] = \begin{vmatrix} e^{2t} & e^{-t} \\ 2e^{2t} & -e^{-t} \end{vmatrix} = 3e^t \neq 0$$

Segue que

$$y(t) = c_1 e^{2t} + c_2 e^{-t}$$

é a **solução geral** da equação diferencial requerida.

Caso 2: Se a equação $b^2 - 4ac = 0$. A equação quadrática tem-se uma única raiz dupla e, portanto, apenas uma solução $e^{\lambda t} = y_1(t)$.

Considerando a seguinte função

$$y_2(t) = y_1(t) v(t)$$

Substituindo $y_2(t)$ na equação homogênea reescrita na forma

$$y''(t) + \frac{b}{a} y'(t) + \frac{c}{a} y(t) = 0,$$

e, calculando a primeira derivada

$$y_2'(t) = y_1'(t) \cdot v(t) + y_1(t) \cdot v'(t)$$

e a segunda derivada,

$$y_2''(t) = y_1''(t) \cdot v(t) + 2y_1'(t) \cdot v'(t) + y_1(t) \cdot v''(t)$$

Logo substituindo para obter

$$y_1''(t) \cdot v(t) + 2y_1'(t) \cdot v'(t) + y_1(t) \cdot v''(t) + \frac{b}{a} [y_1'(t) \cdot v(t) + y_1(t) \cdot v'(t)] + \frac{c}{a} [y_1(t) \cdot v(t)] = 0$$

Agrupando

$$y_1(t) \cdot v''(t) + \left[2y_1'(t) + \frac{b}{a} y_1(t) \right] v'(t) + \left[y_1''(t) + \frac{b}{a} y_1'(t) + \frac{c}{a} y_1(t) \right] v(t) = 0$$

E pelo fato que $y_1(t)$ é solução da equação, temos,

$$y_1(t) \cdot v''(t) + \left[2y_1'(t) + \frac{b}{a} y_1(t) \right] v'(t) = 0,$$

Tem-se que $y_2(t)$ é solução da equação se $v(t)$ for solução da equação diferencial

$$y_1(t) \cdot v''(t) + \left[2y_1'(t) + \frac{b}{a} y_1(t) \right] v'(t) = 0$$

Sendo $v'(t) = u(t)$ e $y_1(t) = e^{\lambda_1 t}$ então:

$$e^{\lambda_1 t} \cdot u'(t) + \left[2\lambda_1 e^{\lambda_1 t} + \frac{b}{a} e^{\lambda_1 t} \right] u(t) = 0$$

Dividindo a equação acima por $e^{\lambda_1 t}$, tem-se:

$$u'(t) + \left[2\lambda_1 + \frac{b}{a} \right] u(t) = 0$$

Tem-se agora, uma equação linear de primeira ordem homogênea com coeficientes constantes, cuja solução é dada por,

$$u(t) = k e^{-\left(2\lambda_1 + \frac{b}{a}\right)t}$$

Para encontrar somente uma solução, faz-se $k = 1$. Então:

$$u(t) = e^{-\left(2\lambda_1 + \frac{b}{a}\right)t} \Rightarrow u(t) = \frac{e^{-\left(\frac{b}{a}\right)t}}{e^{2\lambda_1 t}} \Rightarrow u(t) = \frac{e^{-\left(\frac{b}{a}\right)t}}{y_1^2(t)}$$

Então, pela relação definida anteriormente,

$$v'(t) = \frac{e^{-\left(\frac{b}{a}\right)t}}{y_1^2(t)}$$

Calculando $v(t)$, após integração, temos

$$y_2(t) = y_1(t) \int \frac{e^{-\left(\frac{b}{a}\right)t}}{y_1^2(t)} dt$$

Exemplo 1. Dada a equação diferencial

$$y''(t) - 4y'(t) + 4y(t) = 0$$

Determine a solução geral desta equação.

Solução. Seja $y(t) = e^{\lambda t}$. Então as derivadas são

$$y'(t) = \lambda e^{\lambda t} \quad \text{e} \quad y''(t) = \lambda^2 e^{\lambda t}.$$

Substituindo na equação dada, tem-se:

$$(\lambda^2 - 4\lambda + 4)e^{\lambda t} = 0$$

Calculando a equação característica $\lambda^2 - 4\lambda + 4 = 0$, tem-se:

$$\lambda_1 = \frac{-(-4) + \sqrt{(-4)^2 - 4(1)(4)}}{2(1)} \quad \text{e} \quad \lambda_2 = \frac{-(-4) - \sqrt{(-4)^2 - 4(1)(4)}}{2(1)}$$

$$\lambda_1 = \frac{4 + \sqrt{0}}{2} = 2 \quad \text{e} \quad \lambda_2 = \frac{4 - \sqrt{0}}{2} = 2$$

Então:

$$y_1(t) = c_1 e^{2t}$$

Para encontrar y_2 , faz-se:

$$y_2(t) = y_1(t) \int \frac{e^{-\left(\frac{b}{a}\right)t}}{y_1^2(t)} dt$$

Então:

$$y_2(t) = e^{2t} \int \frac{e^{-\left(\frac{-4}{1}\right)t}}{(e^{2t})^2} dt = e^{2t} \int \frac{e^{4t}}{e^{4t}} dt = e^{2t} \int 1 dt$$

Logo temos,

$$y_2(t) = t e^{2t}$$

Além disso,

$$W[y_1(t), y_2(t)] = \begin{vmatrix} e^{2t} & t e^{2t} \\ 2e^{2t} & e^{2t} + 2t e^{2t} \end{vmatrix} = e^{4t} + 2e^{4t} - 2e^{4t} = e^{4t} \neq 0$$

Segue que

$$y(t) = c_1 e^{2t} + c_2 t e^{2t}$$

é a **solução geral** da equação diferencial dada.

Caso 3: Se a desigualdade $b^2 - 4ac < 0$. Desta forma, tem-se duas raízes complexas. E, pelo que foi visto até então, $y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$, onde $y(t)$ seria complexo. Entretanto, é necessário encontrar duas raízes reais. Para se contornar esse problema, usa-se um artifício:

Seja $y(t) = u(t) + iv(t)$, uma solução da equação homogênea com valores complexos.

Então:

$$a[u''(t) + iv''(t)] + b[u'(t) + iv'(t)] + c[u(t) + iv(t)] = 0$$

Identificando parte real e imaginária,

$$[au''(t) + bu'(t) + cu(t)] + i[av''(t) + bv'(t) + cv(t)] = 0$$

desta forma, o número complexo resultante é nulo. No entanto, para que um número complexo seja nulo, sua parte real e imaginária devem ser iguais a zero. Então:

$$au''(t) + bu'(t) + cu(t) = 0 \quad \text{e} \quad av''(t) + bv'(t) + cv(t) = 0$$

Neste caso, cada solução complexa dessa equação homogênea, dá origem a duas soluções reais.

Considerando a solução complexa $e^{\lambda_1 t}$. Define-se:

$$\theta = \frac{\sqrt{4ac - b^2}}{2a}$$

então

$$e^{\lambda_1 t} = e^{\left(\frac{b}{2a} + i\theta\right)t} = e^{\frac{b}{2a}t} e^{i\theta t} = e^{\frac{b}{2a}t} [\cos(\theta t) + i \operatorname{sen}(\theta t)]$$

Assim sendo

$$e^{\lambda_1 t} = e^{\frac{b}{2a}t} \cos(\theta t) + i e^{\frac{b}{2a}t} \operatorname{sen}(\theta t)$$

Portanto, tem-se as soluções reais,

$$y_1(t) = e^{\frac{b}{2a}t} \cos(\theta t) \quad \text{e} \quad y_2(t) = e^{\frac{b}{2a}t} \operatorname{sen}(\theta t)$$

A seguir calculamos seu wronskiano de ambas soluções, definindo previamente $\sigma = -b/2a$

$$W[y_1, y_2](t) = W[e^{\sigma t} \cos \theta t, e^{\sigma t} \operatorname{sen} \theta t] = \theta e^{\sigma t} \neq 0$$

decorre do fato que $\theta > 0$ definida anteriormente.

Assim sendo, a solução geral é dada por

$$Y = c_1 e^{\sigma t} \cos(\theta t) + c_2 e^{\sigma t} \sin(\theta t)$$

Podemos imaginar que deixamos de lado a segunda solução complexa, pois teríamos mais duas soluções reais, porém não acontece acréscimo de soluções reais.

Considerando a solução complexa $e^{\lambda_2 t}$. Define-se:

$$\theta = \frac{\sqrt{4ac - b^2}}{2a}$$

então

$$\begin{aligned} e^{\lambda_2 t} &= e^{(\sigma - i\theta)t} = e^{\sigma t} e^{-i\theta t} = e^{\sigma t} [\cos(-\theta t) + i \sin(-\theta t)] \\ &= e^{\sigma t} \cos(\theta t) - i e^{\sigma t} \sin(\theta t) \end{aligned}$$

Assim sendo

$$e^{\lambda_2 t} = e^{\frac{b}{2a}t} \cos(\theta t) + i \left(-e^{\frac{b}{2a}t} \sin(\theta t) \right)$$

Portanto, tem-se as soluções reais,

$$y_1(t) = e^{\frac{b}{2a}t} \cos(\theta t) \quad \text{e} \quad y_2(t) = -e^{\frac{b}{2a}t} \sin(\theta t)$$

portanto temos novamente as mesmas soluções.

Exemplo 1. Dada a equação diferencial

$$y''(t) + 4y'(t) + 13y(t) = 0$$

Determine a solução geral desta equação.

Solução. Seja $y(t) = e^{\lambda t}$. Então:

$$y''(t) = \lambda^2 e^{\lambda t}, \quad y'(t) = \lambda e^{\lambda t}.$$

Substituindo na equação dada, tem-se:

$$(\lambda^2 + 4\lambda + 13)e^{\lambda t} = 0$$

Calculando a equação característica $\lambda^2 + 4\lambda + 13 = 0$, tem-se:

$$\lambda_1 = \frac{-(4) + \sqrt{(4)^2 - 4(1)(13)}}{2(1)} \quad \text{e} \quad \lambda_2 = \frac{-(4) - \sqrt{(4)^2 - 4(1)(13)}}{2(1)}$$

portanto

$$\lambda_1 = \frac{-4 + \sqrt{-36}}{2} = -2 + 3i \quad \text{e} \quad \lambda_2 = \frac{-4 - \sqrt{-36}}{2} = -2 - 3i$$

Desta forma, tem-se que $e^{(-2+3i)t}$ é uma solução com valores complexos da equação dada. Mas,

$$\begin{aligned} e^{(-2+3i)t} &= e^{-2t} (\cos 3t + i \operatorname{sen} 3t) \\ &= e^{-2t} \cos(3t) + i e^{-2t} \operatorname{sen}(3t) \end{aligned}$$

Neste caso, tem-se que as soluções da equação homogênea são:

$$y_1(t) = e^{-2t} \cos(3t) \quad \text{e} \quad y_2(t) = e^{-2t} \operatorname{sen}(3t)$$

Além disso,

$$W[y_1(t), y_2(t)] = \begin{vmatrix} e^{-2t} \cos(3t) & e^{-2t} \operatorname{sen}(3t) \\ -2e^{-2t} \cos(3t) - 3e^{-2t} \operatorname{sen}(3t) & -2e^{-2t} \operatorname{sen}(3t) + 3e^{-2t} \cos(3t) \end{vmatrix}$$

$$\begin{aligned} W[y_1(t), y_2(t)] &= -2e^{-4t} \cos(3t) \operatorname{sen}(3t) + 3e^{-4t} \cos^2(3t) \\ &\quad + 2e^{-4t} \cos(3t) \operatorname{sen}(3t) + 3e^{-4t} \operatorname{sen}^2(3t) \\ &= 3e^{-4t} \cos^2(3t) + 3e^{-4t} \operatorname{sen}^2(3t) = 3e^{-4t} \end{aligned}$$

Então, obtemos que o wroskiano é não nulo,

$$W[y_1(t), y_2(t)] = 3e^{-4t} \neq 0$$

Segue que

$$y(t) = c_1 e^{-2t} \cos(3t) + c_2 e^{-2t} \operatorname{sen}(3t)$$

é a **solução geral** da equação diferencial requerida.

Após o estudo das equações lineares de segunda ordem homogêneas, passa-se ao estudo das equações lineares de segunda ordem não-homogêneas.

Considerando

$$ay''(t) + by'(t) + cy(t) = k(t)$$

como uma equação linear de segunda ordem não-homogênea, com coeficientes constantes. Para obter a solução dessa equação, em virtude da sua linearidade, pode-se somar à equação geral da solução homogênea, uma solução particular da equação diferencial completa.

Método da Variação dos Parâmetros ou Constantes.

Sejam $y_1(t)$ e $y_2(t)$ soluções da equação homogênea, com $W[y_1(t), y_2(t)] \neq 0$, então a seguinte expressão

$$Y(t) = c_1 y_1(t) + c_2 y_2(t)$$

é a solução geral da equação homogênea.

Partindo dessa equação, será encontrada uma solução particular da equação completa da forma

$$\varphi(t) = \mu_1(t)y_1(t) + \mu_2(t)y_2(t)$$

Calculando a derivada primeira

$$\begin{aligned}\varphi'(t) &= \mu_1'(t)y_1(t) + \mu_1(t)y_1'(t) + \mu_2'(t)y_2(t) + \mu_2(t)y_2'(t) \\ &= [\mu_1(t)y_1'(t) + \mu_2(t)y_2'(t)] + [\mu_1'(t)y_1(t) + \mu_2'(t)y_2(t)]\end{aligned}$$

e a derivada segunda de $\varphi(t)$ tem-se:

$$\begin{aligned}\varphi''(t) &= [\mu_1'(t)y_1'(t) + \mu_1(t)y_1''(t) + \mu_2'(t)y_2'(t) + \mu_2(t)y_2''(t)] \\ &\quad + [\mu_1''(t)y_1(t) + \mu_1'(t)y_1'(t) + \mu_2''(t)y_2(t) + \mu_2'(t)y_2'(t)]\end{aligned}$$

Percebe-se que

$$[\mu_1'(t)y_1(t) + \mu_2'(t)y_2(t)] = 0$$

desta forma, na derivada segunda de $\varphi(t)$, não se encontrará as derivadas segundas das funções μ_1 e μ_2 .

Para encontrar essas derivadas, modifica-se um pouco a equação

$$ay''(t) + by'(t) + cy(t) = k(t).$$

dividindo a equação dada por pela constante $a \neq 0$, tem-se:

$$\frac{a}{a}y''(t) + \frac{b}{a}y'(t) + \frac{c}{a}y(t) = \frac{1}{a}k(t) \quad \text{logo} \quad y''(t) + \alpha y'(t) + \beta y(t) = h(t)$$

onde $\alpha = \frac{b}{a}$, $\beta = \frac{c}{a}$ e $h(t) = \frac{1}{a}k(t)$

Substituindo $\varphi(t)$ na equação $y''(t) + \alpha y'(t) + \beta y(t) = h(t)$ e utilizando a condição

$$[\mu_1'(t)y_1(t) + \mu_2'(t)y_2(t)] = 0, \text{ tem-se:}$$

$$\left[\mu_1' y_1' + \mu_1 y_1'' + \mu_2' y_2' + \mu_2 y_2'' \right] + \alpha \left[\mu_1 y_1' + \mu_2 y_2' \right] + \beta \left[\mu_1 y_1 + \mu_2 y_2 \right] = h(t)$$

agrupando

$$\mu_1' y_1' + \mu_2' y_2' + \mu_1 \left[y_1'' + \alpha y_1' + \beta y_1 \right] + \mu_2 \left[y_2'' + \alpha y_2' + \beta y_2 \right] = h(t)$$

Pelo fato de y_1 e y_2 soluções da equação homogênea, temos,

$$\mu_1' y_1' + \mu_2' y_2' = h(t)$$

Desta forma, $\varphi(t) = \mu_1 y_1 + \mu_2 y_2$ é solução da equação diferencial dado se resolvemos

$$\mu_1' y_1 + \mu_2' y_2 = 0 \quad (2.1.17)$$

$$\mu_1' y_1' + \mu_2' y_2' = h(t) \quad (2.1.18)$$

Multiplicando (2.1.17) por y_2' e (2.1.18) por y_2 tem-se:

$$\mu_1' y_1 y_2' + \mu_2' y_2 y_2' = 0 \quad (2.1.19)$$

$$\mu_1' y_1' y_2 + \mu_2' y_2' y_2 = h(t) y_2 \quad (2.1.20)$$

Fazendo (2.1.19) menos (2.1.20), tem-se:

$$\mu_1' y_1 y_2' + \mu_2' y_2 y_2' - (\mu_1' y_1' y_2 + \mu_2' y_2' y_2 - h y_2) = 0$$

Agrupando convenientemente

$$\mu_1' y_1 y_2' + \mu_2' y_2 y_2' - \mu_1' y_1' y_2 - \mu_2' y_2' y_2 + h y_2 = 0$$

temos

$$\mu_1' \left[y_1 y_2' - y_1' y_2 \right] = -h y_2$$

Mas, ao se multiplicar (2.1.17) por y_1' e (2.1.18) por y_1 tem-se:

$$\mu_1' y_1 y_1' + \mu_2' y_2 y_1' = 0 \quad (2.1.21)$$

$$\mu_1' y_1' y_1 + \mu_2' y_2' y_1 = h(t) y_1 \quad (2.1.22)$$

Fazendo (2.1.21) menos (2.1.22), tem-se:

$$\mu_1' y_1 y_1' + \mu_2' y_2 y_1' - (\mu_1' y_1' y_1 + \mu_2' y_2' y_1 - h y_1) = 0$$

Agrupando apropriadamente

$$\mu_1' y_1 y_1' + \mu_2' y_2 y_1' - \mu_1' y_1' y_1 - \mu_2' y_2' y_1 + h y_1 = 0$$

obteremos

$$\mu_2' \left[y_1 y_2' - y_1' y_2 \right] = -h y_1$$

Segue então que:

$$\mu_1' = \frac{-hy_2}{y_1 y_2' - y_1' y_2} \quad \text{e} \quad \mu_2' = \frac{-hy_1}{y_1 y_2' - y_1' y_2}$$

Então, após integração

$$\mu_1 = \int \frac{-hy_2}{y_1 y_2' - y_1' y_2} dt \quad \text{e} \quad \mu_2 = \int \frac{-hy_1}{y_1 y_2' - y_1' y_2} dt$$

Exemplo 1. Encontre a solução geral de

$$y'' - y = t^2$$

Solução. A equação homogênea, associada é dada por

$$y'' - y = 0$$

Esta equação homogênea tem a seguinte equação característica:

$$\lambda^2 - 1 = 0$$

Encontrando as raízes de $\lambda^2 - 1 = 0$, tem-se:

$$\lambda_1 = 1 \quad \text{e} \quad \lambda_2 = -1$$

Desta forma, a solução da equação homogênea será:

$$Y(t) = c_1 e^t + c_2 e^{-t}$$

Neste caso, substituindo as constantes por funções, temos,

$$\varphi(t) = \mu_1(t)e^t + \mu_2(t)e^{-t}$$

Calculando a derivada primeira de $\varphi(t)$, tem-se:

$$\varphi'(t) = [\mu_1 e^t - \mu_2 e^{-t}] + [\mu_1' e^t + \mu_2' e^{-t}]$$

Impondo a condição $\mu_1' e^t + \mu_2' e^{-t} = 0$ e calculando a segunda derivada, obtemos

$$\varphi''(t) = \mu_1 e^t + \mu_2 e^{-t} + \mu_1' e^t - \mu_2' e^{-t}$$

Substituindo $\varphi(t)$ e $\varphi''(t)$ na equação dada, tem-se:

$$\mu_1 e^t + \mu_2 e^{-t} + \mu_1' e^t - \mu_2' e^{-t} - (\mu_1 e^t + \mu_2 e^{-t}) = t^2$$

Agrupando

$$\mu_1 e^t + \mu_2 e^{-t} + \mu_1' e^t - \mu_2' e^{-t} - \mu_1 e^t - \mu_2 e^{-t} = t^2$$

e eliminando termos semelhantes com sinal oposto,

$$\mu_1' e^t - \mu_2' e^{-t} = t^2$$

Então, o sistema que falata resolver esta dado por,

$$\begin{aligned} \mu_1' e^t + \mu_2' e^{-t} = 0 \\ \mu_1' e^t - \mu_2' e^{-t} = t^2 \end{aligned} \Rightarrow \begin{aligned} \mu_1' e^t = -\mu_2' e^{-t} \\ \mu_1' e^t - \mu_2' e^{-t} = t^2 \end{aligned}$$

Resolvendo por substituição, da segunda equação anterior obtemos

$$-\mu_2' e^{-t} - \mu_2' e^{-t} = t^2$$

portanto

$$-2\mu_2' e^{-t} = t^2 \Rightarrow \mu_2' = \frac{-t^2}{2e^{-t}} \Rightarrow \mu_2' = \frac{-1}{2}t^2 e^t$$

Integrando μ_2' para encontrar μ_2 , tem-se:

$$\mu_2 = -\left(1 - t + \frac{1}{2}t^2\right)e^t$$

Substituindo μ_2' no sistema linear para encontrar μ_1'

$$\mu_1' e^t + \mu_2' e^{-t} = 0 \text{ implica } \mu_1' e^t + \left(-\frac{1}{2}t^2 e^t\right)e^{-t} = 0$$

assim

$$\mu_1' e^t - \frac{1}{2}t^2 = 0 \Rightarrow \mu_1' e^t = \frac{1}{2}t^2 \Rightarrow \mu_1' = \frac{1}{2}t^2 e^{-t}$$

Integrando μ_1' para encontrar μ_1 , tem-se:

$$\mu_1 = -\left(1 + t + \frac{1}{2}t^2\right)e^{-t}$$

Então $\varphi(t)$ será formado por

$$\begin{aligned} \varphi(t) &= -\left(1 + t + \frac{1}{2}t^2\right)e^{-t}e^t - \left(1 - t + \frac{1}{2}t^2\right)e^t e^{-t} \\ &= -\left(1 + t + \frac{1}{2}t^2\right) - \left(1 - t + \frac{1}{2}t^2\right) \end{aligned}$$

Finalmente o método de Lagrange está concluído:

$$\varphi(t) = -1 - t - \frac{1}{2}t^2 - 1 + t - \frac{1}{2}t^2 = -2 - t^2$$

Desta forma, a solução geral da equação homogênea será dada pela expressão,

$$y(t) = Y(t) + \varphi(t)$$

Isto é, com a soma da solução geral da equação diferencial homogênea e a solução particular obtida pelo método de variação de parâmetros,

$$y(t) = c_1 e^t + c_2 e^{-t} - 2 - t^2$$

SÉRIE DE TAYLOR

Definição. Se uma função f admite uma representação em séries de potências

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$$

com raio de convergência $r > 0$, então a derivada superior $f^{(k)}(c)$ existe para todo inteiro positivo k e os coeficientes são dados por $a_n = \frac{f^{(n)}(c)}{n!}$. Assim, podemos representar a função f pela seguinte expansão,

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \dots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \dots \quad \text{Ver Swokowski [18] (1994)}$$

Observação. Para o caso especial $c=0$, a série de Taylor se reduz à série de Maclaurin.

Exemplo 1. Determine a série que represente a função exponencial natural $f(x) = e^x$ para todo número real x .

Solução. Como a função exponencial natural é de classe $C^\infty(\mathbb{R})$, então é possível calcular todas as derivadas da função e logo podem ser avaliadas no ponto zero. Utilizando a série de Maclaurin, tem-se:

$$e^x = e^0 + e^0 x + \frac{e^0}{2!} x^2 + \dots + \frac{e^0}{n!} x^n + \dots$$

Logo obtemos a expansão desejada

$$\exp(x) = e^x = 1 + x + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \dots + \frac{1}{n!} x^n + \dots = \sum_{n \geq 0} \frac{x^n}{n!}, \quad x \in \mathbb{R} \quad (2.1.23)$$

2.2. Noções de Análise Matricial

Definição. Chama-se *matriz de ordem m por n* a um quadro de $m \times n$ elementos (números, polinômios, funções, etc.) dispostos em m linhas e n colunas: Ver Steinbruch & Winterle [17] (1987)

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \cdots & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

Cada elemento da matriz é representado por a_{ij} , onde i indica a linha e j indica a coluna a que o elemento pertence.

Tipos de Matrizes.

- Matriz nula: é qualquer matriz que possui todos os elementos iguais à zero.
- Matriz linha: possui somente uma linha
- Matriz coluna: possui somente uma coluna
- Matriz quadrada: número de linhas é igual ao número de colunas.

Observação. Na matriz quadrada, os elementos a_{ij} , quando $i = j$, compõem a diagonal principal da matriz. Nesse caso, a matriz quadrada ainda por ser subdividida em:

- Matriz identidade: os elementos da diagonal principal são iguais a 1 e os demais são iguais a 0.
- Matriz triangular: os elementos acima ou abaixo da diagonal principal são iguais a zero.
- Matriz diagonal: na diagonal principal existe pelo menos um elemento diferente de zero e os elementos que não pertencem à diagonal são iguais a zero.

Definição. Duas matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$, de ordem $m \times n$ são iguais se, e somente se, as entradas satisfazem $a_{ij} = b_{ij}$. Ver Steinbruch & Winterle [17] (1987)

Exemplo 1. Sejam as seguintes matrizes,

$$\begin{bmatrix} 1 & -2 & 1 \\ 3 & 5 & 0 \\ 4 & 7 & -5 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 5 & 0 \\ 4 & 7 & -5 \end{bmatrix}$$

Observamos que possuem as mesmas entradas, logo pela definição, são iguais.

Definição. Dada uma matriz $A = [a_{ij}]_{m \times n}$, pode-se obter outra matriz $A^T = [b_{ij}]_{n \times m}$, cujas linhas são as colunas de A , isto é, $b_{ij} = a_{ji}$. O símbolo A^T é denominado transposta da matriz A . Ver Boldrini [2] (1986)

Exemplo 1. Seja $A = \begin{bmatrix} 1 & -2 \\ 3 & 1 \\ 4 & 5 \end{bmatrix}$. Determine a transposta da matriz A .

Solução. Identificando as linhas da matriz dada A , seguindo a definição, obtemos uma nova matriz chamada de transposta,

$$A^T = \begin{bmatrix} 1 & 3 & 4 \\ -2 & 1 & 5 \end{bmatrix}$$

Algumas operações são possíveis de ser efetuadas ao se trabalhar com matrizes.

Operações com Matrizes

Definição. A soma de duas matrizes ordem, $m \times n$, $A = [a_{ij}]$ e $B = [b_{ij}]$, é uma matriz de ordem $m \times n$, que se denota por $A + B$, cujos elementos são somas dos elementos correspondentes de A e B . Isto é,

$$A + B = [a_{ij} + b_{ij}]_{m \times n} \text{ Ver Boldrini [2] (1986)}$$

Exemplo 1. Seja $A = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$, calcule $A + B$.

Solução. Primeiro identificamos o tamanho da matriz, logo efetuamos a adição entrada por entrada correspondentes,

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix} + \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} 1+2 & 3+4 \\ 5+6 & 7+8 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 11 & 15 \end{bmatrix}$$

Exemplo 2. Seja $A = \begin{bmatrix} 10 & 8 \\ 6 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 7 \\ 5 & 2 \end{bmatrix}$. Calcule $A - B$.

Solução. Fazemos uso da soma de matrizes para reescrever a operação de subtração,

$$A + B = A + (-B)$$

Então:

$$\begin{bmatrix} 10 & 8 \\ 6 & 4 \end{bmatrix} + \begin{bmatrix} -1 & -7 \\ -5 & -2 \end{bmatrix} = \begin{bmatrix} 10-1 & 8-7 \\ 6-5 & 4-2 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$$

A multiplicação, em matrizes, é dividida em multiplicação de uma matriz por um escalar e multiplicação de matrizes.

Definição. Seja $A = [a_{ij}]_{m \times n}$ de tamanho $m \times n$ e um número, k , então define-se uma nova matriz

$$k \cdot A = [ka_{ij}]_{m \times n} \text{ Ver Boldrini [2] (1986)}$$

Exemplo 3. Seja $A = \begin{bmatrix} 5 & -4 \\ 2 & 10 \end{bmatrix}$ e $k = 3$, calcule $k \cdot A$.

Solução. O efeito na matriz desta multiplicação pela constante $k=3$, é dada pela perturbação em cada entrada da matriz afetada,

$$3 \begin{bmatrix} 5 & -4 \\ 2 & 10 \end{bmatrix} = \begin{bmatrix} 3 \cdot 5 & 3 \cdot -4 \\ 3 \cdot 2 & 3 \cdot 10 \end{bmatrix} = \begin{bmatrix} 15 & -12 \\ 6 & 30 \end{bmatrix}$$

Definição. Sejam $A = [a_{ij}]_{m \times n}$ e $B = [b_{rs}]_{n \times p}$. Define-se o produto de matrizes, $AB = [c_{uv}]_{m \times p}$,

onde

$$c_{uv} = \sum_{k=1}^n a_{uk} b_{kv} = a_{u1} b_{1v} + \dots + a_{un} b_{nv} \text{ Ver Boldrini [2] (1986)}$$

Observação. Só podemos efetuar o produto de duas matrizes $A_{m \times n}$ e $B_{n \times p}$ se o número de colunas da primeira for igual ao número de linhas da segunda, isto é, $n=l$. Além disso, a matriz-resultado $C = AB$ será de ordem $m \times p$. Ver Boldrini [2] (1986)

Exemplo 4. Seja $A = \begin{bmatrix} 3 & 2 \\ 5 & 0 \\ 1 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$. Calcule AB .

Solução. Inicialmente, verifica-se se a multiplicação entre essas matrizes é possível, analisando a ordem das matrizes: $A_{3 \times 2}$ e $B_{2 \times 2}$.

Como o número de colunas de A é igual ao número de linhas de B , então é possível efetuar a multiplicação destas matrizes. O resultado será uma matriz C de ordem 3×2

$$AB = \begin{bmatrix} 3 \cdot 3 + 2 \cdot 6 & 3 \cdot 1 + 2 \cdot 2 \\ 5 \cdot 3 + 0 \cdot 6 & 5 \cdot 1 + 0 \cdot 2 \\ 1 \cdot 3 + 4 \cdot 6 & 1 \cdot 1 + 4 \cdot 2 \end{bmatrix} = \begin{bmatrix} 9 + 12 & 3 + 4 \\ 15 + 0 & 5 + 0 \\ 3 + 24 & 1 + 8 \end{bmatrix} = \begin{bmatrix} 21 & 7 \\ 15 & 5 \\ 27 & 9 \end{bmatrix}$$

$$\text{Então, } C = AB = \begin{bmatrix} 21 & 7 \\ 15 & 5 \\ 27 & 9 \end{bmatrix}$$

Determinante

Definição. Chama-se *determinante de uma matriz quadrada* à soma algébrica dos produtos que se obtém efetuando todas as permutações dos segundos índices do termo principal, fixados os primeiros índices, e fazendo-se preceder os produtos do sinal + ou -, conforme a permutação dos segundos índices seja de classe par ou de classe ímpar. Ver Steinbruch & Winterle [17] (1987)

O determinante de uma matriz será representado da seguinte maneira:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \vdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Para calcular o determinante de uma matriz quadrada de ordem n , com $n \geq 2$, será utilizado o processo da triangulação. Dessa forma, a partir de uma matriz quadrada, serão efetuados alguns cálculos de modo a tornar essa matriz dada em uma matriz triangular. O determinante desta nova matriz (o determinante de uma matriz triangular é produto dos elementos da diagonal principal), com as necessárias compensações, será o mesmo determinante da matriz original.

Para se transformar a matriz original em uma matriz triangular, poderão ser necessárias as seguintes hipóteses: Ver Steinbruch & Winterle [17] (1987)

1. O elemento é igual a zero. Nesse caso, deve-se proceder à operação de troca de linhas e multiplicar o $\det(A)$ por -1 , como compensação, isto é, para o $\det(A)$ conserve seu valor;
2. O elemento é igual a k . Nesse caso, deve-se multiplicar todos os elementos da linha por $\frac{1}{k}$, com o que se obtém o número 1 como elemento da diagonal principal dessa linha. Por outro lado, para compensar, isto é, para que $\det(A)$ mantenha seu valor, deve-se multiplicá-lo pelo inverso de $\frac{1}{k}$, isto é, por k .
3. O elemento é igual a 1. Nesse caso, nada a fazer no que diz respeito à diagonal principal.

Exemplo 1. Calcular o seguinte determinante,

$$\det A = \begin{vmatrix} 2 & 1 & 7 \\ 1 & 3 & 2 \\ 5 & 3 & 4 \end{vmatrix}$$

Solução. Utilizando o processo de triangulação para calcular o determinante da matriz A:

$$\det A = \begin{vmatrix} 2 & 1 & 7 \\ 1 & 3 & 2 \\ 5 & 3 & 4 \end{vmatrix} \xrightarrow[L_1 \leftrightarrow L_2]{-1 \det A} \det A = -1 \cdot \begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 7 \\ 5 & 3 & 4 \end{vmatrix} \xrightarrow[L_3 \rightarrow -5L_1 + L_3]{L_2 \rightarrow -2L_1 + L_2} \det A = -1 \cdot \begin{vmatrix} 1 & 3 & 2 \\ 0 & -5 & 3 \\ 0 & -12 & -6 \end{vmatrix}$$

$$\xrightarrow[L_2 \leftrightarrow L_3]{-1 \det A} \det A = -1 \cdot -1 \cdot \begin{vmatrix} 1 & 3 & 2 \\ 0 & -12 & -6 \\ 0 & -5 & 3 \end{vmatrix} \xrightarrow[-12 \det A]{L_2 \rightarrow -\frac{1}{12}L_2} \det A = -1 \cdot -1 \cdot -12 \cdot \begin{vmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{1}{2} \\ 0 & -5 & 3 \end{vmatrix}$$

$$\xrightarrow[L_3 \rightarrow 5L_2 + L_3]{-1 \det A} \det A = -1 \cdot -1 \cdot -12 \cdot \begin{vmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \frac{11}{2} \end{vmatrix} \det A = -1 \cdot -1 \cdot -12 \cdot \left[1 \cdot 1 \cdot \frac{11}{2} \right]$$

Então:

$$\det A = -66.$$

Matriz inversa

Definição. Dada uma matriz quadrada A , de ordem n , se existir uma matriz quadrada B , de mesma ordem que satisfaça a condição:

$$AB = BA = I$$

B é inversa de A e se representa por A^{-1} :

$$AA^{-1} = A^{-1}A = I \text{ Ver Steinbruch \& Winterle [17] (1987)}$$

Operações Elementares.

Definição. Denominam-se *operações elementares* de uma matriz as seguintes: Ver Steinbruch & Winterle [17] (1987)

1. Permutação de duas linhas (ou de duas colunas).
2. Multiplicação de todos os elementos de uma linha (ou coluna) por um número real diferente de zero.
3. Substituição dos elementos de uma linha (coluna) pela soma deles com os elementos correspondentes de outra linha (coluna) previamente multiplicados por um número real diferente de zero”

Para calcular a matriz inversa de uma matriz dada, seguem-se os seguintes passos:

1. Escreve-se ao lado da matriz A , a matriz I , separada por um traço vertical;
2. Transforma-se, através das operações elementares (também chamado de escalonamento), a matriz A na matriz identidade.
3. A matriz que se formará no lugar da matriz identidade, escrita originalmente, será a matriz inversa de A .

Exemplo 1. Dado $A = \begin{bmatrix} 1 & 0 & 5 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix}$, calcule a matriz inversa de A .

Solução. Calculando a inversa de A , através do escalonamento, utilizando as operações de setas

$$\begin{pmatrix} 1 & 0 & 5 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{L_2 \rightarrow -2L_1 + L_2} \begin{pmatrix} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 2 & -9 & -2 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{L_2 \leftrightarrow L_3}$$

$$\begin{pmatrix} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \\ 0 & 2 & -9 & -2 & 1 & 0 \end{pmatrix} \xrightarrow{L_3 \rightarrow -2L_2 + L_3} \begin{pmatrix} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \\ 0 & 0 & -15 & -2 & 1 & -2 \end{pmatrix} \xrightarrow{L_3 \rightarrow -\frac{1}{15}L_3}$$

$$\begin{pmatrix} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix} \xrightarrow{\begin{matrix} L_1 \rightarrow -5L_3 + L_1 \\ L_2 \rightarrow -3L_3 + L_2 \end{matrix}} \begin{pmatrix} 1 & 0 & 1 & \frac{5}{15} & \frac{5}{15} & -\frac{10}{15} \\ 0 & 1 & 0 & -\frac{6}{15} & \frac{3}{15} & \frac{9}{15} \\ 0 & 0 & 1 & \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Portanto, da última operação, temos

$$\begin{pmatrix} 1 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ 0 & 0 & 1 & \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Desta forma,

$$A^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Posto de uma Matriz:

Definição. O posto de uma matriz A é o número de linhas não nulas de qualquer matriz triangular obtida a partir de A através de operações elementares sobre linhas.

Exemplo 1. Calcule o posto da matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & 1 \\ 6 & -3 & -3 \end{bmatrix}$.

Solução. A partir da definição de posto de uma matriz, calcula-se o número de linhas não nulas na matriz triangular obtida a partir de A .

Calculando a matriz triangular:

$$\begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & 1 \\ 6 & -3 & -3 \end{bmatrix} \begin{array}{l} L_2 \rightarrow 2L_1 + L_2 \\ L_3 \rightarrow -6L_1 + L_3 \end{array} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 7 \\ 0 & -15 & -21 \end{bmatrix} \quad L_3 \rightarrow 3L_2 + L_3$$

Logo após a última operação, temos

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & 0 \end{bmatrix}$$

Desta forma, $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & 0 \end{bmatrix}$ é a matriz triangular de A .

Como a matriz triangular possui duas linhas não nula, então o posto desta matriz é 2.

Exemplo 2. Calcule o posto da matriz $B = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 3 & -1 & 5 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & -1 & 3 \end{bmatrix}$.

Solução. Calculando a matriz triangular de B :

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ 3 & -1 & 5 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & -1 & 3 \end{bmatrix} \begin{array}{l} L_2 \rightarrow -3L_1 + L_2 \\ L_4 \rightarrow -2L_1 + L_4 \end{array} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & -4 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & -2 & -5 & 3 \end{bmatrix} \quad L_2 \rightarrow -\frac{1}{4}L_2$$

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1/4 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & -2 & -5 & 3 \end{bmatrix} \begin{array}{l} L_3 \rightarrow -2L_2 + L_3 \\ L_4 \rightarrow 2L_2 + L_4 \end{array} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1/4 & 0 \\ 0 & 0 & -1/2 & 1 \\ 0 & 0 & -9/2 & 3 \end{bmatrix} L_3 \rightarrow -2L_3$$

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1/4 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -9/2 & 3 \end{bmatrix} L_4 \rightarrow \frac{9}{2}L_3 + L_4 \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1/4 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & -6 \end{bmatrix}$$

Desta forma, $\begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1/4 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & -6 \end{bmatrix}$ é a matriz triangular de B .

Como a matriz triangular possui quatro linhas não nula, então o posto desta matriz é 4.

2.3. Independência e Dependência Linear de Vetores

Seja V um espaço vetorial sobre \mathbb{R} .

Definição. Dizemos que um conjunto $L = \{u_1, u_2, \dots, u_n\} \subset V$ é *linearmente independente*

(L.I.) se, e somente se, uma igualdade do tipo

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0$$

com os α_i em \mathbb{R} , só for possível para $\alpha_1 = \dots = \alpha_n = 0$ Ver Callioli [3] (1990)

Definição. Dizemos que um conjunto $L = \{u_1, u_2, \dots, u_n\} \subset V$ é *linearmente dependente* (LD)

se, e somente se, L não é (LI), ou seja, é possível uma igualdade do tipo

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0$$

sem que os escalares α_i seja todos iguais ao número zero. Ver Callioli [3] (1990)

Exemplo 1. Determine se o conjunto de vetores

$$L = \{(1, 1, 0, 0); (0, 2, 1, 0); (0, 0, 0, 3)\} \subset \mathbb{R}^4$$

é linearmente independente ou linearmente dependente.

Solução. Devemos encontrar os valores dos coeficientes na seguinte expressão,

$$a(1, 1, 0, 0) + b(0, 2, 1, 0) + c(0, 0, 0, 3) = (0, 0, 0, 0)$$

Utilizando a propriedade de multiplicação por um escalar obtemos,

$$(a, a + 2b, b, 3c) = (0, 0, 0, 0)$$

Da igualdade de vetores, tem-se:

$$\begin{aligned} a &= 0 \\ a + 2b &= 0 \\ b &= 0 \\ 3c &= 0 \end{aligned} \Rightarrow a = b = c = 0$$

Como $a = b = c = 0$, são todas nulas, tem-se que

$$L = \{(1, 1, 0, 0); (0, 2, 1, 0); (0, 0, 0, 3)\} \subset \mathbb{R}^4$$

é linearmente independente (LI).

Exemplo 2. Determine se os seguintes vetores tridimensionais $v_1 = (1, 1, 1)$, $v_2 = (2, -1, 3)$ e $v_3 = (1, -5, 3)$ são linearmente dependentes ou linearmente independentes.

Solução. Os vetores em questão serão linearmente dependentes ou linearmente independentes conforme a seguinte equação vetorial,

$$xv_1 + yv_2 + zv_3 = 0$$

tenha uma solução não nula ou somente a solução zero. Substituindo e manipulando algebricamente, o sistema resultante escalonado tem uma variável livre, logo uma solução não trivial. Assim sendo, os vetores originais são linearmente dependentes.

2.4. Teoria Espectral Elementar

Os problemas que se apresentam podem ser descritos da seguinte forma: se referem a uma equação diferencial do tipo

$$y'(x) = f(x, y(x))$$

onde $f : I \times \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função dada e se encontra definida em alguma região $I \times \Omega$ e nosso objetivo é encontrar explicitamente uma função $y(x) : I \rightarrow \Omega$ com sua derivada $y'(x) = f(x, y(x))$, $x \in I$ e algumas condições adicionais ou, pelo menos, abastecer com a informação que seja possível sobre essa função.

O seguinte problema de encontrar uma função $y(x) : I \subset \mathbb{R} \rightarrow \mathbb{R}$ e que tenha n derivadas em I e satisfaça

$$y^{(n)}(x) = g(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \quad (2.4.1)$$

junto com algumas condições adicionais sendo $g : I \times G \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ pode ser reduzido ao problema anterior. Com efeito, se fazemos a seguinte mudança de variáveis, obtemos

$$\begin{aligned} y(t) &= x_1(t) \\ y'(t) &= x_2(t) \\ y''(t) &= x_3(t) \\ &\vdots \\ y^{(n-1)}(t) &= x_n(t) \end{aligned}$$

onde trocamos a variável independente x pelo tempo t e variável dependente por $x_j = x_j(t)$.

Como no caso de sistemas de equações algébricas, o sistema de equações diferenciais lineares ordinárias também pode ser representado na forma de matricial. Assim sendo, devemos primeiro estabelecer a derivada de uma matriz. Dado um intervalo aberto $I \subset \mathbb{R}$, uma função vetorial $X : I \rightarrow \mathbb{R}^n$ é determinada por n funções $x_j : I \rightarrow \mathbb{R}$, denominadas componentes ou coordenadas da função X e definidas pela seguinte expressão

$$X = [x_1(t), x_2(t), \dots, x_n(t)]^T, \quad t \in I \subset \mathbb{R}$$

Algumas propriedades são determinadas pelas funções coordenadas x_j . Como ilustração, a função $X(t)$ é contínua em $t_1 \in I$ se, as funções coordenadas $x_j(t)$ forem contínuas no ponto t_1 . Com relação à derivação e integração temos um resultado semelhante e são válidas as seguintes representações,

$$X'(t) = [x'_1(t), x'_2(t), \dots, x'_n(t)]^T, \quad t \in I \subset \mathbb{R}$$

Também

$$\int_a^b X(t) dt = \left[\int_a^b x_1(t) dt, \int_a^b x_2(t) dt, \dots, \int_a^b x_n(t) dt \right]^T, \quad t \in I \subset \mathbb{R}; \quad a, b \in I.$$

Essas propriedades funcionam para matrizes $n \times n$. Assim sendo, derivar uma matriz significa derivar cada entrada da matriz, de maneira semelhante para a integração.

Portanto para resolver a equação diferencial proposta (2.4.1) devemos encontrar uma função

$$X = [x_1(t), x_2(t), \dots, x_n(t)]^T, \quad t \in I \subset \mathbb{R}$$

e com valores em \mathbb{R}^n , também com derivada $X'(t): I \rightarrow \mathbb{R}^n$ tal que

$$X'(t) = f(t, X(t))$$

onde, para o elemento $\xi = [\xi_1, \xi_2, \xi_3, \dots, \xi_n]^T$ se define o valor $f(t, \xi)$ por intermédio de

$$f(t, \xi) = \begin{bmatrix} \xi_2 \\ \xi_3 \\ \vdots \\ g(t, \xi_1, \xi_2, \dots, \xi_n) \end{bmatrix}$$

Esta redução da equação diferencial linear ordinária escalar de ordem n ,

$$y^{(n)}(x) = g(x, y(x), y'(x), \dots, y^{(n-1)}(x))$$

para a equação vetorial de dimensão n dada por

$$X'(t) = f(t, X(t))$$

permite restringir as nossas considerações para esta equação.

A seguinte equação matricial, pode também representar uma equação de diferencial de primeira ordem. Felizmente, a forma da solução é equivalente ao processo anterior sempre que se tenha conhecimento de uma exponencial de uma matriz com entradas constantes.

Seja a matriz $X \in M_{3 \times 1}(\mathbb{R})$ e a matriz com entradas constantes $A \in M_3(\mathbb{R})$. Então o vetor coluna

$$X = X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}, \quad t \in I \subseteq \mathbb{R},$$

possui derivada se, as funções componentes sejam diferenciáveis e sua derivada é dada pela derivada de cada componente, isto é,

$$\frac{d}{dt} [X(t)] = \begin{bmatrix} \frac{d}{dt} x_1(t) \\ \frac{d}{dt} x_2(t) \\ \frac{d}{dt} x_3(t) \end{bmatrix}, \quad t \in I \subseteq \mathbb{R}$$

e pode ser representada por,

$$X'(t) = \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \end{bmatrix}$$

Assim, podemos construir uma equação diferencial matricial,

$$X'(t) = f(t, X(t)) = AX(t), \quad t \in I \subseteq \mathbb{R} \quad (2.4.2)$$

A vantagem principal de estudar a equação diferencial matricial acima é que existem métodos matriciais simples de resolver o sistema, pelo menos quando a matriz A possuir entradas constantes. Pela semelhança com um EDO linear escalar no caso homogêneo, podemos imaginar uma solução geral de (2.4.2), da forma

$$X(t) = Ce^{At}$$

onde $C = X(0)$ é uma matriz constante

$$X(t) = X(0)e^{At} = \begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \end{bmatrix}^T e^{At}, \quad t \in I$$

onde a exponencial possui uma matriz como expoente, que se tornará regra a partir do momento em que dermos sentido à matriz e^{At} , e esta *exponencial matricial* será definida por meio da expansão em séries de Maclaurin da função $\exp(x)$ dada em (2.1.23)

Dada a matriz quadrada A , definimos a exponencial da matriz tA como a matriz e^{tA} , de mesma ordem que A , escrita da seguinte maneira,

$$\exp(tA) = e^{tA} = I + \frac{tA}{1!} + \frac{t^2 A^2}{2!} + \frac{t^3 A^3}{3!} + \dots + \frac{t^k A^k}{k!} + \dots = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} \quad (2.4.3)$$

onde I representa a matriz quadrada identidade. Para cada matriz quadrada A de números reais e a para cada $t \in \mathbb{R}$, a série de potências de matrizes acima converge absolutamente para $\exp(tA)$

A exponencial matricial goza das seguintes propriedades básicas.

1. Para qualquer matriz quadrada $A \in M(n, \mathbb{R})$ temos,

$$\exp[(t+s)A] = \exp(tA)\exp(sA); \quad s, t \in \mathbb{R} \quad (2.4.4)$$

2. Sendo as matrizes A e B quadradas da mesma ordem e que comutam, $AB = BA$, temos,

$$\exp[t(A+B)] = \exp(tA)\exp(tB); \quad t \in \mathbb{R} \quad (2.4.5)$$

3. Para toda matriz quadrada A e para todo $t \in \mathbb{R}$, temos

$$A \exp(tA) = \exp(tA) A$$

4. Para qualquer matriz quadrada A com entradas constantes, temos

$$\frac{d}{dt} \exp(tA) = A \exp(tA) = \exp(tA) A \quad (2.4.6)$$

5. Se a matriz Z é nula ou zero, então

$$\exp(Z) = I$$

6. Para cada $t \in \mathbb{R}$, existe a inversa da matriz exponencial $\exp(tA)$ dada pela expressão

$$\exp^{-1}(tA) = \exp(-tA)$$

7. O conjunto dado por

$$B = \{\exp(tA) : t \in \mathbb{R}\}$$

munido com a operação de multiplicação de matrizes determina uma estrutura de grupo abeliano, isto que dizer o seguinte,

- a.) O par (B, \cdot) é associativo
- b.) O par (B, \cdot) possui elemento neutro
- c.) Cada elemento de B tem inverso no próprio conjunto B
- d.) O par (B, \cdot) é comutativo.

Pode-se calcular a $\exp(tA)$ em alguns casos particulares da matriz A com relativa facilidade a partir da fórmula (2.4.3). Mas é um método pouco eficiente, mesmo para matrizes quadradas de ordem dois. A seguir apresentamos um método prático para o cálculo da matriz $\exp(tA)$ a partir dos autovalores da matriz A .

A exponencial de uma matriz real tA de ordem n , pode ser obtida por várias maneiras distintas. Como preparação, vamos fazer uma revisão dessas formas,

1. Uma série infinita de potências de tA da forma,

$$\exp(tA) = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}$$

2. O método espectral, dos autovalores, isto é, devemos primeiro diagonalizar a matriz tA , logo segue facilmente da expressão anterior

$$\exp(tA) = P \exp(E) P^{-1}$$

onde a matriz diagonal $E = E(t\lambda_1, t\lambda_2, \dots, t\lambda_n)$ tem em sua diagonal os autovalores da matriz tA .

3. Pelo teorema de Hamilton-Cayley,

$$\exp(tA) = \sum_{j=0}^{n-1} \alpha_j t^j A^j$$

sendo os escalares α_j obtidos a partir da expressão,

$$\exp(t\lambda) = \sum_{j=0}^{n-1} \alpha_j (t\lambda)^j$$

para cada autovalor λ

Em EDO é preferível utilizar o Teorema de Hamilton-Cayley que simplifica muitas operações. Nesta monografia trabalharemos somente com matrizes quadradas A de ordem n que possuem autovalores reais.

Antes de prosseguir, vejamos o que entendemos por autovalores e autovetores. Considere o seguinte problema, seja a matriz $A \in M_n(\mathbb{R})$ e a seguinte equação matricial,

$$AX = \lambda X, \quad \lambda \in \mathbb{R}, \quad (2.4.7)$$

Objetivo. Encontrar os valores de λ e X que satisfazem a equação (2.4.7). Desta forma, quer-se encontrar os autovalores λ e autovetores X associados aos autovalores que satisfazem a equação (2.4.7).

Definição. Seja $T:V \rightarrow V$ um operador linear e $\lambda \in \mathbb{C}$. Dizemos que λ é um *autovalor* de T se existe $v \in V$, $v \neq 0$ tal que

$$Tv = \lambda v$$

isto é, se o operador $T - \lambda I$ não é invertível.

Se λ é autovalor de T , qualquer $v \in V$ tal que $Tv = \lambda v$ se denomina autovetor de T correspondente ao autovalor λ . Ver Boldrini [2] (1986)

O conjunto de todos os autovalores de T se denomina espectro de T e é denotado por

$$\sigma(T)$$

Os números $\lambda \in \mathbb{C} - \sigma(T)$ se denominam valores regulares do operador T , isto é, $\lambda \in \mathbb{C}$ é um valor regular de T sempre que $T - \lambda I$ seja invertível.

O espectro $\sigma(T)$ de um operador linear qualquer que seja T nunca é vazio nem possui mais de n pontos distintos, sendo n a dimensão de V .

O conjunto de autovetores de T correspondente ao autovalor λ é um subespaço de V que se denomina de autoespaço de T (correspondente a λ) e denotamos por

$$N(\lambda, 1) = \{v \in V : (T - \lambda I)^1 v = 0\}$$

Um vetor v pode ser tal que $(T - \lambda I)v = 0$ porém $(T - \lambda I)^2 v \neq 0$. Para evitar isso, introduzimos o seguinte subespaço de V para $k = 0, 1, 2, \dots$ e qualquer $\lambda \in \mathbb{C}$

$$N(\lambda, k) = \{v \in V : (T - \lambda I)^k v = 0\}$$

Verifica-se claramente a seguinte inclusão,

$$N(\lambda, k+1) \supset N(\lambda, k)$$

Além disso, se temos a igualdade

$$N(\lambda, k+1) = N(\lambda, k)$$

então, para todo $h = 2, 3, \dots$, temos a seguinte igualdade,

$$N(\lambda, k+h) = N(\lambda, k)$$

Manipulando a equação matricial

$$AX = \lambda X \quad \text{obtemos} \quad AX - \lambda X = 0$$

Colocando X em evidência e utilizando a matriz identidade I

$$(A - \lambda I)X = 0 \quad (2.4.8)$$

Em (2.4.8) encontra-se a fórmula para encontrar os autovetores de uma matriz. Sendo

$$A_\lambda = A - \lambda I \quad (2.4.9)$$

Substituindo (2.4.9) em (2.4.8), tem-se:

$$A_\lambda X = 0$$

O polinômio característico de uma matriz quadrada A de ordem n é o polinômio mônico, definido por

$$p(\lambda) = \det[A_\lambda]$$

A equação característica da matriz A é dada por

$$p(\lambda) = \det[A_\lambda] = 0$$

e para cada escalar λ que é um zero do polinômio característico é chamado de autovalor da matriz A .

Seja o vetor não nulo v e λ um autovalor da matriz A tal que $Av = \lambda v$, então o vetor v é chamado de autovetor de A associado ao autovalor λ , sendo que λ pode ser nulo.

Exemplo 1. Seja uma matriz $A \in M_2(\mathbb{R})$ com entradas constantes e $v \in \mathbb{R}^2$ onde

$$A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}_{2 \times 2} \quad \text{e} \quad v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{2 \times 1} \neq 0$$

Encontre $\lambda \in \mathbb{R}$ tal que $Av = \lambda v$

Solução. Substituindo em $Av = \lambda v$ os valores fornecidos,

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Efetando-se a multiplicação de matrizes:

$$\begin{pmatrix} 8-2 \\ 2+1 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 2\lambda \\ \lambda \end{pmatrix}$$

Pela igualdade de matrizes, tem-se:

$$\begin{aligned} 2\lambda &= 6 \\ \lambda &= 3 \end{aligned} \Rightarrow \lambda = 3$$

Desta forma, $\lambda = 3 \in \mathbb{N} \subseteq \mathbb{R}$ é um autovalor da matriz A e $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ é o autovetor associado ao autovalor $\lambda = 3$. Podemos também confirmar o cálculo acima utilizando o polinômio característico da matriz A logo, a sua *equação característica*.

Notação. Utilizamos o seguinte símbolo para representar um autopar

$$(\lambda, X) = \left(\lambda, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(3, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)$$

Sejam $\alpha, \beta \in \mathbb{R}$ e X um vetor, $X \neq 0$

$$X \in V \Rightarrow \alpha X, \beta X \in V; \alpha X \neq 0, \beta X \neq 0$$

Afirmção. Seja (λ, X) um autopar da matriz A , então,

$$A(\beta X) = \lambda(\beta X) \quad \text{para } \beta \in \mathbb{R}$$

De fato,

$$A(\beta X) = \beta AX = \beta \lambda X = \lambda \beta X = \lambda(\beta X)$$

temos que βX é um autovetor associado a λ , isto é, $(\lambda, \beta X)$ é também um autopar.

Ilustração. Seja $\beta \in \mathbb{R}$, então

$$\beta v = \beta \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\beta \\ \beta \end{bmatrix}$$

Do enunciado do problema anterior tinha-se que

$$A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}_{2 \times 2}.$$

Fazendo $A(\beta v) = \lambda(\beta v)$, tem-se:

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2\beta \\ \beta \end{pmatrix} = \lambda \begin{pmatrix} 2\beta \\ \beta \end{pmatrix}$$

Efetuando as multiplicações,

$$\begin{pmatrix} 6\beta \\ 3\beta \end{pmatrix} = \begin{pmatrix} 2\lambda\beta \\ \lambda\beta \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$\begin{aligned} 6\beta &= 2\beta\lambda \\ 3\lambda &= \beta\lambda \end{aligned} \Rightarrow \lambda = 3$$

Desta forma, $\lambda = 3 \in \mathbb{N} \subseteq \mathbb{R}$ é um autovalor da matriz A e $\beta v = \begin{pmatrix} 2\beta \\ \beta \end{pmatrix}$, que é múltiplo de $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ é também um autovetor associado ao autovalor $\lambda = 3$.

Exemplo 2. Encontre os autovalores e seus correspondentes autovetores da matriz

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix}$$

Solução. Para calcular os autovalores da matriz, calcula-se o polinômio característico:

$$p(\lambda) = \det[A_\lambda] = \det[A - \lambda I]:$$

Substituindo os valores correspondentes,

$$p(\lambda) = \det \left[\begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \det \left[\begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right] = \det \left[\begin{pmatrix} 1-\lambda & 3 \\ 4 & 3-\lambda \end{pmatrix} \right]$$

Calculando o determinante:

$$p(\lambda) = (1-\lambda)(3-\lambda) - 12 = 3 - \lambda - 3\lambda + \lambda^2 - 12 = \lambda^2 - 4\lambda - 9$$

Resolvendo a equação característica da matriz em questão, através da fórmula de Bháskara,

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Substituindo os correspondentes valores dos coeficientes,

$$\lambda = \frac{-(-4) \pm \sqrt{(-4)^2 - 4(1)(-9)}}{2(1)} \Rightarrow \lambda_1 = 2 + \sqrt{13}; \quad \lambda_2 = 2 - \sqrt{13}$$

Desta forma, os autovalores da matriz A são:

$$\lambda_1 = 2 + \sqrt{13} \quad \text{e} \quad \lambda_2 = 2 - \sqrt{13}$$

Calculando os autovetores associados a estes autovalores:

$$A_\lambda X = 0 \Rightarrow (A - \lambda I)X = 0$$

Substituindo os valores,

$$\left[\begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

temos

$$\left[\begin{pmatrix} 1-\lambda & 3 \\ 4 & 3-\lambda \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2.4.10)$$

Sendo $\lambda_1 = 2 + \sqrt{13}$, então, substituindo em (2.4.10), tem-se:

$$\left[\begin{pmatrix} 1-(2+\sqrt{13}) & 3 \\ 4 & 3-(2+\sqrt{13}) \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Logo obtemos

$$\left[\begin{pmatrix} -1-\sqrt{13} & 3 \\ 4 & 1-\sqrt{13} \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Efetuada a multiplicação:

$$\begin{pmatrix} (-1-\sqrt{13})x_1 + 3x_2 \\ 4x_1 + (1-\sqrt{13})x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$\begin{aligned} -(1+\sqrt{13})x_1 + 3x_2 &= 0 \\ 4x_1 + (1-\sqrt{13})x_2 &= 0 \end{aligned} \quad (2.4.11)$$

Isolando x_2 na primeira equação de (2.4.11), tem-se:

$$3x_2 = (1+\sqrt{13})x_1 \Rightarrow x_2 = \frac{1}{3}(1+\sqrt{13})x_1$$

Se $x_1 = 1$, temos os seguintes vetores

$$x_1 = \frac{-(1-\sqrt{13})}{4} \quad \text{e} \quad x_2 = \frac{(1+\sqrt{13})}{3}$$

Desta forma, tem-se o autovetor associado ao autovalor $\lambda_1 = 2 + \sqrt{13}$:

$$\left(\lambda_1, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(2 + \sqrt{13}, \begin{bmatrix} 1 \\ \frac{1+\sqrt{13}}{3} \end{bmatrix} \right)$$

Sendo $\lambda_2 = 2 - \sqrt{13}$, então, substituindo em (2.4.10), tem-se:

$$\left[\begin{pmatrix} 1-(2-\sqrt{13}) & 3 \\ 4 & 3-(2-\sqrt{13}) \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Obtemos

$$\begin{bmatrix} -1+\sqrt{13} & 3 \\ 4 & 1+\sqrt{13} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Efetuada a multiplicação:

$$\begin{pmatrix} (-1+\sqrt{13})x_1 + 3x_2 \\ 4x_1 + (1+\sqrt{13})x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$\begin{aligned} (-1+\sqrt{13})x_1 + 3x_2 &= 0 \\ 4x_1 + (1+\sqrt{13})x_2 &= 0 \end{aligned} \tag{2.4.12}$$

Isolando x_2 na primeira equação de (2.4.12), tem-se:

$$3x_2 = -(-1+\sqrt{13})x_1 \quad \text{logo} \quad x_2 = \frac{(1-\sqrt{13})x_1}{3}$$

Se $x_1 = 1$, então,

$$x_2 = \frac{1-\sqrt{13}}{3}$$

Desta forma, tem-se o autovetor associado ao autovalor $\lambda_2 = 2 - \sqrt{13}$:

$$\left(\lambda_2, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(2-\sqrt{13}, \begin{bmatrix} 1 \\ \frac{1-\sqrt{13}}{3} \end{bmatrix} \right)$$

Assim, a matriz X , formada pelos autovetores associados aos autovalores λ_1 e λ_2 é dada por

$$X = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{13}}{3} & \frac{1-\sqrt{13}}{3} \end{pmatrix}$$

Diagonalização de Matrizes

Definição. Seja V um espaço vetorial de dimensão finita. Um operador $T: V \rightarrow V$ se diz *diagonalizável* se existe uma base de V formada por vetores próprios de T . Ver Callioli [3] (1990)

Teorema. Se $T : V \rightarrow V$ é diagonalizável, então existem matrizes D e M tal que

$$[T]_{can} = X \cdot D \cdot X^{-1}$$

onde D é uma matriz diagonal formada pelos autovalores de T e X é uma matriz formada pelos autovetores de T .

Exemplo 1. $A \in M_2(\mathbb{R})$, $A = \begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix}$. Diagonalizar a matriz A .

Solução. Pela definição, para diagonalizar a matriz A , tem-se que:

$$A = XDX^{-1},$$

onde D é uma matriz diagonal formada pelos autovalores de A e X é uma matriz formada pelos autovetores de A . Anteriormente, calculou-se os autovalores e autovetores de A , de modo que:

$$D = \begin{pmatrix} 2 + \sqrt{13} & 0 \\ 0 & 2 - \sqrt{13} \end{pmatrix} \text{ e } X = \begin{pmatrix} 1 & 1 \\ \frac{1 + \sqrt{13}}{3} & \frac{1 - \sqrt{13}}{3} \end{pmatrix}$$

Para conseguir diagonalizar a matriz A , basta calcular a inversa de X :

Calculando a matriz inversa de X através do método de escalonamento:

$$\left(\begin{array}{ccc|cc} 1 & 1 & \vdots & 1 & 0 \\ \frac{1 + \sqrt{13}}{3} & \frac{1 - \sqrt{13}}{3} & \vdots & 0 & 1 \end{array} \right) L_2 \rightarrow -\left(\frac{1 + \sqrt{13}}{3} \right) L_1 + L_2 \left(\begin{array}{ccc|cc} 1 & 1 & \vdots & 1 & 0 \\ 0 & \frac{-2\sqrt{13}}{3} & \vdots & \frac{-1 - \sqrt{13}}{3} & 1 \end{array} \right)$$

$$L_2 \rightarrow -\frac{3}{2\sqrt{13}} L_2 \left(\begin{array}{ccc|cc} 1 & 1 & \vdots & 1 & 0 \\ 0 & 1 & \vdots & \frac{13 + \sqrt{13}}{26} & -\frac{3}{2\sqrt{13}} \end{array} \right) L_1 \rightarrow -L_2 + L_1 \left(\begin{array}{ccc|cc} 1 & 0 & \vdots & \frac{13 - \sqrt{13}}{26} & \frac{3}{2\sqrt{13}} \\ 0 & 1 & \vdots & \frac{13 + \sqrt{13}}{26} & -\frac{3}{2\sqrt{13}} \end{array} \right)$$

Desta forma, a função inversa de X é:

$$X^{-1} = \begin{pmatrix} \frac{13-\sqrt{13}}{26} & \frac{3}{2\sqrt{13}} \\ \frac{13+\sqrt{13}}{26} & -\frac{3}{2\sqrt{13}} \end{pmatrix}$$

Calculando o determinante de X :

$$\det X = \det \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{13}}{3} & \frac{1-\sqrt{13}}{3} \end{pmatrix} = \frac{1-\sqrt{13}}{3} - \left(\frac{1+\sqrt{13}}{3} \right) = -\frac{2\sqrt{13}}{3}$$

Como o determinante é diferente de zero, então, tem-se um sistema linearmente independente. Desta forma, pode-se diagonalizar a matriz A , ou seja, escrever A na forma:

$$A = XDX^{-1},$$

onde D é uma matriz diagonal formada pelos autovalores de A . Então:

$$A = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{13}}{3} & \frac{1-\sqrt{13}}{3} \end{pmatrix} \begin{pmatrix} 2+\sqrt{13} & 0 \\ 0 & 2-\sqrt{13} \end{pmatrix} \begin{pmatrix} \frac{13-\sqrt{13}}{26} & \frac{3}{2\sqrt{13}} \\ \frac{13+\sqrt{13}}{26} & -\frac{3}{2\sqrt{13}} \end{pmatrix}$$

Exemplo 2. Seja a matriz $A \in M_2(\mathbb{R})$, dada por

$$A = \begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}.$$

Diagonalizar a matriz A . Calcule A^3 .

Solução. Calculando os autovalores de A :

$$p(\lambda) = \det[A - \lambda I] = \det \left[\begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \det \left[\begin{pmatrix} 2-\lambda & -3 \\ 2 & -5-\lambda \end{pmatrix} \right]$$

Logo o polinômio característico,

$$p(\lambda) = (2-\lambda)(-5-\lambda) + 6 = -10 - 2\lambda + 5\lambda + \lambda^2 + 6$$

Assim sendo,

$$p(\lambda) = \lambda^2 + 3\lambda - 4$$

Calculando $\lambda^2 + 3\lambda - 4 = 0$, onde $a = 1$, $b = 3$ e $c = -4$, tem-se:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-3 \pm \sqrt{3^2 - 4(1)(-4)}}{2(1)} = \frac{-3 \pm \sqrt{9+16}}{2}$$

portanto

$$\lambda = \frac{-3 \pm \sqrt{25}}{2} \quad \text{e} \quad \lambda = \frac{-3 \pm 5}{2}$$

Portanto,

$$\lambda_1 = \frac{-3+5}{2} = \frac{2}{2} = 1 \quad \text{e} \quad \lambda_2 = \frac{-3-5}{2} = \frac{-8}{2} = -4$$

Desta forma, os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = -4$.

Calculando os autovetores associados aos autovalores:

$$A_\lambda X = 0 \quad \Rightarrow \quad (A - \lambda I)X = 0$$

Substituindo

$$\left[\begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \left[\begin{pmatrix} 2-\lambda & -3 \\ 2 & -5-\lambda \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Substituindo λ por $\lambda_1 = 1$, tem-se:

$$\begin{bmatrix} 2-1 & -3 \\ 2 & -5-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Efetuada as multiplicações, tem-se:

$$\begin{bmatrix} x_1 - 3x_2 \\ 2x_1 - 6x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{matrix} x_1 - 3x_2 = 0 \\ 2x_1 - 6x_2 = 0 \end{matrix} \quad \Rightarrow \quad x_1 = 3x_2$$

O autovetor associado a $\lambda_1 = 1$ é:

$$\left(\lambda_1, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(1, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right)$$

Substituindo λ por $\lambda_2 = -4$, tem-se:

$$\begin{bmatrix} 2-(-4) & -3 \\ 2 & -5-(-4) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 6 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Efetuada as multiplicações, tem-se:

$$\begin{bmatrix} 6x_1 - 3x_2 \\ 2x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{matrix} 6x_1 - 3x_2 = 0 \\ 2x_1 - x_2 = 0 \end{matrix} \quad \text{logo} \quad x_2 = 2x_1$$

O autovetor associado a $\lambda_2 = -4$ é:

$$\left(\lambda_2, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(-4, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right)$$

Desta forma, a matriz X , formada pelos autovetores de A é:

$$X = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$$

Calculando o determinante de X :

$$\det X = 6 - 2 = 4$$

Como o determinante de X é diferente de zero, então a matriz X possui inversa. Calculando a inversa de X , através do escalonamento:

$$\begin{pmatrix} 3 & 1 & : & 1 & 0 \\ 1 & 2 & : & 0 & 1 \end{pmatrix} L_1 \leftrightarrow L_2 \quad \begin{pmatrix} 1 & 2 & : & 0 & 1 \\ 3 & 1 & : & 1 & 0 \end{pmatrix} L_2 \rightarrow -3L_1 + L_2$$

$$\begin{pmatrix} 1 & 2 & : & 0 & 1 \\ 0 & -5 & : & 1 & -3 \end{pmatrix} L_2 \rightarrow -\frac{1}{5}L_2 \quad \begin{pmatrix} 1 & 2 & : & 0 & 1 \\ 0 & 1 & : & -\frac{1}{5} & \frac{3}{5} \end{pmatrix} L_1 \rightarrow -2L_2 + L_1 \quad \begin{pmatrix} 1 & 0 & : & \frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & : & -\frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

Desta forma,

$$X^{-1} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

Então a matriz A pode ser escrita da forma $A = XDX^{-1}$. Então:

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

Para calcular A^2 , tem-se:

$$A^2 = XDX^{-1}XDX^{-1}$$

Mas $X^{-1}X = I$. Então:

$$A^2 = XDIX^{-1}$$

Mas $DI = D$. Então:

$$A^2 = XDDX^{-1}, \quad A^2 = XD^2X^{-1}$$

Para calcular A^3 , tem-se:

$$A^3 = A^2A$$

$$A^3 = XD^2X^{-1}XDX^{-1} = XD^2IDX^{-1} = XD^2DX^{-1}$$

logo a potência cúbica é,

$$A^3 = XD^3X^{-1}$$

Por indução matemática,

$$A^n = XD^nX^{-1}, \quad n \in \mathbb{N}, \quad (2.4.13)$$

Mas,

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Então, D^2 , será:

$$D^2 = DD = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \Rightarrow D^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}$$

Por indução matemática,

$$D^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \quad (2.4.14)$$

Substituindo (2.4.14) em (2.4.13), tem-se:

$$A^n = X \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} X^{-1}$$

Pode-se, ainda escrever a matriz D como:

$$D = X^{-1}AX$$

Calculando A^3 :

$$A^3 = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} (1)^3 & 0 \\ 0 & (-4)^3 \end{pmatrix} \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -64 \end{pmatrix} \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix} \quad \text{logo} \quad A^3 = \begin{pmatrix} 3 & -64 \\ 1 & -128 \end{pmatrix} \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

Portanto

$$A^3 = \begin{pmatrix} \frac{6}{5} + \frac{64}{5} & -\frac{3}{5} - \frac{192}{5} \\ \frac{2}{5} + \frac{128}{5} & -\frac{1}{5} - \frac{384}{5} \end{pmatrix} \Rightarrow A^3 = \begin{pmatrix} 14 & -39 \\ 26 & -77 \end{pmatrix}$$

Exemplo 3. Seja a matriz $A \in M_3(\mathbb{R})$, dada por

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 2 & 0 & -2 \\ 2 & -1 & 1 \end{pmatrix}.$$

Diagonalizar a matriz A , logo calcule A^3 .

Solução. Calculando os autovalores de A :

$$p(\lambda) = \det[A - \lambda I]$$

Substituindo os correspondentes valores,

$$p(\lambda) = \det \left[\begin{pmatrix} 3 & -1 & 2 \\ 2 & 0 & -2 \\ 2 & -1 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] = \det \left[\begin{pmatrix} 3-\lambda & -1 & 2 \\ 2 & -\lambda & -2 \\ 2 & -1 & 1-\lambda \end{pmatrix} \right]$$

Calculando o determinante:

$$p(\lambda) = -\lambda(3-\lambda)(1-\lambda) + 4 - 4 - [-4\lambda + 2(3-\lambda) - 2(1-\lambda)]$$

$$p(\lambda) = -\lambda(3-3\lambda-\lambda+\lambda^2) - [-4\lambda + 6 - 2\lambda - 2 + 2\lambda]$$

$$p(\lambda) = -3\lambda + 4\lambda^2 - \lambda^3 + 4\lambda - 4 = \lambda^3 + 4\lambda^2 - \lambda - 4$$

Calculando $\lambda^3 + 4\lambda^2 - \lambda - 4 = 0$, tem-se:

$$\lambda_1 = 1 \quad \lambda_2 = -1 \quad \lambda_3 = 4$$

Desta forma, os autovalores de A são $\lambda_1 = 1$, $\lambda_2 = -1$ e $\lambda_3 = 4$

Calculando os autovetores associados aos autovalores:

$$A_\lambda X = 0 \Leftrightarrow (A - \lambda I)X = 0$$

Substituindo as matrizes correspondentes

$$\left[\begin{pmatrix} 3 & -1 & 2 \\ 2 & 0 & -2 \\ 2 & -1 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \left[\begin{pmatrix} 3-\lambda & -1 & 2 \\ 2 & -\lambda & -2 \\ 2 & -1 & 1-\lambda \end{pmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Substituindo λ por $\lambda_1 = 1$, tem-se:

$$\begin{bmatrix} \begin{pmatrix} 3-1 & -1 & 2 \\ 2 & -1 & -2 \\ 2 & -1 & 1-1 \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{bmatrix} \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & -2 \\ 2 & -1 & 0 \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Efetuando as multiplicações, tem-se:

$$\begin{bmatrix} 2x_1 - x_2 + 2x_3 \\ 2x_1 - x_2 - 2x_3 \\ 2x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{matrix} 2x_1 - x_2 + 2x_3 = 0 \\ 2x_1 - x_2 - 2x_3 = 0 \\ 2x_1 - x_2 = 0 \end{matrix}$$

Resolvendo o sistema, tem-se:

$$x_3 = 0 \quad \text{e} \quad x_2 = 2x_1$$

O autovetor associado a $\lambda_1 = 1$ é,

$$\left(\lambda_1, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \left(1, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right)$$

Substituindo λ por $\lambda_2 = -1$, tem-se:

$$\begin{bmatrix} \begin{pmatrix} 3-(-1) & -1 & 2 \\ 2 & -(-1) & -2 \\ 2 & -1 & 1-(-1) \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{bmatrix} \begin{pmatrix} 4 & -1 & 2 \\ 2 & 1 & -2 \\ 2 & -1 & 2 \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Efetuando as multiplicações, tem-se:

$$\begin{bmatrix} 4x_1 - x_2 + 2x_3 \\ 2x_1 + x_2 - 2x_3 \\ 2x_1 - x_2 + 2x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{matrix} 4x_1 - x_2 + 2x_3 = 0 \\ 2x_1 + x_2 - 2x_3 = 0 \\ 2x_1 - x_2 + 2x_3 = 0 \end{matrix}$$

Resolvendo o sistema, tem-se:

$$x_1 = 0 \quad x_2 = 2x_3$$

O autovetor associado a $\lambda_2 = -1$ é:

$$\left(\lambda_2, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \left(-1, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right)$$

Substituindo λ por $\lambda_3 = 4$, tem-se:

$$\begin{bmatrix} \begin{pmatrix} 3-4 & -1 & 2 \\ 2 & -4 & -2 \\ 2 & -1 & 1-4 \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{bmatrix} \begin{pmatrix} -1 & -1 & 2 \\ 2 & -4 & -2 \\ 2 & -1 & -3 \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Efetuada as multiplicações, tem-se:

$$\begin{bmatrix} -x_1 - x_2 + 2x_3 \\ 2x_1 - 4x_2 - 2x_3 \\ 2x_1 - x_2 - 3x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{logo} \quad \begin{aligned} -x_1 - x_2 + 2x_3 &= 0 \\ 2x_1 - 4x_2 - 2x_3 &= 0 \\ 2x_1 - x_2 - 3x_3 &= 0 \end{aligned}$$

Resolvendo o sistema, tem-se:

$$x_1 = 5x_2 \quad x_3 = 3x_2$$

O autovetor associado a $\lambda_2 = 4$ é:

$$\left(\lambda_2, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \left(4, \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} \right)$$

Desta forma, a matriz X , formada pelos autovetores de A é:

$$X = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Calculando o determinante de X :

$$\det X = 6 + 0 + 10 - (0 + 1 + 0) = 15$$

Como o determinante de X é diferente de zero, então a matriz X possui inversa. Esta inversa já foi calculada anteriormente, tendo obtido como resposta

$$X^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Então a matriz A pode ser escrita da forma $A = XDX^{-1}$. Assim:

$$A = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Calculando A^3 ,

$$A^3 = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} (1)^3 & 0 & 0 \\ 0 & (-1)^3 & 0 \\ 0 & 0 & (4)^3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 64 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

Efetuada as multiplicações:

$$A^3 = \begin{pmatrix} 1 & 0 & 320 \\ 2 & -2 & 64 \\ 0 & -1 & 192 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{15} & -\frac{1}{15} & \frac{2}{15} \end{pmatrix} = \begin{pmatrix} \frac{1}{3}+0+\frac{640}{15} & \frac{1}{3}+0-\frac{320}{15} & -\frac{2}{3}+0+\frac{640}{15} \\ \frac{2}{3}+\frac{4}{5}+\frac{128}{15} & \frac{2}{3}-\frac{2}{5}-\frac{64}{15} & -\frac{4}{3}-\frac{6}{5}+\frac{128}{15} \\ \frac{2}{5}+\frac{384}{15} & -\frac{1}{5}-\frac{192}{15} & -\frac{3}{5}+\frac{384}{15} \end{pmatrix}$$

Logo a potência três da matriz é dada por

$$A^3 = \begin{pmatrix} 43 & -21 & 42 \\ 10 & -4 & 6 \\ 26 & -13 & 25 \end{pmatrix}$$

Exponencial de uma Matriz

Calculando e^A , onde A é uma matriz diagonalizável, $A \in M_2(\mathbb{R})$.

Se $e^x = f(x)$, pela série de Taylor,

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^k}{k!} + \dots$$

Então:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$e^x = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{x^k}{k!}$$

Pensando-se desta mesma forma para e^A , onde A é uma matriz diagonalizável, $A \in M_2(\mathbb{R})$:

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots + \frac{A^k}{k!} + \dots$$

Então:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

$$e^A = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{A^k}{k!}$$

Se $A = D$, então:

$$e^D = I + \frac{D}{1!} + \frac{D^2}{2!} + \frac{D^3}{3!} + \dots + \frac{D^k}{k!} + \dots \quad (2.4.15)$$

Substituindo (2.4.14) em (2.4.15), tem-se:

$$e^D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + \begin{pmatrix} \frac{\lambda_1^2}{2!} & 0 \\ 0 & \frac{\lambda_2^2}{2!} \end{pmatrix} + \begin{pmatrix} \frac{\lambda_1^3}{3!} & 0 \\ 0 & \frac{\lambda_2^3}{3!} \end{pmatrix} + \dots + \begin{pmatrix} \frac{\lambda_1^k}{k!} & 0 \\ 0 & \frac{\lambda_2^k}{k!} \end{pmatrix} + \dots$$

Efetuada a soma de matrizes:

$$e^D = \begin{pmatrix} 1 + \frac{\lambda_1}{1!} + \frac{\lambda_1^2}{2!} + \frac{\lambda_1^3}{3!} + \dots + \frac{\lambda_1^k}{k!} + \dots & 0 \\ 0 & 1 + \frac{\lambda_2}{1!} + \frac{\lambda_2^2}{2!} + \frac{\lambda_2^3}{3!} + \dots + \frac{\lambda_2^k}{k!} + \dots \end{pmatrix}$$

Aplicando limite, através da Fórmula de Taylor, tem-se:

$$e^D = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}$$

Se $A \in M_n(\mathbb{R})$ e diagonalizável, existe X inversível tal que $A = XDX^{-1}$. Logo:

$$e^A = I + \frac{1}{1!} XDX^{-1} + \frac{1}{2!} XD^2X^{-1} + \frac{1}{3!} XD^3X^{-1} + \dots + \frac{1}{k!} XD^kX^{-1} + \dots$$

$$e^A = X \left[I + \frac{D}{1!} + \frac{D^2}{2!} + \frac{D^3}{3!} + \dots + \frac{D^k}{k!} + \dots \right] X^{-1}$$

$$e^A = Xe^DX^{-1}$$

Exemplo Calcular e^A , $A \in M_2(\mathbb{R})$, onde

$$A = \begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix}$$

Solução.

Calculando os autovalores de A:

$$p(\lambda) = \det A_\lambda$$

$$p(\lambda) = \det[A - \lambda I]$$

$$p(\lambda) = \det \left[\begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right]$$

$$p(\lambda) = \det \begin{pmatrix} -2-\lambda & -6 \\ 1 & 3-\lambda \end{pmatrix}$$

$$p(\lambda) = (-2-\lambda)(3-\lambda) + 6$$

$$p(\lambda) = -6 + 2\lambda - 3\lambda + \lambda^2 + 6$$

$$p(\lambda) = -\lambda + \lambda^2$$

Resolvendo $-\lambda + \lambda^2 = 0$, tem-se:

$$\lambda_1 = 1 \text{ e } \lambda_2 = 0$$

Desta forma, $\lambda_1 = 1$ e $\lambda_2 = 0$ são os autovalores da matriz A.

Calculando os autovetores da matriz A:

$$(A - \lambda I)X = 0$$

$$\begin{pmatrix} -2-\lambda & -6 \\ 1 & 3-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (2.4.16)$$

Substituindo, em (2.4.16), λ por $\lambda_1 = 1$, tem-se:

$$\begin{pmatrix} -2-1 & -6 \\ 1 & 3-1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -3 & -6 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -3x_1 - 6x_2 \\ x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$\begin{cases} -3x_1 - 6x_2 = 0 \\ x_1 + 2x_2 = 0 \end{cases}$$

Então:

$$x_1 = -2x_2$$

Desta forma, tem-se o autovetor associado ao autovalor $\lambda_1 = 1$:

$$\left(\lambda_1, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(1, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right)$$

Substituindo, em (2.4.16), λ por $\lambda_2 = 0$, tem-se:

$$\begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -2x_1 - 6x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$\begin{cases} -2x_1 - 6x_2 = 0 \\ x_1 + 3x_2 = 0 \end{cases}$$

Então:

$$x_1 = -3x_2$$

Desta forma, tem-se o autovetor associado ao autovalor $\lambda_2 = 0$:

$$\left(\lambda_2, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \left(0, \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right)$$

Assim, existe X inversível tal que $A = XDX^{-1}$, onde

$$X = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix}$$

e determinante de X é diferente de zero, o que significa que o sistema é linearmente independente. Desta forma, calculando-se a inversa de X, através do escalonamento, tem-se:

$$\begin{pmatrix} -2 & -3 & : & 1 & 0 \\ 1 & 1 & : & 0 & 1 \end{pmatrix}$$

Fazendo $L_2 \leftrightarrow L_1$, tem-se:

$$\begin{pmatrix} 1 & 1 & : & 0 & 1 \\ -2 & -3 & : & 1 & 0 \end{pmatrix}$$

Fazendo $L_2 \rightarrow 2L_1 + L_2$, tem-se:

$$\begin{pmatrix} 1 & 1 & : & 0 & 1 \\ 0 & -1 & : & 1 & 2 \end{pmatrix}$$

Fazendo $L_2 \rightarrow -L_2$

$$\begin{pmatrix} 1 & 1 & : & 0 & 1 \\ 0 & 1 & : & -1 & -2 \end{pmatrix}$$

Fazendo $L_1 \rightarrow -L_2 + L_1$

$$\begin{pmatrix} 1 & 0 & : & 1 & 3 \\ 0 & 1 & : & -1 & -2 \end{pmatrix}$$

Desta forma, a inversa de X é:

$$X^{-1} = \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

Assim, sendo

$$A = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

Desta forma, e^A , será:

$$e^A = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^1 & 0 \\ 0 & e^0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

$$e^A = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

Efetuada as multiplicações:

$$e^A = \begin{pmatrix} -2e & -3 \\ e & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

$$e^A = \begin{pmatrix} -2e+3 & -6e+6 \\ e-1 & 3e-2 \end{pmatrix}$$

Afirmação: Os autovalores de At são os mesmos autovalores de A .

$$X'(t) = AX(t), \quad X \in \mathbb{R}^n, \quad A \in M_n(\mathbb{R})$$

$$X(0) = x_0$$

$$\frac{d}{dt} X = X'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix} = \begin{bmatrix} \frac{d}{dt} x_1(t) \\ \frac{d}{dt} x_2(t) \\ \vdots \\ \frac{d}{dt} x_n(t) \end{bmatrix}$$

Resultado:

$$X(t) = x_0 e^{At},$$

onde $e^{At} \in M_n(\mathbb{R})$ é uma função matriz e definida por uma série de Taylor:

$$e^{At} = e^{tA} = I + \frac{tA}{1!} + \frac{t^2 A^2}{2!} + \frac{t^3 A^3}{3!} + \dots + \frac{t^k A^k}{k!} + \dots$$

Objetivo: Calcular a função matriz exponencial e^{At} em termos dos “autovalores” e “autovetores” da matriz quadrada A .

Observação: Os autovalores de At são os mesmos autovalores da matriz A .

Teorema de Hamilton-Cayley

“Toda matriz quadrada A e um zero do seu polinômio característico”.

Uma forma prática de lembrar o teorema de Hamilton-Cayley é dada pela seguinte sequência

$$A \Rightarrow p(\lambda) \Rightarrow p(A) = 0$$

O Teorema garante que se A é uma matriz de ordem n e se o seu polinômio característico é dado pela expressão,

$$p(\lambda) = \sum_{j=0}^n c_j \lambda^j$$

onde $c_n = 1$, logo substituindo o símbolo λ pela matriz A teremos um polinômio de matrizes nulo,

$$p(A) = \sum_{j=0}^n c_j A^j = 0, \quad \text{onde } A^0 = I, \quad A^{r+1} = AA^r$$

Ilustração. Ilustre o teorema de Hamilton-Cayley considerando a matriz quadrada

$$A = \begin{bmatrix} 1 & 1 \\ 9 & 1 \end{bmatrix}$$

Solução. Construindo o polinômio característico,

$$p(\lambda) = \det(A - \lambda I) = \lambda^2 - 2\lambda - 8 = (\lambda + 2)(\lambda - 4)$$

E, portanto a partir da equação característica $p(\lambda) = 0$ obtemos que os autovalores são dados por $\lambda_1 = -2$ e $\lambda_2 = 4$. Pelo teorema de Hamilton-Cayley construímos o polinômio nulo na matriz A , isto é,

$$p(A) = A^2 - 2A - 8I = 0$$

A partir dessa relação podemos isolar a matriz de potência máxima em termos das outras, isto é,

$$A^2 = 2A + 8I$$

O resultado anterior pode ser usado para obter as demais potências de A como combinações lineares das matrizes I e A

Explicitamente, as potências da matriz A acima de dois serão feitas da seguinte maneira>

$$\begin{aligned} A^3 &= A \cdot A^2 = A(2A + 8I) \\ &= 2A^2 + 8A = 2(2A + 8I) + 8A \\ &= 12A + 16I. \end{aligned}$$

Podemos seguir em frente com mais potências,

$$\begin{aligned} A^4 &= A \cdot A^3 = A(12A + 16I)^2 \\ &= 12A^2 + 16A = 12(2A + 8I) + 16A \\ &= 40A + 96I. \end{aligned}$$

As potências obtidas acima garantem que todas as potências da matriz A com expoente n inteiro não negativo, podem ser escritas como combinações lineares das matrizes I e A .

Algoritmo de uma Matriz Exponencial.

Para obter a exponencial de uma matriz por intermédio de um algoritmo dado pelo teorema de Hamilton-Cayley, consideremos a matriz quadrada de ordem dois, com entradas constantes

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Os autovalores dessa matriz são as duas raízes λ_1 e λ_2 da equação característica dada pela expressão $p(\lambda) = \det(A_\lambda) = 0$, isto é, raízes do polinômio

$$\lambda^2 - (a_{11} + a_{22})\lambda + \det(A) = 0$$

e essas raízes podem ser reais distintas, reais iguais ou complexas conjugadas. Para uma matriz A de ordem dois, deduziremos passos importantes.

Temos observado que podemos escrever a exponencial da matriz como uma combinação linear finita das matrizes I e A , o que significa que devem existir coeficientes α_0 e α_1 tal que,

$$\exp(tA) = \alpha_0 I + \alpha_1 tA$$

Obteremos os coeficientes α_0 e α_1 utilizando as propriedades dos autovalores da matriz tA , para obter os autovalores da matriz exponencial $\exp(tA)$ sem utilizar os autovetores da matriz tA .

Se $t\lambda$ é um autovalor da matriz tA , então $\lambda^n t^n$ é um autovalor para a matriz $t^n A^n$. Portanto, é possível verificar que $\exp(t\lambda)$ é um autovalor para a matriz $\exp(tA)$. Isto garante que existe um autovetor da matriz exponencial $v_j \neq 0$ tal que

$$\exp(tA)v_j = \exp(t\lambda_j)v_j$$

onde os $t\lambda_j$ são os autovalores da matriz tA .

Por outro lado, substituindo na expressão anterior obtemos,

$$(\alpha_0 I + \alpha_1 tA)v_j = (\alpha_0 + \alpha_1 t\lambda_j)v_j$$

Pelo fato que os autovetores são nulos, podemos afirmar que,

$$\exp(t\lambda_j) = \alpha_0 + \alpha_1 t\lambda_j; \quad j = 1, 2$$

para cada autovalor $t\lambda_j$ da matriz tA . Isto é, os coeficientes α_0 e α_1 devem satisfazer um sistema com duas equações

$$\begin{aligned}\exp(t\lambda_1) &= \alpha_0 + \alpha_1 t\lambda_1 \\ \exp(t\lambda_2) &= \alpha_0 + \alpha_1 t\lambda_2\end{aligned}\tag{2.4.17}$$

Se $t\lambda_j$ é um autovalor de multiplicidade $k=2$, além do sistema (2.4.17) os coeficientes α_0 e α_1 também satisfazem à condição

$$\exp(t\lambda_j) = \alpha_1; \quad j = 1, 2\tag{2.4.18}$$

A relação anterior é equivalente a calcular a seguinte identidade,

$$\exp(\omega) = \frac{d}{d\omega}[\alpha_0 + \alpha_1\omega]; \quad \omega = t\lambda\tag{2.4.19}$$

no autovalor $t\lambda$ da matriz tA .

Exemplo 1. Seja a matriz A com entradas constantes reais,

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

Calcular a sua matriz exponencial e^{tA} .

Solução. Os autovalores da matriz dada são,

$$p(\lambda) = \det(A - \lambda I) = (\lambda + 3)(\lambda - 2) = 0$$

Aplicando a relação (2.4.17) para cada autovalor obtemos o sistema

$$\begin{aligned}\exp(-3t) &= \alpha_0 - 3t\alpha_1 \\ \exp(2t) &= \alpha_0 + 2t\alpha_1\end{aligned}$$

de onde segue que as constantes possuem as seguintes expressões

$$\alpha_0 = \alpha_0(t) = \frac{3}{5}e^{2t} + \frac{2}{5}e^{-3t} \quad \text{e} \quad \alpha_1 = \alpha_1(t) = \frac{1}{5t}(e^{2t} - e^{-3t})$$

Utilizando a relação da exponencial de uma matriz proposta pelo teorema de Hamilton-Cayley,

$$\exp(tA) = \alpha_0 I + \alpha_1 tA = \begin{bmatrix} \alpha_0 + 2t\alpha_1 & 0 \\ 0 & \alpha_0 - 3t\alpha_1 \end{bmatrix} = \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{-3t} \end{bmatrix}.$$

Generalização do Algoritmo de uma Matriz Exponencial

Podemos generalizar de forma natural para sistemas de ordem superior. Para ilustrar consideramos um sistema de ordem três,

$$\begin{aligned}x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3\end{aligned}$$

Neste caso, a matriz exponencial solução $\exp(tA)$ está dada pela expressão,

$$\exp(tA) = \alpha_0 I + \alpha_1 tA + \alpha_2 t^2 A^2$$

onde os coeficientes $\alpha_0, \alpha_1, \alpha_2$ satisfazem o sistema

$$\exp(t\lambda) = \alpha_0 + t\lambda\alpha_1 + t^2\lambda^2\alpha_2 \quad (2.4.20)$$

para cada autovalor $t\lambda$ da matriz tA . Caso o autovalor tenha multiplicidade $k=2$, então os coeficientes $\alpha_0, \alpha_1, \alpha_2$ além do sistema anterior devem satisfazer também

$$\exp(t\lambda) = \alpha_1 + 2t\lambda\alpha_2 \quad (2.4.21)$$

Se a multiplicidade do autovalor $t\lambda$ for $k=3$, os coeficientes $\alpha_0, \alpha_1, \alpha_2$ devem satisfazer as relações (2.4.20), (2.4.21) e também a,

$$\exp(t\lambda) = 2\alpha_2 \quad (2.4.22)$$

Para sistemas de ordem três a multiplicidade k de um autovalor $t\lambda$ pode ser um para uma raiz simples, dois para uma raiz dupla ou três para uma raiz de terceira ordem e as relações satisfeitas pelos coeficientes $\alpha_0, \alpha_1, \alpha_2$ podem ser obtidas pela identidade,

$$\exp(\omega) = \frac{d^r}{d\omega^r} [\alpha_0(t) + \omega\alpha_1(t) + \omega^2\alpha_2(t)]$$

calculada no valor próprio $\omega = t\lambda$ da matriz $\exp(tA)$ onde o índice $j=0, 1, 2$ dependendo da multiplicidade do autovalor $t\lambda$. Para um autovalor simples utilizamos $j=0$ e obtemos a relação (2.4.20); para um autovalor duplo fazemos $j=0$ e $j=1$ e obteremos as relações (2.4.20) e (2.4.21), respectivamente; finalmente, para um autovalor de terceira ordem, fazemos sucessivamente $j=0, 1, 2$, obtendo (2.4.20), (2.4.21) e (2.4.22), respectivamente.

Dessa maneira, podemos estabelecer um método para calcular a matriz exponencial e^{tA} de uma matriz quadrada de qualquer ordem.

Sistemas de EDO Homogêneos

Seja A uma matriz quadrada de ordem n , a correspondente equação vetorial toma a forma

$$X' = AX$$

onde o vetor $X(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ é uma função diferenciável em relação a variável independente.

A solução da EDO em questão neste caso é da forma,

$$X(t) = C \exp(tA)$$

onde $C = [C_1, C_2, \dots, C_n]^T$ é uma matriz de constantes.

Sistemas de EDO não Homogêneos

Consideremos o sistema não homogêneo

$$x'_1 = a_{11}x_1 + a_{12}x_2 + b_1(t)$$

$$x'_2 = a_{21}x_1 + a_{22}x_2 + b_2(t)$$

com coeficientes constantes com funções b_1 e b_2 contínuas. Na forma matricial o sistema acima é dado por,

$$X' = AX + B(t) \tag{2.4.23}$$

onde $B(t)$ é uma matriz coluna com entradas as funções b_1 e b_2 . Considerando uma condição inicial $X(t_0) = X^0$ e procedendo de maneira equivalente ao caso escalar unidimensional, multiplicamos a EDO (2.4.23) pelo fator integrante $\exp(-tA)$ e obtemos

$$\exp(-tA) X' - \exp(-tA) AX = \exp(-tA) B(t)$$

Pelas regras da derivação, resulta que o primeiro membro da relação acima é,

$$\frac{d}{dt} [\exp(-tA) X] = \exp(-tA) B(t) \tag{2.4.24}$$

Integrando a relação (2.4.24) de t_0 até t , obtemos

$$\exp(-tA) X = \exp(-t_0A) X^0 + \int_{t_0}^t \exp(-sA) B(s) ds$$

Portanto a matriz,

$$X(t) = e^{(t-t_0)A} X^0 + \int_{t_0}^t e^{(t-s)A} B(s) ds \tag{2.4.25}$$

é a solução do PVI

$$X' = AX + B(t)$$

$$X(t_0) = X^0$$

Para conseguir a fórmula (2.4.25), em nenhum momento fizemos referência à ordem das matrizes A e B . Isso quer dizer que podemos usar a fórmula para sistemas de EDOs de ordem superior a dois.

CAPÍTULO III

CONTROLE EXATO DE SISTEMAS DE DIMENSÃO FINITA

3.1. Sistemas de Controle

Um trem de massa m se encontra num ponto A de seu percurso retilíneo de A para B deslocando-se com uma velocidade v . Sua força motriz máxima é F e a sua força de frenagem máxima é f . Deseja-se colocá-lo em repouso na posição B no menor tempo possível. Qual a estratégia a seguir, isto é, qual deve ser a força μ , de deslocamento ou frenagem, que devemos aplicar em cada ponto do percurso?

Na formulação anterior podemos analisar os elementos fundamentais que aparecem num problema de controle. Considera-se um sistema S o processo mecânico, elétrico, químico, industrial, social, etc., cuja operação e funcionamento dependem de alguns controles $(\mu_1, \mu_2 \dots \mu_k)$ ou parâmetros variáveis a nosso arbítrio com algumas restrições. Tratamos de levar o processo de um estado inicial E a um estado final ou meta M .

Os diversos modos de funcionamento do sistema que o levam de E até M correspondentes aos distintos valores que fornecemos aos controles, não são iguais. Existem os que proporcionam maiores vantagens. Esta qualidade dos diversos modos de funcionamento do processo é quantificada pela correspondência de um número para cada modo de funcionamento. Esta correspondência é o *índice de funcionamento* J .

Nossa ênfase é gerenciar os controles de maneira que, ao levar o sistema de E até M , o índice seja ótimo.

Na formulação inicial, supondo que o trem esteja governado pela lei de Newton,

$$\mu(t) = m X''(t)$$

onde $\mu(t)$ é a força aplicada no instante t , positiva se é de deslocamento e negativa se é de frenagem e $X(t)$ é a distância no instante t de A à posição atual do trem no sentido de A a B . O controle é $\mu(t)$, que está à nossa disposição em cada instante. O estado inicial é a

posição A com velocidade v e a meta é a posição B com velocidade $v=0$. O *índice de funcionamento* é o tempo T que tarda o trem de A até B aplicando o controle $\mu(t)$ e se deseja escolher este controle de maneira que T seja mínimo.

O índice de funcionamento J é uma função do controle $\mu(t)$, que é, por sua vez, uma função do tempo. Portanto é um funcional chamado de funcional de *custo*. Mede a qualidade, de acordo a critérios prefixados, do controle que aplicamos.

Neste trabalho, o sistema está descrito por uma equação diferencial ordinária. Sistemas mais complicados podem ser governados por uma equação ou um sistema de equações em derivadas parciais, equações estocásticas, em diferenças finitas, equações de argumentos com retardamento, etc. O importante é conhecer a formulação matemática do processo que estudamos.

A minimização de um funcional, isto é, a procura da função ou funções u dentre o conjunto de funções \mathfrak{A} que fazem mínimo o valor de algum funcional $J : \mathfrak{A} \rightarrow \mathbb{R}$, é um problema clássico, objeto de um ramo da análise matemática denominada Cálculo de Variações, idealizada nos inícios do Cálculo Diferencial e Integral.

Os problemas considerados classicamente podem ser escritos de forma unificada e simples como segue. Dado o funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

dependente da função $y = y(x)$, onde a, b são valores fixos e F é uma função dada de três variáveis, determine se a função y faz o funcional $J(y)$ mínimo.

Os problemas da teoria do controle possuem o mesmo estilo, sendo ainda mais versáteis em sua forma. Podemos considerar a teoria do controle como continuação do cálculo de variações. Muitas técnicas do Cálculo de variações são básicas na teoria de controle.

Porém, existem problemas na teoria de controle motivados por objetivos práticos desta teoria. O primeiro que podemos examinar é o da controlabilidade do sistema. No exemplo do trem sem atrito, se o trem se desloca da posição A com uma velocidade v para posição B e a força máxima de frenagem f é muito fraca não poderemos alcançar nosso objetivo, a menos que nos permitamos atravessar B , parar mais à frente e voltar para B . Existem, estados iniciais que não são controláveis com relação à meta e um tipo de controle prefixado.

Um segundo problema muito importante, é o problema de síntese, proposto pela teoria de automação relacionado com os conceitos de controle de ciclo aberto e controle de ciclo fechado ou de retroalimentação.

Como exemplo, consideremos um mecanismo que contém uma caldeira de vapor, cuja pressão é necessário controlar, mantendo-a entre os limites a, b . Uma maneira de consegui-la é a seguinte: Um operário que opera uma torneira, observa a pressão em cada instante, abre ou fecha a torneira quando a pressão P é maior que a ou menor que b . O operário vem assim a fechar o ciclo da caldeira.

Obviamente, perante uma tarefa simples; abrir ou fechar uma torneira surge a ideia de automatizar o processo. O regulador de Watt é um mecanismo simples onde a pressão da caldeira abre a torneira se é maior que b e fecha quando é menor que a . O sistema de controle assim obtido é um *controle de ciclo fechado* ou de *retroalimentação*. A mesma variável que se deseja controlar retroalimenta o regulador ou dispositivo de controle.

Alternativamente poderíamos pensar em programar a caldeira da seguinte maneira: suponhamos que dispomos de um mecanismo de relógio, em cada instante t se regula a priori a posição $e(t)$ (aberta ou fechada) da torneira, isto é,

$$e(t) = \begin{cases} 1 & \text{aberta} \\ -1 & \text{fechada} \end{cases}$$

Suponhamos calculada a pressão P no tempo T como função da pressão no instante $t=0$ e da história da torneira (isto é, da função $e(t)$ de $t=0$ até $t=T$). Então, podemos programar $e(t)$ de maneira que a pressão se conserve sempre entre a e b . A caldeira assim programada é um sistema de controle de ciclo aberto. Não existe a comparação anterior, retroalimentação, do valor atual com o valor desejado.

Este último procedimento, tanto matematicamente como mecanicamente é complicado comparado com a simplicidade do regulador de Watt. No sistema de ciclo aberto tem-se introduzido a variável t no problema. Para saber se temos que abrir ou fechar a torneira não é necessário conhecer a hora mas sim a pressão. Por outro lado, se acontece na caldeira uma perturbação não prevista, o mecanismo do relógio continuará programado pra uma situação diferente da real, com resultados possivelmente catastróficos.

O regulador de Watt é ao invés, auto-regulado. A forma de controle exercida pelo regulador é diretamente determinada pela mesma variável passiva de regulação. O problema de síntese consiste em obter o valor do controle ótimo como uma função do estado do

sistema. Esse problema motivado por conveniências da automação constitui um dos problemas interessantes na teoria de controle cuja solução completa ainda está aberta.

Os sistemas ou processos de controle podem ser representados de maneira conveniente por meio de diagramas funcionais nos quais se visualiza o papel de cada um dos componentes do sistema. Na seguinte figura,

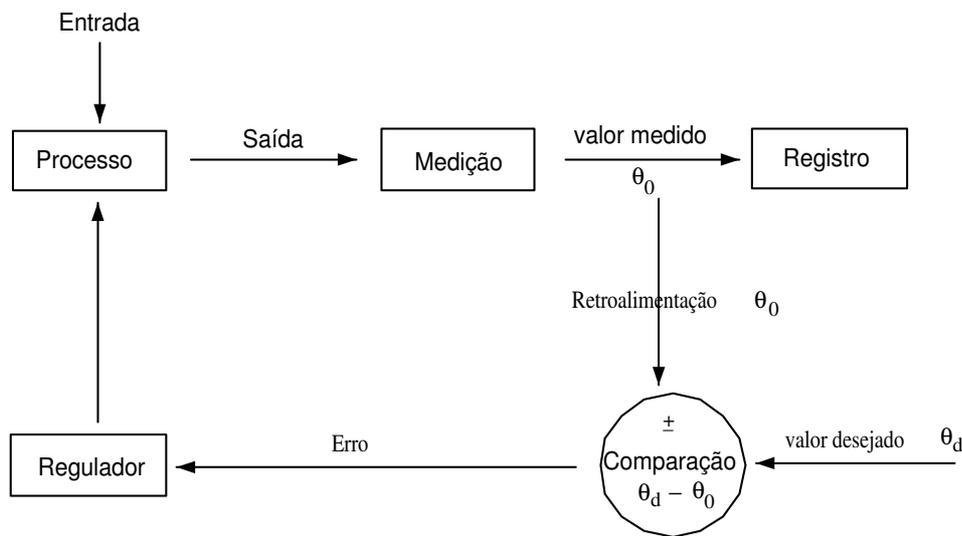


Figura 1: Processo com entrada previsível com certas restrições. (Fonte: GUZMAN, 1980, p.239)

ilustra-se um processo mecânico, físico, etc., com uma entrada previsível com certas restrições, não exata e uma saída desejável θ_d . O valor real da saída θ_0 é detectado por um componente de medida que envia um sinal para outro componente de comparação. Este mede a diferença ou erro $\theta_d - \theta_0$ e transmite um sinal ao componente controlador, que atua sobre o sistema ou processo de forma apropriada a fim de anular o erro produzido.

Uma qualidade importante e desejável de um sistema é a sua estabilidade. Dito de outro modo é necessário que as perturbações que efetuamos nos controles a fim de corrigir os erros de desvio na saída não provoquem uma alteração excessiva no sentido contrário a tal desvio. Assim sendo, o erro do processo passaria alternativamente de um sentido para outro, desvirtuando o sistema de controle.

Um sistema de controle instável pode ser visto num aprendiz de ciclista. Um pequeno erro inicial de direção e equilíbrio é corrigido com intensidade crescente, acabando implacavelmente o percurso com uma queda.

3.2. O ambiente da Teoria do Controle

A teoria de controle abrange disciplinas como automação, teoria de servomecanismos, a cibernética, o estudo da inteligência artificial nas aplicações práticas e num nível abstrato e matemático o cálculo de variações, a programação dinâmica e a disciplina de otimização.

O elemento comum dos problemas modernos propostos para a matemática pela tecnologia atual, a biologia, a física, a química, as ciências humanas e sociais, psicologia, economia, etc. é a sua enorme complexidade comparada com os problemas propostos à matemática clássica, geralmente extraídos da mecânica. Esse fato juntamente com a evolução das ciências da computação está mudando a ciência matemática atual. O desenvolvimento atual da análise numérica e da teoria da aproximação, também a expansão rápida da teoria das probabilidades e seus correspondentes ramos afins. Isto tem repercussão nos campos da matemática que, como a teoria de controle, testam apreender uma realidade tecnológica, social, econômica, muito complexa em sua formulação matemática para ser analisada com métodos da matemática clássica.

A teoria de controle apela a uma variedade de métodos, uns analíticos, outros de aproximação numérica, outros estocásticos a fim de formular e resolver os problemas propostos.

3.3. Os métodos da Teoria de Controle

Os métodos da teoria de controle são variados, dependendo da natureza dos problemas. Esses admitem uma primeira classificação em dois grupos: problemas determinísticos e problemas estocásticos.

São problemas estocásticos aqueles em cuja formulação intervêm elementos regidos por leis aleatórias. Assim, se se trata de formular um novo sistema econômico para o futuro, é claro que o sistema terá aspectos imprevisíveis regidos por leis aleatórias em grande parte. Quem formula tal plano deve supor, baseado em seu conhecimento prévio da trajetória evolutiva de outros sistemas semelhantes, a situação deste e suas reações possíveis a esta ou àquela manipulação.

Os problemas determinísticos admitem uma grande variedade. O processo que se estuda pode ser governado por um sistema de equações diferenciais ordinárias ou parciais ou em diferenças finitas. A formulação exata de um processo real conduz com frequência a

problemas matemáticos complexos por sua excessiva complicação. Cabem então duas atitudes. Pode-se fazer uma simplificação do problema de maneira que possa ser resolvido ou bem renunciar ao tratamento analítico e passar para um tratamento numérico tentando obter uma solução aproximada utilizando computadores e os métodos da análise numérica.

Os problemas que a teoria de controle encara são, com frequência, não lineares. Isto motiva a necessidade da utilização da análise não linear, ramo que se encontra em plena expansão.

O modo de proceder nos problemas do cálculo de variações tem se conservado nos modernos problemas de controle. Dentre as possíveis funções que são necessárias considerar se atinge, primeiramente descartando um conjunto, por intermédio de condições necessárias que há de satisfazer a solução ou soluções para o problema. Além disso, esta primeira seleção permite, em muitos casos, estabelecer a existência da solução ou soluções do sistema do problema. No cálculo de variações, a equação de Euler era uma dessas condições.

Na teoria de controle moderna o *princípio de otimalidade* e o *princípio do máximo de Pontryagin* fazem um papel semelhante.

O *princípio da otimalidade* é a formulação do fato óbvio de que se uma estratégia compreende vários estágios e é ótima, o é necessariamente em cada um de seus estágios. Caso contrário, poderia ser substituída por uma estratégia globalmente melhor. Em outras palavras, se para ir de A a C existe um caminho denotado por G , que é o melhor possível e sabemos que passa por B é claro que o melhor caminho possível para ir de A a B é o que segue G .

O *princípio máximo de Pontryagin* é de formulação mais complicada. Refere-se a sistemas regidos por equações diferenciais ordinárias e garante que o controle ótimo, se existe, há de fazer máxima uma expressão de significado nada óbvia. O princípio é de muita utilidade e permite em muitas situações resolver o problema em questão.

O *princípio de bang-bang* refere-se a sistemas de controle lineares regidos por equações diferenciais ordinárias onde o índice de funcionamento é o tempo. Esses problemas são chamados de problemas lineares de controle ótimo do tempo. Um controle é chamado de comutação quando os valores que assume são $+1$ e -1 alternativamente. Pode ser executado por intermédio de um comutador elétrico que admite somente duas posições.

Um controle de comutação é chamado na literatura de controle bang-bang. O princípio de bang-bang afirma que *qualquer efeito executável por um controle mensurável pode ser executado por um de comutação no mesmo tempo. Portanto, se existe um controle de tempo ótimo, existe um controle de comutação ótimo. Também, se um controle de comutação é de*

tempo ótimo com relação aos outros controles de comutação, então é ótimo. Ver LaSalle et al [8] (1961).

Aplicado à formulação do trem, proposto inicialmente, o princípio do bang-bang garante que, se existir um controle ótimo, que certamente existe neste caso, este será de comutação, isto é, o tempo ótimo se obtém acelerando umas vezes, aqui somente uma vez, com toda a força possível e freando outras vezes, aqui de novo somente uma vez, com toda a força possível.

O princípio de bang-bang como se pode observar, possui grande importância prática, pois um controle de comutação é mais simples de executar ao invés de um controle contínuo.

O estudo dos sistemas regidos por equações em derivadas parciais obtém seus princípios fundamentais da análise funcional, especialmente da análise funcional não linear. Os métodos variam muito segundo o tipo de equações que seja necessário estudar: elípticas, hiperbólicas, parabólicas, etc. São muitos os casos que exigem procedimentos próprios concebidos em particular para o problema em questão, frequentemente engenhosos procedimentos de cálculo que permitem uma solução aproximada viável e efetiva, renunciando ao tratamento analítico completo.

3.4. Controle de Tempo Mínimo para um Trem sem Atrito

Considere a seguinte situação: um trem de massa $m=1$ anda sem atrito por uma via retilínea horizontal. No instante $t_0 = 0$, o trem encontra-se no ponto A , com uma velocidade $v(0) = v_0$. O trem vai em direção ao ponto B , que fica a uma distância d de A . A força motriz do trem pode tomar qualquer valor entre 0 e 1, assim como a força de frenagem pode assumir estes mesmos valores. Deseja-se parar o trem no mínimo tempo possível. Que estratégia deve assumir o condutor para alcançar este objetivo?

Neste problema simples, teremos a oportunidade de examinar de maneira informal e não rigorosa alguns dos componentes e métodos de um problema típico de controle.

A lei matemática que rege este problema pode ser escrita da seguinte forma:

$$m x''(t) = u(t),$$

onde $m=1$, a função $x(t)$ representa a distância do trem no instante t ao ponto B , com distância negativa no ponto A , $x(0) = -d$ e a função $u(t)$ representa a força aplicada no instante t , positiva se a força motriz é na direção de A para B e negativa se vai na direção

oposta. Pelas condições impostas, ainda deve-se verificar $|\mu(t)| \leq 1$ em todo instante t e se sabe que $x'(0) = v$.

Para resolver o problema, elege-se uma função de controle $u(t)$ que leve o trem a parar em B . Desta forma, tem-se um determinado tempo $T(u)$ que tarda o trem em chegar de A ao ponto B . No entanto, o problema pede que se encontre o controle $u^*(t)$ tal que $|T(u^*)| \leq T(u)$, para qualquer outro controle $u(t)$ que satisfaça as condições dadas. Ou seja, o problema pede para se encontrar o controle que faça o trem parar em B no menor tempo possível.

Chamando $\dot{x}(t) = y(t)$, o movimento do trem passa a ser regido pelas seguintes equações,

$$\begin{aligned}x'(t) &= y(t) \\ y'(t) &= \mu(t)\end{aligned}$$

com as condições iniciais $x(0) = -d$, $y(0) = v$ e com a restrição $|\mu(t)| \leq 1$. Colocando na forma matricial, como nos casos anteriores, as matrizes correspondentes

$$X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

estão envolvidas, logo a equação matricial do sistema de controle é dado por,

$$\begin{aligned}X'(t) &= AX(t) + B\mu(t) \\ X(0) &= \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} -d \\ v \end{pmatrix}\end{aligned}$$

A solução deste sistema é calculada através da fórmula,

$$X(t) = e^{At} X^o + \int_0^t e^{A(t-s)} B\mu(s) ds$$

que depende da função controle.

Para mostrar a sua forma escalar, fazemos as seguintes substituições

$$X(t) = \begin{pmatrix} 0 & e^t \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} + \int_0^t \begin{pmatrix} 0 & e^{t-s} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \mu \end{pmatrix} ds$$

Integrando componente a componente,

$$X(t) = \begin{pmatrix} e^t y(0) \\ 0 \end{pmatrix} + \begin{pmatrix} \int_0^t \mu e^{t-s} ds \\ \int_0^t 0 ds \end{pmatrix} \Rightarrow X(t) = \begin{pmatrix} e^t y(0) \\ 0 \end{pmatrix} + \begin{pmatrix} -\mu e^0 - (-\mu e^t) \\ 0 \end{pmatrix}$$

Assim sendo, obtemos

$$X(t) = \begin{pmatrix} e^t y(0) \\ 0 \end{pmatrix} + \begin{pmatrix} -\mu + \mu e^t \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} e^t y(0) - \mu + \mu e^t \\ 0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se cada componente,

$$\begin{aligned} x_1(t) &= y(0)e^t + \mu(e^t - 1) \\ x_2(t) &= 0 \end{aligned}$$

Então, somente a primeira componente depende da função de controle. Desta forma, percebe-se que a função depende do controle $\mu(t)$. Agora, tem-se que encontrar esse controle.

Em primeiro lugar, é necessário escolher um $\mu(t)$, tal que existe T de modo que,

$$X(T) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ou seja, devemos considerar controles que satisfaçam a condição do problema, que é o trem parar em B . Obviamente, o tempo T de chegada em B depende da escolha do controle.

Trata-se, portanto, de localizar a função μ^* . Dentro do caso estudado, deve-se questionar: elegendo cada uma das possibilidades de controle, onde chega-se no tempo t ? Em outras palavras, se \mathcal{A} é o conjunto de funções $\mu(s)$ tal que $|\mu(s)| \leq 1$, como é o conjunto do plano xy

$$K(t) = \left\{ X(t) : X(t) = e^{At} X(0) + e^{At} \int_0^t e^{-As} B \mu(s) ds, \mu \in \mathcal{A} \right\} ?$$

O conjunto $K(t)$ denomina-se conjunto de acessibilidade no tempo t . O conjunto $K(t)$ contém os pontos $X(t)$ de todas as trajetórias possíveis para o trem, mediante todas as funções de controle a nossa disposição.

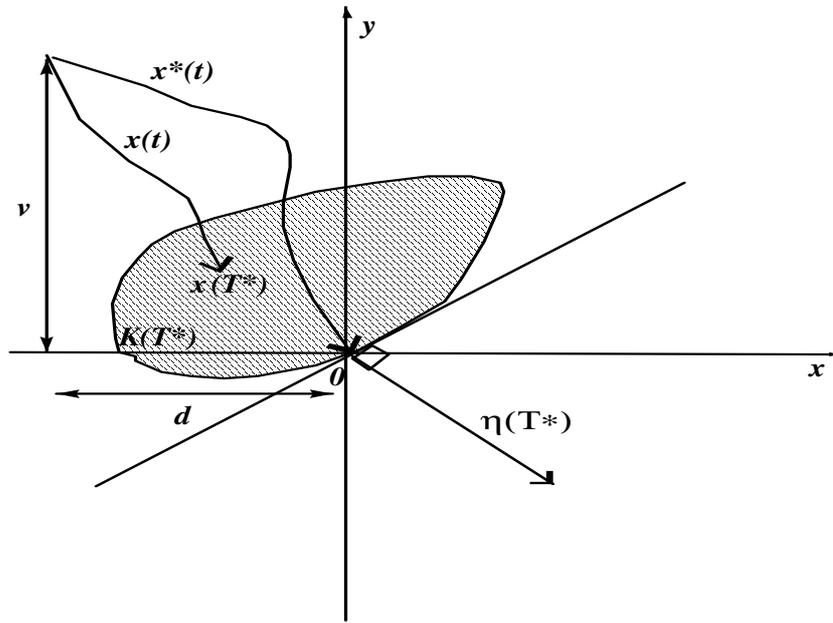


Figura 2: Trajetórias do trem (Fonte: GUZMAN, 1980, p.245)

Observando o gráfico, percebe-se que, se a origem 0 pertence ao conjunto $K(t)$, então, existe uma trajetória $X(s)$ que passa pela origem 0 , para $s = t$, determinada por um certo controle $\mu(s)$. Assim, o tempo do controle ótimo T , que procuramos, será menor ou igual a t .

Através dos conjuntos $K(t)$ é fácil assinalar qual será o tempo ótimo: aquele T tal que $0 \in K(T)$ e tal que, para qualquer $t < T$, temos $0 \notin K(t)$.

Questiona-se: como utilizar esses dados a fim de obter o controle ótimo $\mu^*(s)$ e a trajetória ótima que interessa?

Demonstra-se que os conjuntos $K(t)$ são convexos e compactos e que se deformam continuamente ao variar o tempo t . Essa leitura permite estabelecer que a origem 0 está na fronteira $\partial K(T^*)$ de $K(T^*)$ e, por este motivo, pode-se traçar uma reta que passe por 0 e deixe todo $K(T^*)$ de um mesmo lado. Desta forma, se é traçado um vetor unitário normal a esta reta de direção contrária de $K(T^*)$, vetor $\eta(T^*)$ da figura e tomando um ponto qualquer $X(T^*)$ de $K(T^*)$, tem-se, denotando por (a, b) o produto escalar dos vetores a e b ,

$$(\eta(T^*), X(T^*)) \leq 0 = (\eta(T^*), X^*(T^*))$$

Desta forma, as condições de otimização impostas implicam que, se existe a trajetória ótima $X^*(t)$, então deve haver um vetor $\eta(T^*)$, tal que $(X^*(T^*), \eta(T^*))$ é o máximo de todos os números $(X(T^*), \eta(T^*))$, obtidos ao fazer X percorrer todas as trajetórias possíveis sob a ação dos diferentes controles.

Esta explicação constitui o “Princípio de Máximo de Pontryagin”. Abaixo, será calculado como obter a trajetória e o controle ótimos.

Escrevendo-se a expressão anterior mais explicitamente:

$$(\eta(T^*), X(T^*)) = \left(\eta(T^*), e^{AT} X(0) + e^{AT} \int_0^T e^{-As} B \mu(s) ds \right)$$

Como o primeiro componente do segundo membro é independente de $\mu(s)$, maximizar $(\eta(T^*), X(T^*))$ é maximizar a expressão:

$$\left(\eta(T^*), e^{AT^*} \int_0^{T^*} e^{-As} B \mu(s) ds \right) = \int_0^{T^*} \left(\eta(T^*), e^{A(T^*-s)} B \mu(s) ds \right)$$

Sabendo-se que

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

E chamando

$$A^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{e} \quad \eta(s) = \begin{pmatrix} \eta_1(s) \\ \eta_2(s) \end{pmatrix} = e^{A^*(T^*-s)} \eta(T^*),$$

Resulta que $\mu^*(s)$ deve ser um número tal que maximize a expressão:

$$\int_0^{T^*} \eta_2(s) \mu(s) ds$$

Neste momento, fica claro como deve ser o controle $\mu(s)$ para maximizar a expressão. Recordar-se que $|\mu(s)| \leq 1$ para todo s . Se é feito o máximo $\eta_2(s) \mu(s)$ para cada s , a integral tenderá ao seu valor máximo. Desta forma:

$$\mu(s) = \text{ sinal de } \eta_2(s) = \begin{cases} +1 & \text{se } \eta_2(s) > 0 \\ -1 & \text{se } \eta_2(s) < 0 \end{cases}$$

podendo tomar $\mu(s)$ valores arbitrários ($|\mu| \leq 1$), quando $\eta_2(s) = 0$. Agora, como

$$\eta_2(s) = \eta_2(T^*) + \eta_1(T^*)(T^* - s),$$

resulta que $\eta_2(s)$ é linear e, assim, troca de sinal como máximo uma vez e é nulo no máximo num ponto. Assim, o controle de tempo ótimo procurado, $\mu^*(s)$, é de “comutação ou controle bang-bang”, o que significa dizer que assume somente os valores + 1 e - 1, neste caso, ambos sendo assumidos apenas uma vez, tendo como representação, um dos dois gráficos abaixo,

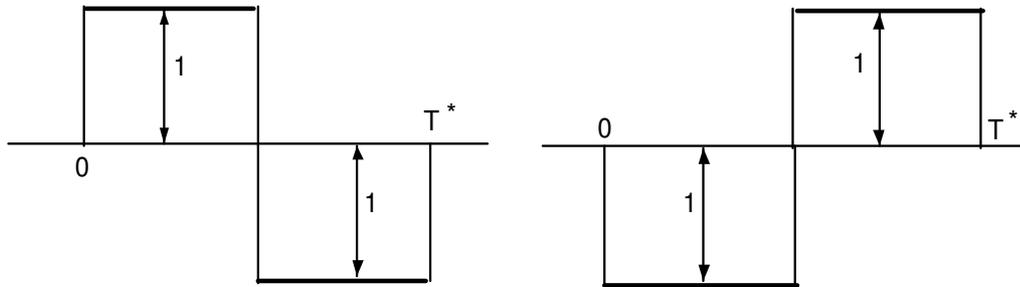


Figura 3: Gráfico do controle no tempo ótimo (Fonte: GUZMAN, 1980, p.247)

O que falta agora é determinar o ponto de troca de sinal. E, a aplicação das condições necessárias (princípio de máximo) reduzi drasticamente os controles a considerar para a situação.

Vendo como são as trajetórias $X(t)$ quando o controle é 1 em todo o momento. O sistema é assim representado:

$$\begin{aligned}x'(t) &= y(t) \\ y'(t) &= 1\end{aligned}$$

e, portanto, a solução da equação acima é dada pela forma:

$$\int x'(t) dt = \int y(t) dt \tag{3.4.1}$$

$$\int y'(t) dt = \int 1 dt \tag{3.4.2}$$

Resolvendo a equação (3.4.2), obteremos

$$\int y'(t) dt = \int 1 dt \Rightarrow y(t) = t + c_1 \tag{3.4.3}$$

Substituindo (3.4.3) em (3.4.1), tem-se:

$$\int x'(t) dt = \int y(t) dt \Rightarrow \int x'(t) dt = \int (t + c_1) dt$$

Logo, obtemos

$$x(t) = \frac{1}{2}t^2 + c_1t + c_2,$$

sendo c_1 e c_2 constantes. Desta forma, a solução é a seguinte:

$$\tau_+ = \begin{cases} x(t) = \frac{1}{2}t^2 + c_1t + c_2 \\ y(t) = t + c_1 \end{cases}$$

Essas trajetórias são parábolas, que são transladadas das seguintes parábolas,

$$\Gamma_+ = \begin{cases} x(t) = \frac{1}{2}t^2, \\ y(t) = t \end{cases},$$

como é indicado na figura seguinte:

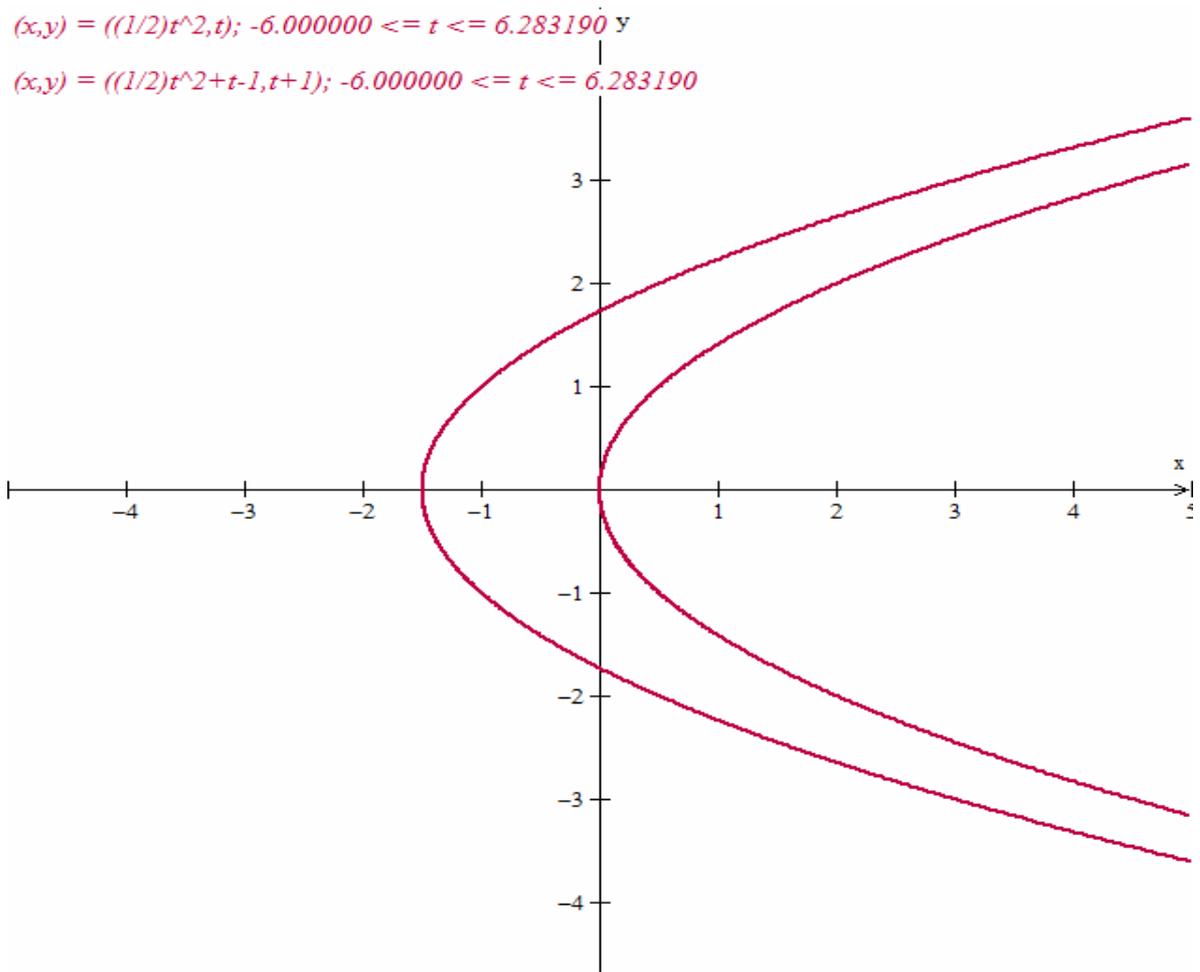


Figura 4: Trajetórias da parábola (do autor)

Analogamente, as trajetórias \mathfrak{S}_- , quando o controle é constantemente o número -1 são como indica a figura a seguir.

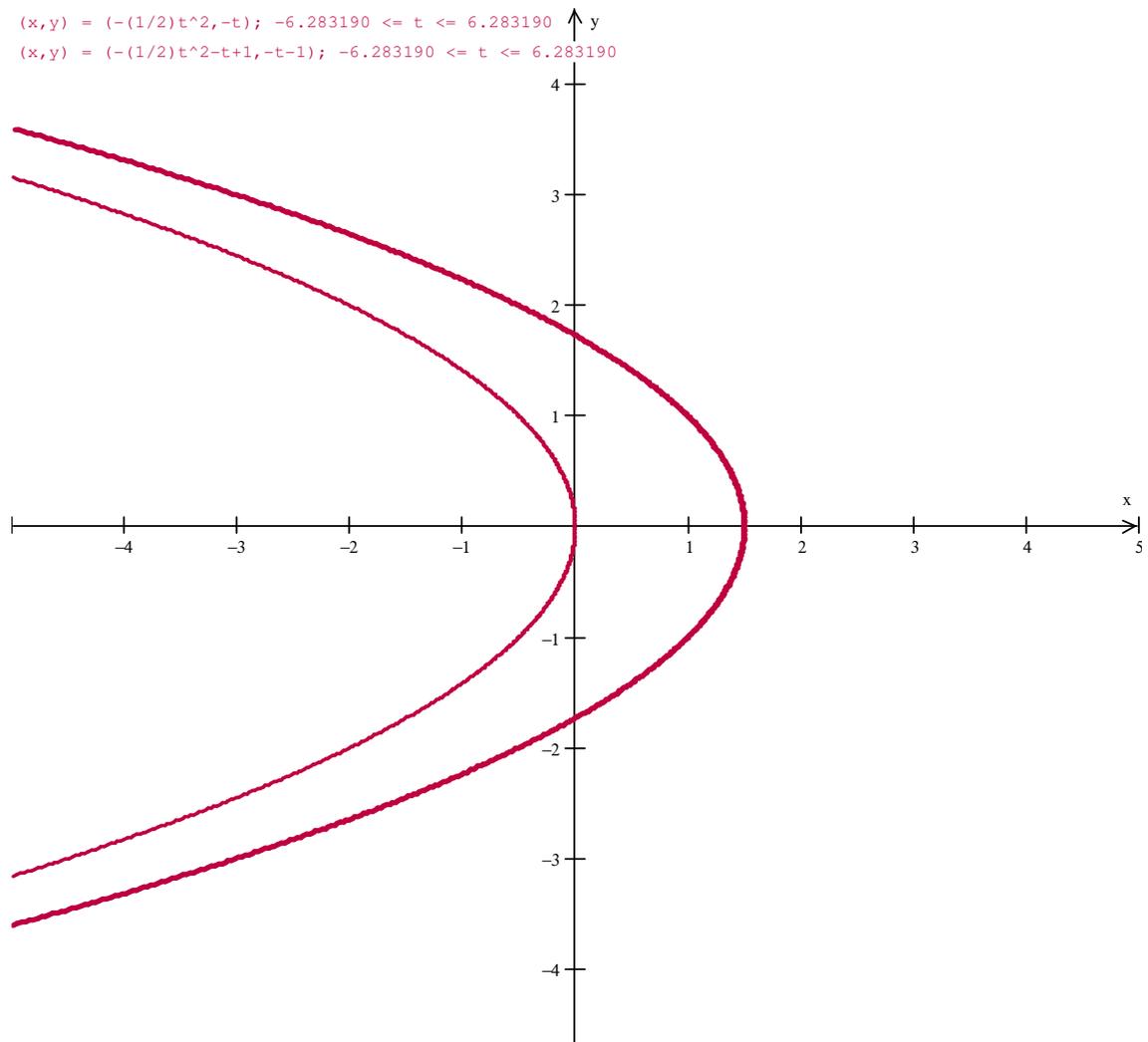


Figura 5: Trajetórias da parábola no sentido negativo (do autor)

Desta forma, sabe-se que a trajetória ótima do problema inicia em um ponto $(-d, v)$ e deve chegar à origem seguindo primeiro uma curva \mathfrak{S}_+ , e logo depois, uma curva \mathfrak{S}_- , ou ao contrário. Posto que deve acabar na origem, é claro que o fará seguindo uma das trajetórias Γ_+ ou Γ_- . Portanto, a solução está bem determinada.

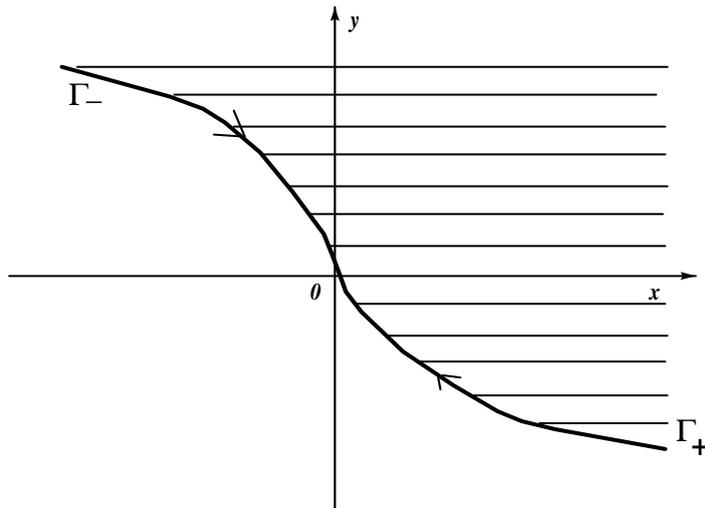


Figura 6: Solução para o problema do trem (Fonte: GUZMAN, 1980, p.248)

Se o ponto $(-d, v)$ pertence à região hachurada, a trajetória ótima segue primeiro a curva do conjunto \mathfrak{S}_- , que passa por $(-d, v)$ até cortar a Γ_+ e logo segue Γ_+ até 0. Se $(-d, v)$ se encontra na região não hachurada, a trajetória ótima segue a curva \mathfrak{S}_+ que passa por este ponto até cortar Γ_- e logo segue até 0.

Desta forma, solucionou-se o problema do trem.

3.5. O problema de Controle de um Sistema Regido por EDOs

Consideramos a seguir os elementos do problema,

- 1.) **A lei do Processo.** Relaciona o estado, resposta ou saída (output) $X(t)$, uma variável de dimensão n , que se pretende controlar com o controle ou entrada (input) $\mu(t)$, uma variável de dimensão m . Supomos que esta lei é dada pela equação (vetorial) diferencial ordinária,

$$X'(t) = f(t, X(t), \mu(t))$$

sendo $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$. A equação pode ser linear ou não linear. O objetivo final consiste em controlar o processo por intermédio de um controle de ciclo fechado e de maneira ótima com relação a um critério.

- 2.) **Estado Inicial e Conjunto Meta.** Indica-se um estado de partida por intermédio de um valor inicial do sistema

$$X^0 = X(t_0)$$

no instante t_0 . As diferentes componentes do estado $X(t)$ que se pretende controlar podem ser magnitudes tais como posição, velocidade, aceleração angular, temperatura, intensidade de corrente, etc.

O conjunto meta $G(t)$ é um conjunto dado do espaço \mathbb{R}^n de estados que variam com o tempo de maneira contínua num sentido especificado. Trata-se de levar o processo do estado X^0 para um estado $X(t_1) \in G(t)$. O conjunto $G(t)$ poderia ser um ponto $g(t)$ e o problema consistirá em controlar o erro $e(t) = X(t) - g(t)$ para zero. Neste caso o problema pode ser simplificado na seguinte representação

$$e(t) = f(t, e(t) + g(t), \mu(t) - g(t)) = f_1(t, e(t), \mu(t))$$

e portanto o conjunto de meta é em todo instante a origem.

3.) **A classe de Controles.** Devemos considerar funções contínuas por seções, definidas sobre intervalos de \mathbb{R} com valores em \mathbb{R}^m . Esta será a nossa classe de controles. Dentre eles chamaremos *controles admissíveis* a aqueles que realizam nosso objetivo fundamental, isto é, são funções contínuas por seções $\mu: [t_0, t_1] \rightarrow \mathbb{R}^m$ para algum $t_1 \geq t_0$, tais que existe uma solução $X(t)$ de

$$X'(t) = f(t, X(t), \mu(t))$$

que verifica as condições,

$$X(t_0) = X^0, \quad X(t_1) \in G(t_1)$$

dentre eles temos de escolher ainda o controle ou controles ótimos. A classe de controles admissíveis será denotada pelo símbolo \mathfrak{A} .

Em geral se imporá à classe de controles outras restrições próprias do sistema que consideramos. Isto geralmente consiste em que os valores dos controles pertencem à algum conjunto compacto Ω de \mathbb{R}^m .

4.) **O índice de Funcionamento ou funcional de custo; controle ótimo.** O funcionamento do sistema sob ação de diferentes controles é naturalmente distinto. Sua qualidade será mensurada por um critério especificado. Em geral consiste em adotar um *custo* dado por um funcional dependente de μ da forma

$$J(\mu) = \int_{t_0}^{t_1} f^0(t, X(t), \mu(t)) dt \quad \text{para todo } \mu \in \mathfrak{A}$$

sendo f^0 uma função escalar. Quando $f^0 \equiv 1$ temos que $J(\mu) = t_1 - t_0$ é um problema de controle de tempo ótimo.

Um controle será melhor que outro quando o custo correspondente seja mais baixo. Assim, um controle será ótimo quando nenhum dentre os controles admissíveis proporciona um custo mais baixo.

3.6. Controle de Sistemas Lineares.

A teoria de controle ótimo de sistemas lineares regidos por equações lineares é de grande interesse nas aplicações e com conteúdo matemático muito interessante.

Apresentamos um resultado sobre a existência e unicidade da solução de uma equação linear onde as funções dadas não são contínuas, porém são contínuas por seções.

3.7. Teorema de Existência e Unicidade.

O tipo de equação que será considerado em geral será,

$$X'(t) = A(t)X(t) + B(t)$$

onde $A: [t_0, t_1] \rightarrow M_n(\mathbb{R})$ é uma função matricial dada, contínua por seções no intervalo de definição. Isto é, existe um número finito $h+1$ de pontos no intervalo $[t_0, t_1]$,

$$t_0 = s_0 < s_1 < \dots < s_{h-1} < s_h = t_1$$

tal que $A(t)$ é contínua em cada intervalo $]s_{j-1}, s_j[$ onde $j = 1, 2, \dots, h$. Além disso, para cada j existe e é finito o limite de $A(t)$ quando t tende para s_j pela direita (se $j \neq h$) e pela esquerda (se $j \neq 0$).

A função vetorial dada,

$$B: [t_0, t_1] \rightarrow \mathbb{R}^n$$

é também contínua por seções. Podemos supor que os pontos de possível descontinuidade $s_0, s_1, \dots, s_{h-1}, s_h$ de $A(t)$ e $B(t)$ são os mesmos.

O problema consiste em encontrar uma função

$$X: [t_0, t_1] \rightarrow \mathbb{R}^n$$

contínua em todo $[t_0, t_1]$ e com derivada contínua $X'(t)$, exceto nos pontos de descontinuidade de $A(t)$, que coincide, ali onde existe, com $A(t)X(t) + B(t)$. Se impõe, além disso, a condição

$$X(t_0) = X^0 \in \mathbb{R}^n.$$

Teorema 1. O problema proposto anteriormente,

$$\begin{aligned} X'(t) &= A(t)X(t) + B(t) \\ X(t_0) &= X^0 \end{aligned}$$

admite uma única solução dada pela fórmula,

$$X(t) = \Phi(t)X^0 + \Phi(t) \int_{t_0}^{t_1} \Phi^{-1}(s)B(s) ds$$

onde a matriz $\Phi(t)$ é única matriz contínua em todo $[t_0, t_1]$ e com derivada $\Phi'(t)$ contínua em $[t_0, t_1]$ tal que,

$$\begin{aligned} \Phi'(t) &= A(t)\Phi(t) \\ \Phi(t_0) &= I \end{aligned}$$

Em particular, se $A(t)$ é constante, então $\Phi(t) = e^{A(t-t_0)}$.

A partir de agora, será explanado um pouco mais sobre o Controle de Sistemas de Dimensão Finita.

Sejam $m, n \in \mathbb{N}$ e $T > 0$

$$\begin{aligned} X' &= AX + Bu, \quad t \in]0, T[\\ X(0) &= X^0 \end{aligned}$$

$$A \in M_{\mathbb{R}}(n), B \in M_{\mathbb{R}}(n \times m), X^0 \in \mathbb{R}^n$$

A função $X : [0, T] \rightarrow \mathbb{R}^n$ referente o estado do sistema; $\mu : [0, T] \rightarrow \mathbb{R}^m$ função “controle” ambas funções vetoriais de “ n ” e “ m ” respectivamente, e dependem da variável t .

Observação: O objetivo é $m \leq n$ isto é, controlar o sistema por intermédio de um número mínimo “ m ” de controles.

3.8. Existência e Unicidade de Soluções em espaços L^2

Dado o valor inicial $X^0 \in \mathbb{R}^n$ e uma função vetorial $u \in L^2(0, T; \mathbb{R}^m)$, o sistema

$$\begin{aligned} X' &= AX + Bu(t) \\ X(0) &= X^0 \end{aligned}$$

possui uma única solução

$$X \in H^1(0, T; \mathbb{R}^n)$$

caracterizada pela fórmula

$$X(t) = e^{At} X^0 + \int_0^t e^{A(t-s)} Bu(s) ds, \quad t \in [0, T]$$

DEFINIÇÃO. O sistema

$$\begin{aligned} X' &= AX + Bu(t), \quad \forall t \in]0, T[\\ X(0) &= X^0 \end{aligned} \tag{3.8.1}$$

é exatamente controlável no tempo $T > 0$ se dado estado inicial e final, X^0 e $X^1 \in \mathbb{R}^n$, existe um controle $u \in L^2(0, T; \mathbb{R}^m)$ tal que a solução de (3.8.1) satisfaz

$$X(T) = X^1$$

OBJETIVO. De acordo com a definição, o processo do controle consiste em conduzir a solução X de (3.8.1) do estado inicial X^0 para o estado final X^1 no tempo T atuando no sistema pelo controle “ u ”.

COMENTÁRIO.

- 1) Recorde que “ m ” é o número de controles agindo no sistema, enquanto “ n ” estabelece o número de componentes do estado a ser controlado.
- 2) Alguns sistemas com grande número de componentes “ n ” podem ser controlados em um único controle (isto é, $m=1$). Mas para que isto seja verdadeiro, o mecanismo de controle, isto é, a matriz B (vetor coluna $m=1$), necessita ser escolhida de uma maneira estratégica dependendo da matriz A .

A condição dada por Kalman sobre o posto de uma matriz, parece uma simples caracterização da controlabilidade permitindo fazer uma escolha apropriada da matriz de controle B .

Exemplo 1. Considere o sistema

$$X' = AX + B\mu, \quad t \in]0, T[\tag{3.8.2}$$

onde

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mu : [0, T] \rightarrow \mathbb{R}$$

$$t \mapsto \mu(t) = \mu$$

Solução: Substituindo os dados em (3.8.2), tem-se:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mu$$

Calculando o sistema através da fórmula:

$$X(t) = e^{Dt} X^o + \int_0^t e^{D(t-s)} B \mu(s) ds$$

Então, substituindo,

$$X(t) = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix} + \int_0^t \begin{pmatrix} e^{t-s} & 0 \\ 0 & e^{t-s} \end{pmatrix} \begin{pmatrix} \mu \\ 0 \end{pmatrix} ds$$

Logo, após a integração de componente a componente,

$$X(t) = \begin{pmatrix} e^t x_1^0 \\ e^t x_2^0 \end{pmatrix} + \begin{pmatrix} \int_0^t \mu e^{t-s} ds \\ \int_0^t 0 ds \end{pmatrix} \Leftrightarrow X(t) = \begin{pmatrix} e^t x_1^0 \\ e^t x_2^0 \end{pmatrix} + \begin{pmatrix} -\mu e^0 - (-\mu e^t) \\ 0 \end{pmatrix}$$

Portanto

$$X(t) = \begin{pmatrix} e^t x_1^0 - \mu + \mu e^t \\ e^t x_2^0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} e^t x_1^0 - \mu + \mu e^t \\ e^t x_2^0 \end{pmatrix}$$

Da igualdade de matrizes, tem-se:

$$x_1(t) = x_1^0 e^t + \mu(e^t - 1)$$

$$x_2(t) = x_2^0 e^t$$

O controle não é válido, pois uma das componentes é insensível ao controle. O controle não age sobre a segunda componente x_2 do estado que é completamente determinado pelo dado inicial x_2^0 . Portanto, o sistema é parcialmente controlável. (no caso, μ não está atuando sobre a componente x_2).

Exemplo 2. Nem todos os sistemas com duas componentes e um controle escalar ($\mu = 2$ e $m = 1$) se comportam mal. Pode-se observar o oscilador harmônico controlado,

$$X'' + X = \mu$$

Solução. Reduzindo de ordem, considere

$$\begin{aligned} x' &= y \\ x'' &= y' \end{aligned} \Leftrightarrow \begin{aligned} x' &= y \\ y' &= \mu - x \end{aligned}$$

Logo, identificando o vetor solução,

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

cuja deriva é dada por,

$$X' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} y \\ \mu - x \end{pmatrix}$$

Encontrando a equação diferencial vetorial

$$X' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ \mu \end{pmatrix} \Rightarrow X' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mu$$

Desta forma, as matrizes A e B são:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Estudando as equações escalares diferenciais

$$\begin{aligned} x' &= y \\ y' &= -x + \mu \end{aligned} \Leftrightarrow \begin{aligned} x'_1 &= x_2 \\ x'_2 &= -x_1 + \mu \end{aligned}$$

Temos um controle “ μ ” para ambas componentes x e y . Agora o controle atua na segunda equação, onde as duas componentes estão presentes. Portanto, não se pode concluir dizendo que o sistema não é controlável.

De fato, é controlável. Certamente, dados algum dado inicial (x_1^0, x_2^0) e final (x_1^1, x_2^1) de outra forma em termos de (x^0, y^0) e (x^1, y^1)

Respectivamente, constrói-se uma função $z = z(t)$ tal que:

$$\begin{aligned} z(0) &= x^0 \\ z'(0) &= y^0 \end{aligned} \Leftrightarrow \begin{aligned} z(t) &= x^0 \\ z'(t) &= y^1 \end{aligned}$$

De fato, existem muitas formas de construir tais funções. Pode-se escolher uma função polinomial cúbica $z(t) = t^3$.

Pode-se definir o controle

$$\mu(t) = z''(t) + z = 6t + t^3$$

Assim sendo o sistema esta dado por,

$$X' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} X + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (6t + t^3)$$

$$X(0) = \begin{pmatrix} x_0^0 \\ y_0^0 \end{pmatrix} = X^0$$

cuja solução é dada pela fórmula

$$X(t) = e^{At} X^0 + \int_0^t e^{A(t-s)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (6s + s^3) ds$$

Esta construção fornece um exemplo de um sistema de duas componentes ($n = 2$) que é controlável por um controle apenas ($m = 1$). Mais do que isto, este exemplo mostra que o controle μ não é único. De fato existem infinitos modos de controlar e diferentes trajetórias de controles de acordo com os controles requisitados. Na prática, escolher o controle e uma otimização (em algum senso para ser mais preciso) é uma importante ferramenta que discutiremos mais tarde.

Observação 1.1. Na definição do controle exato qualquer dado inicial X^0 é necessário que seja conduzido para qualquer dado final X^1 . Pelo menos, pela linearidade do sistema, sem perda de generalidade, podemos supor que $X^1 = 0$. De fato, se $X^1 \neq 0$ podemos resolver

$$\begin{cases} y' = Ay, & t \in (0, T) \\ y(T) = x^1 \end{cases} \quad (3.8.3)$$

voltando no tempo e definindo o novo estado por $z = x - y$ que satisfaz

$$\begin{cases} z' = Az + Bu \\ z(0) = x^0 - y(0) \end{cases} \quad (3.8.4)$$

Notamos que $x(T) = x^1$ se e somente se $z(T) = 0$. Portanto, levando a solução x de (3.8.2) do dado x^0 para x^1 é equivalente a conduzir a solução z de (3.8.4) de o valor inicial $z^0 = x^0 - y(0)$ para zero.

A observação nos dá motivos para a definição:

Definição 1.2. O sistema (3.8.2) é dito como *controlável nulo* no tempo $T > 0$ se dado qualquer valor inicial $x^0 \in \mathbb{R}^n$ então existe $u \in L^2(0, T; \mathbb{R}^m)$ tal que $x(T) = 0$.

A *controlabilidade nula* vale se, e somente se, $0 \in R(x^0, T)$ para qualquer $x^0 \in \mathbb{R}^n$. Por outro lado, a Observação 1.1. anterior mostra que o controle exato e controle nulo são propriedades equivalentes no caso de sistemas lineares com dimensão finita ou para sistemas em dimensão infinita irreversíveis no tempo. Mas isto não é necessariamente o caso de sistemas não-lineares. A equação do calor é um bom exemplo de um sistema com controlabilidade nula e que não é exatamente controlável.

3.9. Propriedade da Observabilidade

A propriedade do controle exato está intimamente relacionada a uma desigualdade do correspondente sistema homogêneo adjunto. A desigualdade é chamada de observação ou desigualdade de observabilidade. Nesta seção vamos introduzir esta noção e mostraremos sua relação com a propriedade do controle exato.

Seja A^* a matriz adjunta de A , isto é, a matriz com a propriedade de que $\langle Ax, y \rangle = \langle x, A^* y \rangle$ para todo $x, y \in \mathbb{R}^n$. Consideramos o sistema adjunto homogêneo de (3.8.2)

$$\begin{cases} -\phi' = A^* \phi, & t \in (0, T) \\ \phi(T) = \phi_T \end{cases} \quad (3.9.1)$$

Lembrando que, para cada $\phi_T \in \mathbb{R}^n$, o sistema (3.9.1) pode ser resolvido voltando no tempo e tem uma única solução $\phi \in C^\omega([0, T], \mathbb{R}^n)$ (o espaço das funções analíticas definidas em $[0, T]$ e com valores em \mathbb{R}^n).

Antes de tudo, vamos deduzir uma condição equivalente para a propriedade da controlabilidade exata.

Lema 1.1. Uma condição inicial $x^0 \in \mathbb{R}^n$ de (3.8.2) é dirigido para zero no tempo T pelo uso do controle $u = v \in L^2(0, T)$ se e somente se

$$\int_0^T \langle u, B^* \phi \rangle dt + \langle x^0, \phi(0) \rangle = 0 \quad (3.9.2)$$

para qualquer $\phi_T \in \mathbb{R}^n$ sendo ϕ a correspondente solução do sistema (3.9.1).

Prova. Seja φ_T arbitrário em \mathbb{R}^n e φ a correspondente solução de (3.9.1). Multiplicando (3.8.2) por φ e (3.9.1) pela variável x , temos que

$$\langle x', \varphi \rangle = \langle Ax, \varphi \rangle + \langle Bu, \varphi \rangle; \quad -\langle x, \varphi' \rangle = \langle A^* \varphi, x \rangle.$$

Então

$$\frac{d}{dt} \langle x, \varphi \rangle = \langle Bu, \varphi \rangle$$

como, após integrando no tempo, temos que

$$\langle x(T), \varphi_T \rangle - \langle x_0, \varphi(0) \rangle = \int_0^T \langle Bu, \varphi \rangle dt = \int_0^T \langle u, B^* \varphi \rangle dt \quad (3.9.3)$$

Nós obtemos então $x(t) = 0$ se e somente se (3.9.2) é verificado para algum $\varphi_T \in \mathbb{R}^n$.

É fácil de ver que (3.9.2) é de fato uma condição ótima para pontos críticos da função quadrática $J: \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$J(\varphi_T) = \frac{1}{2} \int_0^T |B^* \varphi|^2 dt + \langle x_0, \varphi(0) \rangle$$

onde φ é a solução do sistema adjunto (3.9.1) com valor inicial φ_T no tempo $t = T$.

Mais precisamente, nós temos o seguinte resultado:

Lema 1.2.

Suponha que J tenha um mínimo $\bar{\varphi}_T \in \mathbb{R}^n$ e seja $\bar{\varphi}$ a solução do sistema adjunto (3.9.1) com os dados iniciais $\bar{\varphi}_T$. Então

$$u = B^* \bar{\varphi} \quad (3.9.4)$$

é um controle do sistema (3.8.2) com dados iniciais x^0 .

Prova. Se $\bar{\varphi}_T$ é um ponto onde J assume o seu valor mínimo, então

$$\lim_{h \rightarrow 0} \frac{J(\bar{\varphi}_T + h \varphi_T) - J(\bar{\varphi}_T)}{h} = 0, \quad \forall \varphi_T \in \mathbb{R}^n.$$

Isto é equivalente a

$$\int_0^T \langle B^* \bar{\varphi}, B^* \varphi \rangle dt + \langle x_0, \varphi(0) \rangle = 0, \quad \forall \varphi_T \in \mathbb{R}^n,$$

como vimos no Lema 1.1, implica que $u = B^* \bar{\varphi}$ é um controle para (3.8.2)

Observação 1.2.

O Lema 1.2 nos dá um método variacional para obter um controle como o mínimo do funcional J . Esta não é a única possibilidade funcional para se conseguir o controle. Modificando ele convencionalmente, outros tipos de controle (pelo momento apenas *bang-bang*) podem ser obtidos. Lembrando que os controles que encontramos são da forma $B^* \varphi$, sendo φ uma solução do problema homogêneo adjunto (3.9.1). No entanto, estas são funções analíticas sobre o tempo.

A noção a seguir será fundamental para a solução de problemas de controle.

Definição 1.3.

O sistema (3.9.1) é dito como *observável* no tempo $T > 0$ se existir um $c > 0$ tal que

$$\int_0^T |B^* \phi|^2 dt \geq c |\phi(0)|^2, \quad (3.9.5)$$

para todo $\phi_T \in \mathbb{R}^n$, sendo φ a correspondente solução de (3.9.1).

No que segue a desigualdade (3.9.5) será chamada de **observação** ou **desigualdade da observação**. Ele é a garantia de que a solução do problema adjunto para $t = 0$ é unicamente determinado por $B^* \varphi(t)$ para $0 < t < T$. Em outras palavras, as informações contidas neste termo caracterizam completamente a solução de (3.9.1).

Observação 1.3.

A observação da desigualdade (3.9.5) é equivalente ao seguinte: então existe $c > 0$ tal que

$$\int_0^T |B^* \phi|^2 dt \geq c |\phi_T|^2, \quad (3.9.6)$$

para todo $\phi_T \in \mathbb{R}^n$, sendo φ a solução de (3.9.1).

De fato, as seguintes equivalências para o fato de conseguirmos associar para todo $\phi_T \in \mathbb{R}^n$ o vetor $\phi(0) \in \mathbb{R}^n$, é uma transformação linear em \mathbb{R}^n com alguma inversa. Nós

iremos usar as formas (3.9.5) ou (3.9.6) da desigualdade da observação dependendo da necessidade que o problema particular precisará.

A seguinte observação é muito importante no contexto de espaços de dimensão finita.

Proposição 1.1.

A desigualdade (3.9.6) é equivalente ao seguinte princípio da continuação única:

$$B^* \phi(t) = 0, \quad \forall t \in [0, T] \Rightarrow \phi_T = 0 \quad (3.9.7)$$

Prova. Uma das implicações segue imediatamente de (3.9.6). Para a outra, vamos definir a semi-norma em \mathbb{R}^n

$$|\phi_T|_* = \left[\int_0^T |B^* \phi|^2 dt \right]^{\frac{1}{2}}$$

Claramente, $|\cdot|_*$ é a norma em \mathbb{R}^n se e somente se (3.9.7) é válido.

Desde que todas as normas em \mathbb{R}^n são equivalentes, a que temos em (3.9.7) é equivalente a (3.9.6). A prova termina usando a Observação 1.3.

Observação 1.4.

Lembrando que (3.9.5) e (3.9.7) que não serão propriedades equivalentes em espaços de dimensão infinita. Eles nos darão diferentes noções de controle (exato e aproximado, respectivamente).

A importância da desigualdade de observação está no fato de que implica o controle exato de (3.8.2). Por este caminho, a propriedade do controle é reduzida ao estudo de uma desigualdade para o sistema homogêneo (3.9.1) o qual pelo menos conceitualmente, é um problema mais simples. Vamos analisar agora a relação entre a controlabilidade e a propriedade da observabilidade.

Teorema 1.1.

O sistema (3.8.2) é controlável exatamente no tempo T se, e somente se, (3.9.1) é observável no tempo T .

Prova. Vamos provar primeiro que a observação implica controle. De acordo com o Lema 1.2, a propriedade do controle exato no tempo T é válida se para qualquer $x^0 \in \mathbb{R}^n$, o funcional J tem um mínimo. Lembre que J é contínuo. Consequentemente, a existência de um mínimo é garantido se o funcional J é também coercivo, isto é,

$$\lim_{|\phi_T| \rightarrow \infty} J(\phi_T) = \infty \quad (3.9.8)$$

A propriedade coerciva (3.9.8) é uma consequência da propriedade da observação no tempo T . De fato, de (3.9.5) nós obtemos que

$$J(\phi_T) \geq \frac{c}{2} |\phi_T|^2 - \left| \langle x^0, \phi(0) \rangle \right|.$$

O lado direito tende ao infinito quando $|\phi_T| \rightarrow \infty$ e J satisfaz (3.9.8). Reciprocamente suponhamos que o sistema (3.8.2) é exatamente controlável no tempo T . Se (3.9.1) não é observável no tempo T , então existe uma sequência $\{\phi_T^k\}_{k \geq 1} \subset \mathbb{R}^n$ tal que $|\phi_T^k| = 1$ para todo $k \geq 1$ e

$$\lim_{k \rightarrow \infty} \int_0^T |\mathbf{B}^* \phi^k|^2 dt = 0 \quad (3.9.9)$$

Isto quer dizer que existe uma subseqüência de $(\phi_T^k)_{k \geq 1}$, denotada de mesmo modo, que converge para $\phi_T \in \mathbb{R}^n$ e $|\phi_T| = 1$. Ainda, se ϕ é uma solução de (3.9.1) com o valor inicial ϕ_T , de (3.9.9) segue que

$$\int_0^T |\mathbf{B}^* \phi|^2 dt = 0 \quad (3.9.10)$$

Como (3.8.2) é controlável, o Lema 1.1 nos dá que, para algum valor inicial $x^0 \in \mathbb{R}^n$, então existe $u \in L^2(0, T)$ tal que

$$\int_0^T \langle u, \mathbf{B}^* \phi_k \rangle dt = - \langle x^0, \phi_k(0) \rangle, \quad (3.9.11)$$

Tomando o limite em (3.9.11) e tomando o fato de (3.9.10), nós obtemos que $\langle x^0, \phi(0) \rangle = 0$. Desde que x^0 seja arbitrário em \mathbb{R}^n , temos que $\phi(0) = 0$ e, consequentemente, $\phi_T = 0$. Isto é uma contradição com o fato de que $|\phi_T| = 1$.

A prova do teorema está completa.

Observação 1.5.

Uma consequência do Teorema 1.1 consiste no fato de que ele reduz a prova do controle exato para o estudo da desigualdade da observação.

3.10. Condições de Kalman

Para analisar o problema da controlabilidade exata existe uma condição necessária e suficiente de natureza algébrica. Ela é chamada de *Condições de Kalman*: O sistema dado em (3.8.2) é controlável em algum tempo $T > 0$ se e somente se

$$\text{posto} [B, AB, \dots, A^{n-1}B] = n \quad (3.10.1)$$

Mais ainda, quando isso é válido, o sistema em questão é controlável para qualquer tempo $T > 0$. Quando o posto dessa matriz é k com $1 \leq k \leq n$ o sistema não é controlável e para cada dado inicial $x^0 \in \mathbb{R}^n$ o conjunto de soluções de (3.8.2) em cada instante $T > 0$ percorre um subespaço afim de \mathbb{R}^n de dimensão k .

Ou em outras palavras, o sistema (3.8.2) é controlável no tempo $T > 0 \iff$ o posto da matriz, ou seja, número de linhas não nulas da matriz $[B, AB, \dots, A^{n-1}B]$ deve ser igual ao número n , elementos do vetor de estado.

Existe uma prova direta deste resultado que utiliza a representação das soluções de (3.8.2) por intermédio da fórmula de variação de constantes.

É pertinente fazer algumas observações,

- O grau de controlabilidade do sistema (3.8.2) está totalmente determinado pelo posto da matriz (3.10.1). Esse mede o número de componentes do sistema que são afetados pelo controle.
- A matriz dada em (3.10.1) é dimensão $(n \times m) \times n$ de maneira que, quando se tem um só controle ($m = 1$) a matriz possui medida $n \times n$. Neste caso é mais difícil que o posto seja n . Isto está de acordo com a intuição que nos indica que o sistema é mais facilmente controlável quando existem mais controles.
- O sistema é controlável em um tempo $T > 0$ se, e somente se, o é em todo o intervalo do tempo. Isso indica que no sistema (3.8.2) a informação se propaga a velocidade infinita, situação que deixa de ser verdade no contexto das EDP's.
- Quando se estuda em modelos de EDP's é mais conveniente utilizar um método alternativo que consiste em *transformar o problema do controle* em um problema de

observabilidade para o sistema adjunto, pois o problema adjunto pode ser resolvido por uma combinação dos métodos existentes para obter estimativas a priori sobre soluções de EDO e EDP.

Concluindo, temos um clássico resultado devido a Rudolf E. Kalman que nos dá uma completa resposta para o problema de controle exato de sistema linear de dimensão finita. Ele mostra, em particular, que o tempo do controle é irrelevante.

Teorema 1.2.

O sistema (3.8.2) é exatamente controlável em algum tempo T se e somente se

$$\text{posto} = \text{número de linhas não nulas do sistema } [B, AB, \dots, A^{n-1}B] = n \quad (3.10.2)$$

Consequentemente, se o sistema (3.8.2) é controlável em algum tempo $T > 0$ então ele é controlável em qualquer tempo.

Observação 1.6.

De agora em diante simplesmente diremos que (A, B) é controlável se (3.10.2) for válido. A matriz $[B, AB, \dots, A^{n-1}B]$ será chamada de *matriz de controlabilidade*.

Exemplo 1. Ilustrar a condição de Kalman no seguinte sistema, onde a variável de estado possui duas componentes ($n = 2$) e um só controle ($m = 1$) e as matrizes do sistema são dadas por,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solução Aplicaremos a condição de Kalman, para isso devemos construir a matriz de controle

$$[B, AB] = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

onde o produto das matrizes AB é dado por

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

O posto da matriz de controle $[B, AB]$ é igual a um, ($1 < n = 2$), pelo Teorema 1.2. segue que o sistema,

$$x' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u$$

não é controlável.

Exemplo 2. Considere $n = 2$ e $m = 1$. Sejam as seguintes matrizes,

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Solução. Construindo a matriz de controle para aplicar as condições de Kalman

$$[B, AB] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

onde cada componente da matriz acima são dadas por,

$$B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Logo posto da matriz de Kalman, $[B, AB]$, é igual a dois, ($2 = n$). Aplicando novamente o Teorema 1.2., concluímos que o sistema,

$$x' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u$$

é controlável.

Prova do Teorema 1.2.

" \Rightarrow " Suponhamos que o número de linhas não nulas de $[B, AB, \dots, A^{n-1}B] < n$. Então as filas da matriz de controle $[B, AB, \dots, A^{n-1}B]$ são linearmente dependentes e existe um vetor $v \in \mathbb{R}^n$, $v \neq 0$ tal que

$$v^* [B, AB, \dots, A^{n-1}B] = 0,$$

onde os coeficientes das combinações lineares são as componentes do vetor v . Desde que

$$v^* [B, AB, \dots, A^{n-1}B] = [v^* B, v^* AB, \dots, v^* A^{n-1}B],$$

$$v^* B = v^* AB = \dots = v^* A^{n-1}B = 0$$

Pelo Teorema de Cayley-Hamilton deduzimos que existem constantes c_1, c_2, \dots, c_n tal que

$A^n = c_1 A^{n-1} + \dots + c_n I$ e então $v^* A^n B = 0$. De fato, segue que $v^* A^k B = 0$ para todo $k \in \mathbb{N}$ e conseqüentemente $v^* e^{At} B = 0$ para todo t . Mas, pela fórmula da variação das constantes, a solução x de (3.8.2) satisfaz

$$x(t) = e^{At} x_0 + \int_0^t e^{A(t-s)} B u(s) ds \quad (3.10.3)$$

Então

$$\langle v, x(T) \rangle = \langle v, e^{AT} x_0 \rangle + \int_0^T \langle v, e^{A(T-s)} B u(s) \rangle ds = \langle v, e^{AT} x_0 \rangle,$$

onde $\langle \cdot \rangle$ denota o produto interno canônico em \mathbb{R}^n . Ainda, $\langle v, x(T) \rangle = \langle v, e^{AT} x_0 \rangle$. Isto mostra que a projeção da solução x no tempo T sobre o vetor v é independente do valor do controle u . E mais, o sistema não é controlável.

Observação 1.7

A propriedade da conservação da quantidade $\langle v, x \rangle$, provou que para qualquer vetor v para o qual $v[B, AB, \dots, A^{n-1}B] = 0$. Assim, se o número de linhas não nulas da matriz $[B, AB, \dots, A^{n-1}B]$ é $n - k$, o conjunto de estados atingíveis tal que $x(T)$ se encontra é um subespaço afim de \mathbb{R}^n de dimensão igual a $n - k$.

" \Leftarrow " Supomos agora que o número de linhas não nulas do sistema é $\left([B, AB, \dots, A^{n-1}B] \right) = n$. De acordo com o Teorema 1.1 é suficiente mostrar que o sistema (3.9.5) é observável. Pela Proposição 1.1, (3.9.9) existe se, e somente se, (3.9.11) é verificado. Ainda, o Teorema é provado se (3.9.11) é válido. De $B^* \varphi = 0$ e $\varphi(t) = e^{A^*(T-t)} \varphi_T$, segue que $B^* e^{A^*(T-t)} \varphi_T \equiv 0$ para todo $0 \leq t \leq T$. Aplicando a derivada para estas funções em $t = T$ nós obtemos que

$$B^* [A^*]^k \varphi_T = 0, \quad \forall k \geq 0.$$

Mas, tendo que o número de linhas não nulas é $\left([B^*, B^* A^*, \dots, B^* (A^*)^{n-1}] \right) = n$ então $\varphi_T = 0$. Logo, (3.9.11) é verificado e a prova do Teorema 1.2 agora está completa.

Observação 1.8.

O conjunto dos pares controláveis (A, B) é aberto e denso. Portanto,

- Se (A, B) é controlável então existe $\varepsilon > 0$ suficientemente pequeno tal que para qualquer (A^0, B^0) com $|A^0 - A| < \varepsilon$, $|B^0 - B| < \varepsilon$ é também controlável. Isto é consequência do fato de que o determinante da matriz depende continuamente das suas entradas.
- Por outro lado, se (A, B) não é controlável, para algum $\varepsilon > 0$, existe um (A^0, B^0) com $|A - A^0| < \varepsilon$, $|B - B^0| < \varepsilon$ tal que (A^0, B^0) é controlável. Isto é consequência de que o determinante da matriz $n \times n$ depende analiticamente de suas entradas e não pode estar em $\mathbb{R}^{n \times n}$.

A seguinte desigualdade mostra que a norma do controle é proporcional à distância entre $e^{AT}x^0$ (o estado onde o sistema fica livre pela ausência do controle, isto é, com $u=0$) e o objetivo x^1 .

Proposição 1.2

Suponha agora que o par (A, B) é controlável no tempo $T > 0$ e seja u o controle obtido através da minimização da função J . Então, existe uma constante $C > 0$, dependendo de T , tal que a seguinte desigualdade é válida

$$\|u\|_{L^2(0,T)} \leq C |e^{AT}x^0 - x^1| \quad (3.10.4)$$

para qualquer dado inicial x^0 e objetivo final x^1 .

Prova. Vamos primeiro provar (3.10.4) para o caso particular $x^1 = 0$. Seja u o controle de (3.8.2) obtido pela minimização da função J . De (3.9.6) segue que

$$\|u\|_{L^2(0,T)}^2 = \int_0^T |B^* \bar{\phi}|^2 dt = -\langle x_0, \bar{\phi}(0) \rangle.$$

Se w é solução de

$$\begin{cases} w'(t) = Aw(t) \\ w(0) = x^0, \end{cases} \text{ para } t \in (0, T) \quad (3.10.5)$$

temos que

$$w(t) = e^{AT}x^0 \quad \text{e} \quad \frac{d}{dt} \langle w, \phi \rangle = 0$$

para todo $\phi_T \in \mathbb{R}^n$, sendo ϕ a correspondente solução de (3.9.5).

Em particular, pegando $\phi_T = \overline{\phi_T}$, o mínimo de J , temos que

$$\langle x_0, \overline{\phi(0)} \rangle = \langle w_0, \overline{\phi(0)} \rangle = \langle w(T), \overline{\phi_T} \rangle = \langle e^{AT} x^0, \overline{\phi_T} \rangle$$

Nós obtemos que

$$\|u\|_{L^2(0,T)}^2 = -\langle x_0, \overline{\phi(0)} \rangle = -\langle e^{AT} x^0, \overline{\phi_T} \rangle \geq |e^{AT} x^0| \cdot |\overline{\phi_T}|$$

Por outro lado, temos que

$$\|\overline{\phi_T}\| \leq c \|B^* \overline{\phi}\|_{L^2(0,T)} = c \|u\|_{L^2(0,T)}$$

Portanto, o controle u verifica

$$\|u\|_{L^2(0,T)} \leq c |e^{AT} x^0|. \quad (3.10.6)$$

Se $x^1 \neq 0$, a Observação 1.1 implica que o controle de u conduz a solução de x^0 para x^1 coincide com outro que leva a solução de $x^0 - y(0)$ para zero, onde y verifica (3.8.2).

Usando (3.10.6), obtemos que

$$\|u\|_{L^2(0,T)} \leq c |e^{TA} (x^0 - y(0))| = c |e^{TA} x^0 - x^1|$$

e (3.10.5) é provado.

Observação 1.9.

Equações escalares lineares de qualquer ordem provém de exemplos de sistemas de dimensões grandes e arbitrárias que são controláveis com apenas um controle. Então, o sistema de ordem k

$$x^{(k)} + a_1 x^{(k-1)} + \dots + a_{k-1} x = u(t)$$

é controlável. Isto pode ser facilmente obtido pela observação de que dados k valores iniciais e finais podemos sempre encontrar uma trajetória z (de fato um número infinito) juntando-os em qualquer intervalo de tempo. Este argumento foi usado no *Exemplo 2* para o caso $k = 2$.

3.11. Estabilização de Sistemas Lineares de Dimensão Finita.

Nesta seção assumiremos que A é uma matriz adjunta simétrica, isto é, $A^* = -A$. Neste caso, $\langle Ax, x \rangle = 0$

Consideremos o sistema

$$\begin{cases} \dot{x} = Ax + Bu \\ x(0) = x^0 \end{cases} \quad (3.11.1)$$

Observação 1.11

O oscilador harmônico, $m\ddot{x} + kx = 0$, fornece um simples exemplo de um sistema com estas propriedades.

Quando $u \equiv 0$ a energia da solução de (3.11.1) é conservada (sistema conservativo). Ou seja, pela multiplicação de (3.11.1) por x , se $u \equiv 0$, teremos

$$\begin{aligned} \langle x, \dot{x} \rangle &= \langle x, Ax \rangle + 0 \\ \frac{d}{dt} |x(t)|^2 &= 0 \end{aligned} \quad (3.11.2)$$

Então,

$$|x(t)| = |x^0|, \quad t \geq 0 \quad (3.11.3)$$

O problema da estabilização pode ser formulado da seguinte maneira. Suponha que o par (A, B) seja controlável. Procuramos uma matriz L tal que a solução do sistema (3.11.1) com o controle *feedback*,

$$u(t) = Lx(t) \quad (3.11.4)$$

tenha um **decaimento uniformemente exponencial**, isto é, existem uma constante $c > 0$ e $\omega > 0$ tal que

$$|x(t)| \leq ce^{-\omega t} |x^0| \quad (3.11.5)$$

para qualquer solução.

Notemos que, de acordo com a equação (3.11.4), o controle u é obtido no tempo real do estado x .

Em outras palavras, estamos procurando a matriz L tal que a solução do sistema

$$\begin{aligned} \dot{x} &= Ax + BLx \\ \dot{x} &= (A + BL)x = Dx \end{aligned} \quad (3.11.6)$$

tenha uma taxa de decaimento uniformemente exponencial.

Observe que não podemos esperar mais de (3.11.5). Ou seja, as soluções de (3.11.6) podem não satisfazer $x(T) = 0$ no tempo finito T . Ou seja, se este for o caso, pela unicidade das soluções de (3.11.6) com o estado final $t=0$ em $t=T$, poderia ser que $x^0 \equiv 0$. Por outro

lado, qualquer que seja L , a matriz D tem n autovalores λ_j com correspondentes autovetores $e_j \in \mathbb{R}^n$. A solução $x(t) = e^{\lambda_j t} e_j$ de (3.11.6) mostra que o decaimento das soluções não pode ser mais rápido que a exponencial.

Teorema 1.3

Se A é uma matriz adjunta simétrica e o par (A, B) é controlável, então $L = -B^*$ estabiliza o sistema, isto é, a solução de

$$\begin{cases} x' = Ax - BB^*x \\ x(0) = x^0 \end{cases} \quad (3.11.7)$$

tem um decaimento uniformemente exponencial (3.11.5)

Prova. Pelo fato que $L = -B^*$ e multiplicando (3.11.7) por $x(t)$, obtemos que

$$\frac{1}{2} \frac{d}{dt} |x(t)|^2 = -\langle BB^*x(t), x(t) \rangle = -|B^*x(t)|^2 \leq 0.$$

Então, a norma da solução decresce no tempo.

Mais ainda,

$$|x(T)|^2 - |x(0)|^2 = -2 \int_0^T |B^*x|^2 dt. \quad (3.11.8)$$

Para provar o decaimento uniformemente exponencial é suficiente mostrar que existe $T > 0$ e $c > 0$ tais que

$$|x(0)|^2 \leq c \int_0^T |B^*x|^2 dt \quad (3.11.9)$$

para qualquer solução x de (3.11.7). Ou seja, de (3.11.8) e (3.11.9) obteremos que

$$|x(T)|^2 - |x(0)|^2 \leq -\frac{2}{c} |x(0)|^2 \quad (3.11.10)$$

e conseqüentemente

$$|x(T)|^2 \leq \gamma |x(0)|^2 \quad (3.11.11)$$

com

$$\gamma = 1 - \frac{2}{c} < 1. \quad (3.11.12)$$

Então,

$$|x(kT)| \leq \gamma^k |x_0|^2 = e^{(\ln \gamma)k} |x_0|^2, \quad \forall k \in \mathbb{N} \quad (3.11.13)$$

Agora, tomando qualquer $t > 0$, escrevemos na forma $t = kT + \delta$, com $\delta \in [0, T)$ e $k \in \mathbb{N}$, portanto obtemos que

$$\begin{aligned} |x(t)|^2 &\leq |x(kT)|^2 \leq e^{-\ln(\gamma)k} |x_0|^2 \\ &= e^{-\ln(\gamma)\left(\frac{t}{T}\right)} e^{\ln(\gamma)\frac{\delta}{T}} |x_0|^2 \leq \frac{1}{\gamma} e^{\frac{|\ln(\gamma)|}{T}t} |x_0|^2. \end{aligned}$$

Logo, segue o decaimento pelo resultado (3.11.5) com

$$c = \frac{1}{\gamma}, \quad \omega = \frac{|\ln(\gamma)|}{T}. \quad (3.11.14)$$

Para provar (3.11.9), decompos a solução x de (3.11.7) como $x = \phi + y$ com ϕ e y soluções dos seguintes sistemas:

$$\begin{cases} \phi' = A\phi \\ \phi(0) = x_0 \end{cases} \quad (3.11.15)$$

e

$$\begin{cases} y' = Ay - BB^*x \\ y(0) = 0 \end{cases} \quad (3.11.16)$$

Observamos que, desde que A seja adjunta simétrica, (3.11.15) é exatamente o sistema adjunto (3.9.5) exceto pelo fato de que o valor inicial tomado foi $t = 0$.

Como vimos na prova do Teorema 1.2, se o par (A, B) está sendo controlável, a seguinte desigualdade da observação é válida para o sistema (3.11.15):

$$|x_0|^2 \leq C \int_0^T |B^* \phi|^2 dt. \quad (3.11.17)$$

Tomando $\phi = x - y$ deduzimos que

$$|x_0|^2 \leq 2C \left[\int_0^T |B^* x|^2 dt + \int_0^T |B^* y|^2 dt \right].$$

Por outro lado, mostramos que a solução y de (3.11.16) satisfaz:

$$\frac{1}{2} \frac{d}{dt} |y|^2 = -\langle B^* x, B^* y \rangle \leq |B^* x| |B^* y| \leq \frac{1}{2} (|y|^2 + |B^*|^2 |B^* x|^2).$$

Pela desigualdade de Gronwall, deduzimos que

$$|y(t)|^2 \leq |B^*|^2 \int_0^t e^{-s} |B^* x|^2 ds \leq |B^*|^2 e^T \int_0^T |B^* x|^2 dt \quad (3.11.18)$$

e conseqüentemente

$$\int_0^T |B^* y|^2 dt \leq |B|^2 \int_0^T |y|^2 dt \leq T |B|^4 e^T \int_0^T |B^* x|^2 dt.$$

Finalmente, obtemos que

$$|x_0|^2 \leq 2C \int_0^T |B^* x|^2 dt + C |B^*|^4 e^T T \int_0^T |B^* x|^2 dt \leq C' \int_0^T |B^* x|^2 dt$$

e a prova do Teorema 1.3 está completa.

Exemplo 1. Consideremos o oscilador harmônico amortecido,

$$mx'' + Rx + kx' = 0, \quad (3.11.19)$$

onde m , k e R são constantes positivas.

Note que (3.11.19) pode ser escrito equivalentemente na forma

$$mx'' + Rx = -kx'$$

que exibe uma força que atua sobre o oscilador, proporcional à velocidade do ponto de massa e de sinal oposto.

Veremos que as soluções desta equação têm a propriedade do *decaimento exponencial*. Ou seja, é suficiente para lembrarmos que as duas raízes características têm partes reais negativas. Isto é,

$$mr^2 + R + kr = 0 \Leftrightarrow r_{\pm} = \frac{-k \pm \sqrt{k^2 - 4mR}}{2m}$$

E então

$$\operatorname{Re}[r_{\pm}] = \begin{cases} -\frac{k}{2m} & \text{se } k^2 \leq 4mR \\ -\frac{k}{2m} \pm \sqrt{\frac{k^2}{4m} - \frac{R}{2m}} & \text{se } k^2 \geq 4mR \end{cases}$$

Vamos provar o decaimento exponencial da solução de (3.11.19) usando o Teorema 1.3. Primeiro, devemos escrever (3.11.19) na forma de (3.11.7). Colocando

$$X = \begin{bmatrix} x \\ \sqrt{\frac{m}{R}} x' \end{bmatrix}$$

a equação conservativa $mx''+kx=0$ corresponde ao sistema:

$$X' = AX, \quad \text{com} \quad A = \begin{pmatrix} 0 & \sqrt{\frac{R}{m}} \\ -\sqrt{\frac{R}{m}} & 0 \end{pmatrix}.$$

Note que A é uma matriz adjunta simétrica. Por outro lado, se escolhermos

$$B = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{k} \end{pmatrix}$$

obtemos que

$$BB^* = \begin{pmatrix} 0 & 0 \\ 0 & k \end{pmatrix}$$

e o sistema

$$X' = AX - BB^*X \tag{3.11.20}$$

é equivalente a (3.11.19).

Agora, é fácil de ver que o par (A, B) é controlável desde que o posto (número de linhas não nulas) da matriz coluna formada por $[B, AB]$ seja 2.

Segue que a solução de (3.11.19) tem a propriedade do decaimento exponencial.

Se (A, B) é controlável, temos que provar a propriedade da estabilização uniforme do sistema (3.11.1) com a hipótese de que A é adjunta simétrica. Entretanto, esta propriedade é válida se A é uma matriz arbitrária. Mais precisamente, temos,

Teorema 1.4.

Se (A, B) é controlável então ela é estabilizável. Ou seja, é possível de prescrever números complexos $\lambda_1, \lambda_2, \dots, \lambda_n$ como os autovalores da matriz de laço fechado $A+BL$ por uma escolha apropriada da matriz de *feedback* L tal que a taxa de decaimento pode ser arbitrariamente rápida.

Na formulação do Teorema usamos o termo clássico de sistema de laço fechado (*closed loop*) para referirmos ao sistema cujo controle é dado na forma de *feedback*.

A prova do Teorema 1.4 é obtida reduzindo o sistema (3.11.1) até a forma canônica de controle. Ver, para maiores informações, Lee & Markus [10] e Russel D. L. [15].

CONSIDERAÇÕES FINAIS

A Teoria do Controle é um ramo da Matemática que, apesar de ser explorado há muito tempo, inclusive pela natureza em sua organização, mesmo sem o estudo do sistema pelo homem, ainda encontra-se em fase inicial de estudos pelos matemáticos e engenheiros.

Sua utilidade para controlar objetos, como pressão e extração de petróleo, aeronaves, máquinas, atrai cada vez mais a atenção do homem que a tudo quer controlar de forma que se tenha um melhor aproveitamento dos objetos com o menor tempo possível. Dessa forma, esse ramo da Matemática está em crescente expansão. E, observa-se que a matemática utilizada para esses cálculos não é uma operação básica simples e envolve bastante conhecimento de álgebra, cálculo, bem como de conceitos matemáticos e até físicos, químicos e biológicos, dependendo do objeto que se pretende controlar.

REFERÊNCIAS

1. APOSTOL. T. M., *Análisis Matemática*, Segunda Edición, Reverté, Barcelona, 1982.
2. BOLDRINI J. L., et. al. *Álgebra Linear*. 3ª ed. São Paulo: Harbra, 1986.
3. CALLIOLI, Carlos A.; DOMINGUES, Hygino H.; COSTA, Roberto C.F. *Álgebra Linear e Aplicações*. 6ª ed. Ver. São Paulo: Atual, 1990.
4. CAMPBELL, S. L., & HABERMAN, R. *Introducción a las ecuaciones diferenciales con problemas de valor de Fronteira*, McGrawHill, México, 1998.
5. GUZMAN, Miguel de. *Equações Diferenciais Ordinárias; Teoria de Estabilidade e Controle*. Madrid: Ed. Alhambra S.A, 1980
6. JOHN, F., *Partial Differential Equations*, 4ª edition. Springer Verlag, New York, Berlin, 1998.
7. KALMAN, R. E., *Optimization, Mathematical Theory of Control Theory*, Encyclopedia Britannica, Fifteenth Ed. 1974, 636-638.
8. LASALLE, J. P. & LEFSCHETZ, S., *Stability by Liapounov's direct Method*, Academic Press, New York, 1961.
9. LEITHOLD, Louis. *O cálculo com geometria analítica*. Tradução: Cyro de Carvalho Patarra; revisão técnica: Wilson Castro Ferreira Jr e Silvio Pregnotatto. 3ª ed. São Paulo: Harbra, 1994.
10. LEE, E. B. & MARKUS, L., *Foundations of Optimal Control Theory*, The SIAM Series in Applied Mathematics, John Wiley & Sons, 1967.
11. LEON, J. S. *Álgebra Linear com Aplicações*, 4ed LTC- Livros Técnicos e Científicos Editora S. A., 1999.
12. LIONS, J.-L., *Controlabilite Exacte, Stabilisation et Perturbations de Systems Distribues*. Tome 1. Controbilite Exacte, Masson, Paris, RMA 8, 1988.
13. LIONS, J.-L., *Exact Controllability, Stabilizability and Perturbations for Distributed Systems*, SIAM Rev., 30 (1988), 1-68.
14. MARKUS, Lawrence. **A Brief history of control**. Ima Preprint Series, 1993.
15. RUSSELL, D. L., *Controllability and Stabilizability for Linear Partial Differential Equations*, Recent Progress and Open Questions, SIAM Rev., 20, (1978), 639-739.

16. SONTAG, E. D., *Mathematical Control Theory. Deterministic Finite-dimensional Systems, Second editions*. Texts in Applied Mathematics, 6. Springer Verlag, New York, 1998.
17. STEINBRUCH, Alfredo; WINTERLE, Paulo. **Álgebra Linear**. 2ª ed. São Paulo: Pearson Makron Books, 1987.
18. SWOKOWSKI, Earl William. **Cálculo com geometria analítica**. Tradução: Alfredo Alves de Faria, com a colaboração dos professores Vera Regina L. F. Flores e Marcio Quintão Moreno; revisão técnica: Antonio Pertence Junio. 2ª ed. São Paulo: Makron Books, 1994.
19. ZUAZUA, E., *Controlabilidad Exacta y Estabilización la Ecuación de Ondas*, Textos de Métodos Matemáticos 23, Universidade Federal do Rio de Janeiro, 1990.