Notas de Aula de Economia Matemática

Rodrigo Nobre Fernandez

Prefácio

Esta apostila é um resumo das notas de aula do curso de Matemática II do mestrado em economia aplicada da Universidade Federal em Pelotas. Em quase sua totalidade essas notas de aula transcrevem literalmente ou resumem o conteúdo do livro Matemática para Economistas de Simon e Blume (2004). Há também alguns trechos baseados em Leonard e Long (1992) e Sala-i-Martin (2000). Destaco que essa apostila não tem fins comerciais, o texto serve exclusivamente como material de apoio as aulas. Aproveito e agradeço aos seguintes alunos que colaboraram para a construção desse material: Andressa Vasconcelos, Caio Rostirolla, Dianifer Leal Borges, Douglas Pivatto, Gustavo Moreira, Jean Duarte, Leonardo Cordeiro, Márcio Taveira, Mariana Moreira, Patricia Colussi, Raquel Pérez, Silvio Paula, Thais Dietrich e Victor Gabriel Buttignon.

Sumário

1 Noções de Lógica

O homem geralmente se expressa através da linguagem. A linguagem corrente, pode ser vaga e ambígua, não é adequada ao tratamento científico. Por isso, necessitamos, para o tratamento da matemática, de uma linguagem mais adequada chamada de linguagem simbólica.

Nesta linguagem, destaca-se o uso do termo (expressão que nomeia ou descreve algum objeto) e do enunciado (expressão que correlaciona objetos, descreve propriedades de objetos etc.)

Exemplos:

$$\text{Termos}\begin{cases} x \\ x+y \\ \emptyset \\ \{3,5,7\} \end{cases} \qquad \text{Enunciados}\begin{cases} x+2=4 \\ a>b \\ 7<x \\ x^2-5x+6=0 \end{cases}$$

O enunciado aberto é qualquer expressão que contém variáveis. Entendemos por variável um elemento que pode assumir qualquer valor dentro de um conjunto de escolhas. Já no enunciado fechado, a variável deve assumir pelo menos um valor. Podemos também, chamar de sentença ou proposição.

1.1 Átomos da Linguagem

Assim, abaixo destacamos algumas partículas fundamentais ou átomos da linguagem:

1. **Funtores:** Formam termos a partir de termos. Exemplos: $+, \times, -, \cap$ e $\cup$ entre outros.

2. **Juntores:** Formam enunciados a partir de enunciados. Exemplos: não; e;ou; se ... então; se, e somente se.

3. **Predicados:** Formam enunciados a partir de termos. Exemplos: $\in, =, \subset, \supset, <$ e $>$ entre outros.

4. **Operadores (Quantificadores):** Formam enunciados a partir de enunciados. Sua principal propriedade é transformar enunciados abertos em enunciados fechados. Exemplos: para todo, qualquer que seja ($\forall$); Não existe($\nexists$) .

O juntor não (~), de um enunciado p, pode formar-se o enunciado ~p, dito negação de p:

p	~p
V	F
F	V

O próximo juntar a ser tratado é o "e". Dados dois enunciados p e q, podemos formar o enunciado "p e q" dito conjunção de p e q:

p	q	p e q
V	F	F
V	V	V
F	V	F
F	F	F

Também temos o juntor "ou", sendo usado no contexto de lógica como não exclusivo:

p	q	p ou q
V	V	V
V	F	V
F	V	V
F	F	F

Finalmente apresentaremos com um exemplo o juntor "se ... então":
Se fizer sol, então André irá a praia.

- Fez sol e André foi à praia, então podemos concluir que a afirmação acima não foi falseada pelo experimento em questão.
- Fez sol e André não foi à praia, então podemos concluir que o enunciado acima é falso.
- Não fez sol. Neste caso, não importa se André foi ou não a praia. Isto é, concluímos que o enunciado acima é verdadeiro.

p	q	$p \Rightarrow q$
V	V	V
V	F	V
F	V	V
F	F	F

Note que p é chamado de condição suficiente para q e o último é chamado de condição necessária para p.

O juntos se, e somente se é dado por $p \Leftrightarrow q$. Em outras palavras, esse enunciado é chamado de bijunção de p e q. Será considerado como verdadeiro quando os constituintes tiverem o mesmo valor lógico. Em seguida apresentamos a tabela de valores lógicos para a bijunção:

p	q	$p \Leftrightarrow q$
V	V	V
V	F	F
F	V	F
F	F	V

Um enunciado atômico é uma sentença declarativa que contém uma ideia que é falsa ou verdadeira, mas não ambas. Um enunciado é chamado de composto se é obtido com base em enunciados atômicos, através do uso de juntores. Um enunciado composto é dito ser uma tautologia se é verdadeiro ao considerarmos todas as possíveis valorações de seus componentes atômicos:

$$p \Rightarrow p; p \text{ ou } \sim p; \sim\sim p \Leftrightarrow p$$

$$(\text{p e q}) \Leftrightarrow \sim (\sim p \text{ ou } \sim q)\,; p \;\, ou \;\, q \Leftrightarrow (\sim p \text{ e } \sim q)$$

Se um enunciado é uma tautologia, podemos substituir todas as ocorrências de um componente por outro enunciado e o enunciado resultante ainda é uma tautologia. Se p é uma tautologia, diz-se que $\sim p$ é uma contradição.

Dizemos que o enunciado p é logicamente equivalente ao enunciado q quando o enunciado $p \Leftrightarrow q$ é uma tautologia. Por exemplo:

1. $\sim (p \;\, e \;\, q)$ é equivalente a $\sim p$ ou $\sim q$;

2. $\sim (p \;\, ou \;\, q)$ é equivalente a $\sim p$ e $\sim q$;

3. $p \Rightarrow q$ é equivalente a $\sim q \Rightarrow p$

Quantificadores:

Seja o conjunto $X = \{1, 3, 5, 7\}$ os enunciados abertos:

- $p\{x\}$: x é um número ímpar;

- $q\{x\}$: x é múltiplo de 3;

- $r\{x\}$: $x \geq 10$

Podemos facilmente observar que:

Todo elemento de X satisfaz p. Existe elemento de X que satisfaz a q. Não existe elemento de X que satisfaz a r.

Essas afirmações podem ser escritas simbolicamente como:

$$\forall x\,(x \in X \Rightarrow p\,\{x\})$$

$$\exists x\,(x \in X \text{ e } q\,\{x\})$$

$$\sim \exists x\,(x \in X \text{ e } r\,\{x\})$$

Equivalências de enunciados quantificados:

Todos o brasileiro é feliz. Equivale a: Não existe brasileiro que não seja feliz.

Simbolicamente, seja B a coleção de brasileiros e F das pessoas felizes:

$$\forall x\,(x \in B \Rightarrow x \in F) \text{ equivale a } \sim \exists x\,(x \in B \text{ e } x \notin F)$$

De modo geral, para o enunciado p valem as seguintes tautologias:

$$\forall x\, p \Leftrightarrow \exists x \sim p$$

$$\forall x\, p \Leftrightarrow \forall x \sim p$$

e em sequência:

$$\sim \exists\, p \Leftrightarrow \forall x \sim p$$

$$\sim \forall p \Leftrightarrow \exists x \sim p$$

2 Conjuntos, Números e Demonstrações

Definição 1. Um conjunto é qualquer coleção bem especificada de elementos. Para qualquer conjunto A, escrevemos $a \in A$ para indicar que a é um elemento de A e $a \notin A$ para indicar que a não é um elemento de A. Um conjunto que não possui elementos é denominado de vazio ($\emptyset$).

Exemplos: O conjunto de todos os números não negativos é escrito assim:

$$R_+ = \{x \in R : \ x \geqslant 0\}$$

cada elemento de R_+ é um elemento de R^1, diz-se que R_+ é um subconjunto de R ou escrevemos $R_+ \subset R$ ou $R \supset R_+$.

2.1 Operações com Conjuntos

Definição 2. União

Sejam A e $B \subset R$, $A \cup B = \{x : x \in A$ ou $x \in B\}$

Definição 3. Interseção

Sejam A e $B \subset R$, $A \cap B = \{x : x \in A$ e $x \in B\}$

Definição 4. Subtração

Sejam A e $B \subset R$, $A - B, ou\ A \backslash B$,$A - B = \{x : x \in A$ e $x \notin B\}$

Nota: Se U é o conjunto universo podemos escrever *U-A* como A^c e dizemos que é o complementar de A. Por exemplo, o complementar de R_+é o conjunto de todos os números negativos:

$$R_+^c = \{x \in R : \ x < 0\}$$

2.2 Números

O primeiro conjunto de números que apresentamos é o dos números Naturais:

$$N = \{1, 2, 3, ...\}$$

A soma e o produto de dois números naturais é outro número natural, mas a diferença não precisa estar em N. Adicionalmente, o conjunto dos números inteiros inclui os números negativos:

$$Z = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$$

A soma e a diferença de qualquer número inteiro resulta em um número inteiro, mas o quociente não. Assim temos o conjunto dos números Racionais:

$$Q = \left\{\frac{a}{b} : \ a, b \in Z; b \neq 0\right\}$$

Podemos destacar algumas propriedades dos números Racionais:

$$a \text{ e } b \in Q \rightarrow a + b, a - b, ab, \frac{a}{b} \ \in Q$$

Todo o número pode ser escrito como o quociente de dois inteiros? Embora não seja imediatamente óbvio, alguns números como $\sqrt{2}$ e π não podem ser escritos como o quociente de inteiros. Os números que não podem ser escritos como a razão ou quociente de inteiros são denominados irracionais. Por exemplo, expansões decimais sem padrão são números irracionais. Por fim, o conjunto dos números racionais e irracionais é o conjunto dos números reais.

Definição 5. Um inteiro n é dito número par se existe um inteiro m tal que *n=2m.* Um inteiro que não é par é dito ímpar.

Definição 6. Um número natural m é dito primo se, sempre que m puder ser escrito como o produto *m=ab* de dois números naturais, então *a=1* ou *b=1* (um ou outro, não simultaneamente). Os seis primeiros números primos são: 1,2,3,5,7 e 11.

2.3 Propriedades da adição e da multiplicação

Sejam a, b e $c \in R$ então teremos:

1. Fechamento: $a + b$ e $ab \ \in R$
2. Comutatividade: $a + b = b + a$ e $a - b = b - a$
3. Associatividade: $(a + b) + c = a + (b + c)$ e $(ab)\, c = a\,(bc)$
4. Identidades: Existe um elemento $0 \in R$ tal que $a + 0 = a \, \forall \, a \in R$. Existe um elemento $1 \in R$ tal que $\forall \, a \in R \; a.1 = a$
5. Inversos: $\forall \, a \in R, \exists \, b \in R, a + b = 0, b = -a \, \forall a \in R$ não nulo $c \in R$ tal que $a.c = 1$ e $c = \frac{1}{a}$.
6. Distributividade: $a.\,(c + b) = a.b + a.c$

Definição 7. Seja $S \subset R$ e $b \in R$.O número b é cota superior de S se $a \leq b \; a \in S$. O número b é cota inferior de S se $b \leq a \; a \in S$.

Definição 8. Se b é cota superior para S e nenhum elemento menor do que b é cota superior de S, então dizemos que b é um supremo de S.Analogamente, se b é cota inferior de S e nenhum elemento maior do que b é uma cota inferior de S, então dizemos que b é um ínfimo de S.

Exemplo:

$$S = \{0.3; 0.33; 0.333; 0.3333; ...\}$$

Zero é cota inferior de S e um é uma cota superior. O supremo é 0.3333 e o ínfimo 0.3.

2.4 Demonstrações

A maneira direta de provar que $A \to B$ é encontrar uma sequência de axiomas e teoremas aceitos de tal forma que $A_i \to A_{i+1} \, \forall \, i = 1, ..., n$ de tal modo que $A_0 = A$ e $A_{n+1} = B$.

$$A = A_0 \to A_1 \to A_2 \to ... \to A_{n-1} \to A_n = B$$

Provas realizadas da forma acima são denominadas diretas e o método usado é o racional dedutivo. Sejam a, b, c e $d \in R$

$$a = b \text{ e } b = c \to a = c$$

$$a = b \to a + c = b + c$$

$$a = b \to a.c = b.c$$

$$c = d \to a.c = b.d$$

$$a = b$$

Teorema 1. Para qualquer x,y e z $\in R$, se x+z=y+z então x=y

Demonstração.

$$x + z = y + z$$

$$\text{então existe} - z \text{ tal que } z + (-z) = 0$$

$$(x + z) + (-z) = (y + z) + (-z)$$

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

$$x + 0 = y + 0$$

$$x = y$$

□

Teorema 2. Para qualquer $x \in R, x.0 = 0$

Demonstração.

$$0 + 0 = 0$$

$$x(0 + 0) = x.0$$

$$(x.0) + (x.0) = x.0$$

$$(x.0) + 0 = x.0$$

$$(x.0) + (x.0) = (x.0) + 0$$

$$x.0 = 0$$

□

Teorema 3. Seja m um inteiro par e p um inteiro qualquer. Então m.p é um inteiro par.

Demonstração.

$$m \text{ é um inteiro par}$$

$$\text{Existe } q \in Z,\ m = 2q$$

$$m.p = 2\,(q.p) \rightarrow 2q.p = 2qp$$

$$mp \text{ é par}$$

□

Definição 9. Recíproca

Considere a proposição P da forma $A \rightarrow B$. Se vale a hipótese A então vale a conclusão B. A recíproca de P é a afirmação $B \rightarrow A$, que troca a hipótese e conclusão de P.

Vamos supor que A é a situação "n é um número primo maior do que 2" e B a situação "n é um número ímpar". É verdade que $A \rightarrow B$, mas não é verdade que B implica A.

Definição 10. Se $A \rightarrow B$ e $B \rightarrow A$ são verdades, dizemos que A vale se, e somente se B vale, ou então que A é equivalente a B, isto é,$A \leftrightarrow B$.

Suponha que A é a afirmação "n é um número primo par" e B é a afirmação "n=2", então ambas $A \rightarrow B$ e $B \rightarrow A$ são verdades e temos que $A \leftrightarrow B$.

Definição 11. A proposição $\sim B \to \sim A$ é denominada contraposição de $A \to B$.

A contraposição afirmar que ($\sim B$) n não é um inteiro ímpar então ($\sim A$) não é um número primo diferente de 2, de outro modo, se n é par, então n é igual a 2 ou não é um número primo.

Teorema 4. Sejam a, b e $c \in Z$ tais que $a.b = c$. Se c é impar e a e b também são.

Corolário. *Seja a um número inteiro. Se a^2 é ímpar, então a é ímpar.*

2.5 Provas Indiretas

Como $A \to B$ é verdadeira se, e somente se sua contraposição é verdadeira, uma maneira de provar $A \to B$ é provar que $\sim B \to \sim A$. Outra forma, para provar que B é verdadeira é considerarmos todas as suas alternativas possíveis. Se cada uma das alternativas de B leva a uma contradição, ou da própria afirmação A, ou de um axioma do sistema, ou de uma proposição previamente provada, então B deve ser verdadeira. Essa linha de raciocínio é denominada de prova indireta ou redução ao absurdo.

Para dar uma prova indireta de $A \to B$, supomos que a situação B não vale e então aplicamos argumentos indutivos rigorosos até alcançar uma contradição.

Teorema 5. Seja a um inteiro. Se a^2é par então, a é par.

Teorema 6. Se $a = \frac{p}{q}$ é um número racional com p e q inteiros, então p e q podem ser escolhidos de tal modo que ambos não são inteiros pares.

Teorema 7. $\sqrt{2}$ é um número irracional

Demonstração. Vamos realizar essa prova por contradição. Suponha que $\sqrt{2}$ não seja um número racional. Desse modo $\sqrt{2}$ pode ser expresso por uma fração simplificada $\frac{a}{b}$ em que a e b são inteiros primos entre si:

$$\frac{a}{b} = \sqrt{2} \to \frac{a^2}{b^2} \to 2 \to a^2 = 2b^2$$

Note que a^2é múltiplo de 2, então a^2 é par. Assim a pode ser escrito como $a = 2k \ \forall \ k \in Z$.

$$(2k)^2 = 2b^2 \to 4k^2 = 2b^2 \to b^2 = 2k^2$$

Pelo mesmo raciocínio b é múltiplo de 2, logo é par. Podemos concluir que a e b são pares o que é um absurdo, pois são primos entre si. Aqui temos uma contradição, portanto $\sqrt{2}$ não é racional. $\square$

2.6 Indução Matemática

As provas indutivas somente podem ser utilizadas para proposições sobre inteiros ou proposições indexadas por números inteiros. Suponha que estamos considerando uma sequência de afirmações indexadas por números naturais de tal modo que a primeira afirmação é $P(1)$ e a segunda $P(2)$ e a enésima é $P(n)$. Supomos que possamos verificar dois fatos a sequência de afirmações:

1. A afirmação $P(1)$ é verdadeira;

2. Sempre que alguma afirmação $P(k)$ for verdadeira para algum k então $P(k+1)$ também é.

Vejamos o próximo teorema:

Teorema 8. A soma dos n primeiros números naturais 1+2+3+...+n=$\frac{1}{2}n(n+1)$

Demonstração. Para qualquer número natural n, seja *P(n)* a afirmação:

$$P(n) : 1+2+3+...+n = \frac{n(n+1)}{2}$$

Tomando *n=1* no lado direito da equação acima, obtemos:

$$1 = \frac{1(1+1)}{2} \rightarrow 1 = 1$$

Agora aplicaremos a hipótese de indução, ou seja, que a afirmação *P(k)* é verdadeira para algum inteiro k:

$$1+2+3+...+k = \frac{k(k+1)}{2}$$

Adicionando *k+1* em ambos lados da equação acima, teremos:

$$1+2+3+...+k+k+1 = \frac{k(k+1)}{2} + k + 1$$

$$= \left(\frac{k}{2}+1\right)(k+1)$$

$$= \frac{(k+2)(k+1)}{2}$$

Note que essa última expressão é exatamente a afirmação *P(k+1).* Mostramos que *P(1)* é verdadeira e que *P(k)* é verdadeira então vale $P(k+1)$ $\forall k$. Pelo princípio da indução concluímos que $P(n)$ $\forall n$. □

Teorema 9. A soma dos n primeiros números ímpares é n^2.

$$1+3+5+7+...+(2n-1) = n^2$$

Demonstração. É trivial notar que a fórmula acima é válida para *n=1*. Assim, vamos ao passo indutivo. Considere que a equação vale para algum inteiro k positivo:

$$1 + 3 + 5 + 7 + ... + (2k - 1) = k^2$$

O próximo número ímpar a ser somado ao lado esquerdo da equação acima é $(2\,(k+1) - 1) = 2k + 1$

$$1 + 3 + 5 + 7 + ... + (2k - 1) + (2\,(k+1) - 1) = k^2 + 2k + 1$$

$$1 + 3 + 5 + 7 + ... + (2k - 1) + (2\,(k+1) - 1) = (k+1)^2$$

Note que ficamos com *k+1* no lugar de *k*. Por indução concluímos que a primeira fórmula vale $\forall\, n \in N$. □

3 Sistemas de Equações Lineares

Alguns modelos econômicos possuem uma estrutura linear natural. Se as relações entre as variáveis em questão são descritas por um sistema de equações não lineares, tomamos a derivada dessas equações para convertê-las num sistema linear aproximante. Dizemos que o sistema é implícito se as equações que descrevem as relações econômicas em questão têm misturadas entre si, num mesmo lado do sinal de igualdade as variáveis exógenas e endógenas.

3.1 Eliminação de Gauss e Gauss-Jordan

Desejamos resolver os seguintes sistemas de equações lineares:

$$\begin{array}{cc} 2x_1 + 3x_2 = 7 \text{ ou} & x_1 + x_2 + x_3 = 5 \\ x_1 - x_2 = 2 & x_2 - x_3 = 0 \end{array} \tag{1}$$

O sistema linear geral de m equações a n incógnitas pode ser escrito como:

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{gathered} \tag{2}$$

Nesse sistema a_{ij} e $b_{ij} \in R$ sendo a_{ij} o coeficiente da incógnita x_j na i-ésima equação. Uma solução do sistema (2) é uma n-upla de números reais $x_1, \ldots, x_n$ que satisfaz cada uma das m equações em (2). Por exemplo, $x_1 = 2$,$x_2 = 1$ resolve o sistema (1) e $x_1 = 5$,$x_2 = x_3 = 0$ resolve o segundo. Para um sistema linear (2) estamos interessados nas três seguintes questões:

1. Existe alguma solução?
2. Quantas soluções existem?
3. Existe um algoritmo que realmente calcula essas soluções?

Basicamente há três métodos de resolução de tais problemas:

1. Substituição;
2. Eliminação de variáveis, e
3. Métodos Matriciais

3.1.1 Eliminação Gaussiana

A eliminação Gaussiana consiste em realizar operações ditas elementares com as linhas (equações) do sistema:

1. Multiplicar uma equação por um escalar não nulo;
2. Somar e subtrair equações.

A eliminação Gauss-Jordan não usa a substituição inversa: Vejamos o exemplo:

$$
\begin{aligned}
x_1 - 0.4x_2 - 0.3x_3 = 130 \\
x_2 - 0.25x_3 = 125 \\
x_3 = 300
\end{aligned}
$$

Some 0.25 vezes a linha (3) da linha (2) para achar $x_2 = 200$.

3.2 Operações Elementares sobre Linhas

O sistema (2) pode ser escrito:

$$
\underset{m\times n}{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \dots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}
$$

Sendo A a matriz dos coeficientes. Se adicionarmos a coluna correspondente ao lado direito do sistema (2) teremos a matriz aumentada:

$$
\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \vdots & \dots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{pmatrix}
$$

Definição. A matriz é dita escalonada por linhas se cada linha subsequente começa com mais zeros que a linha anterior.

Exemplo:

$$H = \begin{pmatrix} 1 & -0.4 & -0.3 & | & 130 \\ 0 & 0.8 & -0.2 & | & 100 \\ 0 & 0 & 0.7 & | & 210 \end{pmatrix}$$

Definição. Dizemos que a linha de uma matriz tem k zeros líderes se os k primeiros elementos da linha são todos zero e o *(k+1)-iésimo* elemento de cada linha é não nulo. Com essa terminologia podemos dizer que a matriz está escalonada por linha se cada linha tem mais zeros líderes do que a que a precede.

.

Definição. Dizemos que uma matriz em forma escalonada por linhas está na forma reduzida por linhas se cada pivô é 1 e cada coluna que contém um pivô não contém outros elementos não nulos.

3.3 Sistemas com muitas soluções ou nenhuma

No caso mais simples mostraremos um sistemas com duas equações e duas incógnitas. No geral, retas no plano são não paralelas e se cruzam em um único ponto. Caso essas retas sejam paralelas ou elas coincidem ou nunca se cruzam. Se elas coincidem o sistema tem infinitas soluções. No último caso, o sistema não apresenta solução.

Definição. Se a *j-ésima* coluna de $\hat{B}$ não contém um pivô, dizemos que x_j é uma variável livre ou não básica.

Definição. Se a *j-ésima* coluna de $\hat{B}$ (escalonada por linhas) contém um pivô, dizemos que x_j é uma variável básica.

$$\begin{array}{cccccc} w & +2x & +y & -z & = & 1 \\ 3w & -x & -y & +2z & = & 3 \\ & -x & +y & -z & = & 1 \\ 2w & +3x & +3y & -3z & = & 3 \end{array}$$

$$\begin{pmatrix} w & +2x & +y & -z & | & 1 \\ 3w & -x & -y & +2z & | & 3 \\ & -x & +y & -z & | & 1 \\ 2w & +3x & +3y & -3z & | & 3 \end{pmatrix}$$

Na forma escalonada reduzida por linhas:

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{3}{11} & | & \frac{12}{11} \\ 0 & 1 & 0 & -\frac{1}{11} & | & -\frac{4}{11} \\ 0 & 0 & 1 & -\frac{12}{11} & | & \frac{7}{11} \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

Então temos que:

$$w = \tfrac{12}{11} - \tfrac{3}{11}z$$
$$x = -\tfrac{4}{11} + \tfrac{1}{11}z$$
$$w = \tfrac{7}{11} + \tfrac{12}{11}z$$

Assim, z é a única variável livre e as demais são variáveis básica.

3.4 Posto o Critério Fundamental

Definição. o posto de uma matriz é o número de linhas não-nulas em sua forma escalonada por linhas.

Fato. *7.1: Sejam A a matriz de coeficientes e $\hat{A}$ a matriz aumentada correspondente. Então,*

(a) posto de $A \leq$posto $\hat{A}$

(b) posto de $A \leq$número de linhas de A

(c) posto de $A \leq$número de colunas de A

Fato. *7.2: Um sistema de equações lineares com matriz de coeficientes A e matriz aumentada $\hat{A}$ possui uma solução se, e somente se, posto de $\hat{A}$ = posto de A*

7.3: Um sistema linear de equações não tem nenhuma solução, ou apenas uma solução, ou infinitas soluções. Assim, se um sistema tiver mais do que uma solução, então terá infinitas soluções.

Fato. *7.4: Se um sistemas tem exatamente uma solução, então a matriz de coeficientes A tem pelo menos tantas linhas quanto colunas. Em outras palavras, um sistema com solução única deve ter pelo menos tantas equações quanto variáveis.*

Fato. *7.5: Se um sistema de equações lineares tem mais incógnitas do que equações, então o sistema não tem nenhuma solução ou tem uma quantidade infinita de soluções.*

$$\begin{array}{ccc} a_{11}x_1+ & ...+ & a_{1n}x_n = 0 \\ . & . & . \\ a_{m1}x_1+ & ...+ & a_{mn}x_n = 0 \end{array}$$

Esse sistema é dito homogêneo, e tem pelo menos uma solução $x_1 = x_2 = ... = x_n = 0$

Fato. *7.6: Um sistema homogêneo de equações lineares com mais incógnitas do que equações necessariamente possui uma infinidade de soluções distintas*

Fato. *7.7: Um sistema de equações lineares com matriz de coeficientes A tem uma solução para cada escolha de lado direito $b_1, ..., b_m$se, e somente se,*

posto$A$ = número de linhas de A

Fato. *7.8: Se um sistemas de equações lineares tem mais equações do que incógnitas, então existe um lado direito tal que o sistema resultante não possui solução*

Fato. *7.9: Qualquer sistema de equações lineares com matriz de coeficientes A tem no máximo uma solução para cada escolha de lado direito $b_1, ..., b_m$se, e somente se, posto A = número de colunas de A.*

7.10: Uma matriz d coeficientes A é não-singular, ou seja, o sistema linear correspondente tem uma, e só uma solução para cada escolha do lado direito $b_1, ..., b_m$se, e somente se, número de linhas de A = número de colunas de A = posto de A

7.11: Considere o sistema linear de equações $Ax = b$

(a) Se o número de equações <o número de incógnitas, então:

(i) $Ax = 0$ tem um número infinito de soluções;

(ii) para qualquer b dado, $Ax = b$ tem 0 ou um número infinito de soluções, e

(iii) se posto A = número de equações, $Ax = b$ tem um número infinito de soluções para cada escolha do lado direito de b.

(b) Se o número de equações > o número de incógnitas, então:

(i) $Ax = 0$ tem um ou um número infinito de soluções

(ii) para qualquer b dado, $Ax = b$ tem 0, uma ou um número infinito de soluções, e

(iii) se posto A = número de incógnitas, $Ax = b$ tem 0 ou uma solução para cada escolha do lado direito b.

(c) Se o número de equações = o número de incógnitas, então

(i) $Ax = 0$ tem uma ou um número infinito de soluções

(ii) para qualquer b dado, $Ax = b$ tem 0 ou um número infinito de soluções, e

(iii) se posto A = número de incógnitas = número de equações, $Ax = b$ tem exatamente uma solução para cada escolha de lado direito b.

4 Álgebra Matricial

Teorema. *Sejam A uma matriz k x m e B uma matriz m x n. Então $(AB)^T = B^T A^T$*

Demonstração. Para isso precisamos da definição de matriz transposta. □

Definição. Matriz transposta$\left((AB)^T\right)_{ij} = (AB)_{ji}$

$$= \sum_h A_{jh} B_{hi}$$

$$= \sum_h \left(A^T\right)_{hj} \left(B^T\right)_{ih}$$

$$= \sum_h \left(B^T\right)_{ih} \left(A^T\right)_{hj}$$

$$= \left(B^T A^T\right)_{ij}$$

Portanto,

$$(AB)^T = B^T A^T \blacksquare$$

4.1 Tipos Especiais de Matrizes

Problemas especiais utilizam tipos especiais de matrizes. Nesta seção descreveremos algumas importantes classes de matrizes k x n que surgem na análise econômica.

- **Matriz quadrada:** $k = n$, número de linhas igual ao número de colunas.
- **Matriz coluna**: $n = 1$.
- **Matriz linha:** $k = 1$.
- **Matriz diagonal:** $k = n$, $\forall i \neq j$ $a_{ij} = 0$.
- **Matriz triangular superior:** $a_{ij} = 0$ se $i > j$ (geralmente quadrada) na qual cada entrada abaixo da diagonal principal é 0. $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$
- **Matriz triangular inferior:** $a_{ij} = 0$ se $i < j$. $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$
- **Matriz simétrica:** $A^T = A$, $a_{ij} = a_{ji} \forall i, j$. Essas matrizes são necessariamente quadradas.
- **Matriz idempotente:** Uma matriz quadrada B tal que $B.B = B$.
- **Matriz de permutação:** Uma matriz quadrada de entradas 0 e 1, na qual cada linha e cada coluna contêm exatamente 1. Ex: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
- **Matriz não-singular:** Uma matriz quadrada cujo posto é igual ao número de linhas (colunas).

4.2 Matrizes Elementares

Recorde que as três operações elementares sobre linhas são utilizadas para trazer uma matriz à forma escalonada por linhas:

1. permutação de linhas,

2. soma de um múltiplo de uma linha a uma outra linha, e

3. multiplicação de uma linha por um escalar não-nulo.

Essas operações podem ser efetuadas em uma matriz A pela multiplicação à esquerda por certas matrizes especiais denominadas *matrizes elementares.* Por exemplo, o seguinte teorema ilustra como permutar as linhas i e j de uma dada matriz A.

Teorema. *Forme a matriz de permutação E_{ij} pela permuta da i-ésima com a j-ésima linha da matriz identidade I. Então, a multiplicação à esquerda de uma matriz A por E_{ij} tem efeito de permutar a i-ésima com j-ésima linha de A.*

Demonstração. Para verificar isso, denotaremos por e_{hk} uma entrada qualquer de E_{ij}:

$$\begin{gathered} e_{ij} = e_{ji} = 0 \\ e_{ii} = e_{jj} = 0 \\ e_{hh} = 1 \text{ se } h \neq i, j \\ e_{kk} = 0 \text{ caso contrário} \end{gathered} \tag{3}$$

O elemento na linha k e coluna n de $E_{ij}A$ é

$$\sum_m e_{km} a_{mn} = \begin{Bmatrix} a_{jn} & k = i \\ a_{in} & k = j \\ a_{kn} & k \neq i, j \end{Bmatrix}$$

□

por (1). Portanto, $E_{ij}A$ é simplismente A com as linhas i e j trocadas entre si. ■

Exemplo. Suponha uma matriz A de tamanho 3 x 3, temos:

$$E_2(5).A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 5a_{21} & 5a_{22} & 5a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$E_{ij}(r) \rightarrow$ é a soma r vezes a linha i pela linha j da matriz I.

$$E_{23}(5).A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 5a_{21} + a_{31} & 5a_{22} + a_{32} & 5a_{23} + a_{33} \end{pmatrix}$$

Definição. As Matrizes E_{ij}, $E_{ij}(r)$ e $E_i(r)$, que foram obtidas executando as operações elementares sobre linhas na matriz identidade, são denominadas **matrizes elementares**.

Teorema. *Seja E uma matriz elementar n x n obitida executando-se uma dada operação elementar sobre linhas na matriz identidade n x n. Se A é uma matriz n x n qualquer, então EA é a matriz obtida executando aquela mesma operação elementar sobre linhas em A.*

Teorema. *Dada qualquer matriz A de tamanho k x n, existem matrizes elementares E_1, $E_2 \cdots, E_m$ tais que o produto matricial $E_m.E_{m-1}.A = U$, onde U está em forma escalonada (reduzida) por linhas.*

4.3 Álgebra de Matrizes Quadradas

Usamos a notação M_n para classe de matrizes quadradas do tipo n x n

Definição. Seja A uma matriz em M_n. Uma matriz B em M_né uma **inversa** para A se $AB = BA = I$.

Se existir a matriz B, dizemos que A é **invertível**.

Teorema. *Uma matriz A de tamanho n x n pode ter, no máximo, uma única inversa.*

Demonstração. Suponha que B e C sejam inversas de A. Então,

$$C = CI = C(AB) = (CA)B = IB = B.\blacksquare$$

$\square$

Definição. Seja A uma matriz de tamanho k x n, a matriz B de tamanho n x k é uma **inversa à direita** de A se $AB = I$. A Matriz C de tamanho n x k é uma **inversa à esquerda** de A se $CA = I$.

Lema. *Se uma matriz A tem uma inversa à direita B e uma inversa à esquerda C, então A é invertível e $B = C = A^{-1}$. A Prova é análoga a do teorema 8.5 .*

Teorema. *Se uma matriz A de tamanho n x n é invertível, então A é não-singular e a única solução do sistema de equações lineares $Ax = b$ é $x = A^{-1}b$.*

Demonstração. Desejamos mostrar que se A é invertivel, então podemos resolver qualquer sistema de equções do tipo $Ax = b$. Multiplique cada lado deste sistema por A^{-1} para resolver em x como segue:

$$Ax = b$$

$$A^{-1}(Ax) = A^{-1}b$$

$$(A^{-1}A)x = A^{-1}b$$

$$Ix = A^{-1}b$$

$$x = A^{-1}b.\blacksquare$$

$\square$

Teorema. *Se uma matriz A de tamanho n x n é não-singular, então A é invertivel.*

Demonstração. Suponha que A é não-singular. Denotamos e_i a i-ésima coluna de I. Sendo A não-singular a equação $AX = e_i$ tem uma única solução $X = c_i$. Seja C a matriz cujas n colunas são as respectivas soluções $c_{1,} \ldots, c_n$. Como multiplicamos cada linha de A pela j-ésima coluna de C para obter a j-ésima coluna de AC, podemos escrever

$$AC = A\,[c_1, \ldots, c_n]$$

$$= [Ac_1, \ldots, Ac_n]$$

$$= [e_1, \ldots, e_n]$$

$$= I \tag{4}$$

Assim C é uma inversa a direita de A. Para ver que A também possui uma inversa a esquerda, use o teorema 8.4 para escrever $EA = U$, onde E é um produto de matrizes elementares e U é a forma escalonada reduzida por linhas de A. Como A é não-singlar U não tem linha de zero e cada coluna contém exatamente 1, $U = I$. Portanto, E é uma inversa a esquerda de A. Como A tem uma inversa à direita e uma inversa à esquerda, A é invertível. $\blacksquare$

Podemos ser mais eficientes aglutinando todas essas informações em uma matriz aumentada gigantesca $(A \mid e_1, \ldots, e_n) = (A \mid I)$ e executar a eliminação de Gauss-Jordan somente uma única vez em vez de n vezes. Nesse processo, a matriz aumentada se reduz a $(I \mid A^{-1})$. $\square$

Exemplo. 8.4

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5}$$

$$(A \mid I) = \begin{pmatrix} a & b & \mid & 1 & 0 \\ c & d & \mid & 0 & 1 \end{pmatrix}$$

Se $a = c = 0$, A é singular. Vamos supor que $a \neq 0$ primeiro somamos $-c/a$ vezes a linha 1à linha 2, para obter a forma escalonada por linhas.

$$\begin{pmatrix} a & b & \mid & 1 & 0 \\ 0 & \frac{ad-bc}{a} & \mid & \frac{-c}{a} & 1 \end{pmatrix} \tag{6}$$

Se $a \neq 0$, A é não-singular se, e somente se, $ad - bc \neq 0$. Multiplique a primeira linha por $1/a$ e a segunda linha por $a/(ad - bc)$.

$$\begin{pmatrix} 1 & \frac{b}{a} & \mid & \frac{1}{a} & 0 \\ 0 & 1 & \mid & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

Some $-b/a$ vezes a linha 2 da 1.

$$\begin{pmatrix} 1 & 0 & \mid & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \mid & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{7}$$

Teorema. *A matriz arbitrária A de tamanho 2 x 2 dada por* (3) *é não-singular (e portanto invertível) se, e somente se,* $ad - bc \neq 0$. *Sua inversa é a matriz* (5).

Teorema. *Para qualquer matriz quadrada A, são equivalentes as seguintes informações:*

(a) A é invertível.

(b) A tem uma inversa à direita.

(c) A tem uma inversa à esquerda.

(d) O sistema $Ax = b$ tem pelo menos uma solução para cada b.

(e) O sistema $Ax = b$ tem no máximo uma solução para b.

(f) A é não-singular.

(g) A tem posto máximo.

Demonstração. Na seção 7.4 vimos a equivalência das afirmações d) a g). Os enunciados e as provas dos teoremas 8.6 e 8.7 garantem que as afirmações a) a d) são equivalentes. □

Teorema. *Sejam A e B matrizes quadradas invertíveis.*

Então,

(a) $(A^{-1})^{-1} = A$

(b) $(A^T)^{-1} = (A^{-1})^T$

(c) AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$

Teorema. *Se A é invertível:*

(a) A^m é invertível para qualquer inteiro m e $(A^m)^{-1} = (A^{-1})A^{-m}$

(b) Para quaisquer inteiros r e s, $A^r A^s = A^{r+s}$, e

(c) para qualquer escalar $r \neq 0$, rA é invertível e $(rA)^{-1} = (1/r)A^{-1}$.

Teorema. *Qualquer matriz pode ser escrita como um produto $A = F_1, \ldots, F_m U$ no qual as F_i são matrizes elementares e U está na forma escalonada reduzida por linhas. Quando A é não-sigunlar $U = I$ e $A = F_1, \ldots, F_m$.*

Lema. *Sejam L e M duas matrizes triangulares inferiores n x n. Então o protudo matricial LM é trianglar inferior. Se L e M têm somente 1 em suas diagonais, então o mesmo ocorre com LM.*

Demonstração. A $(i,j)-$ésima entrada do produto LM é o produto da $i-$ésima linha de L com a $j-$ésima coluna de M. Usando a hipótese que $l_{ik} = 0$ para $k > i$ e $m_{hj} = 0$ para $h < j$, escrevemos esse produto como:

$$(LM)_{ij} = (l_{i1}, \ldots, l_{i,i-1}, l_{ii}, 0 \ldots 0) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{jj} \\ m_{j+1,j} \\ \vdots \\ m_{nj} \end{pmatrix} \tag{8}$$

Se $i < j$, cada uma das i possivelmente não-nulos entradas no começo da $i-$èsima linha de L será multiplicada pelas i entradas zero do começo da $j-$ésima coluna de M. O resultado é uma entrada zero em LM. Portanto LM é triangular inferior.

A partir de (6) a $(i,i)-$ésima entrada na diagonal de LM é $l_{ii} = m_{ii} = 1$. $\square$

Teorema. *Seja A uma matriz arbitrária k x n suponha que não é necessário efetuar permuta de linhas para reduzir A à sua forma escalonada por linhas. Então A pode ser escrita como um produto LU, onde L é uma matriz trinagular inferior k x k com entradas 1 na diagonal e U é uma matriz triangular superior k x n.*

4.4 Decomposição LU

Vamos resolver o sistema $Ax = b$ da forma $LUx = b$. Primeiro tome $Ux = Z$ e $LZ = b$ e então resolva.

$UX = Z$

$$\begin{pmatrix} 2 & 4 & 0 \\ 4 & 6 & 3 \\ -6 & -10 & 0 \end{pmatrix} = \overset{L}{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{pmatrix}} \overset{U}{\begin{pmatrix} 2 & 4 & 0 \\ 0 & -2 & 3 \\ 0 & 0 & 3 \end{pmatrix}} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \overset{b}{\begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix}}$$

$LZ = b$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix} \Rightarrow \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ 3 \end{pmatrix}$$

$UX = Z$

$$\begin{pmatrix} 2 & 4 & 0 \\ 0 & -2 & 3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ -3 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

4.5 Produto de Kronecker

Seja A_{mxp} e B_{nxq} então

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1p}B \\ \vdots & \ddots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mp}B \end{pmatrix}$$

$$A \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}_{2x3}$$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{2x2}$$

$$I \otimes A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 & a_{21} & a_{22} & a_{23} \end{pmatrix}$$

$$A \otimes I = \begin{pmatrix} a_{11}0 & a_{12}0 & a_{13}0 \\ 0a_{11} & 0a_{12} & 0a_{13} \\ a_{21}0 & a_{22}0 & a_{23}0 \\ 0a_{21} & 0a_{22} & 0a_{23} \end{pmatrix}$$

Propriedades

(1) $(A \otimes B)' = A' \otimes B'$

(2) $(A \otimes B)(C \otimes D) = AC \otimes BD$

(3) $A \otimes (B + C) = A \otimes B + A \otimes C$

(4) $(A + C) \otimes A = B \otimes A + C \otimes A$

(5) $A \otimes (B \otimes B) = (A \otimes B) \otimes B$ Implica que A e B são quadradas e não-singulares.

(5) e (2) implica que $(A \otimes B)(A^{-1} \otimes B^{-1}) = AA^{-1} \otimes BB^{-1} = I_{mm}xI_{nn} = I_{mn_xmn}$.

4.6 Vetorização de Matrizes

Seja A_{mxn} então $VEC(A) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$

Exemplo. $A \begin{pmatrix} 12 & -6 \\ 3 & 15 \end{pmatrix}$ $VEC(A) = \begin{pmatrix} 12 \\ -6 \\ 3 \\ 15 \end{pmatrix}$

5 Determinantes

5.1 Menores de uma Matriz

Definição. Seja A uma matriz $n\ x\ n$. Seja A_{ij}a submatriz $(n-1) \times (n-1)$ obtida suprimindo-se a i-ésima linha e a j-ésima coluna de A. O escalar $M_{ij} \equiv detA_{ij}$ é denominado o (i,j)-ésimo **menor** de A e o escalar $C_{ij} \equiv (-1)^{i+j} M_{ij}$ é denominado o (i,j)-ésimo **co-fator** de A. Observe que $M_{ij} = C_{ij}$ se (i+j) é par e $M_{ij} = -C_{ij}$ se (i+j) é ímpar.

Exemplo. Uma matriz 3 x 3:

$$\det \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

$$det \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Definição. O **determinante** de uma matriz A de tamanho n x n é dado por:

$$detA = a_{11}C_{11} + a_{12}C_{12} + ... + a_{1n}C_{1n}$$

$$detA = a_{11}M_{11} - a_{12}M_{12} + ... + (-1)^{n+1}a_{1n}M_{1n}$$

Teorema. ***9.1***

O determinante de uma matriz triangular inferior ou triangular superior é simplesmente o produto de suas entradas diagonais.

Teorema. ***9.2***

Seja A uma matriz de tamanho n x n e seja R sua forma escalonada por linhas. Então

$$detA = \pm detR$$

Se não tiverem sido usadas permutações de linhas para obter R de A, então $detA = detR$.

5.2 Matriz Adjunta

Teorema. ***9.3***

Uma matriz quadrada é não-singular se, e somente se, seu determinante é não nulo.

Definição. Para qualquer matriz A de tamanho $n \times n$, seja C_{ij}o (i, j)-ésimo *co-fator de A, ou seja,* $(-1)^{i+j}$vezes o determinante da submatriz obtida suprimindo a i-ésmia linha e a j-ésima coluna de A. A matriz de tamanho n x n cuja (i, j)-ésima entrada é C_{ij}, o (i,

j)-ésimo co-fator de A é denominada matriz **adjunta** de A e é denotada por **Adj**A. Em outras palavras, a matriz adjunta é a matriz transposta da matriz de cofatores.

Teorema. ***9.4***

Seja A uma matriz não-singular. Então,

(a) $A^{-1} = \frac{1}{detA} \cdot adjA$

(b) *A única solução* $x = (x_1, ..., x_n)$ *do sistema* $Ax = b$ *de tamanho* n x n *é*

$x_i = \frac{detB_i}{detA}$, *para* $i = 1, ..., n$. *Onde* B_i*é a matriz A com o lado direito* ***b*** *substituindo a i-ésima coluna de A.*

Exemplo. $A = \begin{pmatrix} 4 & 5 \\ 1 & 1 \end{pmatrix}$ $detA = -1$

$adjA = \begin{pmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{pmatrix}$ $C_{11} = 1; C_{12} = -1; C_{21} = -5; C_{22} = 4$

$A^{-1} = \frac{1}{-1} \cdot \begin{pmatrix} 1 & -5 \\ -1 & 4 \end{pmatrix} \Rightarrow A^{-1} = \begin{pmatrix} -1 & 5 \\ 1 & -4 \end{pmatrix}$

Teorema. ***9.5***

Seja A uma matriz quadrada. Então,

(a) $detA^T = detA$

(b) $det(A \cdot B) = det(A) \cdot det(B)$, *e*

(c) $det(A + B) \neq detA + detB$, *em geral.*

6 Espaços Euclidianos

6.1 Comprimento e Distância

Se P e Q são dois pontos no $\mathbb{R}^n$, escrevemos $\overline{PQ}$ para o seguimento ligando P e Q e $\overrightarrow{PQ}$ para o vetor de P a Q. O comprimento do segmento de reta $\overline{PQ}$ é denominado pelo símbolo $\left\|\overline{PQ}\right\|$, valor absoluto na reta. Sejam P e Q $\epsilon\mathbb{R}^2$, para calcularmos o comprimento do segmento ℓ ligando esses dois pontos $P\,(a_1, b_1)$ e $Q\,(a_2, b_2)$ marcamos um ponto intermediário $R\,(a_2, b_1)$. Seja m o segmento de reta (horizontal) de $P\,(a_1, b_1)$ a $R\,(a_2, b_1)$ e n o segmento de reta (vertical) de $Q\,(a_2, b_2)$ a $R\,(a_2, b_1)$. O triângulo correspondente PRQ é um triângulo retângulo cuja hipotenusa é o segmento de reta ℓ.

$$\ell^2 = |a_1 - a_2|^{\,2} + |b_1 - b_2|^{\,2}$$

$$\left\|\overline{PQ}\right\| = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$$

$\|x - y\|$ é o vetor ligando os pontos x e y respectivamente e seu cmprimento é $\|x - y\|$ que é o mesmo que a distância entre esses dois pontos.

$$\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$$

Em particular se y é $(0, ..., 0)$ a distância do ponto $x = (x_1, x_2, ..., x_n)$ à origem ou o comprimento do vetor x é $\|x\| = \sqrt{x_1^2 + ... + x_n^2}$

Teorema. *10.1* $\|r.v\| = \|r\| . \|v\| \forall r \epsilon \mathbb{R}$ *e* $v \epsilon \mathbb{R}^n$

Demonstração.

$$\|r\,(v_1, ..., v_n)\| = \|(rv_1, ..., rv_n)\|$$

$$= \sqrt{r^2\,(v_1)\,^2 + ... + r^2\,(v_n)^2}$$

$$|r|\,\sqrt{v_1^2 + ... + v_n^2}$$

$$\|r\| . \|v\|$$

□

Soponha que V seja um vetor de deslocamento não-nulo. Ocasionalmente precisamos encontrar um vetor w que tenha a mesma direção e sentido que v, mas possua o comprimento 1. Tal vetor w é denominado vetor unitário na direção de v. Para obter tal vetor w, simplesmente multiplique v pelo escalar $r = \frac{1}{\|v\|}$.

$$\left\|\frac{1}{v}.v\right\| = \left|\frac{1}{\|v\|}\right| . \|v\| = \frac{1}{\|v\|} . \|v\| = 1$$

Ex:

$$\|(1, -2, 3)\| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}$$

$$\frac{1}{\sqrt{14}}\,(1, -2, 3) = \left(\frac{1}{\sqrt{14}}, \frac{-2}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right)$$

6.2 Produto Interno (Produto Escalar)

Definição. Sejam $u = (u_1, ..., u_n)$ e $v = (v_1, ..., v_n)$ dois vetores em $\mathbb{R}^n$. O produto interno euclidiano de u e v, denotado por $u.v$ é o número

$$u.v = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

Teorema. *10.2 Sejam u,v e w vetores aleatórios em $\mathbb{R}^n$ e r um escalar. Então,*

(a) $u.v = v.u$

(b) $u.(v+w) = u.v + u.w$

(c) $u.(rv) = r(u.v) = (r.u).v$

(d) $u.u \geq 0$

(e) $u.u = 0 \longrightarrow u = 0$, *e*

(f) $(u+v).(u+v) = u.u + 2(u.v) + v.v$

$u.u = u_1^2 + ... + u_n^2$

$\|u\| = \sqrt{u.u}$

$\|u-v\| = \sqrt{(u-v).(u-v)}$

Teorema. *10.3 Sejam u e v dois vetores em $\mathbb{R}^n$. Seja θ o ângulo entre dois vetores. Então,*

$$u.v = \|u\|.\|v\|.cos\theta$$

Lembramos que o cosseno de θ está entre -1e 1

$cos\theta > 0$ se θ é agudo

$cos\theta < 0$ se θ é obtuso

$cos\theta = 0$ se θ é reto

Demonstração. Suponha que u e v sejam vetores com ponto inicial na origem 0; Digamos que $u = \vec{O}P$ e $v = \vec{O}Q$. Seja ℓ a reta pelo vetor v, ou seja, a reta pelos pontos 0 e Q. Seja m o segmento de reta perpendicular do ponto P à reta ℓ. Seja R o ponto em que m encontra ℓ. Como R está em ℓ, $\vec{O}R$ é um múltiplo escalar de $v = \vec{O}R$. Escreva $\vec{O}R$ como $t.v$. Como u, $t.v$ e o segmento m são os três lados do triângulo retângulo POR, podemos escrever m como o vetor $u - tv$. Como u é a hipoensa desse triângulo retângulo,

$$cos\theta = \frac{\|tv\|}{\|u\|} = \frac{t\|v\|}{\|u\|}(4)$$

Por outro lado, pelo teorema de Pitágoras e o teorema 10.2, o quadrado do comprimento da hipotenusa é: □

$$\|u\|^2 = \|t.v\|^2 + \|u - t.v\|^2$$

$$= t^2\|v\|^2 + (u-tv).(u-tv)$$

$$= t^2\|v\|^2 + u.u - 2u.(tv) + (tv).(tv)$$

$$= t^2 \|v\|^2 + \|u\|^2 - 2t.(uv) + t^2 \|v\|^2$$

Segue que

$$2t.(uv) = 2t^2 \|v\|^2$$

$$t = \frac{u.v}{\|v\|^2} \; (5)$$

Substituindo (5) em (4) resuta:

$$cos\theta = \frac{u.v}{\|v\|^2} . \frac{\|v\|}{\|u\|} = \frac{u.v}{\|u\| . \|v\|}$$

Exemplo. 10.3

Considere um cubo em $\mathbb{R}^3$ com cada aresta de comprimento c. Posicion este cubo em $\mathbb{R}^3$, de modo natural, ou seja, com os vértices em $O(0,0,0)$, $P_1(c,0,0)$, $P_2(0,c,0)$ e $P_3(0,0,c)$. Escreva u_i para o vetor $\overrightarrow{O}P_i$ com $i = 1,2,3$. Então, a diagonal d é $u_1 + u_2 + u_3$, que é o vetor (c,c,c).

O ângulo θentre u_1e d satifaz

$$cos\theta = \frac{u_1.d}{\|u_1\| . \|d\|} = \frac{(c,0,0).(c,c,c)}{c.\sqrt{c^2+c^2+c^2}} = \frac{c^2}{c.3\sqrt{c^2}} = \frac{1}{\sqrt{3}}$$

Teorema. *10.4 O ângulo entre os vetores u e v em $\mathbb{R}^n$ é:*

(a) agudo se $u.v > 0$

(b) obtuso se $u.v < 0$

(c) reto se $u.v = 0$

Quando esse ângulo é reto dizemos que u e v são ortogonais. Dizemos que u e v são ortogonais se, e somente se, $u.v = u_1v_1 + ... + u_nv_n = 0$

Teorema. *10.5 Para quaisquer vetores u e v em$\mathbb{R}^n$,*

$$\|u+v\| \leq \|u\| + \|v\|$$

Essa regra afirma que qualquer lado de um triângulo é mais curto do que a soma dos comprimentos dos outros dois lados.

Demonstração. $\frac{u.v}{\|u\|\|v\|} = cos\theta \leq 1$ A função cosseno está definida da seguinte forma: $f : \mathbb{R} \to [-1,1]$

Pelo teorema 10.3 $u.v \leq \|u\| . \|v\|$→Multiplique por 2 e some por $\|u\|^2 + \|v\|^2$

$$\|u\|^2 + 2(u.v) + \|v\|^2 \leq \|u\|^2 + 2\|u\| . \|v\| + \|v\|^2$$

$$u.u + u.v + v.u + v.v \leq (\|u\| + \|v\|)^2$$

$$(u+v) \,.\, (u+v) \leq (\|u\| + \|v\|)^2$$

$$\|u+v\|^2 \leq (\|u\| + \|v\|)^2$$

$$\|u+v\| \leq (\|u\| + \|v\|)$$

□

Teorema. *10.6 Para quaisquer dois vetores x e y em $\mathbb{R}^n$,* $|\|x\| - \|y\|| \leq \|x - y\|$

Demonstração. Aplique o teorema 10.5 com $u = x - y$ e $v = y$ em (6) para obter a desigualdade

$$\|x\| \leq \|x-y\| + \|y\| \text{ ou } \|x\| - \|y\| \leq \|x-y\| \; (7)$$

Aplique o teorema 10.5 com $u = y - x$ e $v = x$ em (6)

$$\|u+v\| \leq \|u\| + \|v\|$$

$$\|y - x + x\| \leq \|y - x\| + \|x\|$$

$$\|y\| - \|x\| \leq \|y - x\| \;\; (8)$$

As desigualdades (7) e (8) implicam

$$|\|x\| - \|y\|| \leq \|x - y\|$$

□

6.3 Propriedades do comprimento Euclidiano: (norma)

1. $\|u\| \geq 0$ e $\|u\| = 0$ somente se $u = 0$
2. $\|ru\| = |r| \, \|u\|$
3. $\|u + v\| \leq \|u\| + \|v\|$

Qualquer associação de números reais e vetores que satisfaça essas 3 propriedades é denominada norma.

Exercício. 10.15 Prove que $\|u-v\|^2 = \|u\|^2 - 2u.v + \|v\|^2$

$\|u-v\|^2 = (u-v).(u-v) = u.u - 2u.v + v.v = \|u\|^2 - 2u.v + \|v\|^2$

Exercício. 10.16 Prove que $\|\|(u_1,u_2)\|\| = |u_1|+|u_2|(1)$ e $\|\|(u_1,u_2)\|\| = max\{|u_1|,|u_2|\}(2)$ são normas no $\mathbb{R}^2$

Prova (1)

$$\|u| \geq 0$$

$$|u_1| + |u_2| \geq 0$$

se, e somente se mbos são iguais a zero.

$$\|ru\| = |ru_1| + |ru_2| = |r|\left(|u_1| + |u_2|\right) = r\,\|u\|$$

$$\|u+v\| = |u_1+v_1| + |u_2+v_2| \leq |u_1| + |u_2| + |v_1| + |v_2| = \|u\| + \|v\|$$

(2)

$$\|u\| = max\{|u_1|,|u_2|\} \geq 0$$

se, e somente se ambos forem iguais a zero

$$\|ru\| = max\{|ru_1|,|ru_2|\} = |r|\,max\{|u_1|,|u_2|\} = |r|$$

$$\|u+v\| = max\{|u_1+v_1|,|u_2+v_2|\} \leq max\{|u_1|+|v_1|,|u_2|+|v_2|\}$$

$$\leq max\{|u_1|,|u_2|\} + max\{|v_1|,|v_2|\} = \|u\| + \|v\|$$

6.4 Produto Vetorial

É uma multiplicação usada para vetores de $\mathbb{R}^3$. Nesta operação o produto de 2 vetores no $\mathbb{R}^3$ é um vetor no $\mathbb{R}^3$.

$$(u_1,u_2,u_3).(v_1,v_2,v_3) = (u_2v_3 - u_3v_2; u_3v_1 - u_1v_3; u_1v_2 - u_2v_1)$$

$$= \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, -\begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right)$$

As propriedades do produto vetorial são as seguintes:

a) $u.v = -v.u$

b) $u.v$ é perpendicular a u

c) $u.v$ é perpendicular a v

d) $(ru).v = r(u.v) = u.(rv)$

e) $(u_1 + u_2).v = (u_1.v) + (u_2.v)$

f) $\|u.v\| = \|u\|.\|v\| \, sen\theta$

g) $\|u.v\|^2 = \|u\|^2.\|v\|^2 - (u.v)^2$

h) $u.u = 0$

i) $u.(v.w) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

6.5 Retas

Os objetos da geometria euclidiana são retas e planos e pontos. Inicialmente trabalhamos com retas no $\mathbb{R}^2$

Por exemplo:

$$x_2 = mx_1 + b$$

Não "sabemos" resolver $x_2$em termos de x_1. Formalmente precisamos de uma representação paramétrica que expressa x_1e $x_2$em termos de t. O ponto x em $\mathbb{R}^2$ "está na reta" se $x = (x_1(t^*), x_2(t^*))$ com $t\epsilon\mathbb{R}$. Uma reta fica claramente determinada por dois aspectos: um ponto $x_0$na reta e uma direção v na qual move-se a patir de x_0.

$$x(t) = x_0 + tv$$

Ex: A reta que passa pelo ponto (4,2) na direção (1,1)

$$x(t) = (x_1(t), x_2(t))$$

$$= (4, 2) + t(1, 1)$$

$$= (4 + t.1, 2 + t.1)$$

$$x_1 = 4 + t.1$$

$$x_2 = 2 + t.1$$

Uma outra maneira de determinar uma reta é identificando dois de seus pontos. Digamos que x e y são dois pontos da reta ℓ. Então, ℓ pode ser vista como a reta que passa por x e aponta na direção $y - x$. Assim uma parametrização para essa reta é

$$x(t) = x + t(y - x)$$

$$= x + ty - tx$$

$$= (1 - t)x + ty$$

Quando $t = 0$ estamos no pontox; e quando $t = 1$ estamos no ponto y. Se $t\epsilon[0, 1]$ estamos em pontos entre x e y.

Definição. Segmento de reta

O segmento de reta de x a y pode ser expresso como $\ell(x, y) = \{(1 - t)x + ty : 0 \leq t \leq 1\}$

Definição. Plano

Seja P um plano em $\mathbb{R}^3$ pela origem. Sejam v e w dois vetores em P, escolha v e w de modo que apontem para direções diferentes, de tal modo que nenhum deles seja um múltiplo escalar do outro. Para quaisquer escalares s e t o vetor $sv + tw$ é denominado combinação escalar de v e w. Pela nossa interpretação geométrica da adição e da multiplicação, fica claro que todas as combinações lineares de v e w estão no plano P.

$$x = sv + tw$$

$$x_1 = sv_1 + tw_1$$

$$x_2 = sv_2 + tw_2$$

$$x_3 = sv_3 + tw_3$$

6.6 Parametrização

Se o plano P não passa pela origem, mas sim pelo ponto $p \neq 0$ e se v e w são dois vetores direcionais linearmente independentes posicionados em p que estão no plano, podemos parametrizar o plano da seguinte for

$$x = p + sv + tw \; \forall s, t\epsilon\mathbb{R}$$

Para encontrarmos a equação paramétrica contendo os pontos p. q e r observa-se que podemos visualizar $q-p$ e $r-p$ como vetores deslocamento localizados nesse plano com ponto inicial em p.

$$x\left(s,t\right)=p+s\left(q-p\right)+t\left(r-p\right)$$

$$=\left(1-s-t\right)p+sq+tr$$

Uma reta em $\mathbb{R}^2$ e um plano em $\mathbb{R}^3$ são exemplos de conjuntos descritos por uma simples equação linear em $\mathbb{R}^n$. Tais espaços são muitas vezes denominados de hiperplanos. Um hiperplano em $\mathbb{R}^n$ pode ser escrito em forma ponto-normal como:

$$a_1x_1+a_2x_2+...+a_nx_x=d$$

O hiperplano pode ser entendido como o conjunto de todos os vetores com cauda em $\left(0,..,0,\frac{d}{a_n}\right)$ que são perpendiculares ao vetor $n=\left(a_1,...a_n\right)\rightarrow$n é um vetor normal ao hiperplano. Um hiperplano que surge frequentemente nas aplicações é o espaço de vetores-probabilidade

$$P_n=\{(p_1,...,p_n):p_i\geq 0ep_1+p_2+...+p_n=1\}$$

Que denominamos de probabilístico.

7 Independência Linear

Daremos uma definição precisa da “dimensão” de espaços vetoriais. O conceito central é o de independência linear.

7.1 Independência Linear

O conjunto de todos os múltiplos escalares de um vetor não-nulo V é uma reta que passa pela origem $\mathcal{L}\left[v\right]\equiv\{rv:r\in R\}$ e dizemos que essa reta é gerada por V. Se $v\left(1,0,...,0\right)$ então $\mathcal{L}\left[v\right]$ é o eixo x_1 em $\mathbb{R}^n$. Se $v\left(1,1\right)$ em $\mathbb{R}^2$, então $\mathcal{L}\left[v\right]$ é uma reta diagonal.

Definição. Sejam v_1e v_2 dois vetores não-nulos (tomados com suas caudas na origem) podemos tomar todas as combinações lineares de v_1e $v_2$para obter o conjunto gerado por v_1e v_2:

$$\mathcal{L}\left[v_1,v_2\right]=\{r_1v_1+r_2v_2:r_1\in R\text{ e }r_2\in R\}$$

Se v_1 é múltiplo de v_2, então $\mathcal{L}\left[v_1,v_2\right]=\mathcal{L}\left[v_2\right]$ é uma reta.

Se v_1 é múltiplo de v_2, ou vice-versa, dizemos que v_1e $v_2$são linearmente dependentes, caso contrário são linearmente independentes.

Se v_1 é um múltiplo de v_2 escrevemos:

$v_1 = r_2v_2$ ou $v_1 - r_2v_2 = 0$; (1)

Se v_2 é múltiplo de v_1 escrevemos $v_2 = r_1v_1$ou $r_1v_1 - v_2 = 0$ para algum escalar r_1(2)

Podemos dizer que v_1e $v_2$são linearmente dependentes se existirem escalares c_1e c_2 não ambos zero tais que $c_1v_1 + c_2v_2 = 0$ com $c_1$ou $c_2$não-nulos (3)

Também podemos dizer que v_1e v_2 são linearmente independentes se não existirem escalares c_1e c_2 com pelo menos um deles não-nulo, tal que (3) vale

$$c_1v_1 + c_2v_2 = 0 \Longrightarrow c_1 = c_2 = 0$$

Os vetores $v_1, v_2, ..., v_k$em $\mathbb{R}^n$ são linearmente dependentes se, e somente se, existirem escalares $c_1, ..., c_k$ não todos zero, tais que

$$c_1v_1 + c_2v_2 + ... + c_kv_k = 0$$

Esses vetores são linearmente independentes se, e somente se $c_1v_1 + c_2v_2 + ... + c_kv_k = 0$ para escalares $c_1, c_2..., c_k$ implica que $c_1 = ... = c_k = 0$.

Para verificarmos a dependência linear considere os seguintes vetores:

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, w_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \text{ e } w_3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$$

Comecemos com a equação:

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + c_3 \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

E resolvemos esse sistema para todos os possíveis valores de c_1, c_2 e c_3

$$\begin{aligned} c_1 + 4c_2 + 7c_3 = 0 \\ 2c_1 + 5c_2 + 8c_3 = 0 \\ 3c_1 + 6c_2 + 9c_3 = 0 \end{aligned}$$

É um sistema de equações lineares nas variáveis c_1, c_2 e c_3. A formulação matricial do sistema é

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Reduzimos a matriz de coeficientes a forma escalonada por linhas:

$$\begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

A solução (não-nula, 1 delas, na verdade possui infinitas)

$$c_1 = 1;\ c_2 = -2\ ;\ c_3 = 1$$

Concluímos que w_1, w_2e $w_3$são linearmente dependentes.

Teorema. *11.1: os vetores $v_1, ..., v_k$ de $\mathbb{R}^n$, são linearmente dependentes se, e somente se, o sistema linear*

$$A \begin{pmatrix} c_1 \\ . \\ c_k \end{pmatrix} = 0$$

Tem uma solução não-nula$(c_1, ..., c_k)$ onde A é a matriz $n \times k$ cujas colunas são vetores $v_1, ..., v_k$:

$$A = (v_1, v_2, ..., v_k)$$

Demonstração. $A. \begin{pmatrix} c_1 \\ . \\ c_k \end{pmatrix} = c_1v_1 + ... + c_kv_k$,então as colunas de A são linearmente dependentes se, e somente se o sistemas de equações de $A.c = 0$ tem uma solução não-nula.

O próximo teorema é uma reformulação do teorema 11.1 para o caso $k = n$ usando o seguinte fato: uma matriz quadrada é não-singular se, e somente se, seu determinante é não-nulo. □

Teorema. *11.2 um conjunto $v_1, ..., v_n$ de n vetores e $\mathbb{R}^n$é linearmente independente se, e somente se,*

$$det\,(v_1v_2...v_n) \neq 0$$

Teorema. *11.3 Se $k > n$, qualquer conjunto de k vetores de $\mathbb{R}^n$ é linearmente dependente.*

Demonstração. Sejam $v_1, ..., v_k$ quaisquer vetores de $\mathbb{R}^n$, com $k > n$ pelo teorema 11.1 os v_i são linearmente dependentes se, e somente se, o sistema

$$Ac = (v_1v_2...v_k)\,. \begin{pmatrix} c_1 \\ . \\ c_k \end{pmatrix} = 0$$

Tem uma solução não-nula. Qualquer matriz A com mais colunas do que linhas possui uma variável livre e portanto $Ac = 0$ tem infinitas soluções, todas as quais, com excessão de uma são nulas. $\square$

7.2 Conjuntos Geradores

Seja $v_1, ..., v_k$ um conjunto fixado de vetores em $\mathbb{R}^n$. Na última seção, falamos do conjunto:

$$\mathcal{L}[v_1, ..., v_k] \equiv \{c_1v_1 + ... + c_kv_k : c_1, ..., c_k \in R\}$$

de todas as combinações lineares de $v_1, ..., v_k$ e o denominamos conjunto gerado por $v_1, ..., v_k$. Suponha que dado um subconjunto V de $\mathbb{R}^n$. É pertinente perguntarmos se existem ou não vetores $v_1, ..., v_k$ em $\mathbb{R}^n$, tais que cada vetor V pode ser escrito como uma combinação linear de $v_1, ..., v_k$

$$V = \mathcal{L}[v_1, ..., v_k]\,(10)$$

Quando ocorre (10) dizemos que $v_1, ..., v_k$ gera V. Cada reta pela origem é gerada por um vetor não-nulo da reta. Por exemplo, o eixo x_1 é gerado por $e_1 = (1, 0, ..., 0)$ e a reta diagonal

$\Delta \equiv \{(a, a, ..., a) \in \mathbb{R}^n : a \in \mathbb{R}\}$ é gerada pelo vetor $(1, 1, ..., 1)$

Teorema. *11.4 Seja $v_1, ..., v_k$ um conjunto de K vetores de $\mathbb{R}^n$e considere a matriz $n \times k$ cujas colunas são os vetores v_j:*

$$A = (v_1v_2...v_k)\ \textit{(11)}$$

Seja b um vetor de $\mathbb{R}^n$. Então, b está no espaço $\mathcal{L}[v_1, ..., v_k]$ gerado por $v_1, ..., v_k$ se, e somente se, o sistema $Ac = b$ tem uma soluçãoc.

Demonstração. Escreva $v_1, ..., v_k$ em coordenadas, assim

$$v_1 = \begin{pmatrix} v_{11} \\ . \\ v_{1n} \end{pmatrix}, ..., v_k \begin{pmatrix} v_{k1} \\ . \\ v_{kn} \end{pmatrix}$$

Então b está em $\mathcal{L}[v_1, ..., v_k]$ se, e somente se, podemos encontrar $c_1, ..., c_k$ tais que $c_1v_1 + ... + c_kv_k = b$

$$\begin{pmatrix} v_{11} & ... & v_{k1} \\ \vdots & \ddots & \vdots \\ v_{1n} & ... & v_{kn} \end{pmatrix} . \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} (12)$$

(12)

Assim, $b \in \mathcal{L}[v_1, ..., v_k]$ se, e somente se, o sistema (12) tem uma solução c. □

Teorema. *11.5 Seja $v_1, ..., v_k$ um conjunto de k vetores de $\mathbb{R}^n$ e considere a matriz A de tamanho $n \times k$, cujas colunas são os vetores v_j, como em (11). Então $v_1, ..., v_k$ era $\mathbb{R}^n$se, e somente se, o sistema $Ax = b$ tem uma solução x para cada lado direito b.*

Demonstração. Para qualquer vetor coluna b se o sistema de equação $Ax = b$ tem uma solução x^*, então $x_1^* v_1 + ... + x_n^* v_n = b$. Consequentemente, se o sistema de equações tem uma solução para cada "lado direito", então cada vetor b pode ser escrito como uma combinação linear de v_i vetores colunas. Por outro lado, se o sistemas de equações falha em ter uma solução para algum b do lado direito, então b não é uma combinação linear de v_i, e $v_i$não gera $\mathbb{R}^n$. □

Teorema. *11.6 Um conjunto que gera $\mathbb{R}^n$ deve conter pelo menos n vetores.*

Demonstração. Pelo teorema 11.5 os vetores $v_1, ..., v_k$ geram $\mathbb{R}^n$ se, e somente se, o sistema (12) tem uma solução e para cada lado direito $b \in \mathbb{R}^n$. O *fato 7.7* nos diz que se o sistema (12) tem uma solução para cada lado direito, então o posto da matriz de coeficientes é igual ao número de linhas, n. O *fato 7.1* estabelece que o posto da matriz de coeficientes é sempre menor ou igual ao número de colunas, k. Portanto, se k vetores geram $\mathbb{R}^n$, então $n \leq k$. □

7.3 Base e Dimensão em $\mathbb{R}^n$

Definição. Seja $v_1, ..., v_k$ um conjunto dado de k vetores $\mathbb{R}^n$. Seja V o conjunto $\mathcal{L}[v_1, ..., v_k]$ gerado por $v_1, ..., v_k$. Se $v_1, ..., v_k$ é linearmente independente, dizemos que $v_1, ..., v_k$ é uma base de V. De modo mais geral, se $w_1, ..., w_m$ é uma base de V se:

(a) $w_1, ..., w_m$ gera V, e

(b) $w_1, ..., w_m$ são linearmente independentes.

Os vetores $e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix}, ..., e_n = \begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$ em $\mathbb{R}^n$são linearmente independentes pois, dados escalares $c_1, ..., c_n$ tais que $c_1 e_1 + c_2 v_2 + ... + c_n v_n = 0$

$$c_1 \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix} + ... + c_n \begin{pmatrix} 0 \\ . \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ . \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ . \\ 0 \end{pmatrix}$$

Implica que $c_1 = c_2 = ... = c_n$.

Note que se pegarmos um vetor $(a_1, ..., a_n)$ de $\mathbb{R}^n$ para todo $a_i \in \mathbb{R}$ $i = 1, ..., n$ podemos escrever:

$$\begin{pmatrix} a_1 \\ a_2 \\ . \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix} + ... + a_n \begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$$

Isto é o espaço euclidiano n-dimensional é gerado por $e_1, ..., e_n$. Deste modo $e_1, ..., e_n$ é uma base de $\mathbb{R}^n$. Por ser tão natural, essa base recebe o nome de canônica.

Teorema. *11.7 cada base de $\mathbb{R}^n$ contém n vetores.*

Demonstração. Pelo teorema 11.3, uma base de $\mathbb{R}^n$ não pode conter mais do que n elementos. Caso contrário, o conjunto em questão não poderia ser linearmente independente. Pelo teorema 11.6, uma base de $\mathbb{R}^n$ não pode ter menos do que n elementos; caso contrário o conjunto em questão não geraria $\mathbb{R}^n$. Portanto uma base de $\mathbb{R}^n$ deve ter exatamente n elementos. □

Teorema. *11.8 Seja $v_1, ..., v_n$ um agrupamento de n vetores de $\mathbb{R}^n$. Seja a matriz A $n \times n$, cujas colunas são as v_j : $A = (v_1 v_2 ... v_n)$. Então, as seguintes afirmações são equivalentes:*
(a) $v_1, ..., v_n$ são linearmente independentes
(b) $v_1, ..., v_n$ gera $\mathbb{R}^n$
(c) $v_1, ..., v_n$ constitui uma base de $\mathbb{R}^n$ e
(d) o determinante de A é não-nulo.

8 Subespaços Associados a uma Matriz

Seja $V = \mathbf{R^n}$ e para quaisquer u, v, e w em V e quaisquer escalares r,s em $\mathbf{R^1}$,

1. $\mathbf{u} + \mathbf{v}$ é um elemento de V sempre que $\boldsymbol{u}$ e $\boldsymbol{v}$ são elementos de V (a adição é fechada),

2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (a adição é comutativa),

3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (a adição é associativa),

4. Existe um elemento $\mathbf{0}$ em V tal que, para qualquer $\mathbf{v}$ em V, $\mathbf{v} + \mathbf{0} = \mathbf{v}$ (elemento neutro da adição),

5. para qualquer $\boldsymbol{v}$ em V , existe um elemento $\boldsymbol{w}$ em V (geralmente denotado por **-v**) tal que $\mathbf{v} + \mathbf{w} = \mathbf{0}$ (elemento inverso da adição),

6. $r.\mathbf{v}$ é um elemento de V sempre que $\boldsymbol{v}$ é um elemento de V (a multiplicação por escalar é fechada),

7. $r.(\mathbf{u} + \mathbf{v}) = r.\mathbf{u} + r.\mathbf{v}$ (a multiplicação por escalar é distributiva),

8. $(r+s).\mathbf{u} = r.\mathbf{u}+s.\mathbf{u}$,

9. $r.(s.\mathbf{u}) = s.(r.\mathbf{u})$, e

10. $1.\mathbf{u} = \mathbf{u}$

Quando trabalhamos com $\mathbf{R^n}$ como um conjunto dotado dessa estrutura condicional, dizemos que $\mathbf{R^n}$é um **espaço vetorial.**

8.1 Subespaços de $\mathbf{R^n}$

As propriedades de existência: 1,4,5, e 6 não valem para qualquer subconjunto de $\mathbf{R^n}$. As propriedades 2, 5, 7, 8, 9, e 10 valem para qualquer subconjunto de $\mathbf{R^n}$.

Exemplo. Seja $V_0 \equiv \{(x_1, x_2, 0) : x_1, x_2 \in R\}$ prove que V_0 é um subespaço de $\mathbf{R^n}$.

Demonstração. Seja $u \subset V_0$e $v \subset V_0$, tendo-se $u \equiv (u_1, u_2, 0)$ e $v \equiv (v_1, v, 0)$
$u + v \equiv (u_1 + v_1, u_2 + v_2, 0)$
$z = (0, 0, 0)$
$w = -v = u \equiv (-v_1, -v_2, 0) \rightarrow v + w = (0, 0, 0)$
$r \in R$, $r.v \in V_0 \rightarrow r.v = u \equiv (rv_1, rv_2, 0)$ □

Definição. Um subespaço de $\mathbf{R^n}$é um subconjunto de $\mathbf{R^n}$que, tal como V_0, satisfaz as propriedades de (1)-(10).
Note que (6)
$r = 0 \rightarrow$ (4) e $r = -1$ implica (5)
Na verdade precisamos verificar (1) e (6)!

Teorema. ***27.1:***
Seja $V \subset \mathbf{R^n}$. Sejam x e $y \in V$, se $(x+y) \in V$ e seja $r \in R$, e se rx e $ry \in V$ então V é um subespaço vetorial, portanto um subespaço de $\mathbf{R^n}$.

8.2 Base e Dimensão de um Subespaço Próprio

Teorema. ***27.2:***
Seja $u_{1,...,}u_m$ uma base de um subespaço V de $\mathbf{R^n}$. Então, qualquer conjunto contendo mais do que m vetores de V é necessariamente linearmente dependente.

Demonstração. Seja $w_1, ..., w_r$ um conjunto de r vetores em V, com $r > m$. Queremos mostrar que $w_{1,...,}w_r$, são linearmente dependentes. Como $u_1, ..., u_m$ gera V, podemos escrever:

$$w_j = \sum_{i=1}^{m} a_{ji}.u_i \tag{9}$$

para um conjunto $\{a_{ij}\}$de escalares, com $j = 1, ..., r$ e $i = 1, ..., m$. Para verificar a independência linear dos w_j, escreva:

$$c_1 w_1 + ... + c_m w_m = 0 \tag{10}$$

que pode ser reescrito como:

$$c_1. \left(\sum_{i=1}^{r} a_{1i}.u_i \right) + ... + c_n. \left(\sum_{i=1}^{r} a_{mi}.u_i \right) = 0 \tag{11}$$

como os u_i são linearmente independentes por hipótese a única ombinação dos u_i que tem soma 0 é a combinação nula. Portanto,

$$\sum_{j=1}^{r} c_j a_{1i} = 0, ..., \sum_{j=1}^{r} c_j a_m \tag{12}$$

O sistema (4) possui m equações homogêneas nas r incógnitas $c_1, ..., c_r$ com $r > m$. Um sistema de equações homogêneas com mais incógnitas do que equações possui variáveis livres e portanto tem infinitas soluções distintas; portanto existe um conjunto de c_j não nulos que satisfaz o sistema (4) e portanto o sistema (2). Assim, $w_1, ..., w_r$ não pode ser linearmente independente. □

Teorema. ***27.3***

Seja V um subespaço de $\mathbf{R^n}$. Duas quaisquer bases de V possuem o mesmo número de vetores.

Demonstração. Se $u_1, ..., u_m$ e $w_1, ..., w_r$ são bases de V, então ambos os conjuntos são linearmente independentes e portanto r deve ser igual a m pelo teorema 27.2. □

Definição. O número de vetores de qualquer base V, é denominado **dimensão** de V.

8.3 Espaço Linha

Para uma matriz A de tamanho $n \times m$ as linhas de A têm m componentes e podem, portanto, ser consideradas vetores de $\mathbf{R^m}$. Escrevemos $Lin(A)$ para o subespaço de $\mathbf{R^m}$ gerado pelas n linhas de A, então:

$$Lin(A) \equiv \zeta \left[a_1, ..., a_n \right]$$

Lema. ***27.1.*** *Sejam $v_1, ..., v_k$ vetores de $\mathbf{R^m}$. Para algum $j > 1$, seja:*

$$w = c_1 v_1 + c_j v_j$$

onde $c_1 \neq 0$

Então, $\zeta\left[v_1, v_2, ..., v_k\right] = \zeta\left[w, v_2, ..., v_k\right]$

Demonstração. Seja u um vetor arbitrário em $\zeta\left[v_1, ..., v_j\right]$. Então,

$$u = d_1 v_1 + d_j v_j$$

$$u = \frac{d_1}{c_1}(c_1 v_1 + c_j v_j) + \left(d_j - \frac{d_1 c_j}{c_1}\right) v_j$$

somando e subtraindo $d_1 c_j v_j / v_1$

$$u = \frac{d_1}{c_1} \cdot w + \left(d_j - \frac{d_1 c_j}{c_1}\right) v_j$$

Assim, u é uma combinação linear dos vetores w e v_j, de modo que $\zeta\left[v_1, v_j\right] \subset \zeta\left[w, v_j\right]$. Analogamente, se x é um vetor arbitrário em $\zeta\left[w, v_j\right]$

$$x = b_1 w + b_2 v_j$$

$$x = b_1\left(c_1 v_1 + c_j v_j\right) + b_2 v_j$$

$$x = b_1 c_1 v_1 + \left(b_1 c_j + b_2\right) v_j$$

Assim, $\zeta\left[w, v_j\right] \subset \zeta\left[v_1, v_j\right]$ e então $\zeta\left[v_1, v_j\right] = \zeta\left[w, v_j\right]$ □

Conclusões:

- Qualquer forma escalonada por linhas de A_r de A tem o mesmo espaço linha de A.

Lema. ***27.2.*** *Sejam* $v_1, ..., v_k$ *vetores não-nulos, tais que cada* v_{i+1} *tem mais 0 líderes do que* v_i*. Então, os vetores* $v_1, ..., v_k$ *são linearmente independentes.*

Intuição:

Considere 3 vetores:

$$\begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}$$

$$c_1 \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}$$

$$\begin{cases} 5c_1 = 0 & \rightarrow c_1 = 0 \\ 4c_1 + 3c_2 = 0 & \rightarrow c_2 = 0 \\ 3c_1 + 2c_2 + 7c_3 = 0 & \rightarrow c_3 = 0 \end{cases}$$

Demonstração. Escreva o vetor v_j em coordenadas como $v_j = (v_{j1}, ..., v_{jn})$. Para provar a independência, precisamos mostrar que a única solução da equação

$$c_1 v_1 + ... + c_k v_k = 0 \tag{13}$$

é $c_1 = c_2 = ... = c_k = 0$. Seja v_{1i*} o primeiro componente não-nulo de v_1. Como cada um dos demais v_j tem mais zeros líderes do que v_1, obtemos $v_{ji*} = 0$ para $j = 2, ..., n$. Escrevedno a i-ésima equação de (5) temos:

$$c_1 v_{1i*} + c_2 0 + ... + c_k 0 = 0$$

$$v_{1i*} \neq 0;$$

de modo que $c_1 = 0$. Agora, o sistema (5) se reduz a:

$$c_2 v_2 + ... + c_k v_k = 0 \tag{14}$$

Seja v_{2j*}o primeiro componente não nulo de v_2. Pelo mesmo argumento $v_{3j*} = ... = v_{nj*} = 0$. Escrevendo a j-ésima equação de (6), temos $c_1 v_{2j*}$de modo que $c_2 = 0$. Continuando dessa maneira até esgotar todas os $c_i$concluímos que a única solução de (5) é $c_1 = c_2 = ... = c_k = 0$ □

Teorema. *27.4*

Seja A_r qualquer forma escalonada por linhas de uma matriz A. Então, o subespaço $Lin(A)$ é o mesmo que o subespaço $Lin(A_r)$. Os vetores linha não nulos de A_r são uma base de $Lin(A)$ e a dimensão de $Lin(A)$ e o posto de A.

Demonstração. A matriz A_r é construída a partir de A efetuando-se um número finito de operações sobre linhas. O Lema 27.1 diz que cada operação sobre linhas deixa inalterado o espaço linha, de modo que A e A_r tem o mesmo espaço linha. O espaço linha de A_r é gerado pelas linhas não nulas de A_r. O Lema 27.2 diz que essas linhas são independentes, de modo que são uma base de $Lin(A_r)$. Finalmente o posto de A é o número de linhas não nulas de A_r, ou seja, o número de vetores dessa base. □

Exemplo. Exercício 27.8. Encontrar as bases, o subespaço linha, e o posto.

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 3 & 5 & 1 & 7 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$$

Somar L3+L4=L4

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 3 & 5 & 1 & 7 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 3 \end{pmatrix}$$

Subtrair L1-L4=L4

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 3 & 5 & 1 & 7 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Multiplicar 3.L1 e subtrair com L2

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

Somar L3 + (-1).L2=L3

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

Subtrair L2 de L3

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

Bases:
$\begin{pmatrix} 1 & 2 & 0 & 3 \end{pmatrix} .a_1$
$\begin{pmatrix} 0 & -1 & 1 & -2 \end{pmatrix} .a_2$
$Lin(A) = \zeta[a_1, a_2]$

$Posto = 2$

8.4 Espaço Coluna

Denote A por $n \times m$, então A tem m colunas $a_1, ..., a_m$ cada uma das quais possui n componentes. O subconjunto de R^n gerado pelas colunas de A é denominado **espaço-coluna** de A:

$$Col(A) = \zeta[a_1, ..., a_m]$$

8.5 Dimensão do Espaço Coluna de A

Seja $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ e $A_r = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$

$Col(A)$ é o subespaço de R^2 gerado por $(1, 2)$.

$Col(A_r)$ é gerado por $(1, 0)$

Note que $Col(A)$ e $Col(A_r)$ são distintos.

Definição. Uma coluna de uma matriz A é uma coluna básica se a coluna correspondente numa forma escalonada por linhas A_r, contém um pivô.

Teorema. ***27.5***

As colunas básicas de A constituem uma base de $Col(A)$.

Teorema. ***27.6***

Para qualquer matriz $n \times m$,

$$dimLin(A) = dimCol(A) = postoA$$

Exemplo. 27.11

$$A = \begin{pmatrix} 2 & 3 & 1 & 4 \\ 2 & 3 & 7 & 9 \\ 2 & 3 & 13 & 14 \end{pmatrix}, A_r = \begin{pmatrix} 2 & 3 & 1 & 4 \\ 0 & 0 & 6 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$Lin(A) \rightarrow Bases(2, 3, 1, 4)e(0, 0, 6, 5)$$

$$Col(A) \rightarrow Bases(2, 2, 2)e(1, 7, 13)$$

Note que a última coluna de A_r é uma combinação linear da 1ª e 3ª coluna:

$$\begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix} = \frac{19}{12} \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + \frac{5}{6} \cdot \begin{pmatrix} 1 \\ 6 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 4 \\ 9 \\ 14 \end{pmatrix} = \frac{19}{12} \cdot \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + \frac{5}{6} \cdot \begin{pmatrix} 1 \\ 7 \\ 13 \end{pmatrix}$$

Teorema. ***27.7***

Seja uma matriz de tamanho $n \times m$:

1. O sistema de equações $Ax = b$ tem uma solução para um particular $b \in R^n$ se, e somente se, b está no espaço-coluna $Col(A)$.

2. O sistema $Ax = b$ tem uma solução para *qualquer* b se, e somente se, $postoA = n$.

3. Se $Ax = b$ tem uma solução para qualquer b, então

$$n = postoA \leq n^{\underline{o}} \text{ de colunas de } A{=}m.$$

8.6 Espaço Nulo

Teorema. ***27.8***

Seja A uma matriz de tamanho $n \times m$, o conjunto V das soluções do sistema de equações $Ax = 0$ é um subespaço de R^m.

Demonstração. Sejam u e v vetores em V e seja $ru + v$ uma combinação linear dos dois vetores. Então,

$$A(ru + v) = A(ru) + Av$$

$$A(ru + v) = rAu + Av$$

$$A(ru + v) = 0$$

pois u e v estão em V. Assim, V é fechado para combinações lineares e é um subespaço de R^m. □

Definição. O subespaço das soluções do sistema homogêneo $Ax = 0$ é denominado **espaço-nulo** de A e é denotado por $Nul(A)$.

Definição. Sejam dados um subespaço V de R^m e um vetor $c \in R^m$.

O conjunto $\{x \in R^m : x = c + v$ para algum v$\in V\}$ é denominado **translação** de V por c e é denotado por $c + V$. Esse conjunto não é um espaço vetorial, a menos que c pertença a V. Subconjuntos de R^m da forma $c + V$, onde V é um subespaço de R^m, são denominados **subespaços afins** de R^m.

Teorema. 27.9

Seja $Ax = b$ um sistema $n \times m$ de equações lineares e $c_0 \in R^m$ uma solução particular deste sistema. Então, qualquer outra solução $c^{'}$ de $Ax = b$ pode ser escrita como $c^{'} = c_0{+}w$, *onde w é um vetor do espaço-nulo de A. Em outras palavras, o conjunto solução de $Ax = b$ é o subespaço afim $c_0 + Nul(A)$.*

Demonstração. Seja $c^{'}$uma solução $Ax = b$. Então,

$$A(c^{'} - c_0) = Ac^{'} - Ac_0 = b - b = 0$$

de modo que $w = c^{'} - c_0$ está em $Nul(A)$. Reciprocamente, se w está em $Nul(A)$, então $A(c_0 + w) = Ac_0 + Aw = b + 0 = b$ □

Exemplo. 27.14

Seja A a matriz $\begin{pmatrix} 1 & 1 \end{pmatrix}$. O conjunto solução de $Ax = 1$ é o conjunto

$$\{(x_1, x_2) : x_1 + x_2 = 1\} \tag{15}$$

Este conjunto é claramente uma translação de subespaço

$$Nul(A) = \{(x_1, x_2) : x_1 + x_2 = 0\}$$

pelo vetor $(1, 0)$. Analiticamente, $x_1 + x_2 = 1$ se, e somente se, $x_1 = 1 - x_2$. Portanto, o conjunto solução de (7) pode ser escrito como:

$$\begin{pmatrix} 1 - x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 - x_2 \\ x_2 \end{pmatrix} = Nul(A) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Teorema. 27.10

Seja A uma matriz $n \times m$. Então,

$$dimNul(A) = m - postoA.$$

Exemplo. *27.16*

$$A = \begin{pmatrix} 1 & 2 & 0 & -2 & 0 \\ 2 & 4 & 1 & -1 & 0 \\ 1 & 2 & 1 & 1 & 1 \end{pmatrix} \text{ e } A_r = \begin{pmatrix} 1 & 2 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{matrix} L_3 - L_2 \to L_3 \\ L_3 + L_1 \to L_3 \\ 2L_1 - L_2 \to L_2 \\ (-1)L_2 \end{matrix} \text{ operações para escalonar } A$$

Variáveis básicas (Colunas que contém um pivô de A_r), x_1, x_3, x_5; livres x_2, x_4

$$\begin{cases} x_1 + 2x_2 - 2x_4 = & 0 \\ x_3 + 3x_4 = & 0 \\ x_5 = & 0 \end{cases}$$

$$x_5 = 0$$

$$x_3 = -3x_4$$

$$x_1 = -2x_2 + 2x_4$$

O vetor x está em $Nul(A)$ se, e somente se,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

Os vetores $(-2, 1, 0, 0, 0)$ e $(2, 0, -3, 1, 0)$ são bases de $Nul(A)$

A dimensão de $Nul(A)$ é o número de variáveis livres (2) que é o número de colunas menos o posto de A.

$$dimNul(A) = 5 - 3 = 2$$

Conclusões:

- Se posto $A = n$, que é número de linhas (equações), então $Ax = b$ tem alguma solução para cada b.
- Se posto $A < n$, então $Ax = b$ somente tem alguma solução para os b em $Col(A)$, que é um subespaço de R^n de dimensão não maior do que n.
- Se posto $A = m$, então $Nul(A) = \{0\}$ e $Ax = b$ tem no máximo uma solução para cada b.
- Se posto $A < m$, então, se $Ax = b$ tem alguma solução, tem um subespaço afim de soluções, de $dimensão = m - postoA$.

8.7 Espaços Vetoriais Abstratos

Definição. Seja V um conjunto tal que existe uma função $+$ que associa a cada par de elementos de V um outro elemento de V e uma outra função $\cdot$ que associa a cada par, consistindo de um elemento de V e de um número real um outro elemento de V . Se, V junto com as operações $+$ e $\cdot$ satsifaz as propriedades (1) a (10) da seção 27.1 para quaisquer u, v, e w de V e r, s de R^1, então dizemos que V é um espaço vetorial e que seus elementos são vetores.

As funções $+$ e $\cdot$ são denominadas operações de V.

Exemplo. 27.19

Seja $F\left\{x \in R; f(x) : R \to R\right\}$

O que precisamos para somar funções?

Simplesmente somamos seus valores para cada x. Suponha, por exemplo, que $u(x) = x^2$ e que $v(x) = lnx$. Então, sua soma é a função $(u+v)(x) = x^2 + lnx$. A multiplicação por escalar: $ru(x) = rx^2$. O elemento neutro na adição é a função nula $w(x) \equiv 0$.

Exemplo. 27.20

$$F_2 \equiv \left\{a_0x^2 + a_1x + a_2 : a_0, a_1, a_2 \in R^1\right\}$$

Adição

$$\left(a_0x^2 + a_1x + a_2\right) + \left(b_0x^2 + b_1x + b_2\right)$$

$$\left(a_0x^2 + a_1x + a_2\right) + \left(b_0x^2 + b_1x + b_2\right) = \left(a_0 + b_0)x^2 + (a_1 + b_1\right)x + (a_2 + b_2)$$

Multiplicação por escalar

$$\left(ra_0x^2 + ra_1x + ra_2\right) = r(a_0x^2 + a_1x + a_2)$$

Nesse caso dizemos que F_2 é um subespaço de F. Note que se fixarrmos nossa atenção para os três coeficientes, então a operação de adição será:

$$(a_0, a_1, a_2) + (b_0, b_1, b_2) = (a_0 + b_0, a_1 + b_1, a_2 + b_2)$$

Existe uma correspondência injetora entre as funções quadráticas de F_2 e os ternos de R^3, que carrega consigo as operações de adição e multiplicação por escalar. Nesse caso dizemos que F_2 e R^3são espaços **isomorfos**.

9 Limites e Conjuntos Abertos

9.1 Sequências de números reais

Definição. Uma sequência de números reais é uma associação de um número real a cada número natural.

Exemplos:

1. $\{1, 2, 3, 4, \ldots\}$
2. $\{1, 1/2, 1/3, 1/4, \ldots\}$

As vezes escrevemos uma sequência qualquer $\{x_1, x_2, x_3, \ldots\}$ como $\{x_n\}_{n=1}^{\infty}$

Definição. Sejam $\{x_1, x_2, x_3, \ldots\}$ uma sequência de números reais e r um número real. Dizemos que r é o limite desta sequência se, para qualquer número positivo (e pequeno) ε, existe um inteiro positivo N tal que, para cada $n \geq N$, x_n está no intervalo de raio ε em torno de r; ou seja,

$$|x_n - r| < \varepsilon$$

Neste caso, dizemos que a sequência converge a r e escrevemos:

$$\lim_{n \to \infty} x_n = r$$

- Um número s está perto de um número r se s está em algum intervalo em torno de r.

Mais precisamente, denotamos por ε um número real positivo e pequeno. Então, o intervalo de raio ε em torno do número r é definido por:

$$I_\varepsilon(r) \equiv \{s \in \mathbb{R} : |s - r| < \varepsilon\}$$

Em notação intervalar:

$$I_\varepsilon(r) = (r - \varepsilon, r + \varepsilon)$$

Exemplo. Sequência que convergem a 0:

$1, 0, 1/2, 0, 1/3, 0, \ldots$

$1, -1/2, 1/3, -1/4, \ldots$

Definição. (Ponto de acumulação) r é um ponto de acumulação (ponto aderente) da sequência $\{x_n\}$ se, dado qualquer ε positivo, existem infinitos elementos da sequência no intervalo $I_\varepsilon(r)$ de raio ε em torno de r.

Teorema. *Uma sequência tem, no máximo um limite.*

Demonstração. Suponha que $\{x_n\}_{n=1}^{\infty}$tem 2 limites r_1e r_2. Tomamos ε como um número menor do que a metade da distância entre r_1e r_2. Digamos que $\varepsilon = \frac{1}{4}|r_1 - r_2|$, de modo que $I_\varepsilon(r_1)$ e $I_\varepsilon(r_2)$ são intervalos disjuntos. Como $x_n \to r_1$, existe N_1tal que todos os x_n estão em $I_\varepsilon(r_1)$, desde que $n \geq N_1$; como $x_n \to r_2$, existe N_2 tal que todos $x_n$estão em $I_\varepsilon(r_2)$, desde que $n \geq N_2$. Portanto, todos os $x_n$estão em ambos $I_\varepsilon(r_1)$ e $I_\varepsilon(r_2)$ para todos $n \geq \max\{N_1, N_2\}$. Mas ponto algum pode estar em ambos intervalos, o que é uma contradição e prova o teorema. $\square$

Definição. (Subsequência.) Seja M um subconjunto infinito dos números naturais. Escreva M como $\{n_1, n_2, n_3, \ldots\}$ onde $n_1 < n_2 < n_3 < \ldots$. Crie uma nova sequência $\{y_n\}$, dado por $y_j = x_{nj}$para $1, 2, 3, \ldots$ Diz-se que essa nova sequência $\{y_n\}_{j=1}^{\infty}$ é uma subsequência da sequência original $\{x_n\}$. Resumindo, construiremos uma subsequência de uma sequência pela escolha de uma coleção infinita das entradas da sequência original na ordem em que estes elementos nela aparecem.

Teorema. *Sejam $\{x_n\}_{n=1}^{\infty}$ e $\{y_n\}_{n=1}^{\infty}$ sequência com limites x e y respectivamente. Então a sequência $\{x_n + y_n\}_{n=1}^{\infty}$ converge ao limite $x + y$.*

Demonstração. Fixe uma constante pequena e positiva ε. Como sabemos que $x_n \to x$ e $y_n \to y$, existe um inteiro N_1 tal que

$$|x_n - x| < \frac{\varepsilon}{2} \text{ para } n \geq N_1$$

e um inteiro N_2 tal que

$$|y_n - y| < \frac{\varepsilon}{2} \text{ para } n \geq N_2$$

Lembre que:

$$|x + y| \leq |x| + |y|$$

Para qualquer x, y

$$||x| - |y|| \leq |x - y|$$

$$\leq |x_n - x| + |y_n - y|$$

Seja $N = \max\{N_1, N_2\}$. Então, para qualquer $n \geq N$, temos:

$$|(x_n + y_n) - (x + y)| = |(x_n - x) + (y_n - y)|$$

$$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

$$= \varepsilon$$

$\square$

Teorema. *Sejam* $\{x_n\}_{n=1}^{\infty}$ *e* $\{y_n\}_{n=1}^{\infty}$ *sequência com limites* x *e* y *respectivamente. Então a sequência* $\{x_n y_n\}_{n=1}^{\infty}$ *converge ao limite* xy*.*

Demonstração. Para mostrar que $|xy - x_n y_n|$ é pequeno quando $|x_n - x|$ e $|y_n - y|$ são pequenos, tente escrever o primeiro em termos dos dois últimos. Conseguimos isso usando o truque dos matemáticos de somar e subtrair o mesmo elemento à dada expressão; com efeito, fazemos isso duas vezes:

$$|xy - x_n y_n| = |xy - xy_n + xy_n - x_n y_n|$$

$$= |x(y - y_n) + (x - x_n)y_n|$$

$$= |x(y - y_n) + (x - x_n)(y_n - y + y)|$$

$$= |x(y - y_n) + (x - x_n)(y_n - y) + (x - x_n)y|$$

$$\leq |x||y - y_n| + |x - x_n||y_n - y| + |x - x_n||y||$$

Sabemos que cada parcela na última expressão tende a zero. Para tornar este processo preciso, proceda exatamente como na prova do teorema 12.2. Escolha e fixe um número positivo pequeno ε, com $\varepsilon < 1$. Como $x_n \to x$, existe um inteiro N_1tal que:

$$n \geq N_1 \Rightarrow |x - x_n| < \frac{\varepsilon}{3(|y| + 1)}$$

Como $y_n \to y$, existe um inteiro N_2 tal que:

$$n \geq N_2 \Rightarrow |y - y_n| < \frac{\varepsilon}{3(|x| + 1)}$$

Não esqueça o seguinte fato: como ε e $|x|$ são números reais fixados, também $\frac{\varepsilon}{3(|x|+1)}$ é um número real fixo. Como há 3 parcelas no final da desigualdade acima, o número três nestas desigualdades tem a seguinte utilidade: para fazer a primeira expressão menor do que ε, faremos cada uma das três parcelas menor do que $\varepsilon/3$; para fazer a primeira parcela, $|x||y - y_n|$, menor do que $\varepsilon/3$, queremos $|y - y_n| < \varepsilon/(3|x|)$.

Acrescentamos 1 adicional ao denominador do lado direito para garantir o ocaso que x é zero. Agora tomamos $N = \max\{N_1, N_2\}$. Então, se $n \geq N$:

$$|xy - x_ny_n| \leq |x||y - y_n| + |x - x_n||y_n - y| + |x - x_n||y|$$

$$\leq |x|\frac{\varepsilon}{3(|x|+1)} + \frac{\varepsilon}{3(|y|+1)}\frac{\varepsilon}{3(|x|+1)} + \frac{\varepsilon}{3(|y|+1)}|y|$$

$$= \frac{|x|}{(|x|+1)}\frac{\varepsilon}{3} + \frac{\varepsilon^2}{3^2}\frac{1}{(|y|+1)(|x|+1)} + \frac{|y|}{(|y|+1)}\frac{\varepsilon}{3}$$

$$\leq \frac{\varepsilon}{3} + \left(\frac{\varepsilon}{3}\right)^2 + \frac{\varepsilon}{3}$$

$$\leq \varepsilon$$

Aqui usamos os seguintes fatos:

$$\frac{|x|}{(|x|+1)} < 1, \tfrac{|y|}{(|y|+1)} < 1, \tfrac{1}{(|x|+1)} < 1, \tfrac{1}{(|y|+1)} < 1 \text{ e}$$

$$\varepsilon < 1 \Rightarrow \left(\frac{\varepsilon}{3}\right) < 1 \Rightarrow \left(\frac{\varepsilon}{3}\right)^2 < \left(\frac{\varepsilon}{3}\right)$$

□

Teorema. *Seja* $\{x_n\}_{n=1}^{\infty}$ *uma sequência convergente com limite* x *e* b *um número tal que* $x_n \leq b \forall n$. *Então* $x \leq b$. *Se* $x_n \geq b \forall n$, *então* $x \geq b$.

Demonstração. Suponha, então que $x_n \leq b\forall n$ e suponha que $x > b$. Escolha ε de tal modo que $0 < \varepsilon < x - b$, então $b < x - \varepsilon$ e $I_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ fica a direita de b na reta numérica. Existe um inteiro N tal que $x_n \in I_\varepsilon(x) \forall n \geq N$. Para esses x_n temos $b < x_n$; Essa é a contradição da hipótese que todos $x_n$são $\leq b$. Agora suponha que $x_n \geq b\forall n$ e que $x < b$. Escolha $\varepsilon > 0$ tal que $0 < \varepsilon < b - x$. Então, $\varepsilon + x < b$ e $I_\varepsilon(b) = (x - \varepsilon, x + \varepsilon)$ está a esquerda do número b na reta numérica. Existe um $N > 0$ tal que $\forall n \geq N$ $x_n \in I_\varepsilon(x)$. Para esses x_n, $x_n < x + \varepsilon < b$, uma contradição a hipótese que $x_n \geq b\forall n$ □

Definição. (Sequência Limitada) Uma sequência é dita limitada se existe um número B tal que $|x_n| \leq B$, para cada n.

Proposição. *Se* $\{x_n\}_{n=1}^{\infty}$ converge a 0 e se $\{y_n\}_{n=1}^{\infty}$é limitada, então a sequência dos produtos converge a zero.

Demonstração. Seja $\varepsilon > 0$. Escolha N tal que $n \geq N$, $|x_n - 0| \leq \varepsilon B$ onde $|y_n| \leq B\forall n$. Então, $n \geq N$

$$|x_ny_n - 0| = |x_ny_n| = |x_n||y_n| \leq \frac{\varepsilon}{B}B = \varepsilon$$

□

9.2 Sequências em $\mathbb{R}^m$

Uma sequência em $\mathbb{R}^m$ é a associação de um vetor de $\mathbb{R}^m$ a cada número natural n : $\{x_1, x_2, x_3, \ldots\}$. Cada vetor possui m coordenadas e a sua posição na sequência é dada por n

Definição. (Bola Aberta) Seja r um vetor em $\mathbb{R}^m$ e ε um número positivo. A bola de raio ε em torno de r é dada por

$$B_\varepsilon(r) \equiv \{x \in \mathbb{R}^m / \, \|x - r\| < \varepsilon\}$$

Intuitivamente, um vetor x de $\mathbb{R}^m$ está próximo de r se x está em alguma $B_\varepsilon(r)$ para um ε pequeno, mas positivo. Quanto menor ε, mais perto x estará de r.

Definição. Uma sequência de vetores $\{x_1, x_2, x_3, \ldots\}$ converge ao vetor x se, para qualquer escolha de um número real positivo ε, existe um inteiro N tal que, para cada $n \geq N$, vale $x_n \in B_\varepsilon(x)$, ou seja

$$\mathrm{d}(x_n, x) = \|x_n - x\| < \varepsilon$$

O vetor x é denominado o limite da sequência. Em outras palavras, $\{x_n\}_{n=1}^{\infty} \to x$ se e somente se $\{\|x_n - x\|\}_{n=1}^{\infty} \to 0$ (sequência das distâncias) em $\mathbb{R}^1$

Teorema. *Uma sequência de vetores em $\mathbb{R}^m$ converge se, e somente se, cada uma das m sequência de seus componentes converge em $\mathbb{R}^1$.*

Demonstração. (se) Seja $\{x_n\}_{n=1}^{\infty}$ uma sequência de vetores em $\mathbb{R}^m$. Escreva $x_n = (x_{1n}, \ldots, x_{mn})$. Suponha que cada uma das m sequências $\{x_{in}\}_{n=1}^{\infty}$ de números, com $i = 1, \ldots, m$ converge a um limite x_i^*. Seja $x^* = (x_1^*, \ldots, x_m^*)$. Escolha e fixe uma pequena constante positiva ε. Para cada i entre 1 e m, existe um inteiro N_i tal que, se $n \geq N_i$ então $|x_{in} - x_i^*| < \varepsilon/\sqrt{m}$. Seja $N = \max\{N_1, \ldots, N_m\}$. Suponha que $n \geq N$. Então,

$$\|x_n - x^*\| = \sqrt{(x_{1n} - x_1^*)^2 + (x_{2n} - x_2^*)^2 + \ldots + (x_{mn} - x_m^*)^2} < \sqrt{\frac{\varepsilon^2}{m} + \ldots + \frac{\varepsilon^2}{m}} = \varepsilon$$

Assim $\{x_n\}_{n=1}^{\infty}$ converge a x^*

(Somente se) Escolha $\varepsilon > 0$. Existe N, tal que, para cada $n \geq N$, vale $\|x_n - x^*\| < \varepsilon$. Mas então, para cada $n \geq N$ e para cada i

$$|x_{in} - x_i^*| \leq \sqrt{(x_{1n} - x_1^*)^2 + \ldots + (x_{mn} - x_m^*)^2}$$

$$|x_{in} - x_i^*| = \|x_n - x^*\| < \varepsilon$$

□

Teorema. *Sejam* $\{x_n\}_{n=1}^{\infty}$ e $\{y_n\}_{n=1}^{\infty}$ sequências convergentes em $\mathbb{R}^m$ com limites x e y, respectivamente, e seja $\{c_n\}_{n=1}^{\infty}$ uma sequência convergente de números reais com limite c. Então, a sequência $\{c_n x_n + y_n\}_{n=1}^{\infty}$ converge ao limite $cx + y$.

Demonstração. Escolha um $\varepsilon > 0$ e note que

$$\|(c_n x_n + y_n) - (cx + y)\| \leq \|c_n x_n - cx\| + \|y_n - y\|$$

Como a sequência dos $y_n \to y$,sabemos que existe um inteiro N_1tal que para cada $n \geq N_1$,vale $\|y_n - y\| < \varepsilon/2$. Por outro lado, para cada componente i, a sequência $\{c_n x_n\}_{n=1}^{\infty} \to cx_i$ pelo Teorema 12.3. Pelo teorema 12.5 isso implica que a sequência $\{c_n x_n\}_{n=1}^{\infty}$ converge a cx. Assim, existe um N_2tal que, para casa $n \geq N_2$,vale $\|c_n x_n - cx\| < \varepsilon/2$. Segue que, para cada $n \geq N = \max\{N_1, N_2\}$vale $\|(c_n x_n + y_n) - (cx + y)\| \leq \varepsilon$ ■ □

Definição. O vetor x é um ponto de acumulação da sequência $\{x_n\}_{n=1}^{\infty}$ se, para cada $\varepsilon > 0$ dado, existem infinitos números inteiros n tais que $\|x_n - x\| < \varepsilon$

Definição. Uma sequência $\{y_j\}_{j=1}^{\infty}$ de vetores em $\mathbb{R}^m$ é uma subsequência da sequência $\{x_i\}_{i=1}^{\infty}$se existe um conjunto infinito crescente $\{n_j\}$ de números naturais $n_1 < n_2 < n_3 \ldots$, tais que $y_1 = x_{n_1}$; $y_2 = x_{n_2}$; $y_3 = x_{n_3} \ldots$

Exemplo. 12.1 c

$\{1, \frac{1}{2}, 4, \frac{1}{8}, 16, \ldots\} \Rightarrow$ os termos pares formam uma subsequência convergente com limite 0

Exemplo. 12.1 d

$$\{0, -\frac{1}{2}, \frac{2}{3}, -\frac{3}{4}, \frac{4}{5}, \ldots\}$$

os termos de índice par formam uma subsequência convergente com limite igual a -1.

Proposição. *Uma sequência convergente em* $\mathbb{R}^m$ só pode ter um limite e, portanto, só um ponto de acumulação.

Demonstração. Suponha que $x \to a$ e $b \neq a$ e um ponto de acumulação. Escolha $\varepsilon = \frac{\|a-b\|}{4}$. Então existe um N tal que $n \geq N$ e então $\|x_n - a\| < \varepsilon$. Como b é um ponto de acumulação, então existe um $m \geq N$ tal que $\|x_m - b\| < \varepsilon$. Então,

$$\|a - b\| = \|a - x_m + x_m - b\|$$

$$\|a - b\| \leq \|x_m - a\| + \|x_m - b\| < \varepsilon + \varepsilon$$

$$\varepsilon + \varepsilon = 2\varepsilon = \frac{\|a-b\|}{2} \Rightarrow \text{contradição!}$$

□

9.3 Conjuntos Abertos

Definição. (Conjunto Aberto) Um conjunto S em $\mathbb{R}^m$ é aberto se para cada $x \in S$ existe uma bola aberta de raio ε em torno de x completamente contida em S:

$$x \in S \Rightarrow \exists \varepsilon > 0, B_\varepsilon(x) \subset S$$

$$B_\varepsilon(x) = \{x \in S : \|y - x\| < \varepsilon\}$$

Um conjunto aberto contendo o ponto x é denominado vizinhança aberta de x.

Teorema. *Bolas abertas são conjuntos abertos*

Demonstração. Seja B a bola aberta $B_\varepsilon(x)$ em torno de x e seja y um ponto arbitrário de B. Queremos mostrar que existe uma bola em torno de y que está completamente contida em B. Seja $\delta \equiv \|x - y\| < \varepsilon$, mostraremos que a bola aberta V de raio $\varepsilon - \delta$ em torno de y que está contida em B. Seja z um ponto arbitrário de V. Então pela desigualdade triangular:

$$\|z - x\| \leq \|z - y\| + \|y - x\| < (\varepsilon - \delta) + \delta = \varepsilon$$

Assim, $V \subset B$

$$\|z - x\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\|$$

□

Teorema. *(a) A união de conjuntos abertos é um conjunto aberto. (b) A intersecção finita de conjuntos abertos é um conjunto aberto*

Demonstração. (a) Seja $S = S_i \bigcup S_j$, e S_i é aberto $\forall i, j = 1, \ldots, n$. Seja $x \in S$, ou seja, $x \in S_i$ ou $x \in S_j$. Então se $x \in S_j$e este conjunto é aberto, existe $B_\varepsilon(x) \subset S_j$.Então $B_\varepsilon(x) \subset S$. (b) Sejam $S_1, \ldots, S_n$ conjuntos abertos e $S = \bigcap_{i=1}^n S_i$.Se $x \in S$, então $x \in$a cada S_i. Como cada S_ié aberto, para cada i existe um ε_i tal que $B_{\varepsilon_i}(x) \subset S_i$. Seja $\varepsilon = \min \varepsilon_i$. A bola $B_\varepsilon(x)$ está contida em cada $B_{\varepsilon_i}(x)$ e portanto está contida em cada S_i.Assim, a bola está contida na intersecção das ,S_i. □

Definição. Dado um subconjunto S de $\mathbb{R}^m$, denotamos por int Sa união de todos os conjuntos abertos contidos em S. O conjunto aberto int Sé denominado interior de S. O interior de um conjunto pode ser considerado o maior conjunto aberto que está contido no conjunto dado.

9.4 Conjuntos Fechados

Definição. Um conjunto S em $\mathbb{R}^m$ é fechado se, sempre que $\{x_n\}_{n=1}^{\infty}$ é uma sequência convergente completamente contida em S, seu limite também está em S.

Teorema. *Um conjunto S em $\mathbb{R}^m$ é fechado se, e somente se, seu complementar $S^C = \mathbb{R}^m - S$ é aberto.*

Demonstração. (Somente se) Seja S fechado. Precisamos mostrar que S^C é aberto. Isto é, $\forall x \in S^C \exists \varepsilon > 0, B_\varepsilon(x) \subset S^C$. Escolhemos um $x \in S^C$e supomos que isto não ocorre, ou seja, que nenhum $B_\varepsilon(x)$ está completamente contido em S^C.Então, para cada $\varepsilon > 0$, temos $B_\varepsilon(x) \bigcap S \neq \emptyset$. Em particular, para cada inteiro positivo n, existe um elemento $x_n$de S em $B_{1/n}(x)$. A sequência $\{x_n\}_{n=1}^{\infty}$ está em S e converge a x, pois $\|x_n - x\| < 1/n$. Como S é fechado, x está em S – uma contradição com nossa escolha de x em S^C.

(Se) Seja S o complementar de S^C. Seja $\{x_n\}_{n=1}^{\infty}$ uma sequência convergente em S com limite x. Para mostrar que S é fechado, precisamos mostrar que $x \in S$. Suponha que isso não ocorra, isto é, suponha que $x \in S^C$. Como S^Cé aberto, existe $B_\varepsilon(x)$ em torno de x contida em S^C. Como a sequência converge a x, temos $x_n \in B_\varepsilon(x)$ de modo que $x_n \in S^C$, novamente uma contradição, pois os $x_n$estão em S, o complementar de S^C. $\square$

Teorema. *(a) Qualquer intersecção de conjuntos fechados é um conjunto fechado. (b) A união finita de conjuntos fechados é um conjunto fechado*

Definição. (Fecho de um Conjunto) Dado um subconjunto S de $\mathbb{R}^m$, denotamos fecho $S(\bar{S})$ a intersecção de todos os conjuntos fechados contendo S. O conjunto fechado S é denominado fecho de S.

Teorema. *Seja S um conjunto em $\mathbb{R}^m$. Então x está no fecho de S se, e somente se, existe uma sequência de pontos de S convergente a x.*

Demonstração. (Se) Seja $\{x_n\}_{n=1}^{\infty}$ uma sequência convergente de pontos de um conjunto S com limite x. Então $\{x_n\}_{n=1}^{\infty} \subset T$ para qualquer conjunto fechado $T \supset S$,assim, $x \in T$. Como isto vale para qualquer conjunto fechado contendo S, temos que $x \in$ fecho S

(Somente se) Suponha que $x \in$ fecho S. Afirmamos que $B_\varepsilon(x) \bigcap S \neq \emptyset$, para cada $\varepsilon > 0$. De fato, se tivéssemos $B_\varepsilon(x) \bigcap S \neq \emptyset$ para alguma $\varepsilon > 0$, então o complementar de $B_\varepsilon(x)$ seria um conjunto fechado contendo S, mas não x, contradizendo que $x \in$ fecho S. Agora construímos a sequência $\{x_n\}_{n=1}^{\infty}$ escolhendo $x_n \in B_{1/n}(x) \bigcap S$. Isto é uma sequência em S com limite x. $\square$

Definição. Um ponto x está na fronteira de um conjunto S se cada bola aberta em torno de x contém tanto pontos de S quanto pontos do complementar de S

Teorema. *Um conjunto de pontos de fronteira de um conjunto S é igual a intersecção fecho $S \bigcap$ fecho $\bar{S}^C$* dos fechos de S e do complementar.

9.5 Conjuntos compactos

Definição. O conjunto S é limitado se existe um número B tal que $\|x\| \leq B$ para cada $x \in S$. Isto é, S está contido em alguma bola de $\mathbb{R}^m$. Qualquer união finita de $\mathbb{R}^m$

Definição. Um conjunto S em $\mathbb{R}^m$ é compacto se, e somente se, é limitado e fechado.

Teorema (Teorema de Bolzano-Weirstrass)**.** *Qualquer sequência definida em um conjunto compacto deve conter uma subsequência convergente.*

Teorema. *Qualquer sequência contida no intervalo fechado e limitado* $[0, 1]$ *tem uma subsquência convergente.*

Teorema. *Seja C um subconjnto compacto de $\mathbb{R}^m$ e seja $\{x_n\}_{n=1}^{\infty}$ uma sequência qualquer em C. Então $\{x_n\}_{n=1}^{\infty}$ tem uma subsequência convergente cujo limite é um ponto de C.*

10 Limites e Conjuntos Compactos

Definição. Uma sequência $\{x_n\}_{n=1}^{\infty}$ é uma sequência de Cauchy se, para qualquer número positivo ε existe um inteiro N tal que $d(x_i, x_j) < \varepsilon$ para quaisquer $i, j \geq N$.

Teorema. *(29.1)*

Toda sequência convergente em R^m é de Cauchy.

Demonstração. Seja $\{x_n\}_{n=1}^{\infty}$ uma sequência convergente a x. Dado $\varepsilon > 0$, escolha N tal que $d(x_n, x) < \frac{\varepsilon}{2}$ para qualquer $n \geq N$. Então, para quaisquer $n, m \geq N$

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Definição. Seja $S \subset R$. S tem uma cota superior se existe um número B tal que cada x em S é menor que B, $\forall x \in S$. O supremo de um tal conjunto S é o número C, que é uma cota superior de S e satisfaz $C \leq B$ para qualquer cota superior B de S. Qualquer conjunto não vazio de números reais que tem uma cota superior tem um supremo.

Definição. Uma sequência de números reais $\{x_n\}_{n=1}^{\infty}$ é monotonamente crescente se cada entrada da sequência é pelo menos tão grande quanto a entrada anterior, ou seja, se $x_{n-1} \leq x \; \forall n$. A sequência é monótona se é monotonamente crescente ou decrescente. Para usarmos o termo estrita(o), utilizamos a desigualdade estrita.

Teorema. *(29.2)*

Toda sequência monótona e limitada converge. (prova como exercício)

Lema. *(29.1)*

Toda sequência de Cauchy é limitada.

Demonstração. Seja $\{x_n\}_{n=1}^{\infty}$ uma sequência de Cauchy em R^1. Fixe $\varepsilon > 0$. Existe um N tal que $|x_i - x_j| < \varepsilon$ para quaisquer $i, j \geq N$. Em particular $|x_n - x_i| < \varepsilon$ para qualquer $i \geq N$. Como:

$$|x_i| - |x_n| \leq |x_i - x_n|$$

$$|x_i| \leq |x_n| + \varepsilon$$

para qualquer $i \geq N$ e portanto $|x_n| + \varepsilon$ é uma cota para todos os termos da sequência, exceto possivelmente os primeiros N-1.

Seja $b = max\left\{ \ |x_1, \quad \dots \quad |x_{n-1}|, \quad |x_n| \ \right\}$. Daí, $|x_i| \leq b + \varepsilon$ para cada x_i da sequência. □

Lema. *(29.2)*

Toda sequência possui uma subsequência monótona.

Lema. *(29.3)*

Se uma sequência $\{x_n\}_{n=1}^{\infty}$ de Cauchy tem uma subsequência monótona convergente a y, então a sequência toda converge a y.

Demonstração. $\varepsilon > 0$. Como a sequência é de Cauchy, existe um N tal que $|x_i - x_j| < \frac{\varepsilon}{2}$ para qualquer $i, j \geq N$. Escolha um x_k da subsequência convergente com $k \geq N$ e k suficientemente grande para ter $|x_k - y| < \frac{\varepsilon}{2}$. Então, para qualquer $i \geq N$,

$|x_i - y| = |(x_i - x_k) + (x_k - y)| \leq |x_i - x_k| + |x_k - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$

Portanto, $\{x_n\}_{n=1}^{\infty}$ converge a y. □

Teorema. *(29.3)*

Qualquer sequência de Cauchy de números reais converge.

1. 29.1 (Lema) Essa sequência é limitada (e suas subsequências também).

2. 29.2 (Lema) Ela contém uma subsequência monótona.

3. 20.3 (Lema) A sequência original também converge.

10.1 Conjuntos Compactos

Teorema. *(29.4)*

Qualquer sequência contida num subconjunto fechado e limitado de R^1 possui uma subsequência convergente.

Teorema. *(29.5)*

Qualquer sequência contida num subconjunto fechado e limitado de R^m possui uma subsequência convergente.

Teorema. *(29.6)*

Seja S um subconjunto de R^m que tem a seguinte propriedade: qualquer sequência contida em S possui uma subsequência convergente com limite em S. Então, S é fechado e limitado.

10.2 Conuntos Conexos

Definição. Dizemos que S é disconexo se existirem conjuntos abertos U_1 e U_2, tais que

(1) $S \subset U_1 \cup U_2$

(2) $S \cap U_1 \neq \emptyset$ e $S \cap U_2 \neq \emptyset$

(3) $U_1 \cap U_2 = \emptyset$

Se um conjunto não é disconexo, ele é conexo.

Teorema. *(29.7)*

Um subconjunto S de R é conexo se, e somente se, dados $x \in S$, $z \in S$ e $x \leq y \leq z$, vale $y \in S$.

10.2.1 Propriedade da Cobertura Finita

Seja S um conjunto em R^m. Seja u uma coleção de conjuntos abertos, tal que cada ponto de S esteja em pelo menos um dos conjuntos $u : S \subset \cup\{U : U \in u\}$. A coleção u é denominada cobertura aberta de S. Dizemos que o conjunto S tem a propriedade da cobertura finita se, de qualquer cobertura aberta u de S, pudermos extrair uma subcoleção finita de conjuntos u que ainda cobre S, ou seja: Existem $U_1, \ldots, U_k \in u$, tais que $S \subset \cup_{k=1}^{k} U_k$

Teorema. *De Heine-Borel*

Os conjuntos com a propriedade da cobertura finita são compactos.

Teorema. *(29.11)*

Se um conjunto S em R^n é fechado e limitado, então S tem a propriedade da cobertura finita.

11 Funções de Várias Variáveis

Definição. Uma função linear de R^k em R^m é uma função f que preserva a estrutura do espaço vetorial:

$$f(x+y) = f(x) + f(y) \text{ e } f(rx) = rf(x)$$

Teorema. *(13.1)*

Seja $F : R^k \to R$ uma função linear. Então existe um vetor $a \in R^k$ tal que $f(x) = a.x$ $\forall x \in R^k$.

Demonstração. Seja $a = (a_1, a_2)$;

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ e } e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Seja $a_i = f(e_i)$. Então, para cada vetor $x \in R^2$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1e_1 + x_2e_2$$

e

$$f(x) = f(x_1e_1 + x_2e_2)$$

$$= x_1f(e_1) + x_2f(e_2)$$

$$= x_1a_1 + x_2a_2$$

$$= a.x$$

□

Teorema. *(13.2)*

Seja $f : R^k \to R^m$ uma função linear. Então, existe uma matriz A de tamanho $m \times k$ tal que $f(x) = Ax$ $\forall x \in R^k$.

$$f(x) = a.x = \begin{pmatrix} a_1 & \dots & a_k \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = Ax$$

Demonstração. Prova é análoga a anterior. □

Definição. Uma forma forma quadrática em R^k é uma função real no formato:

$$Q\begin{pmatrix} x_1 & \dots & x_k \end{pmatrix} = \sum_{ij=1}^{k} a_{ij}x_ix_j$$

Teorema. *(13.3)*

A forma quadrática geral

$$Q\begin{pmatrix} x_1 & \dots & x_k \end{pmatrix} = \sum_{i\leq j} a_{ij}x_ix_j$$

pode ser escrita como:

$$\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} & \dots & \frac{1}{2}a_{1n} \\ \frac{1}{2}a_{21} & a_{22} & \dots & \frac{1}{2}a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ \frac{1}{2}a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ou seja, como x^TAx, A é simétrica

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} \\ \frac{1}{2}a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Definição. Uma função $f : R^k \to R$ é um monômio se pode ser escrita como

$$f\begin{pmatrix} x_1 & \dots & x_k \end{pmatrix} = cx_1^{a1}x_2^{a2}\dots x_k^{ak}$$

$a_1, \dots, a_k$ são inteiros não negativos. $\sum_{i=1}^{k} a_i$ é denominada grau do monômio.

Definição. Uma função $f : R^k \to R$ é um polinômio se é a soma finita de monômios de R^k. O grau mais alto que ocorre entre os monômios é denominado grau do polinômio. Uma função $f : R^k \to R^m$ é um polinômio se cada uma de suas funções componentees é um polinômio a valores reais.

Definição. Se $R^k \to R^m$ é um polinômio de grau 1, então cada componente de f tem a forma

$f_i(x) = a_ix + b_i$

Assim, a própria f também tem a forma $f(x) = Ax + b$ para alguma matriz A de tamanho mxk e algum vetor b de R^m. Tal função é denominada afim.

11.1 Funções Contínuas

Definição. Seja f uma função de R^nem R^k. Seja x_0 um vetor em R^k e $y = f(x_0)$ sua imagem. A função f é contínua em x_0 se, dada qualquer sequência $\{x_n\}_{a=1}^{\infty}$ em R^k que converge a x_0 vale que a sequência $\{f(x_n)\}_{n=1}^{\infty}$ das imagens em R^m converge a $f(x_0)$. A função f é dita contínua se é contínua em cada ponto do seu domínio.

Teorema. *(13.4)*

Sejam f e g funções de R^k em R^m. Suponha que f e g são contínuas em x, então: f+g, f-g e f.g são contínuas em x.

Teorema. *(13.5)*

Seja $f = (\ f_1 \quad \ldots \quad f_n)$ uma função de R^k em R^m. Então f é contínua se, e somente se, cada $f_i : R^k \to R^m$ é contínua em x.

Definição. Dizemos que $f : A \to B$ é sobrejetora ou aplica A sobre B se para $b \in B$, existe um $a \in A$ tal que $f(a) = b$; Em outras palavras se a imagem é igual ao contradomínio de *f*.

Definição. Em geral, dizemos que $f : A \to B$ é injetora em um conjunto C de A se, e somente se, para quaisquer x,y em C

$f(x) = f(y) \Rightarrow x = y$

$C \subset A$ se para cada $b \in f(C)$ é a imagem de precisamente um elemento de c.

11.2 Funções Inversas:

Definição. Seja $f : A \to B$ injetora num conjunto $C \subset A$, existe uma função natural de $f(C)$ de volta para C que associa a cada $b \in f(C)$ um único ponto de C que foi levado em b. Essa aplicação é denominada a inversa de *f* em C e é denotada por:

$f^{-1} : f(C) \to C$

11.3 Função Composta

Definição. Sejam $f : A \to B$ e $g : C \to D$ duas funções. Suponha que B, o contradomínio de *f*, é um subconjunto de C, o domínio de g. Então, a função composição $g \circ f : A \to D$ com g é definida por:

$$(g \circ f)(x) = g(f(x)) \forall x \in A$$

Teorema. *(13.7)*

Sejam $f : R^k \to R^m$ uma função contínua em $x \epsilon R^k$ e $g : R^m \to R^n$ uma função contínua em $f(x) \in R^m$. Então, a função composta $g \circ f : R^k \to R^n$ é uma função contínua em x.

Demonstração. Seja $\{x_n\}_{n=a}^{\infty}$ uma sequência em R^k convergente a x. Por continuidade de f em x, $\{f(x_n)\}_{n=1}^{\infty}$ converge a $f(x)$. Por continuidade de g em $f(x)$, a sequência $\{g(f(x_n))\}_{n=1}^{\infty}$ converge a $g(f(x))$. Assim, $g \circ f$ é contínua em x. $\square$

12 Cálculo a várias variáveis

Seja $(x_1(t), \cdots, x_n(t))$ para $t\in[a,b]$. Estamos interessados em: $g(t = f(x_1(t), \cdots, x_n(t)))$ para $t \in [a,b]$. $g'(t)$ da a variação de f ao longo da curva $X(t)$.

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

$$\frac{dg}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(t)) \cdot \frac{\partial x_i}{\partial t}$$

Definição. Uma função $f : R^n \to R$ é continuamente diferenciável (ou C^1) em um conjunto aberto $U \subset R^n$, se, e somente se, para cada i, a derivada parcial $\left(\frac{\partial f}{\partial x_i}\right)$ existe em cada X de U é contínua em X. Analogamente uma curva $X : (a,b) \to R^n$ é continuamente diferenciável (C^1) se cada uma de suas funções componentes $x_i(t)$é continumente diferencíavel.

Teorema. *(Regra da Cadeia I)*

Se $X(t) = (x_1(t), \cdots, x_n(t))$ é uma curva C^1 num intervalo em torno de t_0 e f é uma função C^1 numa bola em torno de $X(t_0)$, então $g(t) \equiv f(x_1(t), \cdots, x_n(t))$ é uma função C^1 em t_0 e

$$\frac{dg}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(t_0)) \frac{\partial x_i}{\partial t}$$

12.1 Derivadas Direcionais

Suponha que desejamos calcular a taxa de variação de uma função $F(x_1, \cdots, x_n)$ em um dado X^* na direção de qualquer vetor $V = (v_1, \cdots, v_n)$ dado. Para parametrizar a direção V a partir do ponto X^*, escreva a equação paramétrica da reta por X^* na direção V:

$$X = X^* + tV$$

Para ver como F varia ao longo dessa reta, inicialmente calculamos F ao longo dessa reta:

$$g(t) = F(X^* + tV) = F(x_2^* + tv_2, \cdots, x_n^* + tv_n)$$

Em seguida usaremos a regra da cadeia para tornar a derivada de g em $t = 0$

$$g'(0) = \frac{\partial F}{\partial x_1}(X^*)\, v_1 +, \cdots, + \frac{\partial F}{\partial x_n}(X^*)\, v_n$$

Derivada de F em X^* na direção V.

$$\frac{dg}{dt} = \left(\frac{\partial F}{\partial x_1}(X^*) \cdots \frac{\partial F}{\partial x_n}(X^*) \right) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = DF_{X^*}.V$$

A matriz coluna $\left(DF_{X^*}^T\right)$ pode ser interpretado como um vetor de R^n com cauda em X^*. Esse vetor é denotado por $\nabla F(X^*)$ ou, as vezes grad $F(X^*)$, é denominado vetor gradiente de F em X^*, ou simplesmente gradiente de F em X.

O comprimento, direção e sentido de um vetor gradiente têm significado. Suporemos que $\| V \| = 1$

Seja $F(K, L) = 4K^{3/4}L^{1/4}$

A derivada de $F(10000, 625)$ na direção $(1, 1)$ é simplesmente:

$$\frac{\partial F}{\partial K}(K^*)\, v_1 + \frac{\partial F}{\partial L}(L^*)\, v_2 = 3K^{-1/4}.1 + L^{-3/4}.1 = 1,5.1 + 8.1 = 9,5$$

Normalizando o comprimento para 1, fazemos:

$$\frac{\partial F(K^*)}{\partial K}.\frac{v_1}{\| V \|} + \frac{\partial F(L^*)}{\partial L}.\frac{v_2}{\| V \|}$$

$\nabla F(X^*).V$ mede a taxa à qual F aumenta ou diminui quando saímos de X^* na direção de V.

$$\nabla F(X^*).V = \| \nabla F(X^*) \| \| V \| \, cos\theta = \| \nabla F(X^*) \| \, cos\theta$$

Teorema. *Seja $F : R^n \to R$ uma função C^1. Em cada ponto X do domínio de F em que $\nabla F(X) \neq 0$, o vetor gradiente de F em X aponta na direção que F cresce mais rapidamente.*

Teorema. *Seja $F : R^n \to R^m$ e $a : R \to R^n$ funções C^1. Então a função composta $g(t) = F(a(t))$ é uma função C^1 de R^1 em R^m e*

$$g\prime_i(t) = \sum_j \frac{\partial f_i}{\partial x_j}(a_1(t), \cdots, a_n(t))\, a\prime_j(t) = Df_i(a(t)).a\prime(t)$$

Juntando todas essas condições de componentes, obtemos a equação vetorial

$$g\prime(t) = D(F \circ a(t)) = DF(a(t)).a\prime(t)$$

Teorema. *Sejam $F : R^n \to R^m$ e $A : R^s \to R^n$ funções C^1 e $s^* \in R^s$ e $x^* = A(s^*) \in R^n$ pontos. Considere a função composta*

$$H = F \circ A : R^s \to R^n$$

Teorema. *Sejam $DF(X^*)$ a matriz jacobiana $m \times n$ das derivadas parciais de F em X^* e $DA(s^*)$ a matriz jacobiana $m \times s$ das derivadas parciais de A em s^*. Então, a matriz jacobiana $DH(s^*)$ é dada pelo produto das matrizes jacobianas:*

$$DH(s^*) = D(F \circ A)(s^*) = DF(x^*) .DA(s^*)$$

Como a multiplicação matricial pode ser vista como a composição das correspondentes aplicações lineares, a regra da cadeia diz que a derivada da aplicação composta é a composição das derivadas das aplicações componentes, calculadas nos pontos certos.

Definição. Funções continuamente diferenciáveis

$$\frac{\partial f}{\partial x_i}(X^*) = \lim_{h \to 0} \frac{f\left(x_1^*, \cdots, x_i^* + h, x_{i+1}^*, \cdots, x_n^*\right) - f(x_1^*, \cdots, x_i^*, \cdots, x_n^*)}{h}$$

Se todas as derivadas de f de ordem $\leq k$ que são contínuas em um intervalo J , dizemos que f é k vezes continuamente diferenciável ou C^k. Se f é C^k para cada k, cada derivada de f de qualquer ordem existe e é contínua e dizemos que f é C^∞.

Exemplo. $f(K, L) = 4K^{\frac{3}{4}}L^{\frac{1}{4}}$

$$f_k = 3K^{-\frac{1}{4}}L^{\frac{1}{4}}$$

$$f_L = \frac{1}{4}K^{\frac{3}{4}}L^{-\frac{3}{4}}$$

$$f_{KK} = -\frac{3}{4}K^{-\frac{5}{4}}L^{\frac{1}{4}}$$

$$f_{LL} = -\frac{3}{4}K^{\frac{3}{4}}L^{-\frac{7}{4}}$$

$$f_{KL} = \frac{3}{4}K^{-\frac{1}{4}}L^{-\frac{3}{4}}$$

$$f_{LK} = \frac{3}{4}K^{-\frac{1}{4}}L^{-\frac{3}{4}}$$

$$D^2 f_{(k,L)} = \begin{pmatrix} -\frac{3}{4}K^{-\frac{5}{4}}L^{\frac{1}{4}} & \frac{3}{4}K^{-\frac{1}{4}}L^{-\frac{3}{4}} \\ \frac{3}{4}K^{-\frac{1}{4}}L^{-\frac{3}{4}} & -\frac{3}{4}K^{-\frac{5}{4}}L^{\frac{1}{4}} \end{pmatrix}$$

$f : R^2 \to R$, 2 variáveis 2^n derivadas.

$$D^2 f_x = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n x_2} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 x_n} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \text{Matriz Hessiana}$$

Teorema. *Suponha que* $Y = f(x_1, \ldots, x_{n)}$ *seja* C^1*em um* $J \subset R^n$. *Suponha que* J *é aberto, então* $\forall$ X *de* J *e para cada par de índices* i *e* j*:*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(X) = \frac{\partial^2 f}{\partial x_j \partial x_i}(X)$$

PS: O teorema de *Young* vale para derivadas de ordens superiores.

13 Cálculo a Várias Variáveis II

Nesse capítulo provaremos dois teoremas importantes:

1. O Teorema de Weierstrass, que afirmar que uma função contínua, cujo domínio é um conjunto compacto, sempre alcança seu valor máximo e seu valor mínimo em seu domínio;

2. O Teorema do Valor Médio, que é útil para quantificar a aproximação de funções diferenciáveis por polinômios de Taylor.

Definição. Um conjunto fechado e limitado de R^n é dito compacto. Esses conjuntos podem ser caracterizados pela seguinte condição: qualquer sequência cujos pontos permanecem num conjunto compacto tem uma subsequência que efetivamente converge a um limite no conjunto compacto.

13.1 Teorema de Weierstrass

Teorema. ***de Weierstrass***

Seja $F : C \to R$ *uma função contínua cujo domínio é um subconjunto* C *compacto de* R^n. *Então, existem pontos* x_m *e* x_M *em* C *tais que* $F(x_m) \leq F(x) \leq F(x_M)$ *para qualquer* $x \in C$ *; ou seja,* $x_m \in C$ *é o mínimo global de* F *em* C *e* $x_M \in C$ *é o máximo global de* F *em* C.

Demonstração. Suponha que F não é limitada. Então existe $\{x_n\}_{n=1}^{\infty}$ em C tal que $F(x_n) \to \infty$ quando $n \to \infty$. Se C é compacto, temos $\{y_n\}_{n=1}^{\infty}$ uma subsequência de $\{x_n\}_{n=1}^{\infty}$ que converge para $y^* \in C$. Como $F(x_n) \to \infty$ e $\{y_n\}_{n=1}^{\infty}$ é uma subsequência de $\{x_n\}_{n=1}^{\infty}$ então $\{F(y_n)\} \to \infty$. Contudo, como F é contínua em C e $y_n \to y^*$ a sequência $\{F(y_n)\} \to F(y^*)$ (**contradição**) . Provamos que F é limitada em C.

Adicionalmente, suponha que F não atinja seu máximo em C. Seja M o conjunto dos valores do supremo de F em C. Pelo argumento do parágrafo anterior, M é limitado. Ainda, existe uma subsequência $\{z_n\}_{n=1}^{\infty}$ tal que $\{F(z_n)\} \to M$. Como C é um conjunto compacto, podemos encontrar uma sequência $\{w_n\}_{n=1}^{\infty}$ de $\{z_n\}_{n=1}^{\infty}$ tal que $\{w_n\}_{n=1}^{\infty} \to w^*$ em C. Como uma sequência convergente possui apenas um limite $F(w_n) = M$ e portanto $w^* \in C$ é o máximo global de F em C. □

Teorema. ***de Rolle***

Seja $F : [a,b] \to R$ uma função contínua em [a,b] e C^1 em]a,b[. Se $f(a) = f(b) = 0$, existe um $c \in [a,b]$ tal que $f'(c) = 0$.

Demonstração. Se f *é* constante em [a,b], então $f'(c) = 0 \forall c \in]a,b[$. Se f não é constante em [a,b] então podemos supor, sem perda de generalidade que *f>0* em algum intervalo de]a,b[. Pelo Teorema de Weierstrass, f atinge seu máximo em algum ponto $c \in [a,b]$. Então, $f'(c) > 0$ e $c \in]a,b[$. O que resulta em $f'(c) = 0$. □

Teorema. ***do valor médio***

Seja $F : U \to R$ uma função C^1 em U de R. $\forall a$ e$b \in U$ existe um c entre a e b tal que:

$$f(b) - f(a) = f'(c)(b-a)$$

$$f'(c) = \frac{f(b) - f(a)}{(b-a)}$$

Esse teorema afirma que se traçarmos qualquer segmento de reta entre quaisquer dois ponto de U, então existe um ponto entre eles no qual a reta tangente ao gráfico de f é paralela a esse segmento de reta.

Demonstração. Construa a função:

$$g_0(x) = f(b) - f(x) + \frac{f(b) - f(a)}{(b-a)}(x-b)$$

É fácil ver que $g_0(a) = 0$ e $g_0(b) = 0$; Pelo Teorema de Rolle existe $c \in]a,b[$ tal que $g_0' = 0$. Derive a expressão acima em relação a x para obter:

$$g_0'(x) = +\frac{f(b) - f(a)}{(b-a)} - f'(x)$$

Fazendo x=c e $g_0' = 0$ teremos:

$$f'(c) = \frac{f(b) - f(a)}{b-a}$$

□

14 Funções Implícitas e suas Derivadas

14.1 Função Implícita

Nas condições de primeira ordem frequentemente temos variáveis exógenas misturadas com as endógenas, como em:

$$G(x_1, \ldots, x_n, y) = 0$$

A equação acima define a variável endógena y como uma função implícita das variáveis exógenas $x_{1,} \ldots, x_n$.

Teorema (Teorema 15.1)**.** *Seja $G(x, y)$ uma função C^1 numa bola em torno de (x_0, y_0) em $\mathbb{R}^2$. Suponha que $G(x_0, y_0) = c$ e considere a expressão:*

$$G(x, y) = c$$

Se $(\partial G/\partial y)(x_0, y_0) \neq 0$ então existe uma função $y = Y(x)$ definida em um intervalo I em torno do ponto x_0 que é C^1tal que:

*1. **

(a) $G(x, y(x)) \equiv c$ para qualquer x em I

(b) $y(x_0) = y_0$

(c) $y'(x_0) = -\frac{\partial G/\partial x(x_0,y_0)}{\partial G/\partial y(x_0,y_0)}$

Considere a função:
$G(x, y) = x^2 - 3xy + y^3 - 7 = 0$ em torno do ponto $(x_0, y_0) = (4, 3)$
$\frac{\partial G}{\partial x} = 2x - 3y = -1$ em $(4, 3)$
$\frac{\partial G}{\partial y} = 3y^2 - 3x = 15$ em $(4, 3)$ $\rightarrow$como $\partial G/\partial y \neq y$, então existe uma função C^1de x em torno de (x_0, y_0).
Além disso, $y'(x_0) = -\frac{\partial G/\partial x(x_0,y_0)}{\partial G/\partial y(x_0,y_0)} = \frac{1}{15}$
$y_1 \approx y_0 + y'(x_0)\Delta x = 3 + \frac{1}{15}(0, 3)$ para $x_1 = 4, 3$.

Teorema (Teorema 15.2)**.** *Seja $G(x_1, \ldots, x_k, y)$ uma função C^1 numa bola em torno do ponto $(x_1^*, \ldots, x_k^*, y^*)$. Suponha também que $(x_1^*, \ldots, x_k^*, y^*)$ satisfaz:*

$$G(x_1^*, \ldots, x_k^*, y^*) = c$$

$$\frac{\partial G}{\partial y}(x_1^*, \ldots, x_k^*, y^*) \neq 0$$

Então existe uma função C^1 $y = y(x_1, \ldots, x_k, y)$ definida numa bola aberta B em torno de $(x_1^*, \ldots, x_k^*)$ tal que:

1. *

   (a) $G(x_1, \ldots, x_{k,} y(x_1, \ldots, x_k)) = c$ para qualquer $(x_1, \ldots, x_k) \in B$

   (b) $y^* = y(x_1^*, \ldots, x_k^*)$

   (c) $\forall i \ \frac{\partial y}{\partial x_i}(x_1^*, \ldots, x_k^*) = -\frac{\partial G/\partial x_i(x_1^*,\ldots,x_k^*)}{\partial G/\partial y(x_1^*,\ldots,x_k^*,y^*)}$

14.2 Função Inversa

Teorema. *Teorema da Função Inversa*

Definição. Uma função $F : \mathbb{R}^n \to \mathbb{R}^m$ é sobrejetora se para cada b em $\mathbb{R}^n$ existe pelo menos um x em $\mathbb{R}^n$ tal que $F(x) = b$. Uma função F é injetora se, para qualquer b em $\mathbb{R}^n$, existe no máximo um x em$\mathbb{R}^n$ tal que $F(x) = b$.

Definição 12. Seja x_0 um ponto no domíno de $F : \mathbb{R}^n \to \mathbb{R}^m$ com $F(x_0) = b_0$. Então, F é localmente sobrejetora em x_0 se dada qualquer bola aberta $B_r(x_0)$ em torno de x_0 em $\mathbb{R}^n$, existe uma bola $B_s(b_0)$ em torno de x_0 em $\mathbb{R}^m$ tal que, para cada b em $B_s(b_0)$, existe pelo menos um x em $B_r(x_0)$ tal que $F(x) = b$. Analogamente, F é localmente injetora em x_0 se existem uma bola $B_r(x_0)$ em torno de x_0 em $\mathbb{R}^n$e uma bola $B_s(b_0)$ em torno de $b_0$em $\mathbb{R}^m$ tais que, $\forall b$ em $B_s(b_0)$ existe no máximo um x em $B_r(x_0)$ tal que $F(x) = b$.

Teorema (Teorema 15.8)**.** *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função C^1 com $F(x^*) = b^*$. Seja DF_{x^*} a matriz jacobiana $m \times n$ de F em x^*:*

1. *

   (a) Se DF_{x^*}é sobrejetora ($n \geq m =$ posto de DF_{x^*}), então F é localmente sobrejetora em x^*

   (b) Se DF_{x^*}é injetora ($m \geq n =$ posto de DF_{x^*}), então F é localmente injetora em x^*

Teorema (Teorema 15.9)**.** *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função C^1 com $F(x^*) = y^*$. Se DF_{x^*} é não singular, então existem uma bola aberta $B_r(x^*)$ em torno de x^* e um conjunto aberto V em torno de y^* tais que F é uma aplicação injetora e sobrejetora de $B_r(x^*)$ em V. A inversa natural $F^{-1} : V \to B_r(x^*)$ é também C^1e*

$$(DF^{-1})_{F(x^*)} = (DF_{x^*})^{-1}$$

Observação. Uma aplicação F contínua e injetora de um conjuto U sobre um conjunto V, que possui uma inversa $F^{-1} : V \to U$ contínua, é denominada homeomorfismo entre U e V. Se F e F^{-1} são C^1, F é dita um difeomorfismo entre U e V.

15 Formas Quadráticas e Matrizes Definidas

Uma forma quadrática sempre assume o valor zero no ponto $x = 0$.

A forma quadrática geral de uma variável é $y = ax^2$. Se $a > 0$, então ax^2 é sempre ≥ 0 e é igual a zero somente quando $x = 0$. Tal forma é chamada positiva; $x = 0$ é um mínimo gloobal. Se $a < 0$, então $ax^2 \leq 0$ é igual a 0 quando $x = 0$.Tal forma é chamada negativa; $x = 0$ é seu máximo global.

- Formas quadráticas positivas ou negativas, são ditas como definidas.
- $Q(x_1,\ x_2) = x_1^2 - x_2^2$, que assumem valores tanto positivos quanto negativos são chamadas indefinidas.

Há dois casos intermediários: uma forma quadrática que é sempre $\geq$, mas pode ser zero em alguns x não nulos, é chamada de não negativa. Veja por exemplo, $Q_4(x_1^2,\ x_2^2) = (x_1 + x_2)^2 = x_1^2 + 2x_1x_2 + x_2^2$, é nunca negativa, mas é zero em pontos não-nulos como $(x_1,\ x_2) = (1, -1)$ ou $(-2, 2)$. Uma forma quadrática como $Q_5(x_1,\ x_2) = -(x_1 + x_2)^2$ que nunca é positiva, mas pode ser zero em alguns x fora da origem é chamada não-positiva. Formas quadráticas não-negativas ou não-positivas são chamadas de semidefinidas.

15.1 Matrizes Simétricas Definidas

Definição. Seja A uma matriz simétrica $n \times n$. Então A é:

(a) *positiva* se $x^T Ax > 0 \forall x \neq 0$ *em* R^n*;*

(b) *não negativa* se $x^T Ax \geq 0$;

(c) *negativa* se $x^T Ax < 0$*;*

(d) *não positiva* se $x^T Ax \leq 0$;

(e) *indefinida* se $x^T Ax > 0$ para alguns x em R^n e $x^T Ax < 0$ para outros x em R^n.

Definição. Menor Principal: Seja A uma matriz $n \times n$. Uma submatriz principal de ordem k de A é uma submatriz de A de tamanho $k \times k$ formada a partir de A suprimindo $n - k$ colunas, digamos, as colunas i_1, i_2, ..., i_{n-k} e as mesmas $n - k$ linhas, ou seja, as linhas i_1,i_2, ... ,i_{n-k}. O determinante de uma submatriz principal $k \times k$ é denominado um menor principal de ordem k de A.

Definição. Seja A uma matriz $n \times n$. A submatriz principal de ordem k de A obtida suprimindo as últimas $n - k$ linhas e as últimas $n - k$ colunas de A é denominada a submatriz principal líder de ordem k de A.Seu determinante é denominado o menor principal líder de ordem k de A. Vamos denotar a submatriz principal líder de ordem k por A_k e o correspondente menor principal líder por $\left|A_k\right|$.

Uma matriz $n \times n$ tem n submatrizes principais líderes.

Teorema. ***16.1:***

Seja A uma matriz $n \times n$ simétrica. Então,

(a) A é positiva se, e somente se, todos os n menores principais líderes são (estritamente) positivos.

(b) A é negativa se, e somente se, os n menores principais líderes de A alternam de sinal, como segue:

$$\left|A_1\right| < 0, \ \left|A_2\right| > 0, \ \left|A_3\right| < 0, \text{ etc.}$$

O k-ésimo menor principal líder deveria ter o mesmo sinal de $(-1)^k$.

(c) Se algum menor principal líder de A de ordem k (ou um par de menores) é não nulo mas não encaixa em nenhum dos dois padrões de sinal acima, então A é indefinida. Este caso ocorre quando A tem um menor principal líder de ordem k negativo com k um inteiro par ou quando A tem um menor principal líder negativo de ordem k e um menor principal líder positivo de ordem l, com k e l dois inteiros ímpares distintos.

Teorema. ***16.2:***

Seja A uma matrizn $\times\, n$ simétrica. Então A é não negativa se, e somente se, todos os menores principais de A são ≥ 0; A é não positiva se, e somente se, cada menor principal de A de ordem ímpar é ≤ 0 e cada menor principal de A de ordem par é ≥ 0.

15.2 Restrições Lineares e Matrizes Orladas

Como a maioria dos problemas em economia envolve restrições sobre as variáveis em estudo, discutiremos as formas quadráticas restritas a subespaços de R^n :

$$Q(x_1,\ x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

s. a

$$Ax_1 + Bx_2 = 0.$$

$x_1 = -\frac{B}{A}x_2$

$$Q\left(-\tfrac{Bx_2}{A}, x_2\right) = a\left(-\tfrac{Bx_2}{A}\right)^2 + 2b\left(-\tfrac{Bx_2}{A}\right)x_2 + cx_2^2.$$

Então Qé positiva no conjunto-restrição se, e somente se $aB^2 - 2bAB + cA^2 > 0$. E Q é negativa se $aB^2 - 2bAB + cA^2 < 0$.

Alternativamente:

$$aB^2 - 2bAB + cA^2 = -det \begin{pmatrix} 0 & A & B \\ A & a & b \\ B & b & c \end{pmatrix}.$$

Teorema. ***16.3:***

A forma quadrática $Q(x_1,\ x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$ é positiva (respectivamente, negativa) no conjunto-restrição $Ax_1 + Bx_2 = 0$ se, e somente se, o determinante:

$$det \begin{pmatrix} 0 & A & B \\ A & a & b \\ B & b & c \end{pmatrix}$$

é negativo (respectivamente, positivo).

Teorema. ***16.4:***

Para determinar a classificação da forma quadrática,

$$Q(x) = x^T A x = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

de n variáveis, restrita ao conjunto-restrição $Bx = 0$ dado por:

$$\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

equações lineares, construa a matriz simétrica orlada H de tamanho $(n+m) \times (n+m)$ colocando os coeficientes B da restrição linear na orla acima e à esquerda de A :

$$H = \begin{pmatrix} 0 & B \\ B^T & A \end{pmatrix}.$$

Confira os sinais dos últimos $n - m$ menores principais líderes de H, começando com o determinante de H mesmo.

(a) Se $det\ H$ tem o mesmo sinal de $(-1)^n$ e se estes últimos $n - m$ menores principais líderes alternam de sinal, então Q é *negativa* no conjunto-restrição $Bx = 0$ e $x = 0$ é um *max* global estrito de Q neste conjunto-restrição.

(b) Se *det* H e estes últimos $n-m$ menores principais líderes têm todos o *mesmo* sinal de $(-1)^m$, então Q é *positiva* no conjunto-restrição $Bx = 0$ e $x = 0$ é um *min* global estrito de Q neste conjunto-restrição.

(c) Se ambas as condições *a)* e *b)* são violadas por menores principais líderes não-nulos, então Q é indefinida no conjunto-restrição $Bx = 0$ e $x = 0$ não é nem um *max* nem um *min* de Q no conjunto-restrição.

Teorema. ***16.5:***

Para determinar a definição de uma forma quadrática $Q(x_1, \ldots, x_n)$ sujeita auma restrição linear, construa a matriz orlada H de tamanho $(n+1) \times (n+1)$ usual como em:

$$H_{n+1} = \begin{pmatrix} 0 & A_1 & \cdots & A_n \\ A_1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_n & a_{1n} & \cdots & a_{nn} \end{pmatrix}.$$

Suponha que $A_1 \neq 0$. Se os últimos n menores principais líderes de H_{n+1} têm o mesmo sinal, então Q é positiva no conjunto-restrição (e $x = 0$ é um min restrito de Q). Se os últimos n menores principais líderes de H_{n+1} alternam o sinal, então Q é negativa no conjunto-restrição (e $x = 0$ é um max restrito de Q).

16 Otimização não condicionada

Definição. Seja $F : U \to \mathbb{R}$ uma função real de n variáveis, cujo domínio é um subconjunto do $\mathbb{R}^n$.

1. Um ponto $x^* \in U$ é máximo de F em U se $F(x^*) \geqslant F(x)$ para cada $x \in U$.
2. Um ponto $x^* \in U$ é um máximo estrito se x^* é um máximo e $F(x^*) > F(x)$ para cada $x \neq x^*$ em U.
3. Um ponto $x^* \in U$ é um máximo local de F se existe uma bola $B_r(x^*)$ em torno de x^* tal que $F(x^*) \geq F(x)$, para cada $x \in B_r(x^*) \cap U$.
4. Um ponto $x^* \in U$ é um máximo local estrito de F se existe uma bola $B_r(x^*)$ em torno de x^* tal que $F(x^*) > F(x)$ para cada $x \neq x^*$ em $B_r(x^*) \cap U$.

Observação. Invertendo as desigualdades, temos as definições de mínimo.

16.1 Condições de primeira ordem

Teorema. *(17.1) Seja $F : U \to \mathbb{R}$ uma função C^1 definida num subconjunto U de $\mathbb{R}^n$. Se x^* é um máximo ou um mínimo local de F em U e se x^* é um ponto interior de U, então:*

$$\frac{\partial F}{\partial x_i}(x^*) = 0, \; \forall \, i = 1, ..., n.$$

Demonstração. (máximo local)

Seja $B = B_r(x^*)$ uma bola em torno de x^* com a seguinte propriedade: $F(x^*) \geq F(x)$ para cada $x \in B$. Como x^* maximiza F em B, x^* também é o máximo de F em qualquer segmento de reta paralelo a um dos eixos que passa por x^* e está contido em B. Em outras palavras, x_i^* maximiza a função de uma variável:

$$x_i \to F(x_1^*, ..., x_{i-1}^*, x_i^*, x_{i+1}^*, ..., x_n^*)$$

para $x_i \in (x_{i-r}^*, x_{i+r}^*)$ e, então: $\frac{\partial F}{\partial x_i}(x^*) = 0, \; \forall \, i = 1, ..., n.$ □

16.2 Condições de segunda ordem

Condições suficientes

Definição. Um ponto n-dimensional x^* é um ponto crítico de uma função $F(x_{1,}..., x_n)$ se x^* satisfaz:

$$\frac{\partial F}{\partial x_i}(x^*) = 0, \; \forall \, i = 1, ..., n.$$

Teorema. *(17.2) Seja $F : U \to \mathbb{R}$ uma função C^2 cujo domínio é o aberto U em $\mathbb{R}^n$. Suponha que x^* é um ponto crítico de F, isto é, satisfaz a definição anterior:*

1. *Se a hessiana $D^2F(x^*)$ é uma matriz simétrica negativa, então x^* é um máximo local estrito de F.*

2. *Se a hessiana $D^2F(x^*)$ é uma matriz simétrica positiva, então x^* é um mínimo local estrito de F.*

3. *Se $D^2F(x^*)$ é indefinida, então x^* não é nem um máximo local nem um mínimo local de F.*

Definição. Um ponto crítico x^* de F para o qual a hessiana $D^2F(x)$ é indefinida é denominado ponto de sela de F.

Teorema. *(17.3) Seja $F : U \to \mathbb{R}$ uma função C^2 cujo domínio é um conjunto aberto U em $\mathbb{R}^n$. Suponha que:*

$$\frac{\partial F}{\partial x_i}(x^*) = 0, \; \forall \, i = 1, ..., n$$

e que os menores principais líderes de $D^2F(x^*)$ alternam de sinal em x^*:

$$\begin{vmatrix} F_{x_1x_1} \end{vmatrix} < 0, \qquad \begin{vmatrix} F_{x_1x_1} & F_{x_2x_1} \\ F_{x_1x_2} & F_{x_2x_2} \end{vmatrix} > 0, \qquad \begin{vmatrix} F_{x_1x_1} & F_{x_2x_1} & F_{x_3x_1} \\ F_{x_1x_2} & F_{x_2x_2} & F_{x_3x_2} \\ F_{x_1x_3} & F_{x_2x_3} & F_{x_3x_3} \end{vmatrix} < 0, \ldots$$

Então, x^* é um máximo local estrito de F.

Teorema. *(17.4) Análogo, os menores principais líderes devem ser todos positivos para que x^* seja um mínimo.*

Teorema. *(17.5) Se os menores principais líderes não respeitam esses critérios, então x^* é um ponto de sela.*

Condições necessárias

A desigualdade fraca substitui a desigualdade estrita. Em resumo, substituimos as condições de negativa e positiva da hessiana de F pela exigência que a hessiana deve ser não positiva (máximo) e não negativa (mínimo).

Nessas condições, os teoremas 17.6 e 17.7 reproduzem os resultados dos teoremas 17.3 e 17.4.

16.3 Máximo e mínimo globais

Máximo global de funções côncavas

Teorema. *(17.8) Seja $F : U \to \mathbb{R}$ uma função C^2 cujo domínio é um subconjunto aberto e convexo de $\mathbb{R}^n$.*

(a) As 3 condições são equivalentes:

1. F é uma função côncava em U;
2. $F(y) - F(x) \leq DF(x)(y - x)$ para quaisquer $x, y \in U$; e
3. $D^2F(x)$ é não positiva para qualquer $x \in U$.

(b) As 3 condições são equivalentes:

1. F é uma função convexa em U;
2. $F(y) - F(x) \geq DF(x)(y - x)$ para quaisquer $x, y \in U$; e
3. $D^2F(x)$ é não negativa para qualquer $x \in U$.

(c) Se F é côncava em U e $DF(x^*) = 0$ para algum $x^* \in U$, então x^* é um máximo global de F em U.

(d) Se F é convexa em U e $DF(x^*) = 0$ para algum $x^* \in U$, então x^* é um mínimo global de F em U.

Exemplo. Seja $F(x, y) = x^3 - y^3 + 9xy$.
As derivadas:

$$F_x = 3x^2 + 9y = 0$$

$$F_y = -3y^2 + 9x = 0$$

Pela razão $\frac{F_x}{F_y}$, temos:
$\frac{x^2}{y^2} = \frac{x}{y}$ e $-y = x$.
Logo:

$$3x^2 + 9x = 0$$

$$3x^2 = -9x$$

$$y = -3 \text{ e } x = 3$$

$$y = 0 \text{ e } x = 0$$

A hessiana: $D^2F(x^*) = \begin{pmatrix} F_{xx} & F_{yx} \\ F_{xy} & F_{yy} \end{pmatrix}$
As derivadas segundas:

$$F_{xx} = 6x$$

$$F_{yy} = -6y$$

$$F_{xy} = F_{yx} = 9$$

E teremos: $\begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$
Como: $F_{xx} > 0$; $F_{xx}F_{yy} - F_{xy}F_{yx} > 0$
$-36xy - 81 > 0$ (mínimo)

$-36xy - 81 < 0$ (máximo)

Então:

$(0,0)$: ponto de sela: $F_{xx} = 0$ e $det = -81$.

$(3,-3)$: mínimo local estrito: $F_{xx} = 18$ e $det = 243$.

$(-3,3)$: máximo global: $F_{xx} = -18$ e $det = 243$.

17 Otimização com restrições I: Condições de Primeira Ordem

Teorema. ***18.1:***

Seja f e h funções C^1de duas variáveis. Suponha $x^* = (x_{1,}^*\ x_2^*)$ é uma solução do problema:

$$max\ f(x_1,\ x_2)\ s.a\ h(x_1,\ x_2) = c$$

Suponha também que (x_1^*, x_2^*) não é um ponto crítico de h. Então existe um número real μ^* tal que $(x_1^*,\ x_2^*,\ \mu^*)$ é um ponto crítico da função lagranngiana. Temos:

$$\tfrac{\partial L}{\partial x_1} = 0;\ \tfrac{\partial L}{\partial x_2} = 0 \text{ e } \tfrac{\partial L}{\partial \mu} = 0.$$

Observação. Se $\frac{\partial h}{\partial x_1}$ e $\frac{\partial h}{\partial x_1}$ fossem zero no máximo a redução de um problema com restrições para um problema sem restrições, isto é, $L(x_1,\ x_2,\ \mu)$ não funcionaria. Essa imposição $\nabla h(x) \neq 0$ chama-se **qualificação da restrição.**

Demonstração. Note que f e h são tangentes em $x*$. Então

$$\nabla f(x^*) = \mu^*.\nabla h(x^*)$$

.

Em seguida consideramos o problema de maximizar uma função $f(x_1,\dots,x_n)$ de n variáveis condicionada por mais de uma, digamos, por m restrições de *igualdade.* Sejam $h_1(x),\ \dots,\ h_m(x)$ as funções que definem o conjunto-restrição. Em outras palavras, queremos

$$max \text{ ou } min\ f(x_1,\dots,x_n)$$

$$\text{sujeito a}$$

$$C_h = \{x = (x_1,\dots,x_n) \mid h_1(x) = a_1,\dots,h_m(x) = a_m\}$$

.

Então:

$$\left(\frac{\partial h}{\partial x_1}(x^*),\ \frac{\partial h}{\partial x_2}(x^*),\ \ldots,\ \frac{\partial h}{\partial x_n}(x^*)\right) \neq (0,\ 0,\ \ldots,\ 0)$$

.

A generalização natural caso estejamos tratando com m funções envolve a derivada Jacobiana:

$$Dh(x^*) = \begin{pmatrix} \frac{\partial h_1(x*)}{\partial x_1} & \cdots & \frac{\partial h_1(x*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m(x*)}{\partial x_1} & \cdots & \frac{\partial h_m(x*)}{\partial x_n} \end{pmatrix}$$

.

De modo geral x^*é um ponto crítico de $h = (h_1,\ \ldots,\ h_m)$ se o posto da matriz $Dh(x^*) < m$. Mais formalmente dizemos que $(h_1,\ \ldots,\ h_m)$ satisfaz **QRND** (Qualificação da Restrição Não Degenerada) em x^* se o posto da matriz Jacobiana $Dh(x^*)$ é igual a m.

□

Teorema. ***18.2:***

Sejam $f,\ h_{1,} \ldots, h_m$ funções C^1 de n variáveis. considere o problema de maximizar (ou minimizar) $f(x)$no conjunto-restrição:

$$C_h = \{x = (x_{1,}\ \ldots,\ x_n):\ h_1(x) = a_1,\ \ldots,\ h_m(x) = a_m\}.$$

*Suponha que $x* \in C_h$ e que $x*$ é um max ou um min (local) de f em C_h. Suponha também que x^*satisfaz a **QRND** acima. Então existem $\mu_1^*,\ \ldots,\ \mu_m^*$ tais que $(x_1^*,\ \ldots,\ x_n^*,\ \mu_1^*,\ \ldots,\ \mu_m^*) \equiv (x^*,\ \mu^*)$ é um ponto crítico do lagrangiano*

$$L(x,\ \mu) = f(x) - \sum_{i=1}^{m} \mu_i[h_i(x) - a_i]$$

$$\nabla L(x^*,\ \mu^*) = 0.$$

Exemplo:$U(x_1, x_2) = kx_1^a x_2^{1-a}$ $s.a$ $\sum_{i=1}^{2} p_i x_i = b$

17.1 Uma Restrição de Desigualdade

$$max\ f(x,\ y)\ s.a\ g(x,\ y) \leq b$$

Algumas considerações:

A restrição é ativa, isto é, se $g(x,\ y) - b = 0$ então$\lambda \geq 0$. A restrição é inativa quando $\lambda = 0$. Tal situação, na qual uma das duas desigualdades deve ser ativa, é denominada condição de folga complementar (*slackness condition).*

Teorema. ***18.3:***

Suponha que f e g são funções C^2 em R^2 e que $(x^*,\ y^*)$ maximiza f no conjunto-restrição $g(x,\ y) \leq b$. Se $g(x^*,\ y^*) = b$, suponha que:

$$\frac{\partial g}{\partial x}(x^*,\ y^*) \neq 0 \text{ ou } \frac{\partial g}{\partial y}(x^*,\ y^*) \neq 0$$

$$L(x,\ y,\ \lambda) = f(x,\ y) - \lambda[g(x,\ y) - b]$$

Então, existe um λ^* tal que:

(a) $\frac{\partial L}{\partial x}(x^*, y^*, \lambda^*) = 0.$

(b) $\frac{\partial L}{\partial y}(x^*, y^*, \lambda^*) = 0$

(c) $\lambda^*[g(x^*, y^*) - b] = 0$

(d) $\lambda \geq 0$

(e) $g(x^*, y^*) \leq b$

Teorema. ***18.4:***

Suponha que $f, g_1, \ldots, g_k$ são funções C^1 de n variáveis. Suponha que $x^* \in R^n$ é um max local de f no conjunto-restrição definido pelas k desigualdades

$$g_1(x_1, \ldots, x_n) \leq b_1, \ldots, g_k(x_1, \ldots, x_n) \leq b_k.$$

Para facilitar a notação, suponha que as primeirask restrições são ativas em x^* e que as últimas $k - k_0$ são inativas. Suponha que a seguinte ***QRND*** *está satisfeita em x^* :*

O posto x^* da matriz Jacobiana

$$\begin{pmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{k_o}(x^*)}{\partial x_1} & \cdots & \frac{\partial g_{k_0}(x^*)}{\partial x_n} \end{pmatrix}$$

das restrições ativas é k_0, ou seja, é o maior possível.

Forme o lagrangiano:

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k) \equiv f(x) - \sum_{i=1}^{m} \lambda_i[g_i(x) - b_i]$$

Então existem multiplicadores $\lambda_1^*, \ldots, \lambda_k^*$ tais que:

(a) $\frac{\partial L}{\partial x_1}(x^*, \lambda^*) = 0,\ \ldots,\ \frac{\partial L}{\partial x_n}(x^*, \lambda^*) = 0$

(b) $\lambda_1[g_1(x^*) - b_1] = 0, \ldots, \lambda_k[g_k(x^*) - b_k] = 0$

(c) $\lambda_1^* \geq 0, \ldots, \lambda_k^* \geq 0$

(d) $g_1(x^*) \leq b_1, \ldots, g_k(x^*) \leq b_k$

Teorema. ***18.5:***

Suponha que $f, g_1, \ldots, g_k, h_1, \ldots, h_m$ são funções C^1 de n variáveis. Suponha que $x^* \in R^n$ é um max local de f no conjunto-restrição definido pelas k desigualdades e pelas m igualdades:

$$g_1(x_{1,} \ldots, x_n) \leq b_1, \ldots, g_k(x_1, \ldots, x_n) \leq b_k$$

$$h_1(x_1, \ldots, x_n) = c_1, \ldots, h_m(x_1, \ldots, x_n) = c_m$$

*Sem perda de generalidade, podemos supor que as k_0 restrições de desigualdade são ativas em x^* e que as outras $k - k_0$ restrições de desigualdade são inativas. Suponha que a seguinte **QRND** está satisfeita em x^*.*

O posto em x^* da matria Jacobiana

$$\begin{pmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{k_0}(x^*)}{\partial x_1} & \cdots & \frac{\partial g_{k_0}(x^*)}{\partial x_n} \\ \frac{\partial h_1(x^*)}{\partial x_1} & \cdots & \frac{\partial h_1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m(x^*)}{\partial x_1} & \cdots & \frac{\partial h_m(x^*)}{\partial x_n} \end{pmatrix}$$

das restrições de igualdade e das restrições de desigualdade ativas *é $k_0 + m$, ou seja, é o maior possível.*

Forme o lagrangeano

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_m) = f(x) - \sum_{i=1}^m \lambda_i[g_i - b_i] + \mu_i[h_i(x) - c_i]$$

Então existem multiplicadores $\lambda_1^*, \ldots, \lambda_k^*, \mu_1^*, \ldots, \mu_m^*$ tais que:

(a) $\frac{\partial L}{\partial x_1}(x^*, \lambda^*) = 0, \ \ldots, \ \frac{\partial L}{\partial x_n}(x^*, \lambda^*) = 0$

(b) $\lambda_1^*[g_1(x^*) - b_1] = 0, \ldots, \lambda_k^*[g_k(x^*) - b_k] = 0$

(c) $h_1(x^*) = c_1, \ldots, h_m(x^*) = c_m$

(d) $\lambda_1^* \geq 0, \ldots, \lambda_1^* \geq 0$

(e) $g_1(x^*) \leq b_1, \ldots, g_k(x^*) \leq b_k$

- O **Teorema *18.6*** para um *min* global é análogo ao ***18.5*** invertendo-se as restrições $g_i(x^*) \geq b_i$.

17.2 Formulação de Kuhn-Tucker

$$Max\ f(x_{1,}\ \ldots,\ ,x_n)\ s.a\ g_1(x_1,\ \ldots,\ x_n) \leq b_1,\ \ldots,\ g_k(x_1,\ \ldots,\ x_n) \leq b_k$$

$$x_1 \geq 0,\ \ldots,\ x_n \geq 0 \tag{16}$$

KUHN e TUCKER trabalharam em um lagrangeano que não inclui as restrições de não negatividade:

$$\mathcal{L}(x,\ \lambda_1,\ \ldots,\lambda_{k,}\ v_1,\ \ldots,\ v_n) = L^{\sim}(x,\ \lambda_1,\ \ldots,\ \lambda_k) + \sum_{i=1}^{n} v_i x_i$$

$\forall\, j = 1,\ \ldots,\ n,$

$\frac{\partial L}{\partial x_j} = \frac{\partial L^{\sim}}{\partial x_j} + v_j = 0\ \ ou\ \ \frac{\partial L^{\sim}}{\partial x_j} = -v_j$

$\frac{\partial L^{\sim}}{\partial x_j} \leq 0\ \ e\ \ x_j \frac{\partial L^{\sim}}{\partial x_j} = 0$

Por outro lado para cada x,

$$\frac{\partial L^{\sim}}{\partial \lambda_i} = \frac{\partial L}{\partial \lambda_j} = b_j - g_j(x) \geq 0$$

Resumidamente:

$$\frac{\partial L^{\sim}}{\partial x_j} \leq 0;\ \ x_j \frac{\partial L^{\sim}}{\partial x_j} = 0;\ \ x_j \geq 0 \tag{17}$$

$$\frac{\partial L^{\sim}}{\partial \lambda_j} \geq 0;\ \ \lambda j \frac{\partial L^{\sim}}{\partial \lambda j} = 0;\ \ \lambda_j \geq 0 \tag{18}$$

Teorema. *18.7:*

Considere o problema de maximização condicionada (1) *sem restrições de igualdade e com uma coleção completa de restrições de não-negatovodade. Forme o kuhntuckeriano* $L^{\sim}$*, e suponha que* x^**é uma solução de* (1) *e que a matriz* $({}^{\partial g_i}/_{\partial x_j})$ *tem posto máximo em* x^**, onde os* i *variam sobre os índices das restrições* g_i *que são ativas em* x^* *e os* j *variam sobre os índices para os quais* $x_j^* > 0$. *Então existem multiplicadores não-negativos* $\lambda_1^*,\ \ldots,\ \lambda_k^*$ *tais que* $x_1^*,\ \ldots,\ x_k^*,\ \lambda_1^*,\ \ldots,\ \lambda_k^*$ *satisfaz o sistema de equações e desigualdades* (2) *e* (3) .

Exemplo. $Max \;\; U(x_1,\, x_2) \;\; s.a \;\; \sum p_i x_i \leq b$

$$L^{\sim}(x_1,\, x_2,\, \lambda) = U(x_1,\, x_2) - \lambda(\sum p_i x_i - b)$$

$$L^{\sim}_{x_1} = Ux_1 - \lambda p_1 \leq 0; \;\; x_1 L_{x_1} = 0; \;\; x_1 \geq 0$$

$$L^{\sim}_{x_2} = Ux_2 - \lambda p_2 \leq 0; \;\; x_2 L_{x_2} = 0; \;\; x_2 \geq 0$$

$$L^{\sim}_{\lambda} = p_1 x_1 + p_2 x_2 - b \geq 0; \;\; \lambda L_{\lambda} = 0; \;\; \lambda \geq 0$$

Suponha $U(x_1,\, x_2) = x_1^{0.5} x_2^{0.5}$ $p^{\rightarrow} = (1, 1)$ e $b = 100$.

$$L^{\sim}_{x_1} = 0.5 x_1^{-0.5} x_2^{0.5} - \lambda \leq 0; \;\; x_1 L_{x_1} = 0; \;\; x_1 \geq 0$$

$$L^{\sim}_{x_2} = 0.5 x_1^{0.5} x_2^{-0.5} - \lambda \leq 0; \;\; x_2 L_{x_2} = 0; \;\; x_2 \geq 0$$

$$L^{\sim}_{\lambda} = x_1 + x_2 - 100 \geq 0; \;\; \lambda L_{\lambda} = 0; \;\; \lambda \geq 0$$

$\lambda > 0 \;\; \rightarrow L_{\lambda} = 0; \;\; x_{1,} x_2 > 0 \;\; \rightarrow L_{x_1} = L_{x_2} = 0$

$$x_1 + x_2 = 100$$

$$\frac{0.5 x_1^{-0.5} x_2^{0.5}}{0.5 x_1^{0.5} x_2^{-0.5}} = \frac{\lambda}{\lambda} \;\; \rightarrow \;\; \left(\frac{x_2}{x_1}\right) = 1$$

$$x_2 = x_1$$

$$x_1 = 50; \;\; x_2 = 50; \;\; \lambda = 0.5$$

Usando:

$x_i = \alpha \frac{m}{p_i} \quad \rightarrow x_1 = 0.5 \frac{100}{1} \quad x_1 = 50.$

18 Otimização com restrições II

18.1 O significado do multiplicador

Os multiplicadores medem a sensibilidade do valor ótimo da função objetivo a variações das restrições e, como uma consequência, os multiplicadores fornecem uma medida natural de recursos escassos em problemas de maximização econômica.

18.1.1 Uma restrição de igualdade

Considere o seguinte problema:

$$max \;\; f(x,y) \;\; s.a \;\; h(x,y) = a$$

Vamos considerar que a varia de problema a problema. Para qualquer a (fixo), escreva $(x^*(a), y^*(a))$ para a solução do problema acima e escreva $\mu^*(a)$ para o multiplicador que corresponde a esta solução. Seja $f(x^*(a), y^*(a))$ a função de valor ótimo. Vamos provar que, sob condições razoáveis que valem para quase todos os problemas de maximização, $\mu^*(a)$ mede a taxa de variação do valor ótimo f em relação ao parâmetro a.

Teorema. *(19.1) Sejam f e h funções C^1 de duas variáveis. Para qualquer a fixo, seja $(x^*(a), y^*(a))$ a solução do problema max $f(x,y)$ s.a $h(x,y) = a$ com o multiplicador correspondente $\mu^*(a)$. Suponha que x^*, y^* e μ^* são funções C^1 de a e que $QRND$ (qualificação da restrição não degenerada) vale em $(x^*(a), y^*(a), \mu^*(a))$. Então:*

$$\mu^*(a) = \frac{d}{da} f(x^*(a), y^*(a))$$

Demonstração.

$$\mathcal{L}(x, y, \mu; a) = f(x,y) - \mu(h(x,y) - a)$$

$$\frac{\partial \mathcal{L}}{\partial x}(x^*(a), y^*(a), \mu^*(a); a) = 0 \;\; \rightarrow \;\; \frac{\partial f}{\partial x}(x^*(a), y^*(a), \mu^*(a)) - \mu^*(a)\frac{\partial h}{\partial x}(x^*(a), y^*(a), \mu^*(a))$$

$$\frac{\partial \mathcal{L}}{\partial y}(x^*(a), y^*(a), \mu^*(a); a) = 0 \;\; \rightarrow \;\; \frac{\partial f}{\partial y}(x^*(a), y^*(a), \mu^*(a)) - \mu^*(a)\frac{\partial h}{\partial y}(x^*(a), y^*(a), \mu^*(a))$$

para cada a. Além disso, temos que: $h(x^*(a), y^*(a))$. Usando a regra da cadeia:

$$\frac{\partial h}{\partial x}(x^*, y^*)\frac{dx^*}{da}(a) + \frac{\partial h}{\partial y}(x^*, y^*)\frac{dy^*}{da}(a) = 1$$

Sabemos que $f^*(x^*(a), y^*(a))$, logo:

$$\frac{d}{da}f(x^*(a),y^*(a))=\frac{\partial f}{\partial x}(x^*(a),y^*(a))\frac{dx^*}{da}(a)+\frac{\partial f}{\partial y}(x^*(a),y^*(a))\frac{dy^*}{da}(a)$$

$$\frac{d}{da}f(x^*(a),y^*(a))=\mu^*\frac{\partial h}{\partial x}(x^*(a),y^*(a))\frac{dx^*}{da}(a)+\mu^*\frac{\partial h}{\partial y}(x^*(a),y^*(a))\frac{dy^*}{da}(a)$$

$$\frac{d}{da}f(x^*(a),y^*(a))=\mu^*\left(\frac{\partial h}{\partial x}(x^*(a),y^*(a))\frac{dx^*}{da}(a)+\frac{\partial h}{\partial y}(x^*(a),y^*(a))\frac{dy^*}{da}(a)\right)$$

$$\frac{d}{da}f(x^*(a),y^*(a))=\mu^*.1$$

□

18.2 Várias restrições de igualdade

Teorema. *(19.2) Sejam $f, h_1, ..., h_m$ funções C^1 de $\mathbb{R}^n$. Seja $a=(a_1,...,a_m)$ uma m-tupla de parâmetros exógenos e considere o problema:*

$$max\ f(x_1,...,x_m)$$

$$s.a\ h_1(x_1,...,x_n)=a_1,...,h_m(x_1,...,x_n)=a_m.$$

Seja $x^*=(x_1^*(a),...,x_n^*(a))$ a solução do problema exposto acima, com correspondentes multiplicadores de Lagrange $\mu_1^*(a),...,\mu_m^*(a)$. Suponha também que x_i^* e μ_j^* são funções diferenciáveis de $(a_1,...,a_m)$ e que vale $QRND$. Então, para cada $j=1,...,m$ temos:

$$\mu_j(a_1,...,a_m)=\frac{\partial}{\partial a_j}f(x_1^*(a_1,...,a_m),...,x_n^*(a_1,...,a_m)).$$

Demonstração. Análoga a anterior. □

18.3 Restrições em desigualdade

Teorema. *(19.3) Seja $a^*=(a_1^*,...,a_k^*)$ uma k-tupla. Considere o problema (Q_a^*) de maximizar $f(x_1,...x_n)$ sujeita às k restrições de desigualdade:*

$$g_1(x_1,...,x_n)\le a_1^*,...,g_k(x_1,...,x_n)\le a_k^*.$$

*Seja $x_1^*(a^*),...,x_n^*(a^*)$ a solução do problema (Q_a^*) e sejam $\lambda_1^*(a),...,\lambda_k^*(a)$ os correspondentes multiplicadores de Lagrange. Suponha que à medida que **a** varia perto de a^*,*

$x_1^*, ..., x_n^*$ e $\lambda_1^*, ..., \lambda_k^*$ são funções diferenciáveis de $(a_1, ..., a_k)$ e que vale a QRND em a^*. Então, para cada $j = 1, ..., k$ temos:

$$\lambda_j(a_1^*, ..., a_k^*) = \frac{\partial}{\partial a_j} f(x_1^*(a_1^*, ..., a_k^*), ..., x_n^*(a_1^*, ..., a_k^*)).$$

Demonstração. Escrevemos a^* como a. Sejam g_j as restrições inativas: $g_j(x^*(a)) < a_j$. Seja também $a_j^{'}$ qualquer número que $g_j(x^*(a)) < a_j^{'} < a_j$ e $C^{'}$ descrito por:

$$g_1(x_1, ..., x_n) \leq a_1^*, ..., g_k(x_1, ..., x_n) \leq a_k^* \text{ com } g_j(x) \leq a_j^{'}.$$

Como $x^*(a)$ maximiza f em C, $C^{'} \subset C$ e $x^*(a) \in C^{'}$, segue que $x^*(a)$ maximiza f em $C^{'}$. Em outras palavras, se $a^{'} = (a_1, ..., a_{j-1}, a_j^{'}, a_{j+1}, ..., a_k)$, então $x^*(a^{'}) = x^*(a)$ e, portanto, $f(x^*(a^{'})) = f(x^*(a))$, de modo que o valor máximo de f não é afetado quando a_j varia um pouco. Isso implica que:

$$\frac{\partial}{\partial a_j} f(x_1^*(a_1, ..., a_m), ..., x_n^*(a_1, ..., a_m)) = 0$$

Como $\lambda_j^*(a) = 0$, a equação

$$\lambda_j(a_1^*, ..., a_k^*) = \frac{\partial}{\partial a_j} f(x_1^*(a_1^*, ..., a_k^*), ..., x_n^*(a_1^*, ..., a_k^*)).$$

vale para para restrições em desigualdade inativas. □

18.4 Teoremas de envoltória

Os teoremas 19.1-19.3 são casos especiais de uma classe de teoremas que descrevem como o valor ótimo da função objetivo num problema de otimização parametrizado se altera quando um dos parâmetros se modifica. Tais teoremas são denominados teoremas de envoltória. Começamos com o teorema da envoltória para problemas sem restrições.

Problemas sem restrições

Teorema. *(19.4) Seja $f(x; a)$ uma função C^1 de $x \in \mathbb{R}^n$ e a escalar a. Para cada escolha do parâmetro a, considere o problema sem restrições: max $f(x; a)$ em relação a x.*

Seja $x^*(a)$ uma solução do problema. Suponha que $x^*(a)$ é uma função C^1 de a. Então:

$$\frac{d}{da} f(x^*(a); a) = \frac{\partial}{\partial a} f(x^*(a); a).$$

Demonstração. Calculamos a regra da cadeia, tal que:

$$\frac{d}{da} f(x^*(a); a) = \sum_i \frac{\partial f}{\partial x_i}(x^*(a); a) \frac{dx_i^*}{da}(a) + \frac{\partial f}{\partial a}(x^*(a); a) = \frac{\partial f}{\partial a}(x^*(a); a)$$

pois $\frac{\partial f}{\partial x_i}(x^*(a); a) = 0, \ \forall\, i = 1, ..., n.$ □

Exemplo. Suponha que $f(x; a) = -x^2 + 2ax + 4a^2$. Desejamos escolher x que maximize essa função.

$$f^{'}(x; a) = -2x + 2a = 0 \quad \rightarrow \quad x^*(a) = a$$

$$f(x^*(a); a) = -a^2 + 2a^2 + 4a^2 = 5a^2 \quad \rightarrow \quad \frac{\partial f}{\partial a}(x^*(a); a) = 10a$$

Se aplicássemos diretamente o teorema de envoltória, poderíamos ter pulado o primeiro passo e encontrado:

Avaliando f em $x^*(a)$, isto é, $f(x^*(a); a) = -x^2 + 2ax^* + 4a^2$:

$$\frac{\partial f}{\partial a}(x^*(a); a) = \frac{dx^*}{dx}\frac{dx}{da} + 2\left(x^* + \frac{dx^*}{dx}\frac{dx}{da}\right) + 8a = 2x^* + 8a = 10a.$$

Exemplo. $\pi(p; \alpha) = \max_x (p\alpha y - c(y))$

Primeiramente, calcule $y(\alpha)$: $p\alpha = c^{'}(y)$

$$\frac{d\pi}{d\alpha} = \frac{\partial}{\partial \alpha}(p\alpha y - c(y)) = py > 0$$

$$\pi^*(p; \alpha) = p\alpha y^*(\alpha) - c^*(y^*(\alpha))$$

$$\frac{d\pi^*}{d\alpha} = p\alpha\frac{dy^*}{dy}\frac{dy}{d\alpha} + py^*(\alpha) - \frac{dc^*}{dy}\frac{dy}{d\alpha} \quad \rightarrow \quad py^* > 0$$

18.5 Problemas com restrições

Teorema. *(19.5) Sejam $f, h_1, ..., h_k : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ funções C^1. Seja $x^*(a) = (x_1^*(a), ..., x_n^*(a))$ a solução do problema de maximizar $x \rightarrow f(x; a)$ no conjunto restrição:*

$$h_1(x; a) = 0, ..., h_k(x; a) = 0$$

para qualquer escolha do parâmetro a. Suponha que $x^*(a)$ e os multiplicadores de Lagrange $\mu_1(a), ..., \mu_k(a)$ são funções C^1 de a e que vale a $QRND$. Então:

$$\frac{d}{da}f(x^*(a); a) = \frac{\partial \mathcal{L}}{\partial a}(x^*(a), \mu(a); a)$$

onde $\mathcal{L}$ é o lagrangeano natural deste problema.

Demonstração. A prova é deixada como exercício. □

18.6 Condições de segunda ordem

Problemas de maximização com restrição

Teorema. *(19.6) Sejam $f, h_1, ..., h_k$ funções C^2 em $\mathbb{R}^n$. Considere o problema de maximizar f no conjunto restrição:*

$$C_h \equiv \{x : h_1(x) = c_1, ..., h_k(x) = c_k\}$$

Forme o lagrangeano:

$$\mathcal{L}(x_1, ..., x_n, \mu_1, ..., \mu_k) = f(x) - \mu_1(h_1(x) - c_1) - ... - \mu_k(h_k(x) - c_k),$$

e suponha que:

(a) *x^* está no conjunto restrição C_h;*

(b) *Há $\mu_1^*, ..., \mu_k^*$ tal que: $\frac{\partial \mathcal{L}}{\partial x_1} = 0, ..., \frac{\partial \mathcal{L}}{\partial x_n} = 0, \frac{\partial \mathcal{L}}{\partial \mu_1} = 0, ..., \frac{\partial \mathcal{L}}{\partial \mu_k} = 0$, para os pontos $(x_1^*, ..., x_n^*, \mu_1^*, ..., \mu_k^*)$;*

(c) *A hessiana de $\mathcal{L}$ com respeito a x em (x^*, μ^*), $D_x^2\mathcal{L}(x^*, \mu^*)$, é negativa no conjunto restrição linear, ou seja, $v \neq 0$ e $Dh(x^*)v = 0 \Rightarrow v^T D_x^2\mathcal{L}(x^*, \mu^*))v < 0$. Então x^* é um máximo condicionado local estrito de f em C_h.*

Primeiramente, vejamos a prova do teorema 19.6 para o problema de maximização condicionada mais simples: 2 variáveis e uma restrição de igualdade.

Teorema. *(19.7) Sejam f e h funções C^2 em $\mathbb{R}^2$. Considere o problema de maximizar f no conjunto restrição $C_h = \{(x, y) : h(x, y) = c\}$. Forme o lagrangeano:*

$$\mathcal{L}(x, y, \mu) = f(x, y) - \mu(h(x, y) - c).$$

Suponha que (x^*, y^*, μ^*) satisfaz:

(a) $\frac{\partial \mathcal{L}}{\partial x} = 0, \frac{\partial \mathcal{L}}{\partial y} = 0, \frac{\partial \mathcal{L}}{\partial \mu} = 0$ *em (x^*, y^*, μ^*); e*

(b) $$det\begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix} > 0 \text{ em } (x^*, y^*, \mu^*).$$

Então, (x^*, y^*) é um máximo local de f em C_h.

Demonstração. A condição (b) implica $\nabla h(x^*, y^*) \neq 0$. Suporemos que $\left(\frac{\partial h}{\partial y}\right)(x^*, y^*) \neq 0$. Então, pelo teorema da função implícita (Teorema 15.1), o conjunto restrição C_h pode ser considerado como o gráfico de uma função $y = \phi(x)$ que é C^1 em torno de (x^*, y^*); em outras palavras: $h(x, \phi(x)) = c$, para qualquer x perto de x^*. Derivando-a, temos:

$$\frac{\partial h}{\partial x}(x, \phi(x)) + \frac{\partial h}{\partial y}(x, \phi(x))\phi^{'}(x) = 0 \quad ou \quad \phi^{'}(c) = -\frac{\frac{\partial h}{\partial x}(x, \phi(x))}{\frac{\partial h}{\partial y}(x, \phi(x))}$$

Seja $F(x) \equiv f(x, \phi(x))$ a função f calculada em C_h que é uma função variável não restrita. Pelas condições usuais de primeira e de segunda ordens para tais funções, se $F^{'}(x^*) = 0$ e $F^{''}(x^*) < 0$, então x^* será um máximo local estrito de F e $(x^*, y^*) = (x^*, \phi(x^*))$ será um máximo local condicionado de f.

$$F^{'}(x) = \frac{\partial f}{\partial x}(x, \phi(x)) + \frac{\partial f}{\partial y}(x, \phi(x))\phi^{'}(x)$$

Multiplicando $\frac{\partial h}{\partial x}(x, \phi(x)) + \frac{\partial h}{\partial y}(x, \phi(x))\phi^{'}(x) = 0$ por -μ^*, somando com a $F^{'}(x)$ acima e calculando ambas em $x = x^*$:

$$F^{'}(x^*) = \left(\frac{\partial f}{\partial x}(x^*, y^*) - \mu^* \frac{\partial h}{\partial x}(x^*, y^*)\right) + \phi^{'}(x^*)\left(\frac{\partial f}{\partial y}(x^*, y^*) - \mu^* \frac{\partial h}{\partial y}(x^*, y^*)\right)$$

$$= \frac{\partial \mathcal{L}}{\partial x}(x^*, y^*) + \phi^{'}(x^*)\frac{\partial \mathcal{L}}{\partial y}(x^*, y^*).$$

Agora, tome mais uma derivada de $F(x)$ em x^*, coloque $y^* = \phi(x^*)$ na equação anterior, então, teremos:

$$F^{''}(x^*) = \frac{\partial^2 \mathcal{L}}{\partial x^2} + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\phi^{'}(x^*) + \frac{\partial^2 \mathcal{L}}{\partial y^2}\phi^{'}(x^*)^2 =$$

$$= \frac{\partial^2 \mathcal{L}}{\partial x^2} + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\left(-\frac{\frac{\partial h}{\partial x}}{\frac{\partial h}{\partial y}}\right) + \frac{\partial^2 \mathcal{L}}{\partial y^2}\left(-\frac{\frac{\partial h}{\partial x}}{\frac{\partial h}{\partial y}}\right)^2 =$$

$$= \frac{1}{\left(\frac{\partial h}{\partial y}\right)^2}\left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\left(\frac{\partial h}{\partial y}\right)^2 - 2\frac{\partial^2 \mathcal{L}}{\partial x \partial y}\frac{\partial h}{\partial x}\frac{\partial h}{\partial y} + \frac{\partial^2 \mathcal{L}}{\partial y^2}\left(\frac{\partial h}{\partial x}\right)^2\right]$$

que é negativa pela hipótese (b) do teorema. Como $F^{'}(x^*) = 0$ e $F^{''}(x^*) < 0$, implica que $x \to F(x) = f(x, \phi(x))$ tem um máximo local em x^* e, portanto, f restrita a C_h tem um máximo local em (x^*, y^*). □

18.7 Problemas de minimização

As CSO para um problema de minimização condicionada envolvem a positividade de $D_x^2\mathcal{L}(x^*, \mu^*)$ no espaço nulo de $Dh(x^*)$, substituindo a condição:

$$v \neq 0 \text{ e } Dh(x^*)v = 0 \Rightarrow v^T D_x^2 \mathcal{L}(x^*, \mu^*))v < 0$$

por:

$$v \neq 0 \text{ e } Dh(x^*)v = 0 \Rightarrow v^T (D_x^2 \mathcal{L}(x^*, \mu^*))v > 0.$$

Lembre que se *det* H tem o mesmo sinal de $(-1)^n$ e se estes últimos $(n-m)$ menores principais líderes alternam de sinal, então H é negativa no conjunto restrição. Se *det* H e esses últimos $(n-m)$ menores principais líderes tem todos o mesmo sinal de $(-1)^m$, então H é positiva no conjunto restrição.

Observação. n é o número de variáveis, m é o número de restrições.

18.8 Restrições em desigualdade

Teorema. *(19.8) Sejam $f, g_1, ..., g_m, h_1, ..., h_k$ funções C^2 em $\mathbb{R}^n$. Considere o problema de maximizar f no conjunto restrição:*

$$C_{g,h} \equiv \{x : g_1(x) \leq b_1, ..., g_m(x) \leq b_m, h_1(x) = c_1, ..., h_k(x) = c_k\}.$$

Forme o lagrangeano:

$$\mathcal{L}(x_1, ..., x_n, \lambda_1, ..., \lambda_m, \mu_1, ..., \mu_k) = f(x) - \lambda_1(g_1(x) - b_1) - ... - \lambda_m(g_m(x) - b_m)$$

$$-\mu_1(h_1(x) - c_1) - ... - \mu_k(h_k(x) - c_k)$$

(a) *Suponha que existam $\lambda_1^*, ..., \lambda_m^*, \mu_1^*, ..., \mu_k^*$ tais que valem as condições de primeira ordem do Teorema 18.5, ou seja, que:*

- $\frac{\partial \mathcal{L}}{\partial x_1} = 0, ..., \frac{\partial \mathcal{L}}{\partial x_n} = 0$ *em* (x^*, λ^*, μ^*);
- $\lambda_1^* \geq 0, ..., \lambda_m^* \geq 0$;
- $\lambda_1^*(g_1(x^*) - b_1) = 0, ..., \lambda_m^*(g_m(x^*) - b_m) = 0$;
- $h_1(x^*) = c_1, ..., h_k(x^*) = c_k$.

(b) *Intencionando simplificar a notação, suponha que $g_1, ..., g_e$ são restrições ativas em x^* e $g_{e+1}, ..., g_m$ inativas. Escreva $(g_1, ..., g_e)$ como g_E. Suponha que a hessiana de $\mathcal{L}$ em relação a x em (x^*, λ^*, μ^*) é negativa no seguinte conjunto restrição linear:$\{v : Dg_E(x^*)v = 0 \text{ e } Dh(x^*)v = 0\}$, ou seja, temos: $v \neq 0$, $Dg_E(x^*)v = 0$, $Dh(x^*)v = 0 \Rightarrow v^T(D_x^2 \mathcal{L}(x^*, \lambda^*, \mu^*))v < 0$.*

Então, x^* é um máximo local estrito condicionado de f em $C_{g,h}$.

18.9 Versão de minimização

1. Troque "maximizar" por "minimizar";
2. Escreva as restrições de desigualdade como $g_i(x) \geq b_i$ na apresentação do conjunto restrição $C_{g,h}$;
3. Troque "negativa" e "< 0" na condição (b) por "positiva" e "> 0";
4. Troque "máximo" por "mínimo" na frase final.

18.10 Dependência suave dos parâmetros

Teorema. *(19.9) Seja $x^*(a)$ a solução do problema parametrizado de otimização condicionada (S_a) e seja $\mu^*(a)$ o correspondente multiplicador de Lagrange. Fixa-se o valor do parâmetro a em $a_0 (a = a_0)$. Se a matriz hessiana é não-singular no ponto $(x^*(a_0), \mu^*(a_0); a_0)$, então:*

(a) $x^*(a)$ *e* $\mu^*(a)$ *são funções* C^1 *de* a *em* $a = a_0$; *e*

(b) $QRND$ *vale em* $(x^*(a_0), \mu^*(a_0); a_0)$.

18.11 Qualificações de restrição

Teorema. *(19.10) Sejam f e h funções C^1 de duas variáveis. Suponha que $x^* = (x_1^*, x_2^*)$ é uma solução do problema $\max f(x_1, x_2)$ no conjunto restrição $\{(x_1, x_2) : h(x_1, x_2) = c\}$. Construa o lagrangeano:*

$$\mathcal{L}(x_1, x_2, \mu_0, \mu_1) \equiv \mu_0 f(x_1, x_2) - \mu_1 \left[h(x_1, x_2) - c\right].$$

Então, existem multiplicadores μ_0^* e μ_1^* tais que:

(a) μ_0^* *e* μ_1^* *não são ambos nulos;*

(b) μ_0^* *é* 0 *ou* 1*; e*

(c) *A quádrupla* $(x_1^*, x_2^*, \mu_0^*, \mu_1^*)$ *satisfaz as equações:*

- $\frac{\partial \mathcal{L}}{\partial x_1} = \mu_0 \frac{\partial f}{\partial x_1}(x_1, x_2) - \mu_1 \frac{\partial h}{\partial x_1}(x_1, x_2) = 0$
- $\frac{\partial \mathcal{L}}{\partial x_2} = \mu_0 \frac{\partial f}{\partial x_2}(x_1, x_2) - \mu_1 \frac{\partial h}{\partial x_2}(x_1, x_2) = 0$
- $\frac{\partial \mathcal{L}}{\partial \mu_1} = c - h(x_1, x_2)$

Demonstração. Suponha que (x_1^*, x_2^*) é uma solução do problema de maximização condicionada. Se (x_1^*, x_2^*) não é um ponto crítico de h, podemos tomar $\mu_0 = 1$ e usar o Teorema 18.1 para deduzir que x_1^*, x_2^*, μ_1^* satisfaz o sistema apresentado em (c). Por outro lado, se (x_1^*, x_2^*) é um ponto crítico de h e, portanto, $\frac{\partial h}{\partial x_1}$ e $\frac{\partial h}{\partial x_2}$ em (c) são nulos em (x_1^*, x_2^*), podemos tomar μ_1^* como sendo qualquer número não-nulo e tomar μ_0^* igual a zero. A quádrupla $x_1^*, x_2^*, \mu_0^*, \mu_1^*$ resultante será uma solução do sistema de (c). □

19 Funções Homogêneas e Homotéticas

Definição 13. Dado um escalar k, dizemos que uma função real $f(x_1, ..., x_n)$ é homogênea de grau k se:

$$f(tx_1, ..., tx_n) = t^k f(x_1, ..., x_n) \qquad \forall x_i \, i = 1, ..., n \text{ e } \forall t > 0$$

De modo geral, em economia trabalhamos com funções homogêneas em R_+^n.

Teorema 10. Seja $y = f(x_1, ..., x_n)$ uma função C^1 num cone aberto de R^n. Se f é homogênea de grau k, suas derivadas parciais de primeira ordem são homogêneas de grau k-1.

Demonstração.

$$f(tx_1, ..., tx_n) = t^k f(x_1, ..., x_n)$$

Tome a derivada em relação a um x_i particular:

$$t \times \frac{\partial f(tx_1, ..., tx_n)}{\partial x_i} = t^k \frac{\partial f(x_1, ..., x_n)}{\partial x_i}$$

Divida ambos os lados por *t:*

$$\frac{\partial f(tx_1, ..., tx_n)}{\partial x_i} = t^{k-1} \frac{\partial f(x_1, ..., x_n)}{\partial x_i}$$

□

Teorema 11. Seja $y = f(x_1, ..., x_n)$ uma função homogênea e C^1 no octante positivo de R^n. Os planos tangentes aos conjuntos de nível de f têm inclinação constante ao longo de raios a partir da origem.

Demonstração. Por simplicidade faremos a demostração para uma função de utilidade no R_+^2. Basicamente desejamos demonstrar que a taxa marginal de substituição (TMS) é constante ao longo de raios saindo da origem. Sejam (x_0, y_0) e $(x_1, y_1) = t\,(x_0, y_0)$ serem duas cestas de consumo originadas no mesmo raio que tem início na orgiem. Escreveremos u_x e u_y como as derivadas parciais da função de utilidade em relação a seus parâmetros.

Então teremos que:

$$\frac{u_x(x_1,y_1)}{u_y(x_1,y_1)} = \frac{u_x(tx_0,ty_0)}{u_y(tx_0,ty_0)} \quad \text{pela definição de } (x_1,y_1)$$

$$\frac{t^{k-1}u_x(x_0,y_0)}{t^{k-1}u_y(x_0,y_0)} \quad \text{pelo } \textbf{Teorema1}$$

$$\frac{u_x(x_0,y_0)}{u_y(x_0,y_0)}$$

□

Teorema 12. (Teorema de Euler) Seja $f(x_1,...,x_n)$ uma função C^1 homogênea de grau k em R^n_+. Então para qualquer x,

$$x_1\frac{\partial f}{\partial x_1} + ... + x_n\frac{\partial f}{\partial x_n} = kf(x_1,...,x_n)$$

$$x\nabla f(x) = kf(x)$$

Demonstração. Defina $\mathbf{x} = x_1,...,x_n$. Derive os dois lados da primeira equação em relação a t obtendo:

$$\frac{df(t\mathbf{x})}{dt} = \sum_{i=1}^{n} \frac{\partial f(t\mathbf{x})}{\partial x_i} x_i$$

$$\frac{d}{dt}\left[t^k f(t\mathbf{x})\right] = kt^{k-1}f(t\mathbf{x})$$

Pela definição de homogeneidade os dois lados esquerdos são iguais. Fazendo *t=1* nos dois lados direitos, obteremos o resultado desejado. □

Em seguida aplicaremos os conceitos de homogeneidade para a teoria econômica. Note que essa característica é interessante para avaliarmos como mudanças nos parâmetros da função representam em mudanças no comportantes dos agentes, sejam eles firmas ou consumidores. Num primeiro momento, vamos nos preocupar com as características da função de utilidade:

- **Utilidade Ordinal:** Depende apenas da forma e da localização dos conjuntos de indiferença do consumidor;

- **Utilidade Cardinal:** Depende da quantidade efetiva de utilidade que a função de utilidade associada a cada conjunto de indiferença. Deste modo, podemos concluir que a homogeneidade é uma propriedade cardinal.

Em seguida enunciaremos outra definição que é importante para nossa análise:

Definição 14. Seja I um intervalo da reta real. Dizemos que $g : I \to R$ é uma transformação monótona de I se g é uma função estritamente crescente em I. Além disso, se g é uma transformação monótona e u é uma função real de n variáveis então dizemos que $g \circ u : x \to g(u(x))$ é uma transformação monótona de u.

Vejamos alguns exemplos:

Seja $u(x, y) = xy$ as seguintes funções podem ser definidas como transformações monótonas de u.

$$lnu(x, y) = lnx + lny$$

$$exp(u(x, y)) = exp(xy)$$

$$3(u(x, y)) + c = 3xy + c$$

$$(u(x, y))^2 = (xy)^2$$

Definição 15. Uma característica de funções é dita ordinal se toda a transformação monótona de uma função com essa característica ainda possui essa característica.

Observação. Propriedades cardinais não são preservadas em transformações monótonas.

19.1 Funções Homotéticas

Em síntese motivaremos o uso dessas funções para a economia:

1. Os conjuntos de nível são expansões e contrações radiais uns dos outros;
2. A inclinação dos conjuntos de nível é constante ao longo de raios a partir da origem.

Definição 16. Uma função $v : R^n_+ \to R$ é denominada homotética se é uma transformação monótona de uma função homogênea, ou seja, se existe uma transformação monótona $z \to g(z)$ de R_+e uma função homogênea $u : R^n_+ \to R_+$ tais que $v(x) = g(u(x))$ para cada x do domínio de v.

Teorema 13. Seja $u : R^n_+ \to R$ uma função estritamente monótona, Então, u é homotética, se, e somente se, para x e y em R^n_+,

$$u(x) \geq u(y) \Leftrightarrow u(\alpha x) \geq u(\alpha y) \;\forall\, \alpha > 0$$

A prova desse teorema é simples e será deixada como exercício.

Teorema 14. Seja u uma função C^1em R^n_+. Se u é homotética, então a inclinação dos planos tangentes aos conjuntos de nível de u é constante ao longo de raios a partir da origem. Em outras palavras, para quaisquer i e j e qualquer x em R^n_+:

$$\frac{\frac{\partial u(t\mathbf{x})}{\partial x_i}}{\frac{\partial u(t\mathbf{x})}{\partial x_j}} = \frac{\frac{\partial u(\mathbf{x})}{\partial x_i}}{\frac{\partial u(\mathbf{x})}{\partial x_j}} \quad \forall\, t > 0$$

Demonstração. Suponha que u seja homotética, então teremos que $u\,(x, y) = \phi\,(h\,(x, y))$ para alguma função homogênea h e também monótona ϕ com $\phi^{'} > 0$.

$$\frac{\frac{\partial u(tx,ty)}{\partial x}}{\frac{\partial u(tx,ty)}{\partial y}} = \frac{\frac{\partial \phi(h(x,y))}{\partial h(x,y)}\frac{\partial h(x,y)}{\partial x}}{\frac{\partial \phi(h(x,y))}{\partial h(x,y)}\frac{\partial h(x,y)}{\partial y}}$$

$$= \frac{t^{k-1}\frac{\partial h(x,y)}{\partial x}}{t^{k-1}\frac{\partial h(x,y)}{\partial y}} = \frac{\frac{\partial h(x,y)}{\partial x}}{\frac{\partial h(x,y)}{\partial y}}$$

$$= \frac{\frac{\partial \phi(h(x,y))}{\partial h(x,y)}\frac{\partial h(x,y)}{\partial x}}{\frac{\partial \phi(h(x,y))}{\partial h(x,y)}\frac{\partial h(x,y)}{\partial y}}$$

$$= \frac{\frac{\partial h(x,y)}{\partial x}}{\frac{\partial h(x,y)}{\partial y}}$$

□

Corolário. *Seja u uma função C^1em R^n_+. Se vale a condição expressa no* ***Teorema 8****, então $\forall\, \alpha$ em $R^n_+, t > 0$ e i e j então u é homotética.*

20 Funções Côncavas e Quase côncavas

Definição 17. Uma função real f definida num subconjunto convexo U de R^n é côncava, se para quaisquer x e y em U e para todo t entre zero e um:

$$f(tx + (1-t)\,y) \geq tf(x) + (1-t)\,f(y)$$

Uma função real g definida num subconjunto convexo U de R^n é convexa, se para quaisquer x e y em U e para todo t entre zero e um, temos:

$$g(tx + (1-t)\,y) \leq tg(x) + (1-t)\,g(y)$$

Observação. Se f é côncava, -f é convexa.

Definição 18. Um conjunto U é um conjunto convexo se dados quaisquer pontos x e y em U, o segmento de reta ligando x a y:

$$l\,(x,y) = \{(tx + (1-t)\,y) : \forall t \in [0,1]\}$$

Teorema 15. Seja f definida num subconjunto convexo U de R^n. Então, f é côncava se, e somente se, sua restrição a qualquer segmento de reta em U é uma função côncava (convexa) de uma variável.

Demonstração. Escolha x e y como 2 pontos arbitrários de U. Seja $g\,(t) = f(tx+(1-t)\,y)$. Por hipótese, g é côncava. Assim, para t entre zero e um temos:

$$f(tx + (1-t)\,y) = g(t)$$

$$= g(t.1 + (1-t)\,.0)$$

$$\geq tg(1) + (1-t)\,g(0)$$

$$= tf(x) + (1-t)\,f(y)$$

consequentemente, f é côncava.

Reciprocamente, suponha que f é côncava. Queremos mostrar que a função é côncava a restrição $g\,(t) = f(tx + (1-t)\,y)$ de f ao segmento de reta contendo x e y. Para fazer isso, fixamos s_1 e s_2 e tomamos um t entre zero e um. Então,

$$g(ts_1 + (1-t)\,s_2) = f\,[(ts_1 + (1-t)\,s_2)x + (1 - (ts_1 + (1-t)\,s_2))\,y]$$

$$= f\,[(ts_1x + (1-t)\,s_1y) + (1-t)\,(s_2x + (1-t)\,s_2y)]$$

$$\geq tf\,(ts_1x + (1-t)\,s_1y) + (1-t)\,f\,(s_2x + (1-t)\,s_2y)$$

$$= tg(s_1) + (1-t)\,g(s_2)$$

Portanto, g é côncava. A prova para funções convexas é praticamente idêntica. □

Teorema 16. Seja f uma função C^1 num subconjunto convexo U *de* R^n. *Então,* f é côncava se, e somente se, para quaisquer x *e* y em U*:*

$$f\,(y) - f\,(x) \leq Df\,(x)\,(y-x)$$

ou seja,

$$f(y) - f(x) \leq \frac{\partial f(x)}{\partial x_1}(y_1 - x_1) + ... + \frac{\partial f(x)}{\partial x_n}(y_n - x_n)$$

Analogamente, f é convexa em U se, e somente se, $f(y) - f(x) \geq Df(x)(y - x)$ para quaisquer x e y em U.

Demonstração. $g_{x,y}(t) = f(ty + (1-t)x)$

Então pela regra da cadeia,

$$g'_{x,y}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + t(y - x))(y_i - x_i)$$

e

$$g'_{x,y}(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)(y_i - x_i) = Df(x)(y - x)$$

Pelos teoremas *1* e *2*, f é côncava se, e somente se, cada uma destas $g_{x,y}$ é côncava se, e somente se, para quaisquer x e y em U:

$$g_{x,y}(1) - g_{x,y}(0) \leq g'_{x,y}(0)(1 - 0) = g'_{x,y}(0)$$

se, e somente se, para quaisquer x e y em U:

$$f(y) - f(x) \leq Df(x)(y - x)$$

□

Teorema 17. Seja f uma função C^2 num conjunto aberto U de R^n. Então, f é uma função côncava em U se, e somente se, a matriz hessiana $D^2 f(x)$ é não positiva para x em U. A função f é uma função convexa em U se, e somente se, $D^2 f(x)$ é não negativa para cada x em U.

Demonstração. Escolha pontos arbitrários x e y de U e seja $g_{x,y}(t) = f(ty + (1-t)x)$. Então, f é côncava em U se, e somente se, cada $g_{x,y}(t)$ é côncava, que é equivalente a cada $g''_{x,y}(t) \leq 0$. Agora, pela equação

$$g'_{x,y}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + t(y - x))(y_i - x_i)$$

e pela regra da cadeia:

$$g''_{x,y}(t) = \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + t(y - x))(y_i - x_i)\right)$$

$$= \left(\sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x + t(y - x))(y_i - x_j)(y_i - x_i)\right)$$

$$= \left(\sum_{j,i=1}^{n} (y_i - x_j) \frac{\partial^2 f}{\partial x_i \partial x_j} (x + t(y - x)) (y_i - x_i) \right)$$

$$= (y - x)^T D^2 f (x + t(y - x)) (y - x)$$

Se cada $D^2 f(z)$ é não positiva, então,

1. cada $g''_{x,y}(t) \leq 0$

2. cada $g_{x,y}(t)$ é côncava, e a própria f é côncava.

Reciprocamente, suponha que f é côncava em U. Seja z um ponto arbitrário em U e seja v um vetor deslocamento arbitrário em R^n. Queremos mostrar que $v^T D^2 f(z) \leq 0$. Como U é aberto, existe um $t_0 > 0$ tal que $y = z + t_0 v$ está em U. Como f é côncava $g_{z,y}(t)$ é côncava e $g''_{z,y}(t) \leq 0$.

$$0 \geq g''_{z,y}(0) = (y - z)^T D^2 f(z) (y - z)$$

$$= (t_0 v)^T D^2 f(z) (t_0 v)$$

$$= t_0^2 v^T D^2 f(z) v$$

Assim, $t_0^2 v^T D^2 f(z) v \leq 0$ e $D^2 f(z)$ é não positiva para cada z em U □

20.1 Propriedades de funções côncavas

Para funções com estas características as seguintes propriedades são válidas:

1. Seus pontos críticos são máximos globais;

2. A soma ponderada de funções côncavas é uma função côncava;

3. Os conjuntos de nível de uma função côncava tem o formato ideal para a teoria do consumo e da produção.

Teorema 18. Seja f uma função côncava (convexa) num subconjunto aberto e convexo U de R^n. Se x_0 é um ponto crítico de f, ou seja, se $Df(x_0) = 0$, então $x_0 \in U$ é um máximo (mínimo) global de f em U.

Teorema 19. Seja f uma função C^1 definida num subconjunto convexo U de R^n. Se f é uma função côncava e se x_0 é um ponto de U que satisfaz $Df(x_0)(y - x_0) \geq 0 \, \forall \, y \in U$, então $x_0 \in U$ é um máximo global de f em U. Se f é uma função convexa e se x_0 é um ponto de U que satisfaz $Df(x_0)(y - x_0) \geq 0 \, \forall \, y \in U$, então $x_0 \in U$ é mínimo global do f em U.

20.2 Funções quase côncavas e quaseconvexas

Definição 19. Uma função definida num subconjunto convexo U de R^n é quasecôncava se, para cada número real a,

$$C_a^+ = \{x \in U : \ f(x) \geq a\}$$

é um conjunto convexo. Analogamente, f é quaseconvexa se, para cada número real a,

$$C_a^- = \{x \in U : \ f(x) \leq a\}$$

é um conjunto convexo.

Teorema 20. Suponha que F é uma função C^1 num subconjunto aberto convexo U de R^n. Então F é quasecôncava em U se, e somente se,

$$F(y) \geq F(x) \rightarrow DF(x)(y - x) \geq 0$$

F é quase convexa em U se, e somente se,

$$F(y) \leq F(x) \rightarrow DF(x)(y - x) \leq 0$$

Observação. Funções côncavas são quase côncavas.

21 Auto vetores e autovalores

Os autovalores de uma matriz de uma matriz $n \times n$ são os n números que resumem as propriedades essenciais daquela matriz. Como esses n números realmente caracterizam a matriz sendo estudada também são denominadas algumas vezes "valores características" ou "valores próprios".

Definição. Seja A uma matriz $n \times n$. Um autovalor de A é um número tal que, se for subtraído de cada entrada na diagonal de A, converte A numa matriz singular. Subtrair um escalar r de cada entrada diagonal de A é o mesmo que subtrair r vezes a matriz identidade I de A. Portanto, r é um autovalor de A se, e somente se, $A - rI$ é uma matriz singular.

Exemplo 1. .

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Subtraindo 2 de cada entrada diagonal A transformamos essa matriz em singular.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Teorema. *As estradas de uma matriz diagonal D são autovalores de D.*

$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ e $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = -\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$ 2 e 3 são autovalores de D.

Teorema. *Uma matriz quadrada A é singular se, e somente se,0 é um autovalor de A.*

$B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \Rightarrow B - rI = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ ou $B - rI = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$ dado que $r = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

Definição. Matriz Singular. Uma matriz A é singular se, e somente se, $detA = 0$.

Nesse caso r é um autovalor de A, ou seja, $A - rI$ é uma matriz singular se, e somente se,

$$det(A - rI) = 0$$

Para $A_{n \times n}$ o lado esquerdo da equação acima é um polinômio de grau n na variável r, denominado polinômio característico de A. O número r é um autovalor de A se, e somente se, r é uma zero do polinômio característico de A.

Seja $A_{2\times 2}$:

$$der(a - rI) = det \begin{bmatrix} a11 - n & a12 \\ a21 & a22 - r \end{bmatrix} = r^2 - (a_{11} + a_{22})r + (a_{11}a_{22} - a_{12}a_{21})$$

Portanto, uma matriz 2×2 tem no máximo dois autovalores e uma matriz $n \times n$ no máximo n autovalores.

Definição. Quando r é um autovalor de A e um vetor não nulo V tal que $(A - rI).V = 0$. Então, denominamos V um autovetor de A associado ao autovalor r.

$$Av - rIV = 0$$
$$Av = rV$$

Teorema. *Seja $A_{n\times n}$ e r um escalar.*

Então, as seguintes afirmações são equivalentes.

a. A subtração de r de cada elemento da diagonal de A transforma A em uma matriz Singular;

b. $A - rI$ é uma matriz Singular;

c. $det(A - rI) = 0$;

d. $det(A - rI)V = 0$ para algum vetor V não nulo;

e. $AV = rV$

Exemplo. Vejamos a seguinte matriz:

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

$$det(A - rI) = det\begin{vmatrix} (-1-r) & 3 \\ 2 & (0-r) \end{vmatrix}$$

$$(-1-r).-r-6->(1+r).r-6 = r^2+r-6 = (r+3)(r-2) = 0$$

As raízes do polinômio característico -3 e 2 (autovalores)

$$r = \frac{-1 \pm \sqrt{1^2.-4.1-6}}{2.1} = \frac{-1\pm\sqrt{25}}{2} = \frac{-1\pm 5}{2} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

Vejamos os autovetores

$$(A - rI)V = 0$$

$$(A - 2I)V = 0$$

$$\left(\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}\right) V = 0$$

$$\begin{bmatrix} -3 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 0$$

$$-3V_1 + 3V_2 = 0 \;\rightarrow\; V_1 = V_2$$

$$2V_1 - 2V_2 = 0->V_2 = V_1$$

$$(A - (-3)rI)V = 0$$

$$\left(\begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \right) V = 0$$

$$\begin{bmatrix} 2 & 3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 0$$

$$2V_1 + 3V_2 = 0 \ \rightarrow \ V_1 = \frac{-3}{2} V_2$$

$$\frac{-6}{2} V_2 + 3V_2 = 0$$

$$r = -3 \ \Rightarrow \ \left(\begin{bmatrix} 1 \\ \frac{-2}{3} \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{-3}{2} \\ 1 \end{bmatrix} \right)$$

$$r = 2 \ \Rightarrow \ \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

Definição. O conjunto unidimensional da equação linear $(a-rI)V = 0$, incluindo $V = 0$, é denominado auto-espaço de A em relação a r.

Exemplo.

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

$$det(B - rI) = det \begin{bmatrix} (1-r) & 0 & 2 \\ 0 & (5-r) & 0 \\ 3 & 0 & (2-r) \end{bmatrix} = (1-r)(5-r)(2-r) - 6(5-r)$$

$$(5-r)\left[(1-r)(2-r) - 6\right] = (5-r)\left[2 - r - 2r + r^2 - 6\right] = (5-r)(r^2 - 3r - 4)$$

$$(r-4)(r+2) = (r^2 + r - 4r)$$

Os autovalores de B são: 5, 4 e -1.

$(5 - r) = 0$
$(r - 4) = 0$
$(r + 1) = 0$

Calculamos o espaço nulo de $(B - 5I)$

$$(B - 5IV) = \begin{bmatrix} -4 & 0 & 2 \\ 0 & 0 & 0 \\ 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} -4V_1 + V_2 \\ 0 \\ 3V_1 - 3V_3 \end{bmatrix}$$

Cuja solução é $V_1 = V_3 = 0$ e V_2livre

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = V_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \rightarrow \text{auto vetor para } r = 5$$

Para $r = -1$

$$(B - (-1)I)V = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 6 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = 0$$

$$2V_1 + 2V_3 = 0$$

$$6V_2 = 0$$

$$3V_1 + 3V_3 = 0$$

Solução: $V_2 = 0$ e $V_1 = -V_3$ $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix}$

Para $r = 4$, $\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$

Em alguns casos é necessário utilizar eliminação gaussiana para solucionar o sistema linear $(A - rI)V = 0$

Teorema. *Os autovalores de uma matriz triangular são as suas entradas diagonais.*

Triangular superior 2×2

$A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}$-> $det(A - rI) = \begin{vmatrix} (a_{11} - r) & a_{12} \\ 0 & (a_{22} - r) \end{vmatrix} = (a_{11} - r)(a_{22} - r) = 0$

Isso ocorre se, e somente se $r = 0$

Teorema. *Seja A uma matriz invertível. Se $(A - rI)V = 0$ então $(A^{-2} - \frac{1}{r}I)V = 0$, isto é, se A é invertível r é seu autovalor se, e somente se, $\frac{1}{r}$ é um autovalor de A^{-1}.*

Demonstração.

$$(A - rI)V = 0$$

$$AV = rV \implies A^{-1}AV = A^{-1}rV \implies IV = A^{-1}rV \implies V = A^{-1}rV$$

$$\frac{V}{r} = A^{-1}V \implies V(A^{-1} - \frac{1}{r}) = 0$$

$$(A^{-1} - \frac{1}{r}I)V = 0$$

□

Exemplo. Equações lineares a diferenças

a.

$$Y_{t+1} = KY_t$$

$$Y_{t+2} = KY_{t+1}$$

$$Y_{t+3} = KY_{t+2}$$

Recursivamente: $Y_{t+3} = K.K.K.Y_t = K^3Y_t^1$

b.

$$Y_{t+1} = (1 + r)Y_t$$

$$Y_{t+2} = (1 + r)Y_{t+1}$$

$$Y_{t+3} = (1+r)Y_{t+2}$$

Recursivamente: $Y_{t+i} = (1+r)^i Y_t$

Exemplo. Modelo de Leslie

$b_1 = 1 b_2 = 4$ $d_1 = 0,5$ (sistema acoplado)
$X_{n+1} = X_n + 4Y_n$
$Y_{n+1} = \frac{1}{2}X_n + 0Y_n$-> $Z_{n+1} = \begin{bmatrix} X_{n+1} \\ Y_{n+1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix} = AZ_n$

obs.: se $b = c = 0$estas equações estão desacopladas.

Usando o método de mudança de coordenadas:
$X = \frac{1}{6}X + \frac{1}{3}Y$
$Y = \frac{1}{6}X + \frac{2}{3}Y$

Cuja transformação inversa

$$X = 4X - 2Y$$

$$Y = X + Y$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} \\ \frac{-1}{6} & \frac{2}{3} \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

As duas matrizes dos coeficientes são inversas uma da outra.

$$X_{n+1} = \frac{1}{6}X_{n+1} + \frac{1}{3}Y_{n+1} = \frac{1}{6}(X_n + 4Y_n) + \frac{1}{3}(\frac{1}{2}X_n)$$

$$Y_{n+1} = \frac{-1}{6}X_{n+1} + \frac{2}{3}Y_{n+1} = \frac{-1}{6}(X_n + 4Y_n) + \frac{2}{3}(\frac{1}{2}X_n)$$

$$X_{n+1} = \frac{1}{3}X_n + \frac{2}{3}Y_n = \frac{1}{3}(4X_n - 2Y_n) + \frac{2}{3}(X_n + Y_n) = 2X_n$$

$$Y_{n+1} = \frac{1}{6}X_n - \frac{2}{3}Y_n = \frac{1}{6}(4X_n - 2Y_n) - \frac{2}{3}(X_n + Y_n) = -Y_n$$

Está facilmente desacoplado

$$X_{n+1} = 2X_n - > X_n = 2^n C_1$$

$$Y_{n+1} = -Y_n - > Y_n = (-1)^n C_2$$

Então:

$$X_n = 4X_n - 2Y_n = 2^n C_1 - 2(-1)^n C_2$$

$$Y_n = X_n + Y_n = 2^n C_1 + (-1)^n C_2$$

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} = \begin{bmatrix} 4.2^n C_1 & -2(-1)^n C_2 \\ 2^n C_1 & (-1)^n C_2 \end{bmatrix} = C_1 2^n \begin{bmatrix} 4 \\ 1 \end{bmatrix} + C_2(-1)^n \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

As constantes C_1e C_2 são determinadas por condições iniciais exógenas X_0e Y_0. Pois, dadas as quantidades inicias X_0e Y_0 com o nº teremos:

$$X_0 = 4C_1 - 2C_2$$

$$Y_0 = C_1 + C_2$$

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$$

21.1 Sistemas Bidimensionais Abstratos

$$Z_{n+1} = AZ_n$$

Vamos reproduzir o exemplo anterior, mas utilizaremos notação matricial abstrata. Escreva P e P^{-1} para as matrizes de mudança de coordenadas:

$$Z = PZ$$

$$Z = P^{-1}Z$$

As variáveis originais são escritas como z e as transformadas como Z.

$$Z_{n+1} = P^{-1}z_{n+1} = P^{-1}(Az_n) = (P^{-1}A)z_n = (P^{-1}A)(PZ_n) = (P^{-1}AP)Z_n$$

Sejam V_1e V_2 as duas colunas da matriz P de tamanho 2×2

$D = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}$, agora a equação $P^{-1}AP = D$ é equivalente a equação:

$$AP = PD$$

Para P Invertível. Escreva a equação como:

$$A\left[V_1V_2\right] = \left[V_1V_2\right]\begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}$$

Teorema. *Seja A uma matriz $k \times k$. Sejam $r_1, ..., r_k$ autovalores de A e $V_1, V_2, ..., V_k$ os autovetores associados. Forme a matriz:*

$$P = [V_1V_2....V_k]$$

Cujas colunas são esses k autovetores. Se P é invertível então,

$$P^{-1}AP = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{bmatrix}$$

Reciprocamente, se $P^{-1}AP$ é uma matriz diagonal D, então as colunas de P são autovetores de A e todas entradas da diagonal D são autovalores de A.

Teorema. *Seja A uma matriz $k \times k$ com h autovalores distintos $r_1, ..., r_h$. Sejam $V_1, ..., V_h$os autovalores. Então $V_1, ..., V_h$ são linearmente independentes, ou seja, nenhum desses vetores pode ser escrito como uma combinação linear dos demais.*

Teorema. *Seja A uma matriz $k \times k$ com k autovalores reais e distintos $r_1, ..., r_k$ e autovetores associados $V_1, ..., V_k$. Então a solução geral do sistema de equações a diferenças $z_{n+1} = Az_n$ é*

$$z_n = C_1 r_1^n V_1 + ... + C_k r_k^n V_k$$

Teorema. *Seja A uma matriz $k \times k$. Suponha que exista uma matriz não singular P tal que:*

$$P^{-1}AP = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{bmatrix}$$

P é a matriz dos autovetores. Uma matriz diagonal, então:

$$A^n = P \begin{bmatrix} r_1^n & 0 & \cdots & 0 \\ 0 & r_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k^n \end{bmatrix} P^{-1}$$

A solução desse sistema de equações a diferenças $z_{n+1} = Az_n$ com vetor inicial z_0 é

$$Z^n = P \begin{bmatrix} r_1^n & 0 & \cdots & 0 \\ 0 & r_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k^n \end{bmatrix} P^{-1} z_0$$

Teorema. *Se a matriz A de tamanho $k \times k$ tem k autovetores reais distintos, então todas as soluções do sistema linear geral de equações a diferenças $z_{n+1} = Az_n$ tendem a zero se, e somente se, todos os autovalores de A têm valor absoluto menor do que* 1.

21.1.1 Propriedades de autovalores

Do ponto de vista prático, os autovalores de uma matriz A de tamanho $k \times k$ são simplesmente os zeros do polinômio característico de A, o polinômio de grau K dado por: $p(r) = det(A - rI)$

De fato, há 3 possibilidades para as raízes de $p(r)$.
1.$p(r)$ tem K raízes reais distintas;

2.$p(r)$ tem algumas raízes repetidas, ou
3.$p(r)$ tem algumas raízes complexas;

21.2 Traço como soma de autovalores

Definição. O traço de uma matriz quadrada é a soma das suas entradas diagonais

$$trA = a_{11} + a_{22} + a_{33} + ... + a_{kk}$$

Teorema. *Seja $A_{k\times k}$ com autovalores $r_1, ..., r_k$. Então,*

$r_1 + r_2 + ... + r_k = trA$, *e*
$r_1.r_2...r_k = detA$

Demonstração. .

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$p_A(r) = det \begin{vmatrix} (a-r) & b \\ c & (d-r) \end{vmatrix} = r^2 - (a+d)r + (ad - bc)$$

$p_A(r) = Br^2 - B^{(r_1+r_2)^r} + Br_1r_2$
$p_A(r) = B(r_1 - r)(r_2 - r)$

Duas formas de dizermos o mesmo

Coeficiente $r^2 : 1 = B$

Coeficiente de $r : -(a + d) = -B(r_1 + r_2)$

Termo constante: $ad - bc = Br_1r_2$ □

Portanto:

$$\beta = 1; trA = (a + d) = r_2 + r_1 \text{ e } detA = ad - bc = r_1r_2$$

Exemplo. Para matrizes markovianas a soma da coluna é sempre 1, logo ele é um autovalor.

$$A = \begin{pmatrix} 0,3 & 0,6 \\ 0,7 & 0,4 \end{pmatrix}$$

$$detA = \left(\frac{3}{10}.\frac{4}{10} - \frac{7}{10}.\frac{6}{10}\right) = \frac{12}{10} - \frac{42}{10} = -0,3$$

$$r_1 r_2 = -0,3 \tag{19}$$

$$r_1 + r_2 = 0,7 \tag{20}$$

$$r_1^2 + r_1 r_2 = 0,7 r_1$$

$$r_1^2 - 0,3 = 0,7 r_1$$

$$r_1^2 - 0,7 r_1 - 0,3 = 0$$

$$r_1 = \frac{+0,7 \pm \sqrt{(-0,7)^2 - 4.1 - 0,3}}{2.1}$$

$$r_1 = \frac{0,7 \pm \sqrt{1,69}}{2}$$

$$r_1 = \frac{0,7 \pm 1,3}{2}$$

$$r_1 = -0,3$$

$$r_1 = 1$$

Se $r_1 = 1$

$$r_2 = -0,3$$

Se $r_1 = -0,3$

$$r_2 = 1$$

21.3 Autovalores repetidos

Definição: Uma matriz A que tem um autovalor de multiplicidade $m > 1$, mas não possui m autovalores independentes associados a esse autovalor, é denominada matriz não diagonalizável ou defectiva.

Teorema. Seja A uma matriz 2×2 com dois autovalores iguais. Então, A é diagonalizável se, e somente se, A já é diagonal.

Demonstração. Se A é diagonalizável pela mudança de variáveis P, então as entradas na diagonal de $P^{-1}AP$ são os autovalores de A. Seja r^* o único autovalor de A, Então, $P^{-1}AP$ deve ser a matriz $\begin{pmatrix} r^* & 0 \\ 0 & r^* \end{pmatrix} = r^*I$:

$$P^{-1}AP = r^*I$$

ou equivalentemente,

$$A = P(r^*I)P^{-1} = r^*PIP^{-1} = rI$$

□

Definição: Seja r^*um autovalor da matriz A. Um vetor (não-nulo) v tal que $(A-r^*I)v \neq 0$ mas $(A - r^*I)^m v = 0$ para algum inteiro $m > 1$ é denominado um autovetor generalizado de A associado a r^*.

Exemplo:

$$A = \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix} \qquad r = 3 \text{ e } 3$$

$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ o autovetor generalizado v_2 é uma solução de $(A - 3I)v_2 = v_1$ou

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Tome, por exemplo, $v_{21} = 1, v_{22} = 0$ e forme

$$P = [v_2 = v_1] = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

checamos que:

$$P^{-1}AP = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

Teorema. Seja A uma matriz 2×2 com dois autovalores iguais $r = r^*$. Então,

(a) ou A tem dois autovetores independentes associados a r^*,e neste caso, A é a matriz diagonal r^*I.

(b) ou A tem somente um autovetor independente, digamos v_1 tal que $(A - r^*I)v_2 = v_1$e, se $P = [v_1 v_2]$ então $P^{-1}AP = \begin{pmatrix} r^* & 1 \\ 0 & r^* \end{pmatrix}$.

21.4 Resolvendo equações a diferenças não diagonalizáveis

Vamos solucionar um sistema de equações a diferenças $z_{n+1} = Az_n$ quando A não é diagonalizável.

$$Z_{n+1} = \begin{pmatrix} X_{n+1} \\ Y_{n+1} \end{pmatrix} = \begin{pmatrix} r & 1 \\ 0 & r \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$$

$$X_{n+1} = rX_n + Y_n$$

$$Y_{n+1} = rY_n$$

Esse sistema está acoplado, porém minimamente. Podemos usar a segunda equação e dizer que:

$$Y_n = c_1 r^n$$

Inserimos essa equação na primeira:

$$X_{n+1} = rX_n + c_1 r^n \quad (34)$$

Agora temos uma equação a diferenças linear homogênea e escalar para resolver. Vamos iterar a equação (34) a partir de $n = 0$ para descobrirmos a solução geral:

$$X_0 = C_0$$

$$X_1 = rX_0 + c_1 \quad n = 0$$

$$X_2 = rX_1 + c_1 r \quad n = 1$$

$$X_2 = r(rX_0 + c_1) + c_1 r$$

$$X_2 = r^2 c_0 + rc_1 + c_1 r$$

$$X_2 = r^2 c_0 + 2rc_1$$

$$X_3 = rX_2 + c_1 r^2 \quad n = 2$$

$$X_3 = r^3 c_0 + 3r^2 c_1$$

$$X_4 = rX_3 + c_1 r^3 \quad n = 3$$

$$X_4 = r^4 c_0 + 4r^3 c_1$$

Em geral:

$$X_n = r^n c_0 + nc_1 r^{n-1} \quad (35)$$

Para ver que (35) é a solução geral de (34), substitua-a em (34):

$$X_{n+1} = r(r^n c_0 + nc_1 r^{n-1}) + c_1 r^n$$

$$X_{n+1} = r^{n+1} c^0 + nc_1 r^n + c_1 r^n$$

$$X_{n+1} = c_0 r^{n+1} + (n+1) c_1 r^n$$

Então a solução geral de (33) é:

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} = \begin{pmatrix} c_0 r^n + nc_1 r^{n-1} \\ c_1 r^n \end{pmatrix}$$

Finalmente usamos a mudança de coordenadas $z = PZ$ para escrever a solução geral do nosso sistema original $z_{n+1} = Az_n$:

$$Z_n = PZ_n = [v_1 v_2] \begin{pmatrix} c_0 r^n + nc_1 r^{n-1} \\ c_1 r^n \end{pmatrix}$$

$$Z_n = (c_0 r^n + nc_1 r^{n-1}) v_1 + c_1 r^n v_2$$

Teorema. Seja A uma matriz 2×2 com um autovalor múltiplo r e somente um autovetor independente v_1. Seja v_2 um autovetor generalizado associado a v_1e r. Então, a solução geral do sistema de equações a diferenças $z_{n+1} = Az_n$ é:

$$z_n = (c_0 r^n + nc_1 r^{n-1}) v_1 + c_2 r^n v_2$$

Teorema. Seja A uma matriz $k \times k$ com entradas reais. Se $r = \alpha + i\beta$ é um autovalor de A, também seu complexo conjugado $\bar{r} = \alpha - i\beta$ é um autovalor. Se $u + iv$ é um autovetor para $\alpha - i\beta$ então $u - iv$ é um autovetor para $\alpha - i\beta$. Se k é ímpar, então A deve possuir pelo menos um autovalor real.

Seja

$$A = \begin{pmatrix} 1 & 1 \\ -9 & 1 \end{pmatrix}$$

$$p(r) = r^2 - 2r + 10$$

cujas raízes são $r = 1 + 3i$ e $1 - 3i$. Um autovetor para $r = 1 + 3i$ é uma solução w de

$$(A - (1 + 3i)I)w = \begin{pmatrix} -3i & 1 \\ -9 & -3i \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Usando a primeira linha dessa matriz, concluímos que um autovetor w é uma solução da equação

$$-3iw_1 + w_2 = 0$$

$w = \begin{pmatrix} 1 \\ 3i \end{pmatrix}$, que escrevemos como $\begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. Pelo teorema 23.13 um autovetor para o autovalor

$1 - 3i$ é:

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - i \begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -3i \end{pmatrix}$$

Formamos uma matriz P de mudança de coordenadas cujas colunas são estes dois autovetores:

$$P = \begin{pmatrix} 1 & 1 \\ 3i & -3i \end{pmatrix}$$

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{6}i \\ \frac{1}{2} & \frac{1}{6}i \end{pmatrix}$$

$$P^{-1}AP = \begin{pmatrix} \alpha + 3i & 0 \\ 0 & \alpha - 3i \end{pmatrix}$$

Teorema. Seja A uma matriz 2×2 real com autovalores complexos $\alpha^* \pm i\beta^*$ com autovetores complexos associados $u^* \pm iv^*$. Escreva os autovalores $\alpha^* \pm i\beta^*$em coordenadas polares como $r^*(cos\theta^* + isen\theta^*)$,onde

$$r^* = \sqrt{\alpha^{*2} + \beta^{*2}} \text{ e } (cos\theta^*, sen\theta^*) = \left(\frac{\alpha^*}{r^*}, \quad \frac{\beta^*}{r^*} \right)$$

Então a solução geral da equação a diferenças $z_{n+1} = Az_n$ é:

$$z_n = r^{*^n}[(c_1 cosn\theta^* - c_2 sennn\theta^*)u^* + (c_2 cosn\theta^* + c_1 sennn\theta^*)v^*]$$

No exemplo 23.17, calculamos que os autovalores de $A = \begin{pmatrix} 1 & 1 \\ -9 & 1 \end{pmatrix}$ são $1 \pm 3i$ com autovetores associados $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ 3 \end{pmatrix}$

Em coordenadas polares,

$$1 + 3i = \sqrt{10}\left(\frac{1}{\sqrt{10}} + i \quad \frac{3}{\sqrt{10}} \right) = \sqrt{10}(cos\theta^* + isen\theta^*)$$

onde $\theta^* = arcocos\left(\frac{1}{\sqrt{10}} \right) \approx 71,56^\circ$ou 1,25 radianos.

A solução geral de:

$$x_{n+1} = x_n + y_n$$

$$y_{n+1} = -9x_n + y_n$$

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = (\sqrt{10})^n[(c_1 cosn\theta^* - c_2 sen\theta^*)\begin{pmatrix} 1 \\ 0 \end{pmatrix}] - [(c_2 cosn\theta^* + c_1 sen\theta^*)\begin{pmatrix} 0 \\ 3 \end{pmatrix}]$$

$$= (\sqrt{10})^n \begin{pmatrix} c_1 cosn\theta^* - c_2 sennn\theta^* \\ -3c_2 cosn\theta^* - 3c_1 sennn\theta^* \end{pmatrix}$$

21.5 Processos de Markov

Definição: Um processo estocástico é uma regra que dá a probabilidade com que o sistema (ou um indivíduo deste sistema) estará no estado i no período $n+1$ sabendo as probabilidades com que esteve nos vários estados em períodos anteriores.

Definição: Um Processo Markov é um processo estocástico se a probabilidade com que o sistema está no estado i no período $n+1$ depende somente do estado em que os sistema esteve no período n; Para processos de Markov somente o passado imediato interessa.

(1) a probabilidade $x^i(n)$ de ocorrer o estado i no *n-ésimo* período de tempo ou, alternativamente, a fração da população em questão que está no estado i no n-ésimo período de tempo e,

(2) as probabilidades de transição m_{ij}, ou seja, as probabilidades com que o processo estará no estado i no tempo $n+1$ se estiver no estado j no tempo n.

É natural agrupar as probabilidades de transição numa matriz, que denominamos matriz de transição, ou matriz estocástica, ou ainda matriz de Markov:

$$M = \begin{pmatrix} m_{11} & \dots & m_{1k} \\ \vdots & \ddots & \vdots \\ m_{k1} & \cdots & m_{kk} \end{pmatrix}$$

Uma matriz de Markov é qualquer matriz (m_{ij}) de entradas não negativas cujas colunas tem soma $\sum_i m_{ij}$ iguais a 1. Estamos considerando que as probabilidades m_{ij} estão fixas e são independentes de n. Para descrever essa hipótese dizemos que o processo é homogêneo no tempo ou que as probabilidades de transição são estacionárias. A dinâmica de Markov pode ser descrita do seguinte modo. Suponha que $x^j(n)$ denota a fração de membros de uma população de tamanho N que está no estado j no período de tempo n. Então, o número total de membros da população no estado j no período n é $x^j(n)N$. Por exemplo, $m_{ij}x^j(n)N$ desses estarão no estado i no período $n+1$. O número total $x^i(n+1)N$ de membros da população no estado i $n+1$ é a soma sobre j dos membros da população que mudaram de j para i:

$$m^i(n+1)N = \sum_{j=1}^{k} m_{ij}x^j(n)N$$

em notação matricial, depois de dividir por N

$$\begin{pmatrix} x^1(n+1) \\ \vdots \\ x^k(n+1) \end{pmatrix} = \underset{\substack{\downarrow \\ M}}{\begin{pmatrix} m_{11} & \dots & m_{1k} \\ \vdots & \ddots & \vdots \\ m_{k1} & \cdots & m_{kk} \end{pmatrix}} \begin{pmatrix} x^1(n) \\ \vdots \\ x^k(n) \end{pmatrix} \qquad (44)$$

Sistema de Markov.

Exemplo. Cada indivíduo da população está empregado ou desempregado. Seja $x^1(n)$ a fração da população que estuda e que está empregada no fim do período de tempo n e $x^2(n)$ denota o total de desempregados. Suponha que uma pessoa empregada possuí 90% de chances de estar empregada no próximo período e, portanto, 10% de chance de estar desempregada no próximo período. Um indivíduo desempregado possuí 40% de probabilidade de se empregar e 60% de chances de se manter desempregado. A dinâmica desse problema é a seguinte:

$$x^1(n+1) = 0,9x^1(n) + 0,4x^2(n)$$

$$x^2(n+1) = 0,1x^1(n) + 0,6x^2(n)$$

$$\begin{pmatrix} x^1(n+1) \\ x^2(n+1) \end{pmatrix} = \begin{pmatrix} 0,9 & 0,4 \\ 0,1 & 0,6 \end{pmatrix} \begin{pmatrix} x^1(n) \\ x^2(n) \end{pmatrix}$$

$r = 1 \rightarrow$a soma das colunas.

O traço é 1,5 então o outro auto valor é 0,5.

$$(A - rI)v = 0$$

$$\begin{pmatrix} -0,1 & 0,4 \\ 0,1 & -0,4 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0,4 & 0,4 \\ 0,1 & 0,1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Pelo teorema 23.6 a solução geral desse sistema é:

$$\begin{pmatrix} x^1(n) \\ x^2(n) \end{pmatrix} = c_1 \begin{pmatrix} 4 \\ 1 \end{pmatrix} .1 + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} 0.5^n$$

como $1^n = 1$ e $\lim\limits_{n\to\infty} 0,5^n = 0$ então a solução de longo prazo da equação acima tende a:

$$w1 = c_1 \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

como o vetor de componentes deve somar 1, tome $c_1$como a recíproca da soma dos componentes de w_1, isto é, $\frac{1}{5}$. Podemos concluir que $w_1$tende a $\begin{pmatrix} 0,8 \\ 0,2 \end{pmatrix}$quando $n \rightarrow \infty$ e nossas pressuposições levam a um nível de desemprego de 20% nessa comunidade.

Definição. Matriz regular de Markov

Seja M uma matriz de Markov, isto é, uma matriz não negativa cuja soma das suas entradas é igual a 1. Então M é chamada de matriz regular de Markov se M^r possui somente entradas positivas para algum inteiro r. Se *r=1*, isto é, se cada entrada de M é positiva, M é chamada de matriz positiva.

Teorema. *Seja M uma matriz regular de Markov, então,*

(a) 1 é um autovalor de multiplicidade M;

(b) Qualquer outro valor de M satisfaz $|r| = 1$;

(c) O autovalor 1 possui um autovetor w_1 *com componentes estritamente positivos;*

(d) Se escrevermos v_1 *por* w_1*dividido pela soma das suas componentes, então* v_1*é um vetor de probabilidade e cada solução* $X\,(n)$*de* $X\,(n+1) = M_X(n)$*tende a* v_1*com* $n \to \infty$.

21.6 Matrizes Simétricas

Teorema. *Seja* A *uma matriz simétrica* $k \times k$ *então,*

(a) Todas as k raízes da equação característica $Det\,(A - rI) = 0$ *são números reais;*

(b) Os autovetores correspondentes a autovalores distintos são ortogonais;

(c) Se A possui múltiplos autovalores, então existe uma matriz não-singular P cujas colunas w_1*,...,* w_n*são autovetores de A tal que*

(i) w_1*,... ,* w_n *são mutuamente ortogonais*

(ii) $P^{-1} = P^T$

(iii) $P^{-1}AP = P^TAP = \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix}$

Definição. Uma matriz P que satisfaça a condição $P^{-1=}P^T$ou antelativamente(equivalentemente) $P^TP = I$ é chamada de matriz ortogonal.

Vetores Ortonormais são vetores que são ortogonais e possuem comprimento igual a 1.

21.7 Formas Quadráticas definidas

Teorema. *Seja A uma matriz simétrica. Então,*

(a) A é positiva definida se,e somente se, todos os autovalores de A são maiores que zero (>0):

(b) A é negativa definida se todos os autovalores de A são menores que zero (<0);

(c) A é positiva semidefinida se todos os autovalores de A são maiores ou iguais a zero $(\geqslant 0)$

(d) A é negativa semidefinida se todos os autovalores de A são menores ou iguais a zero$(\leqslant 0)$

(e) A é indefinida se A possui um autovalor positivo e outro negativo.

Demonstração. Seja x um vetor arbitrário não zero em R^k e seja $Y = P^{-1}x = P^Tx$. Então, Y é não zero e

$$x^TAx = y^TP^TAPy$$

$$= y^T \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix} (55)$$

$$= r_1 y_1^2 + \cdots + r_k y_k^2$$

onde pelo menos um dos y_i^2 é positivo. Se todos os $r\acute{}is$ são positivos, então $x^T A x > 0$ e A é positiva definida.

Se todos os $r_1\acute{}is$ são $(\geqslant 0)$ então $x^T A x \geqslant 0$ e A é positiva semidefinida.

Se $r_1 > 0$ e $r_2 < 0$, por exemplo, seja $e_1 = (1, \cdots, 0)^T$ e $e_2 = (0, 1, 0, \cdots, 0)^T$. Seja $x_1 = Pe_1$ e $x_2 = Pe_2$. Então,

$$x_1^T A x_1 = e_1^T P^T A P e_1 = r_1 > 0$$

$$x_2^T A x_2 = e_2^T P^T A P e_2 = r_2 < 0$$

e A é indefinida. Reciprocamente se A é indefinida, então deve levar um $e_j$negativo e um $r_i$positivo em (55).Ou seja, A tem um autovalor positivo e outro negativo. □

Teorema. *Seja A uma matriz simétrica. Então as seguintes condições são equivalentes:*

(a) A é positiva definida;

(b) Existe uma matriz não singular B tal que $A = B^T B$

(c) Existe uma matriz não singular Q tal que $Q^T A Q = I$

Demonstração. Como A é uma matriz simétrica podemos escrever

$$P^T A P = \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix} \textit{ou } A = P \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix}$$

onde $r_1, \cdots, r_k$são os autovalores de A e $P = (v_{1,\cdots,}v_k)$ é uma matriz independente de autovetores de A.

Se A é positiva definida, então $r_1, \cdots, r_k$ são todos (> 0).

Para $a \Rightarrow b$

$$B = \begin{pmatrix} \sqrt{r_1} & 0 & \dots & 0 \\ 0 & \sqrt{r_2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{r_k} \end{pmatrix}$$

Então teremos que:

$$B^T B = \begin{pmatrix} \sqrt{r_1} & 0 & \dots & 0 \\ 0 & \sqrt{r_2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{r_k} \end{pmatrix}^T \begin{pmatrix} \sqrt{r_1} & 0 & \dots & 0 \\ 0 & \sqrt{r_2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{r_k} \end{pmatrix} P^T = A$$

Para$b \Rightarrow a$

Por outro lado, se $A = B^T B$ para uma matriz não singular B, então para qualquer x não nulo de R^x,

$$x^T A x = x^T B^T B = \|B_x\|^2 > 0$$

Como B é não singular então $B_x \neq 0$

Para $a \Rightarrow c$

Suponha que A é positiva definida, então os autovalores de A são positivos. Seja

$$Q = (\frac{1}{\sqrt{r_1}} v_1, \cdots, \frac{1}{\sqrt{r_k}} v_k) = P \begin{pmatrix} \frac{1}{\sqrt{r_1}} & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{r_k}} \end{pmatrix}$$

Então, $Q^T A Q = I$.

Para $c \Rightarrow a$

Note que se a condição vale, seja x um vetor arbitrário não nulo e então $y = P^{-1} x$:

$$x^T A x = (Qy)^T A (Qy) = y^T (Q^T A Q) y$$

$$= y^T I y = y^T y = \|y\|^2 > 0$$

Então A é positiva definida. $\square$

22 Equações Diferenciais Ordinárias

Pense no crescimento de fundos aplicados na cardeneta de poupança. Suponha que esse investimento cresça a uma taxa fixa r que satisfaça a seguinte equação em diferenças:

$$\frac{y(t+1)-y(t)}{y(t)} = r \text{ ou } y(t+1) = (1+r)y(t) \tag{21}$$

Imagine que essa taxa de juros é paga a cada variação no tempo $\triangle t$, então teríamos:

$$\frac{y(t+\triangle t)-y(t)}{\triangle t} = ry(t)$$

$$\lim_{\triangle t \to 0} y(t) = \frac{y(t+\triangle t)-y(t)}{\triangle t} = \frac{dy(t)}{dt} = ry(t) \tag{22}$$

Por conveniência escreveremos $\dot{y} = \frac{dy(t)}{dt} = ry(t)$. Se aplicarmos o ln em (2) e tomarmos a derivada em relação ao tempo teremos:

$$ln(\dot{y}) = ln(r) + ln(y(t))$$

$$\frac{1}{\dot{y}}\ddot{y} = \frac{1}{y}\ddot{y}$$

$$\frac{\ddot{y}}{\dot{y}} = \frac{\dot{y}}{y}$$

$$g_{\ddot{y}} = g_{\dot{y}}$$

Definição. Equação diferencial ordinária. Uma equação diferencial ordinária é uma operação que descreve um relacionamento entre uma função de uma varável e sua derivada.

Exemplo. Veja a seguinte equação:

$$\dot{y}\,(t) = 2y\,(t) \text{ ou } \dot{y} = 2\dot{y}$$

$$\frac{dy}{dt} = 2y$$

$$\frac{dy}{y} = 2dt$$

$$\int \frac{dy}{y} = 2\int dt$$

$$ln(y) = 2t + c$$

$$e^{ln(y)} = e^{2t}$$

$$y = e^{2t}$$

Definição. De modo mais geral $y(t) = ke^{2t}$ para qualquer constante K. Essa constante é determinada pelo valor inicial y, isto é, $y(t_0)$.

Definição. Equação Diferencial Parcial. As equações diferenciais que descrevem um relacionamento entre uma função de várias variáveis e suas derivadas parciais são chamadas de equações diferenciais parciais.

Exemplo:

$$\dot{y} = y^2$$

$$\frac{dy}{dt} = y^2$$

$$\int dy y^{-2} = \int dt$$

$$-\frac{1}{y} = t + c$$

$$-\frac{1}{t} = y$$

Definição. Uma equação diferencial ordinária é uma equação $\dot{y} = F(y,t)$ entre a derivada de uma função desconhecida $y(t)$ e uma expressão $F(y,t)$ envolvendo y e t, isto é, se a equação pode ser escrita como $\dot{y} = F(y)$, nós a chamamos de autônomo ou tempo independente.

Retomando a equação $\dot{y} = ry$, ela também poderia ser usada para descrever o tamanho de uma população a uma taxa de crescimento constante r. Esse pressuposto de crescimento constante é às vezes chamado de lu de Malthus. A solução geral para esta equação é $y(t) = ke^{rt}$. No entanto, podemos especificar um valor inicial para o tamanho populacional. Com uma constante k pré determinada. Ao fazermos $y(0) = y_0$,(um valor constante) o problema de encontrarmos uma função $y(t)$ que satisfaça a essas condições é chamado de problemas de valor inicial.

Exemplo. Vejamos agora o exemplo da equação logística

$$\dot{y} = y(100 - 2y)$$

A função constante $y = 50\,\forall t$ é a solução que faz ambos os lados da equação serem zero. Essa solução constante, muitas vezes é chamada de solução estacionária ou de estado estacionário. Outro ponto importante é definirmos que uma solução parametrizada $y(t,k)$ de uma equação diferencial $\dot{y} = F(y,t)$ échamda de *solução* geral.

22.1 Soluções explícitas

22.1.1 Equações lineares de primeira ordem

Seja $\dot{y} = ay$ sendo a uma constante. A solução geral para esta equação é

$$y(t) = ke^{at} \text{ (1)}$$

$$\frac{dy}{dt} = kae^{at}$$

$$\dot{y} = ke^{at}a$$

(2) Seja $a, b > 0$ a solução geral para a equação:

$$\dot{y} = ay + b$$

$$y = \frac{-b}{a} + ke^{at} \longrightarrow ay = (-b + ke^{at})$$

Para verificar isso insira a solução candidata dentro da equação:

$$\dot{y}(t) = ake^{at}$$

$$ay(t) + b = (-b + ake^{at}) + b = ake^{at}$$

Note que $y(t) = \frac{-b}{a}$ é uma solução de estado estacionário. Agora examinaremos versões não autônomas dessas duas classes: $\dot{y} = a(t)y$ (3)

A solução geral é:

$$y(t) = ke^{\int^t a(r)dr} \text{ ou } y = ke^{\int^t a}$$

.

$$\dot{y} = (a(t))y + b(t) \quad (4)$$

$$y(t) = \left[k + \int^t b(s)e^{-\int^s a(u)du}ds\right]e^{\int^t a(s)ds} \quad (10)$$

Para encontrar (10) como solução, escreva a equação diferencial como $\dot{y} - a(t)y = b(t)$ e multplique cada termo por $exp\left(-\int^t a(s)ds\right)$

$$\dot{y}(t)e^{-\int^t a(s)ds} - a(t)y(t)e^{-\int^t a(s)ds} = b(t)e^{-\int^t a(s)ds} \quad (11)$$

Como o lado esquerdo de (11) é precisamente a derivada da expressão $y(t)exp\left(-\int^t a(s)ds\right)$então (11) pode ser escrita como:

$$\frac{d}{dt}\left(y(t)e^{-\int^t a(s)ds}\right) = b(t)e^{-\int^t a(s)d(s)} \quad (12)$$

Integrando ambos os lados de (12) e multiplicando por $e^{\int^t a}$ para obtermos (10). Chamamos a expressão $e^{-\int^t a(s)ds}$de fator integrante.

Definição. Uma equação diferencial $\dot{y} = F(y,t)$ é chamada separável se $F(y,t)$ puder ser escrita como o produto: $F(y,t) = g(y)h(t)$.

Exemplos:

$$\dot{y} = y^2 + t^2;\ \dot{y} = a(t) + b(t)\ e\ \dot{y} = ty + t^2y^2$$

Solução geral da equação separável:

$$\dot{y} = g(y)h(t)$$

$$\frac{dy}{dt} = g(y)h(t)$$

$$\frac{d(y)}{g(y)} = h(t)dt$$

$$\int \frac{dy}{g(y)} = \int h(t)dt$$

Outro:

$$\dot{y} = t^2y$$
$$\frac{dy}{dt} = yt^2$$
$$\int \frac{dy}{y} = \int t^2dt$$

$$ln(y) = \frac{t^3}{3} + c$$
$$y = e^{\frac{t^3}{3}+c}$$

22.2 Equações lineares de segunda ordem

Nos concentraremos em equações lineares de segunda ordem que são homogêneas e possuem coeficientes constantes. Escreveremos essas equações do seguinte modo:

$$a\ddot{y} + b\dot{y} + cy + 0 \quad (23)$$

Se $a = 0 \rightarrow y = e^{rt}$

Para encontrearmos soluções, plugue $y = e^{rt}, \dot{y} = re^{rt}$ e $\ddot{y} = r^2e^{rt}$e teremos:

$$ar^2e^{rt} + bre^{rt} + ce^{rt} = 0$$

$$e^{rt}(ar^2 + br + c) = 0$$

Como e^{rt} é não nulo, então $y = e^{rt}$ é uma solução de (23) se, e somente se, r satisfaz a equação:

$$ar^2 + br + c = 0$$

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

(1) $\triangle > 0 \rightarrow 2$ raízes reais

(2) $\triangle < 0 \rightarrow 2$ raízes complexas

(3) $\triangle = 0 \rightarrow 1$ raíz

Raízes reais distintas da equação característica

Teorema. *Se o polinômio característico* (24) *da equação diferencial de segunda ordem* (23) *possui duas raízes reais distintas* r_1*e* r_2*, então a solução geral de* (23) *é* $y(t) = k_1e^{r_1t} + k_2e^{r_2t}$

Demonstração. Como r_1e r_2 são as raízes do polinômio característico de $ar^2 + br + c = 0$, saberemos que ambas $y(t) = k_1e^{r_1t}$ e $y(t) = k_2e^{r_2t}$ são soluções da equação diferencial $a\ddot{y} + by + cy = 0$ (23). Uma consequência advinda da linearidade dessa equação é que se $y_1(t)$ e $y_2(t)$ são duas soluções de (23) então a soma $y_1(t) + y_2(t)$ aimda é uma solução de (23), já que

$$\Big(a(y_1 + y_{2)}\Big)^{\cdot\cdot} + b(y_1 + y_2)^{\cdot} + c(y_1 + y_2) = ((a\ddot{y} + b\dot{y}_1 + cy_1) + (a\ddot{y}_2 + b\,\dot{}_1 + cy_2) \dot{=} 0 + 0 = 0$$

$$y(t) = k_1e^{r_1t} + k_2e^{r_2t} \quad (26)$$

$\square$

Portanto (26) é uma solução de (23). Finalmente, mostraremos que (26) é a solução geral, ou seja:

$$(a\ddot{y} + b\dot{y} + cy) = 0$$

$$y(t_0) = y_0$$

$$\dot{y}(t_o) = z_0 \ (27)$$

Dado qualquer problema de valor inicial existe uma única escolha de constantes de integração k_1^* e k_2^* tais que $y(t) = k_1^* e^{r_1 t} + k_2^* e^{r_2 t}$ é uma solução de (27). Para provar, considere $t_0 = 0$ e substitua o valor inicial (27) na solução (26)

$$y_1(0) = k_1 e^{r_1 .0} + k_2 e^{r_2 .0} = k_1 + k_2$$

$$y_2(0) = r_1 k_1^{r_1 .0} + r_2 k_2^{r_2 .0} = r_1 k_1 + r_2 k_2$$

o que leva à equação matricial

$$\begin{pmatrix} 1 & 1 \\ r_1 & r_2 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}$$

Como $r_1 \neq r_2$a matriz de coeficientes é não singular, dados quaisquer y_0 e z_0. Esse sistema possui uma única solução para k_1 e k_2. Isso prova que, dado qualquer problema de valor inicial (27), podemos mostrar k_1e k_2 tais que a solução (26) de (23) satisfaz as condições iniciais. Portanto, (26) é Solução geral de (23).

Exemplos:

$$\ddot{y} - y = 0 \ y(0) = \dot{y}(0) = 1$$

$$a\ddot{y} - cy = 0$$

$$r^2 - 1 = 0$$

$$r = \pm 1$$

$$y(t) = k_1 e^t + k_2 e^{-t}$$

$$y(0) = 1 \longrightarrow 1 = k_1e^0 + k_2e^{-0} \ (1)$$

$$\dot{y}(0) = 1 \longrightarrow 1 = k_1e^0 + k_2e^{-0}(2)$$

Somando (1) e (2)

$$2 = 2k_1$$

$$1 = k_1 \ (3)$$

Inserindo (3) em (2)

$$1 = 1 - k_2$$

$$k_2 = 0 \ (4)$$

A solução geral é:

$$y(t) = k_1e^{t.r_1} + k_2e^{-r_2} \longrightarrow y(t) = 1e^t$$

b)

$$\ddot{y} - 5\dot{y} + 6y = 0$$

$$y(0) = 3 \rightarrow \quad 3 = k_1 + k_2 \ (1)$$

$$\dot{y}(0) = 7 \rightarrow \quad 7 = 3k_1 + 2k_2 \ (2)$$

$$a = 1$$

Multiplique (1) por 2 e subtraia de (2)

$$b = -5$$

$$1 = k_1(3)$$

$$c = 6$$

Inserindo k_1 em (3)

$$r^2 - 5r + 6 = 0$$

$$r = \frac{5 \pm \sqrt{25-24}}{2}$$

$$k_2 = 2$$

$$r_1 \to 3$$

$$r_2 \longrightarrow 2$$

$$y = e^{3t} + 2e^{2t} \text{ (solução geral)}$$

$$y(0) = 3$$

$$y(3) = k_1 e^{3t} + k_2 e^{2t}$$

$$\dot{y}(0) = 7$$

$$7 = 3k_1 e^{3t} + 2k_2 e^{2t}$$

22.2.1 Raízes reais e iguais

Teorema. *Se o polinômio característico da equação diferencial linear de segunda ordem possui raízes iguais então a solução geral é*

$$y(t) = k_1 e^{r_1 t} + k_2 t e^{r_1 t}$$

Exemplo. $\ddot{y} - 4\dot{y} + 4y = 0$ $y(0) = 2$; $\dot{y}(0) = 5$

$a = 1; \quad b = -4; \quad c = 4.$

$$y = k_1 e^{r_1 t} + k_2 t e^{r_1 t}$$

$$y = k_1 e^{2t} + k_2 t e^{2t}$$

$$y = e^{2t}(k_1 + tk_2)$$

$$\dot{y} = 2k_1e^{2t} + k_2e^{2t} + tk_2e^{2t}$$

$$r = \frac{4 \pm \sqrt{(-4)^2 - 4.1.4}}{2}$$

$$r = 2$$

$$\dot{y} = e^{2t}(2k_1 + k_2(1 + 2t))$$

$$y(0) = 2 \longrightarrow 2 = k_1$$

$$\dot{y}(0) = 5 \longrightarrow 5 = 2k_1 + k_2 \longrightarrow k_2 = 1$$

$$y = e^{2t}(2 + t) \longrightarrow \text{Solução}$$

Teorema. *Se o polinômio característico da equação diferencial linear de segunda ordem tem raízes complexas $\alpha \pm i\beta$, ou seja, se $b^2 - 4ac < 0$ então a solução geral é $y(t) = e^{at}(c_1 cos\beta t \pm c_2 sen\beta t)$.*

22.3 Equações não homogêneas de segunda ordem

$$a\ddot{y} + b\dot{y} + cy = g\,(t) \quad (38)$$

A função $g(t)$ representa, especialmente em problemas mecânicos uma força externa, sem a qual a equação é autônoma. Para encontrarmos a solução geral de uma equação não autônoma, basta acharmos a solução geral da equação homogênea e uma solução particular da equação não homogênea.

Teorema. *Seja $y_p(t)$ uma solução particular qualquer da equação diferencial não homegênea $a\ddot{y} + b\dot{y} + cy = g(t)$. Seja $k_1y_1(t) + k_2y_2(t)$ uma solução geral da equação homogênea $a\ddot{y} + b\dot{y} + cy = 0$ correspondente. Então, uma solução geral de (38) é $y(t) = k_1y_1(t) + k_2y_2(t) + y_p(t)$.*

22.3.1 Método dos coeficientes indeterminados

Nesse método procuramos uma solução particular de $a\ddot{y} + b\dot{y} + cy = g(t)$ que possua o mesmo formato de $g(t)$. Se $g(t)$ é uma função constante $g(t) = g(0)$ então uma solução particular da equação diferencial é a função constante $y(t) = g_0/c$.Se $g(t)$ é um polinômio t de grau j procuramos uma solução particular que se adeque a essa forma.

Exemplo. $\ddot{y} - 2\dot{y} - 3y = 9t^2(40)$

A solução geral é

$$y(t) = k_1e^{3t} + k_2e^{-t}$$

$$a = 1;b = -2;\ c = -3;$$

$$y(t) = k_1e^{3t} + k_2e^{-t}$$

$$r = \frac{2 \pm \sqrt{(-2)^2 - 4.1. - 3}}{2.1}$$

$$r = \frac{2 \pm \sqrt{16}}{2}$$

$$r_1 = 3$$

$$r_2 = -1$$

$$y_p(t) = At^2 + Bt + c$$

Derive o candidato e substitua em (40)

$$\dot{y}_p = 2At + B$$

$$\ddot{y}_p = 2A$$

$$9t^2 = \ddot{y}_p - 2\dot{y}_p - 3y_p$$

$$9t^2 = 2A - 2(2At + B) - 3(At^2 + Bt + c)$$

$$9t^2 = 2A - 4At - 2B - 3At^2 - 3Bt - 3c$$

$$9t^2 = t^2(-3A) + t(-4A - 3B) + 2A - 2B - 3C$$

Suponha que:

$$9t^2 + (0)t + (0) = (-3A)t^2 + (-4A - 3B)t + (2A - 2B - 3C)$$

Como os lados esquerdo e direito dessa equação são iguais para qualquer t, os coeficientes de cada potência t devem ser iguais:

$$9 = 3A \ (1)$$

$$0 = -4A - 3B \ (2)$$

$$0 = 2A - 2B - 3C \ (3)$$

Por (1) $A = -3$; Inserindo (1) em (2) $B = 4$. Usando essas duas informações em (3)

$$0 = -6 - 8 - 3c \longrightarrow c = -\frac{14}{3}$$

Portanto, uma solução particular é:

$$y_p(t) = -3t^2 + 4t - \frac{14}{3}$$

A solução geral:

$$y(t) = k_1 e^{3t} + k_2 e^{-t} - 3t^2 + 4t - \frac{14}{3}$$

Se um termo candidato da equação não homogênea for igual a um termo da solução geral da equação homogênea dizemos que o sistema está em ressonância. Em geral, multiplicamos o candidato natural $y_p(t)$ por t ou às vezes até por t^2, para encontrar um candidato com chances de ser uma solução particular da equação não homogênea.

22.3.2 Existência de soluções

Teorema. *Considere o problema de valor inicial $\dot{y} = f(t, y)$, $y(t_o) = y_o$ (42). Suponha que f é contínua no ponto (t_o, y_o). Então, existe uma função $I \to \mathbb{R}$ que é C^1 num intervalo aberto $I = (t_o - a,\, t_o + a)$ em torno de t_o e é tal que $y(t_o) = y_o$e $\dot{y}(t) = f(t, y(t))$ para cada $t \in I$, ou seja, $y(t)$ é uma solução do problema de valor incial (42). Além disso, se f é C^1 em $(t_o,\, y_o)$ então a solução $y(t)$ é única. Quaisquer duas soluções de (42) devem ser iguais entre si na intersecção de seus domínios.*

22.3.3 Retratos de Fase e equilíbrios em $\mathbb{R}^n$

O retrato de fase nos mostra como as soluções da equação diferencial evoluem ao longo do tempo. Vejamos o seguinte exemplo:

$$\dot{y} = y - y^3$$

Iremos verificar os valores de estado estacionário dessa equação diferencial, isto é, $\dot{y} = 0$

$$0 = y - y^3$$

$$0 = y - y^3$$

$$y = 0$$

$$y = \pm 1$$

Após encontrarmos as raízes (interceptos com o eixo x) veremos as taxas de crescimento e descrescimento da função $f(y) = y - y^3$, como segue:

$$f''(y) = 1 - 3y^2$$

$$f'(y) = 0 \to 1 - 3y^2 = 0 \to \frac{1}{3} = y^2 \to y = \pm\frac{1}{\sqrt{3}} \to \pm 0,58 = y$$

Gráfico 1

Vejamos $f'"(y)$:

$$f'"(y) = 0 \to -6y = 0 \to y = 0$$

Temos que zero é um ponto de inflexão. O próximo passo é fazermos o gráfico desta função:

Gráfico 2

Note que se $-1 < y < 0$ então $f(y)$ é negativa. Isso também ocorre no intervalo (1,∞). Nos intervalos ($-\infty$, -1) e (0,1) $f(y)$ é positiva. Para finalizarmos o gráfico devemos inserir as setas que dão a ideia de movimento ou convergência. Faremos isso passo a passo:

1. Se $y \in]-\infty,1]$ então $f(y) > 0$ desse modo, colocamos uma seta para a direita.

2. Se $y \in]-1,0]$ então $f(y) < 0$ e então colocamos uma seta para a esquerda.

3. Se $y \in]0,1[$, $f(y) > 0, \dot{y} > 0$, então colocamos uma reta para a direita.

4. Se $y \in]1,+\infty]$ então $f(y) < 0, \dot{y} < 0$ e adicionamos uma seta para a esquerda.

Agora podemos olhar para o nosso gráfico e verificarmos que -1 e $+1$ são dois equilíbrios estáveis no estado estacionário. O teorema abaixo resume e explica esse resultado:

Teorema. *Seja y_o um ponto de equilíbrio de uma equação diferencial $\dot{y} = f(y)$ que é C^1na reta, de modo que $f(y_o) = 0$. Se $f'(y_o) < 0$, então y_o é um equilíbrio assintoticamente estável. Se $f'(y_o) > 0$, então y_o é um equilíbrio estável.*

22.4 Modelo de Solow

$$\dot{k} = s.f(k) - (n+g+d).k$$

Lembrando que $\dot{k} = \dot{K}/LA$

$$f(k) = k^\alpha \to f(K,L) = K^\alpha (AL)^{1-\alpha} \to f(K,L)/AL = K^\alpha.(AL)^{1-\alpha}/AL = K^\alpha/AL^\alpha = K^\alpha$$

$$\dot{k} = 0$$

Note que:

$$0 = sk^\alpha - (n+g+d).k$$

$$(n+g+d).k = s.k^\alpha$$

$$k = (\frac{s}{n+g+d}).k^\alpha$$

se $k = 0 \to 0 = 0$

$$k^{\frac{1}{\alpha}} = (\frac{s}{n+g+d})^{\frac{1}{\alpha}} k$$

Chamaremos $F(k) = sk^\alpha - (n+d+g)k$

$$k^{\frac{1-\alpha}{\alpha}} = (\frac{s}{n+g+d})^{\frac{1}{\cancel{\alpha}}\cdot\frac{\cancel{\alpha}}{1-\alpha}}$$

seja $0 < \alpha < 1$ então se $\alpha s k^{\alpha-1} > (n+g+d)$

$$k* = (\frac{s}{n+g+d})^{\frac{1}{1-\alpha}}$$

$$F'(k) = \alpha s k^{\alpha-1} - (n+d+g) \to F'(k) > 0$$

$$F"(k) = \alpha(\alpha-1)sk^{\alpha-2} < 0$$

23 Equações Diferenciais Ordinárias: Sistemas de Equações

O sistema geral de duas equações diferenciais pode ser escrito como:

$$\begin{aligned} \dot{x} &= F(x,y,t) \\ \dot{y} &= G(x,y,t) \end{aligned} \tag{23}$$

Uma Solução de (1) é um par $x*(t)$ e $y*(t)$ de funções de t tais que, para cada t, valem:

$$\begin{aligned} \dot{x}(t)^* &= F(x^*(t), y^*(t), t) \\ \dot{y}(t)^* &= G(x^*(t), y^*(t), t) \end{aligned} \tag{24}$$

O sistema (3) é um sistema de 1ª ordem.

Se F e G não dependem explicitamente de t dizemos que o sistema é autônomo ou tempo independente. Caso contrário, o sistema é não autônomo. O modo usado para encontrarmos uma solução geral e uma particular, por exemplo, é análogo ao que fizemos anteriormente.

Fato. *Todo o sistema de segunda ordem a uma variável pode ser escrito naturalmente como um sistema de primeira ordem a duas variáveis.*

Exemplo. $v = \dot{y}$, $\dot{v} = f(v,y,t)$ e $\ddot{y} = f(\dot{y},y,t)$

Definição. O par posição/velocidade $(y, \dot{y})$ é denominado <u>variável de estado</u>.

Definição. O conjunto de todas essas possíveis variáveis de estado é denominado espaço de estados.

Fato. *Toda equação diferencial não-autônoma* $\dot{y} = f(y,)$ *em* y *pode ser escrita como um sistema autônomo de duas equações diferenciais em* (y, r)*:*

$$\dot{y} = f(y, r)$$
$$\dot{r} = 1$$

Note que: $\frac{dr}{dt} = 1 \rightarrow \int dr = \int dt \rightarrow r(t) = t + c$

Teorema. *Existência e Unicidade de Soluções: Se* F *e* G *são funções contínuas numa vizinhança de* (x_0, y_0, t_0) *então existem funções* $x^*(t)$ *e* $y^*(t)$*, definidas e contínuas num intervalo aberto* $I = (t_0 - \varepsilon, t_0 + \varepsilon)$ *em torno de* t_0*, tais que* $x(t_0) = x_0$ *e* $y(t_0) = y_0$ *e* $\dot{x}^*(t) = F(\dot{x}(t), \dot{y}(t), t)$ *e* $\dot{y}^*(t) = G(\dot{x}(t), \dot{y}(t), t)$ *vale para qualquer* t *e* I*. Além disso, se* F *e* G *são funções C¹, a solução do problema de valor incial é única.*

23.1 Sistemas Lineares por meio de Autovalores

O sistema geral de equações diferenciais lineares pode ser escrito como:

$$\begin{aligned} \dot{x} &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + \ldots + a_{nn}x_n \end{aligned} \tag{25}$$

Ou simplesmente como $\dot{x} = Ax$. Se A é uma matriz diagonal, ou seja, se $a_{ij} = 0$ para $i \neq j$ em (3), então (3) separa em n equações independentes $\dot{x}_i = a_{ii}x_i$ entre as quais não há interações. Nesse caso o sistema pode ser resolvido como:

$$x_1(t) = c_1 e^{a_{11}t}, ..., x_n(t) = c_n e^{a_{nn}t}$$

Se algum dos a_{ij} fora da diagonal não é nulo, de modo que as equações são relacionadas entre si, podemos usar autovalores de A para transformar o sistema (3) de n equações mais ou menos independentes, exatamente como procedemos com sistemas de equações a diferenças lineares.

23.1.1 Autovalores Reais Distintos

Suponha que A possui n autovalores reais distintos $r_1, ..., r_n$ com autovetores associados $v_1, ..., v_n$:

$$Av_i = r_i v_i \tag{26}$$

Seja P a matrix $n \times n$ cujas colunas são estes n autovetores:

$$P = [v_1 v_2 ... v_n] \tag{27}$$

Então, podemos escrever as equações (4) como:
$AP = PD$ onde:

$$D \equiv \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{pmatrix} \tag{28}$$

Como autovetores associados a autovalores distintos são linearmente independentes, P é não singular e, portanto, invertível; podemos escrever:

$$P^{-1}AP = D \tag{29}$$

Usando mudança linear de coordenadas $y = P^{-1}x$, com inversa $x = Py$, para transformar o sistema (3) num sistema nas variáveis $y_1, ..., yn$:

$$\begin{aligned} &\dot{y} = P^{-1}\dot{x} \\ &\dot{y} = P^{-1}Ax \ (pois\ \dot{x} = Ax) \\ &\dot{y} = P^{-1}Ay \ (pois\ x = Py) \\ &\dot{y} = Dy \end{aligned}$$

Como D é uma matriz diagonal, o sistema $\dot{y} = Dy$ pode ser escrito como:

$$\dot{y}_1 = r_1 y_1, ..., \dot{y}_n = r_n y_n$$

Sua solução:

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} c_1 e^{r_1 t} \\ \vdots \\ c_n e^{r_n t} \end{pmatrix}$$

$$Y = \beta_0 + \beta_1 X + u$$

Finalmente, utilize a transformação $x = Py$ para voltar às coordenadas originais $x_1, ..., x_n$:

$$\begin{aligned} x(t) &= P(y(t)) \\ x(t) &= [v_1, ..., v_n] \begin{pmatrix} c_1 e^{r_1 t} \\ \vdots \\ c_n e^{r_n t} \end{pmatrix} \\ x(t) &= c_1 e^{r_1 t} + \ldots + c_n e^{r_n t} \end{aligned}$$

Teorema. *Seja A uma matriz de tamanho $n \times n$ com n autovalores reais distintos $r_1, ..., r_n$ e autovetores associados $v_1, ..., v_n$. Então, a solução geral do sistema linear $\dot{x} = Ax$ de equações diferenciais é dada por:*

$$x(t) = c_1 e^{r_1 t} v_1 + \ldots + c_n e^{r_n t} v_n$$

Teorema. *Seja A uma matriz real 2×2 com autovalores complexos $\alpha \pm i\beta$ e autovetores associados $u \pm iw$. Então a solução geral do sistema linear de equações diferenciais $\dot{x} = Ax$ é dada por*

$$x(t) = e^{\alpha t} cos\beta t(c_1 u - c_2 w) - e^{\alpha t} sen\beta t(c_2 u - c_1 w)$$

Teorema. *Seja A uma matriz de tamanho 2×2 com autovalores iguais $r_1 = r_2 = r$ e somente um autovetor independente v. Seja w um autovetor generalizado para A. Então, a solução geral do sistema linear de equações diferenciais $\dot{x} = Ax$ é dada por:*

$$x(t) = e^{rt}(c_1 v + c_2 w) + te^{rt}(c_2 v)$$

23.1.2 Resolvendo Sistemas por Substituição

Essa técnica não utiliza autovalores ou autovetores. Escreveremos um sistema de primeira ordem a duas variáveis como um sistema de segunda ordem a uma variável. Vejamos o exemplo a seguir:

$$\begin{aligned}
\dot{y}_1 &= a_{11}y_1 + a_{12}y_2 \\
\dot{y}_2 &= a_{21}y_2 + a_{22}y_2
\end{aligned}$$

Veja que os parâmetros do sistema são arbitrários, entao faremos uma escolha apenas para fins ilustrativos.

$$\dot{y}_1 = y_1 + 4y_2 \tag{30}$$
$$\dot{y}_2 = y_1 + y_2 \tag{31}$$

$$y_2 = \frac{1}{4}(\dot{y}_1 - y_1) \rightarrow \textit{equação } (8)$$

Derive essa expressão em relação ao tempo:

$$\dot{y}_2 = \frac{1}{4}(\ddot{y}_1 - \dot{y}_1)$$

Inserindo essa expressão em (9)

$$\frac{1}{4}(\ddot{y}_1 - \dot{y}_1) = y_1 + y_2$$

$$\frac{1}{4}(\ddot{y}_1 - \dot{y}_1) = y_1 + \frac{1}{4}(\dot{y}_1 - y_1)$$

$$\frac{1}{4}(\ddot{y}_1 - \dot{y}_1) - \frac{1}{4}(\dot{y}_1 - y_1) - y_1 = 0$$

$$\ddot{y}_1 - \dot{y}_1 - \dot{y}_1 + y_1 - 4y_1 = 0$$

$$\ddot{y}_1 - 2\dot{y} - 3y_1 = 0 \tag{32}$$

Encontrando o polinômio característico:

$$r^2 - 2r - 3 = 0$$

$$r = \frac{2 \pm \sqrt{(-2)^2 - 4.1.(-3)}}{2.1}$$

$$r = \frac{2 \pm \sqrt{16}}{2}$$

$$r_1 = 3 \ ; \ r_2 = -1$$

A solução geral para y_1 é : $y_1(t) = c_1 e^{3t} + c_2 e^{-t}$

Derive em relação a t : $\dot{y}_1 = 3c_1 e^{3t} - c_2 e^{-t}$

Substitua essas duas equações em (9)

$$y_2 = \frac{1}{4}(\dot{y}_1 - y_1)$$

$$y_2 = \frac{1}{4}(3c_1 e^{3t} - c_2 e^{-t} - c_1 e^{3t} - c_2 e^{-t}$$

$$y_2 = \frac{1}{4}(2c_1 e^{3t} - 2c_2 e^{-t})$$

$$y_2 = \frac{1}{2}(c_1 e^{3t} - c_2 e^{-t})$$

Em notação matricial:

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = c_1 e^{3t} \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 1 \\ -1/2 \end{pmatrix}$$

Definição. Uma solução estacionária y^* do sistema $\dot{y} = F(y)$ é dita globalmente assintóticamente estável se praticamente qualquer solução $\dot{y} = F(y)$ tende a y^* quando $t \to \infty$. Mais precisamente, o equilíbrio y^* é globalmente assintoticamente estável se para qualquer y_0(exceto possivelmente para um conjunto de dimensão inferior de y_0) a solução do problema de valor inicial $\dot{y} = F(y)$, $y(0) = y_0$ tende a y^* quando $t \to \infty$.

Definição. Uma solução estacionária $y*$do sistema $\dot{y} = F(y)$ é dita neutramente estável se não é localmente assintóticamente estável e se todas as soluções que começam perto $y*$ permanecem perto de $y*$quando $t \to \infty$.

Definição. Se um equilíbrio $y*$ de $\dot{y} = F(y)$ é assintoticamente estável ou neutramente estável, dizemos que é estável.

Teorema. *A solução constante $x = 0$ é sempre um estado estacionário do sistema linear de equações diferenciais (3): $\dot{x} = Ax$*

a) Se cada autovalor real de A é negativo e cada autovalor complexo de A tem parte real negativa, então $x = 0$ é um estado estacionário globalmente assintoticamente estável: qualquer solução tende a zero quanto $t \to \infty$.

b) Se A tem um autovalor real positivo ou um autovalor complexo com parte real positiva, então $x = 0$ é um estado estacionário instável: praticamente qualquer solução afasta-se da origem quando $t \to \infty$.

c) Se A tem um autovalor que não tem um conjunto completo de autovetores independentes e que é nulo ou puramente imaginário, então $x = 0$ é um estado estacionário *instável.*

d) Se A tem autovalores reais iguais a zero ou autovalores complexos puramente imaginários, tais que cada um desses autovalores tem um conjunto completo de autovetores independentes, e se todos os outros autovalores têm parte real negativa, então a origem é um estado estacionário neutramente estável do sistema.

$$\begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

23.1.3 Estabilidade de Sistemas Lineares

Critérios de estabilidade para um sistema de equações diferenciais de primeira ordem:

(1) Assintoticamente estável se $f(y^*) = 0$ e $f'(y^*) < 0$, e
(2) instável se $f(y^*) = 0$e $f'(y^*) > 0$

Teorema. *Seja y^* um estado estacionário do sistema de equações diferenciais $\dot{y} = F(y)$ de primeira ordem em $\mathbb{R}^n$, onde F é uma função C^1 de $\mathbb{R}^n$ em $\mathbb{R}^n$.*

(a) Se cada autovalor da matriz jacobiana $DF(y^*)$ de F em y^* é negativo ou tem parte real negativa, então y^* é um estado estacionário assintoticamente estável de $\dot{y} = F(y)$.

(b) Se $DF(y^*)$ tem pelo menos um autovalor real positivo ou um autovalor complexo com parte real positiva, então y^* é um estado estacionário instável de $\dot{y} = F(x)$.

Se o teste do teorema 6 não for aplicável, ou seja, se $DF(y^*)$ tiver algum autovalor puramente imaginário ou um autovalor nulo, mas nenhum autovalor positivo e nenhum autovalor com parte real positiva, então não podemos usar a matriz jacobiana em y^* para determinar a estabilidade de y^*.

Definição. Estabilidade Assintótica. Escreva $x(t; y_0)$ para a solução $x(t)$ do sistema autônomo de equações diferenenciais $\dot{x} = f(x)$ que satisfaz a condição inicial $x(0) = y_0$. Um estado estacionário x^* desse sistema é assintoticamente estável se existe $\varepsilon > 0$ tal que para qualquer y_0 com $\parallel y_0 - x_0 \parallel < \varepsilon$, $x(t; y_0) \to x^*$ quando $t \to \infty$.

Considere o seguinte sistema linear

$$\begin{aligned}\dot{y}_1 &= a_{11}y_1 + a_{12}y_2 \\ \dot{y}_2 &= a_{21}y_1 + a_{22}y_2\end{aligned}$$

Esse sistema pode ser expresso como:

$\dot{x} = Ax$ em $\mathbb{R}^2$

Vamos mostrar que $(0, 0)$ é um estado estacionário assintoticamente estável se, e somente se, o traço de A é negativo e o determinante de A é positivo.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} - r & a_{12} \\ a_{21} & a_{22} - r \end{pmatrix}$$

Então o polinômio característico de A:

$$(a_{11} - r)(a_{22} - r) - (a_{12}a_{21})$$

$$a_{22}a_{11} - ra_{11} - ra_{22} + r^2 - a_{12}a_{21}$$

$$r^2 - r(a_{11} + a_{22}) + (a_{22}a_{11} - a_{12}a_{21})$$

$$r^2 - r(a_{11} + a_{22}) + detA = 0$$

A soma dos autovalores $(a_{11} + a_{22})$ é igual ao traço de A. O produto dos autovalores é igual ao determinante de A. Note que se os autovalores de A são ambos negativos sua soma resulta em um número < 0 e seu produto é > 0. Se os autovalores são complexos $a \pm ib$ com $a < 0$, então sua soma $2a$ é negativa e seu produto $a^2 + b^2$é positivo. Por outro lado, se seu produto é positivo e sua soma é negativa os autovalores são números complexos com uma parte real negativa ou eles são números reais negativos.

Se $detA < 0$, os autovalores possuem sinais opostos; ou algum deles é positivo. Se o traço de $A > 0$, os autovalores (um deles) deve ser positivo, ou ambos podem ser complexos com a parte real positiva. Finalmente, se $detA > 0$ e o $traçoA = 0$ os autovalores devem ser números complexos com a parte real igual a zero, isto é, são puramente números imaginários $\pm ib$. A solução geral nesse caso é $x = (cosbt)u + (senbt)v$. Nesse caso, todas as órbitas são periódicas e $(0, 0)$ é uma solução estável.

Exemplo. .
$\dot{x} = 2xy - 2y^2$
$\dot{y} = x + y^2 - 2$

Encontrando os pontos no estado estacionário

$$\dot{x} = 0$$
$$0 = 2y(x - y)$$

Então, $y = 0$ ou $x = y$

$$\dot{y} = 0$$
$$x = -y^2 + 2$$

Se $y = 0$ então $x = 2$

Se $y = x$ então

$$y = -y^2 + 2$$

$$0 = -y^2 - y + 2$$

$$y = \frac{1 \pm \sqrt{(-1)^2 - 4.(-1)2}}{2(-1)}$$

$y = 1$ ou $y = -2$

Temos as seguintes pontos de estado estacionário: $(1, 1)$, $(-2, -2)$ e $(2, 0)$
Montando a matriz jacobiana:

$$D(f, g)(x, y) = \begin{pmatrix} 2y & 2x - 4y \\ 1 & -4y \end{pmatrix}$$

Em $(2, 0)$:

$$D(f, g)(2, 0) = \begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}$$

As raízes do polinômio característico alternam de sinal, isto é, $r_1 = 2$ e $r_2 = -2$ logo $(2, 0)$ é um equilbrio instável.

Em $(1, 2)$:

$D(f, g)(1, 2) = \begin{pmatrix} 2 & -2 \\ 1 & -4 \end{pmatrix}$ As raízes são $-1, 65$ e $3, 65$. Então $(1, 2)$ é um equilíbrio instável.

Em $(-2, -2)$:

$$D(f, g)(-2, -2) = \begin{pmatrix} -4 & 4 \\ 1 & 8 \end{pmatrix}$$

As raízes são $8, 32$ e $-4, 3$, ou seja, mais um equilíbrio instável.

23.2 Retratos de fase de sistemas planares

Suponha que tenhamos o seguinte sistema planar:

$$\dot{x} = f(x, y)$$

$$\dot{y} = g(x, y)$$

A vizualização geométrica desse sistema é um vetor:

$$(\dot{x}, \dot{y}) = (f(x, y), g(x, y))$$

Em cada ponto (x, y) a vizualização geométrica correspondente de uma solução $(x^*(t), y^*(t))$ é uma curva no plano tangente em cada ponto ao campo de vetores.

Definição. Diagrama de fase. Quando a curva $(x^*(t), y^*(t))$ passa pelo ponto (x, y), seu vetor tangente $(\dot{x}^*(t), \dot{y}^*(t))$ deveria apontar na direção e sentido de $(f(x, y), g(x, y))$. O conjunto de todas essas curvas é denominado retrato de fase ou diagrama de fase.

Exemplo. Veja o sistema desacoplado

$$\dot{x} = 2x$$

$$\dot{y} = -2y$$

$$\begin{array}{c|c}
\frac{dx}{dt} = 2x & \dot{y} = -2y \\
\frac{dx}{x} = 2dt & \frac{dy}{dt} = -2y \\
\int \frac{dx}{x} = 2\int dt & \int \frac{dy}{y} = -2\int dt \\
lnx = 2t + c & lny = -2t + c \\
x = c_1 e^{2t} & y = c_2 e^{-2t}
\end{array}$$

Para encontrarmos as soluções não parametrizadas resolvemos cada equação por e^{2t} e então eliminamos t igualando as expressões:

$$\frac{x}{c_1} = e^{2t} \text{ e } e^{2t} = \frac{c_2}{y}$$

$$\frac{x}{c_1} = \frac{c_2}{y}$$

$$y = c_1 c_2 x^{-1}$$

Podemos expressar algumas dessas hipérboles $y = \frac{K}{x}$
[Gráfico]

23.3 Retratos de fase sistemas lineares

Faremos o diagrama de fase para o seguinte sistema:

$$\dot{x} = x(4 - x - y)$$

$$\dot{y} = y(6 - y - 3x)$$

Passo 1. Encontre os valores de equilíbrio igualando o lado direito de cada equação a zero. A soluções são $(0,0)$, $(0,6)$, $(4,0)$ e $(1,3)$.

Passo 2. Use o teorema 25.5 para determinar a estabilidade local de cada um desses equilíbrios. Note que

$(0,0)$ e $(1,3)$ são instáveis;

e $(0,6)$ e $(4,0)$ são estáveis.

Passo 3. Esboce as isóclinas

Um vetor $(\dot{x},\dot{y})$é vertical se $\dot{x} = 0$ e $\dot{y} \neq 0$.Se $\dot{y} > 0$ o vetor aponta para cima se $\dot{y} < 0$ aponta para baixo. Como $\dot{x} = f(x,y)$, o conjunto dos pontos (x,y) nas quais $\dot{x} = 0$ é dado pela curva $f(x,y) = 0$.Uma tal curva ao longo da qual o campo de vetores sempre aponta na mesma direção e sentido, é denominada isóclina do sistema.

$\dot{x} = 0 \rightarrow x = 0$ e $4 - x - y = 0$

[**Gráfico**]

Vejamos para $\dot{y} = 0 \rightarrow y = 0$ e $6 - y - 3x = 0$

[**Gráfico**]

Passo 4. Complete as setas nas isóclinas e nos setores entre as isóclinas. As isóclinas dividem o plano em regiões denominadas setores, e constituem as fronteiras desses setores.

[**Gráfico**]

Seja $\quad \dot{x} = -4x + 16$ e $\dot{y} = -5y + 15$

Se $\dot{x} = 0$ então $x = 4$

Se $\dot{y} = 0$ então $y = 3$

Trace as isóclinas e divida o quadrante positivo em 4 setores:

$(x^*, y^*) = (4,3)$

Vamos determinar o movimento entre os setores:

Se $x < 4 \rightarrow \dot{x} > 0$ (seta direita)

Se $x > 4 \rightarrow \dot{x} < 0$ (seta esquerda)

Se $y > 3$, $\dot{y} < 0$ (seta para baixo)

Se $y < 3$, $\dot{y} > 0$ (seta para cima)

[**Gráfico** ao lado]

Note que o ponto $(4, 3)$ é um equilíbrio estável.

24 Introdução a otimização dinâmica

Esse capítulo tem por objetivo motivar o leitor ao estudo de modelos dinâmicos. Para isso usamos um modelo bastante simples da determinação da renda, como segue:

$$C = cY \tag{33}$$

$$C + I = Y \tag{34}$$

A renda de pleno emprego Y depende do nível de estoque de capital, s que se dá por meio de uma função de produção :

$$Y \leq Y = f(s) \tag{35}$$

Suponha que $I = (1 - c)Y$ (no pleno emprego).

Para obtermos um modelo de crescimento formalmente igualamos a taxa líquida de mudança no estoque de capital ao investimento:

$$\dot{s} = I \tag{36}$$

Usando as equações anteriores podemos escrever uma nova versão de (4):

$$I = (1 - c)Y$$

$$Y = f(s)$$

$$\mathring{s} = (1 - c)f(s) \tag{37}$$

Essa equação diferencial pode ser solucionada para uma forma específica de $f(s)$ e uma dada condição inicial $s(0) = s_0$.

24.1 Empréstimo Ótimo

Consideraremos a possibilidade de empréstimo o que permite o aumento da quantidade inicial do estoque de capital. considere L como o tamanho do empréstimo, esse empréstimo é continuamente composto a taxa de juros r. Podemos dizer que Le^{rT}. Temos que $Y = \dfrac{s^{\alpha}}{1-\alpha}$, $\alpha \in (0,1)$ e a condição inicial $s(0) = s_0$. Considere que o preço do capital é $1 \forall t$, desse modo podemos começar com $s_0 + L$ unidades de capital. Lé otimamente escolhido para que possamos maximizar o montante do capital disponível ao tempo T, após o empréstimo ter (ser sido) repago (pago novamente). Note que a equação diferencial (5) toma a seguinte forma:

$$\dot{s} = (1-c)\frac{s^{\alpha}}{(1-\alpha)} \tag{38}$$

$$\int \frac{\partial s}{\partial t} dt = \int \frac{(1-c)}{(1-\alpha)} s^{\alpha} dt$$

$$\int \frac{\partial s}{s^{\alpha}(1-\alpha)^{-1}} = \int (1-c) dt$$

$$\int ds\, s^{-\alpha}(1-\alpha) = \int (1-c) dt$$

$$\frac{s^{-\alpha+1}}{1-\alpha}(1-\alpha) = (1-c)t + A$$

Lembre que $(1-c) \in (0,1)$

$$s^{1-\alpha} = (1-c)t + A \tag{39}$$

Se a quantidade L é tomada por empréstimo, teremos que:

$$(s_0 + L)^{1-\alpha} = A \qquad \rightarrow s(0) = s_0 \tag{40}$$

Assim teremos que:

$$s(t) = \left[(1-c)\,t + (s_0 + L)^{1-\alpha}\right]^{\frac{1}{1-\alpha}} \tag{41}$$

Essa equação é a solução para $t \in [0, T]$. Ao tempo T o empréstimo é pago e a recompensa é:

$$\left[(1-c)T + (s_0 + L)^{1-\alpha}\right]^{\frac{1}{1-\alpha}} - Le^{rT} \tag{42}$$

Desejamos escolher L^*que maximize (10)

$$\left[(1-c)T+(s_0+L)^{1-\alpha}\right]^{\frac{1}{1-\alpha}-1}\cdot\cancel{(1-\alpha)}\cdot\frac{1}{\cancel{(1-\alpha)}}\cdot(s_0+L)^{1-\alpha-1}-e^{rT}=0$$

$$\left[(1-c)T+(s_0+L)^{1-\alpha}\right]^{\frac{\alpha}{\alpha-1}}(s_0+L)^{-\alpha}=e^{rT}$$

$$\left[\frac{(1-c)T+(s_0+L)^{1-\alpha}}{(s_0+L)^{1-\alpha}}\right]=e^{rT\frac{(1-\alpha)}{\alpha}}$$

$$(1-c)T(s_0+L)^{\alpha-1}+1=e^{rT\frac{(1-\alpha)}{\alpha}}$$

$$(1-c)T(s_0+L)^{\alpha-1}=e^{rT\frac{(1-\alpha)}{\alpha}}-1$$

$$(s_0+L)=[(1-c)T]^{\frac{1}{1-\alpha}}(e^{\frac{-rT}{\alpha}}-1)$$

obs.: $1^{\frac{1}{1-\alpha}}$ ou $1^{\frac{1}{\alpha-1}}$ se $\alpha\in(0,1)$ $\alpha-1\prec 0$

$$L^*=\frac{[(1-c)T]^{\frac{1}{1-\alpha}}}{e^{\frac{rT}{\alpha}}-1}-s_0 \tag{43}$$

Substituindo (11) em (10)

$$s^*(T)=e^{rT}\left[s_0+[(1-c)T]^{1-\alpha}[e^{rT\frac{(1-\alpha)}{\alpha}}-1]^{\frac{-\alpha}{(1-\alpha}}\right]$$

24.2 Política Fiscal

Suponha que possamos adicionar o governo no modelo anterior

$$C=c(1-\theta)Y \tag{44}$$

$$G=\theta Y \tag{45}$$

$$Y=s^{\alpha}q/(1-\alpha) \tag{46}$$

$$C+I+G=Y \tag{47}$$

$$\dot{s}=I+G \tag{48}$$

θ é o imposto de renda, $(1-\theta)\,Y$ e a renda disponível e G é a receita do governo ou investimento público. Temos que I é o investimento privado . Por simplicidade tome

$\alpha = 0,5.$

$$I = Y - G - C$$

$$I + G = Y - C$$

$$I + G = [1 - c(1 - \theta)]\,Y$$

$$\mathring{s} = I + G$$

$$\mathring{s} = [1 - c(1 - \theta)]\,\frac{s^{0,5}}{0,5}$$

$$\mathring{s} = 2s^{\frac{1}{2}}\,[1 - c(1 - \theta)]$$

$$\frac{\partial s}{\partial t} = 2s^{\frac{1}{2}} = 2\,[1 - c(1 - \theta)]$$

$$\frac{\partial s}{dt} s^{\frac{-1}{2}} = 2\,[1 - c(1 - \theta)]$$

$$\int \frac{\partial s}{\partial t} s^{\frac{-1}{2}} dt = 2\,[1 - c(1 - \theta)] \int dt$$

$$\frac{s^{\frac{1}{2}}}{\frac{1}{2}} = 2\,[1 - c(1 - \theta)]\,t + c$$

para $s(0) = 0$

$$s_o = \left(\frac{c}{2}\right)^2$$

$$s_0^{\frac{1}{2}} = \frac{c}{2}$$

$$s(t) = \left[(1 - c + c\theta)t + s_0^{\frac{1}{2}}\right]^2$$

Suponha que o governo deseja maximizar o consumo no horizonte e tempo [0,T]. Em outras palavras, o governo dece escolher θ tal que:

$$W = \int_0^T C(t)\partial t = \int_0^T 2c(1 - \theta)\left[(1 - c + c\theta)\,t + s_0^{\frac{1}{2}}\right] dt$$

Note que$\frac{\partial s}{\partial \theta} = 2\left[\left((1-c)+c\theta\right)t + s_0^{\frac{1}{2}}\right].ct$

$$C = c(1-\theta)2\underline{s}^{0,5}$$

$$C = c(1-\theta)2\cdot\left[1 + c\left(\theta - 1\right)t + s_0^{\frac{1}{2}}\right] \rightarrow \text{esse é o termo dentro da integral}$$

$$W = 2c(1-\theta)\left[1 - c(1-\theta)\frac{T^2}{2} + s_0^{\frac{1}{2}}T\right]$$

$$W = c(1-\theta)\left[1 - c(1-\theta)T^2 + 2s_0^{\frac{1}{2}}T\right]$$

$$= -\theta\left[1 - c(1-\theta)T^2 + 2s_0^{\frac{1}{2}}T\right] + c\left[...\right]$$

$$= c\left[-\theta T^2 + (\theta c - c\theta^2)T^2 - \theta 2s_0^{\frac{1}{2}}T\right] + c\left[...\right]$$

$$\frac{\partial W}{\partial \theta} = c\left[-1T^2 + cT^2 - 2c\theta T^2 - 2s_0^{\frac{1}{2}}T\right] = 0$$

$$-1T^2 + cT^2 - 2s_0^{\frac{1}{2}}T^2 = 2c\theta T^2$$

$$\frac{-1T^2 + cT^2 - 2s_0^{\frac{1}{2}}T^2}{2cT^2} = \theta$$

$$\theta = 1 - \left(\frac{2s_0^{\frac{1}{2}}T^{-1} + 1}{2c}\right)$$

$$\theta = 1 - \frac{2s_0^{\frac{1}{2}}T^{-1} + 1}{2c}$$

Qual a influência da duração do mandato do governo no parâmetro de política fiscal? Note que θaumenta quando T cresce. Atribuímos $s_0 = 1$ e $c = 0,75$ e deixamos T assumir uma gaama de valores.

Temos que:

$$\theta = 1 - \frac{2+T}{2\cdot 0,75T}$$

$$\theta = 1 - \frac{2+T}{1,5T}$$

$$\theta = 1 - \frac{4+2T}{3T}$$

$$\theta = \frac{3T - 4 - 2T}{3T}$$

$$\theta = \frac{T - 4}{3T}$$

$$Y(t) = 2\left[(1 - c + c\theta)t + s_0^{\frac{1}{2}}\right] = t(T-2)/T + 2$$

$$\mathring{s}(1 - c + c\theta)Y = Y(T-2)/2T$$

Para um caso extremo note que se $T \to \infty$, $\theta \to \frac{1}{3}$ e $Y(t) = t + 2$.Isto indica uma alta taxa de imposto e um rápido crescimento da renda. As poupanças decaem em favor do investimento em capital que ira resultar numa renda superior e assim propiciará um consumo mais elevado no futuro.

24.3 Caminho subótimo de consumo

Voltamos a um modelo sem governo e escolhemos $f(s) = 4s$; e $s(0) = 1$. Então a equação (5) assume a seguinte forma:

$$\mathring{s} = 4(1 - s)s \quad s(0) = 1$$

Nosso objetivo é maximizar a utilidade em algum horizonte temporal $\int_0^T U\,(c(t))\,dt$. Tomando $U\,(c(t)) = lnC(t)$ e T=1 devemos maximizar:

$$V = \int_0^1 lnC(t)dt$$

$$s.a. \quad C(t) = 4s(t).c$$

Podemos solucionar $\mathring{s} = 4(1 - c)s$

$$\frac{\mathring{s}}{s} = 4(1 - c)$$

$$\int \frac{\partial s}{\partial t} \cdot \frac{1}{s} \cdot dt = 4(1 - c) \int dt$$

$$ln\, s = 4(1 - c)t + constante$$

$$s = e^{4(1-c)t} + constante$$

Ignorando a constante temos que:

$$C(t) = 4s \cdot c$$

$$C(t) = 4c \cdot e^{4(1-c)t}$$

$$V = \int_0^1 \left[ln\,4c + 4(1-c)t \right] dt$$

$$V = \left[ln4ct + 2(1-c)t^2 \right]_0^1 = ln4c + 2(1-c)$$

$$\frac{\partial V}{\partial c} = \frac{1}{\cancel{4}c} \cdot \cancel{4} - 2 = 0 \quad \rightarrow c^* = \frac{1}{2}$$

Então teremos que $s^* = e^{4(1-\frac{1}{2})t} = e^{2t}$ e $C^*(t) = 2e^{2t}$

$$V^* = ln\,4 \cdot \frac{1}{2} + 2\left(1 - \frac{1}{2}\right) \cong 1,7$$

25 Princípio do Máximo

Esse conteúdo foi desenvolvido por Pontryagin *et al* (1962). Começaremos estudando os casos mais simples para esse problema, sem nos preocuparmos com algumas questões relacionados a regularidade. Começaremos apresentando um problema simples de controle ótimo e usaremos o princípio do máximo para solucioná-lo:

$$V = \int_0^T V\left(s(t), c(t), t\right) dt \tag{49}$$

$$s.a$$

$$\dot{s} = f\left(s\left(t\right), c\left(t\right), t\right) \tag{50}$$

$$s\left(0\right) = s0 \quad e \quad s\left(t\right) = st \tag{51}$$

Normalmente a variável que está na restrição exposta por uma equação diferencial é denominada a variável de estado. Em outras palavras, representa o quanto essa variável em particular oscila de acordo com uma variação no tempo. Para o nosso exemplo a variável de controle (ou escolha) é $c(t)$. Temos também que$t$ denota o horizonte temporal. Como nos problemas de otimização estática, usaremos uma variável denominada de coestado $\pi\left(t\right)$que terá um funcionamento similar (ou análogo ao nosso multiplicador de Lagrange. Desse modo, construiremos uma função chamada de Hamiltoniano, para o problema expresso pelas equações de (1) a (3).

$$H(s\left(t\right),c(t),\pi\left(t\right),t) = V\left(s\left(t\right),c\left(t\right),t\right)+\pi\left(t\right)f\left(s\left(t\right),c\left(t\right),t\right) \tag{52}$$

Teorema: *Princípio do máximo.*

Uma solução ótima para o problema naterior é a tripla $\left(s\left(t\right),c(t),\pi\left(t\right)\right)$ que deve satisfazer as seguintes condições:

(i) $c(t)$ maximiza $H(s\left(t\right),c(t),\pi\left(t\right))$, tal que:

$$\frac{\partial H}{\partial c\left(t\right)} = 0; \tag{53}$$

(ii) As variáveis de estado e coestado satisfazem o par de equações diferenciais:

$$\dot{s} = \frac{\partial H}{\partial \pi\left(t\right)} \tag{54}$$

$$\dot{\pi} = -\frac{\partial H}{\partial s\left(t\right)} \tag{55}$$

Com as condições “terminais” expressas em (3) usando a definições de Hamiltoniano (4) e as equações (5)-(7) teremos:

$$\frac{\partial V}{\partial c\left(t\right)}+\pi\left(t\right)\frac{\partial f}{\partial c\left(t\right)} = 0 \tag{56}$$

$$\dot{s}\left(t\right) = f\left(s\left(t\right),c\left(t\right),t\right) \tag{57}$$

$$\dot{\pi}\left(t\right) = -\frac{\partial V}{\partial s\left(t\right)}-\pi\left(t\right)\frac{\partial f}{\partial s\left(t\right)} \tag{58}$$

25.1 Derivação das Condições de Primeira Ordem

Temos o seguinte problema de maximização intertemporal:

$$\underset{c_t}{max} V(0) = \int_0^T u(c_t, k_t, t)dt$$

$$s.a.$$

$$(a)\ \dot{k_t} = g[k, c_t, t]$$

$$(b)\ k_0 \ \ \text{(condição inicial)}$$

$$(c)\ k_T e^{-\bar{r}_T T} \geq 0$$

A condição *(c)* indica a quantidade de capital deixada no último momento é positiva ou zero.

$V(0)\ \rightarrow$ Função valor em $t = 0$

$\bar{r_T}\ \rightarrow$ taxa de desconto

$\dot{k_t}\ \rightarrow$ variável de estado

$c_t\ \rightarrow$ variável de controle

Imagine que os multiplicadores de Lagrange pudessem ser utilizados para solucionarmos o problema de maximização anterior:

$$L = \int_0^T u(c_t, k_t, t)dt + \int_0^T [\lambda_t(g[k_t, c_t, t] - \dot{k_t})]dt + \psi k_T e^{-\bar{r}_T T}$$

λ_t é o multiplicador associado com 1.a e ψ o multiplicador associado com 1.c. Se tivéssemos o Lagrangeano convencional, tomaríamos as derivadas parciais e igualaríamos a zero. O problema surgirá quando for necessário tomar a derivada de $\dot{k_t}$ em relação a k_t. Um modo mais simples de solucionar este problema é integrar por partes $\int_0^T v_t k_t dt$:

$$\frac{\partial kv}{\partial t} = \dot{k}v + \dot{v}k$$

Se integrarmos ambos os lados dessa expressão encontramos

$$\int_0^T \frac{\partial kv}{\partial t} = \int_0^T \dot{k}vdt + \int_0^T \dot{v}kdt$$

$$\int_0^T \frac{\partial kv}{\partial t} = k_T v_T - k_0 v_0$$

Rearranjando a expressão anterior:

$$k_T v_T - k_0 v_0 - \int_0^T \dot{v}kdt = \int_0^T \dot{k}vdt$$

Então o Lagrangeano toma a seguinte forma:

$$L = \int_0^T (u(c_t, k_t, t) + \lambda_t g[k_t, c_t, t])dt - k_T\lambda_T + k_0\lambda_0 + \int_0^T \dot{\lambda_t} k_t dt + \psi k_T e^{-\overline{r_T} T}$$

Definindo $\quad H(c_t, k_t, t) = u(c_t, k_t, t) + \lambda_t g[k_t, c_t, t]$

$$L = \int_0^T [H(c_t, k_t, t) + \dot{\lambda_t} k_t]dt - k_T\lambda_T + k_0\lambda_0 + \psi k_T e^{-\overline{r_T}T}$$

As CPOs:

$$\frac{\partial H}{\partial c} = \frac{\partial u}{\partial c} + \lambda \frac{\partial g}{\partial c} = 0$$

$$\frac{\partial H}{\partial k} = \frac{\partial u}{\partial k} + \lambda \frac{\partial g}{\partial k} + \dot{\lambda} = 0$$

$$\frac{\partial H}{\partial k} = -\dot{\lambda} = \frac{\partial u}{\partial k} + \lambda \frac{\partial g}{\partial k}$$

$$\frac{\partial L}{\partial k_T} = -\lambda_{T+}\psi e^{-\overline{r_T}T} = 0$$

$$\lambda_T = \psi e^{-\overline{r_T}T}$$

A condição de folga do Kuhn-Tucker nos diz que

$$k_T\lambda_T = 0$$

$$k_T\psi e^{-\overline{r_T}T} = 0$$

$$\lambda_T k_T = 0$$

No instante T o valor do capital é igual a zero, porque seu preço sombra vale zero $\lambda_{T.}$ Para o horizonte infinito a derivação é análoga trocando T por ∞. No entanto a condição de folga de Kuhn-Tucker deve ser substituída por

$$\lim_{t\to\infty} \lambda_t k_t = 0$$

25.2 Alguns exemplos de aplicação do teorema

Veremos um exemplo numérico para apresentarmos o princípio do máximo:

$$V = \int_0^1 ln\,[64s]\,dt \qquad s.a \qquad \dot{s} = 4s\,(1-c)$$

$$com \quad s\,(0) = 1 \quad e \quad s\,(1) = e^2$$

$$H\left(s,c,\pi\right)=ln4cs+\pi\left[4s\left(1-c\right)\right]$$

$$\frac{\partial H}{\partial c}=\frac{1}{4cs}4s-4s\pi=0 \tag{59}$$

$$\frac{\partial H}{\partial s}=-\dot{\pi}=\frac{1}{4cs}4c+\pi4\left(1-c\right) \tag{60}$$

$$\frac{\partial H}{\partial \pi}=\dot{s}=4s\left(1-c\right) \tag{61}$$

Por (11) teremos que $c=\frac{1}{4\pi s}$substituindo esse resultado nas outras equações obteremos:

$$\dot{\pi}=-\frac{1}{s}-4\pi\left(1-\frac{1}{4\pi s}\right) \tag{62}$$

$$\dot{s}=4s\left(1-\frac{1}{4\pi s}\right) \tag{63}$$

simplificando a equação (14) teremos:

$$\dot{\pi}=-\frac{1}{s}-4\pi\left(\frac{4\pi s-1}{4\pi s}\right)$$

$$\dot{\pi}=-\frac{1}{s}-\frac{4\pi s}{s}+\frac{1}{s}$$

$$\dot{\pi}=-4\pi \quad (14')$$

integrando (14') teremos:

$$\int\frac{\partial\pi}{\partial t}\frac{\partial t}{\pi}=-4\int dt$$

$$ln\pi=-4t$$

$$\pi=ce^{-4t} \quad (15)$$

$$\pi\left(t\right)=\pi\left(0\right)e^{-4t}$$

Então teremos

$$\dot{s}=4s-\frac{1}{\pi}$$

$$\dot{s} = 4s - \frac{e^{4t}}{\pi\left(0\right)}$$

$$\dot{s} = 4s = -\frac{e^{4t}}{\pi\left(0\right)}$$

$$a\left(t\right) = -4$$

$$b\left(t\right) = -\frac{e^{4t}}{\pi\left(0\right)}$$

Passe todos os termos de s para o lado esquerdo e multiplique pelo fator integrante e^{-4t}:

$$\dot{s}e^{-4t} - 4se^{-4t} = \frac{1}{\pi\left(0\right)}$$

Note que a derivada de se^{-4t}é:

$$\dot{s}e^{-4t} - 4se^{-4t}$$

Portanto a integral da expressão acima deve ser:

$$se^{-4t}$$

$$se^{-4t} = \frac{-t}{\pi\left(0\right)} + c$$

$$\frac{-t}{\pi\left(0\right)} + c = \int \frac{1}{\pi\left(0\right)} dt$$

sabemos que $s\left(0\right) = 1$

$$se^{-4(0)} = -\frac{0}{\pi\left(0\right)} + c = 1$$

$$1 = c$$

$$s\left(1\right) = e^{2}$$

$$s\left(1\right) = 2^{4}\left(-\frac{1}{\pi\left(0\right)} + 1\right)$$

$$e^{-2} = 1 + \frac{1}{\pi\left(0\right)}$$

Temos que :

$$\frac{1}{e^{2}} - 1 = \frac{1}{\pi\left(0\right)}$$

$$\pi\left(0\right) = \frac{1}{e^{-2} - 1} \approx 1.156$$

$$s = e^{4t} - 0,865te^{4t}$$

Sabemos que

$$c(t) = \frac{1}{4\pi s} = \frac{1}{4.62 - 4t}$$

$$V = \int_{0}^{1} ln\left(4cs\right)dt$$

$$= \int_{0}^{1} ln\left(\frac{4}{4\pi s}s\right)dt$$

$$= \int_{0}^{1} ln\left(\frac{1}{\pi\left(0\right)e^{-4t}}\right)dt$$

$$V = \int_{0}^{1} \left[ln\left(0.865\right) + 4t\right]dt = \left[tln\left(0,865\right) + 2t^{2}\right]_{0}^{1} \approx 1.855$$

Exemplo. Consumo ótimo com fator de desconto.

Desejamos maximizar

$$V = \int_{0}^{T} e^{-\delta t}lnc\left(t\right)dt \quad s.a \quad \dot{s}\left(t\right) = rs - c\left(t\right)$$

$$s\left(0\right) = s0 \quad e \quad s\left(T\right) = sT$$

Teremos que:

$$H\left(s\left(t\right), c\left(t\right), t\right) = e^{-\delta t}lnc\left(t\right) + \pi\left(t\right)\left[rs\left(t\right) - c\left(t\right)\right]$$

$$\frac{\partial H}{\partial c} = e^{-\delta t}\frac{1}{c} - \pi = 0 \tag{64}$$

$$\frac{\partial H}{\partial s} = -r\pi = \dot{\pi} \tag{65}$$

$$\frac{\partial H}{\partial \pi} = rs\left(t\right) - c\left(t\right) = \dot{s} \tag{66}$$

Usaremos a equação (17) para encontrarmos uma solução geral para π:

$$\int \frac{\partial \pi}{\partial t}\frac{1}{\pi} = \int -rdt$$

$$ln\pi = -rt + c_1$$

$$\pi = c_1 e^{-rt} \tag{67}$$

$$c_1 = \pi\left(0\right)$$

Usando (19) em (16) teremos

$$\frac{e^{-\delta t}}{c} = \pi\left(0\right) e^{-rt}$$

$$e^{+t(r-\delta)} \pi\left(0\right)^{-1} = c\left(t\right) = (5)$$

Em (18)

$$\dot{s} = rs - e^{t(r-\delta)} \left(\pi\left(0\right)\right)^{-1}$$

$$\dot{s} - rs = -e^{t(r-\delta)} \left(\pi\left(0\right)\right)^{-1}$$

$$\dot{s}e^{-st} - e^{-rt} rs = -\left(\pi\left(0\right)\right)^{-1} e^{-t\delta}$$

$$se^{-st} = \frac{e^{-\delta t}}{\delta \pi\left(0\right)} + A$$

Usando as condições iniciais teremos:

$$s\left(0\right) = s0 \qquad s0 = A + \left(\pi\left(0\right)\right)^{-1}$$

$$s\left(T\right) = sT \qquad sT = Ae^{rT} + \frac{e^{(r-\delta)T}}{S\pi\left(0\right)}$$

$$A = \frac{sTe^{-rT} - s0e^{-\delta T}}{1 - e^{\delta T}}$$

Para não termos valores negativos para o consumo, é necessário que

$$s0e^{rT} > sT$$

Exemplo. Modelo de Ramsey-Cass-Koopmans de crescimento usando uma função de utilidade da família *CARA*.

$$u(c_t) = -(\frac{1}{\alpha})e^{-\alpha c_t}$$

$$max\ u_0 = \int_0^{\infty} u(c_t)e^{-\theta_t}dt$$

$$s.a.$$

$$\dot{k} = f(k_t) - c_t - nk_t$$

$$H_t = u(c_t)e^{-\theta_t} + \mu_t[f(k_t) - nk_t - c_t]$$

Fazendo $\quad \lambda_t = \mu_t e^{\theta_t}$

$$H_t = [u(c_t) + \lambda_t(f(k_t) - nk_t - c_t)]e^{-\theta_t}$$

$$H_c = (-\frac{1}{\alpha})e^{-\alpha c}(-\alpha)e^{-\theta_t}(-\lambda_t)e^{-\theta_t} = 0$$

$$e^{-\alpha c} = \lambda_t \ \ (1)$$

$$\lambda_t = \mu_t e^{\theta_t}$$

$$(derivando\ em\ relação\ ao\ tempo)$$

$$\dot{\lambda} = \dot{\mu}e^{\theta_t} + \mu\theta e^{\theta_t}$$

→ Utilizando a condição do hamiltoniano

$$\dot{\mu} = -H_k$$

$$\dot{\mu} = e^{-\theta_t}\lambda_t[f'(k_t) - n]$$

→ Substituindo na equação de $\dot{\lambda}$

$$\dot{\lambda} = e^{-\theta_t}\lambda_t[-f'(k_t) + n]e^{\theta_t} + (\frac{\lambda_t}{e^{\theta_t}})\theta e^{\theta_t}$$

$$\dot{\lambda} = \lambda_t[-f'(k_t) + n] + \lambda_t\theta$$

$$\dot{\lambda} = \lambda_t[\theta + n - f'(k_t)] \;\; (2)$$

11) Controle ótimo:

→ Condição de transversalidade

$$\lim_{k\to\infty} k_t u'(c_t)e^{-\theta_t} = 0 \;\; (3)$$

→ Combinando as equações (1) e (2) para eliminar a variável de co-estado λ_t

→ Derivando a equação (1) em relação ao tempo

$$e^{-\alpha c} = \lambda_t \;\; (1)$$

$$-\alpha\dot{c}e^{-\alpha c} = \dot{\lambda}_t \;\; (1')$$

→ Substituindo estas relações na equação (2)

$$-\alpha\dot{c}e^{-\alpha c} = e^{-\alpha c}[\theta + n - f'(k_t)] \;\; (2')$$

→ BOX → Elasticidade de substituição

$$\sigma(c_t) = -\frac{u'(c_s)/u'(c_t)}{c_s/c_t}\left[\frac{\partial\,[u'(c_s)/u'(c_t)]}{\partial[(c_s/c_t)]}\right]^{-1}$$

Se tomarmos o limite da expressão acima, como s converge para t teremos:

$$\sigma(c_t) = -\frac{u'(c_t)}{u''(c_t)}c_t$$

→ Utilizando esta relação em $(2')$, no entanto devemos primeiramente multiplicar a equação por $\frac{1}{c_t}$:

$$-\alpha\dot{c}e^{-\alpha c}\frac{1}{c_t} = e^{-\alpha c}\frac{1}{c_t}[\theta + n - f'(k_t)]$$

$$\frac{\dot{c}}{c_t} = \frac{e^{-\alpha c}}{-\alpha e^{-\alpha c}c_t}(-[f'(k_t) - \theta - n]) \quad (2'')$$

$$\frac{\dot{c}}{c_t} = \sigma(c_t)[f'(k_t) - \theta - n]$$

11) A partir da equação $(2'')$ podemos fazer algumas simplificações:

$$\frac{\dot{c}}{c_t} = \frac{e^{-\alpha c}}{-\alpha e^{-\alpha c}c_t}(-[f'(k_t) - \theta - n])$$

$$\dot{c}\frac{\cancel{c_t}}{\cancel{c_t}} = \frac{\cancel{e^{-\alpha c}}}{\cancel{e^{-\alpha c}}(-\alpha)}(-[f'(k_t) - \theta - n])$$

$$\dot{c} = \alpha^{-1}[f'(k_t) - \theta - n]$$

→ Esta é a equação de Euler.

25.2.1 Diagrama de Fase (problemas de controle ótimo)

Vejamos um exemplo simples de controle ótimo.

$$Max\, V = \int_0^T e^{-0,05t} lnc\, dt$$

$$s.a$$

$$\dot{s} = 2s^{0,5} - c$$

$$s\,(0) = s0, \; s\,(T) = sT$$

*Obs: Não colocarei as variáveis dependentes do tempo na forma $c(t)$. Escreverei apenas c.

Teremos que:

$$H\left(s, c, \pi, t\right) = e^{-0,05t} lnc + \pi \left[2s^{0,5} - c\right]$$

$$\frac{\partial H}{\partial c} = e^{-0,05t} c^{-1} - \pi = 0 \tag{68}$$

$$\frac{\partial H}{\partial s} = -\dot{\pi} = s^{-0,5} \pi \tag{69}$$

$$\frac{\partial H}{\partial \pi} = 2s^{0,5} - c = \dot{s} \tag{70}$$

Nesse caso não conseguiremos solucionar πpor (21). Devemos diferenciar a equação (20) em relação ao tempo.

$$-e^{-0,05t} \dot{c} c^{-2} - 0,05 e^{-0,05t} c = \dot{\pi}$$

$$-e^{-0,05t} \left(\frac{\dot{c}}{c^2} + \frac{0,05}{c}\right) = \dot{\pi}$$

$$-e^{-0,05t} \left(\frac{\dot{c}}{c^2} + \frac{0,05}{c}\right) = -\pi s^{-0,5} \qquad (usando\,(21))$$

$$Por\ \ (20) \qquad \pi = \frac{e^{-0,05t}}{c}$$

$$-e^{-0,05} \left(\frac{\dot{c}}{c^2} + \frac{0,05}{c}\right) = \frac{-e^{-0,05t}}{c} s^{-0,5}$$

$$-\left(\frac{\dot{c}}{c^2} + q^{0,5}\right) = -s^{-0,5}$$

$$+\,\dot{}\, + 0,05c \dot{=} +cs^{-0,5}c$$

$$+\dot{c} = c\left(s^{-\frac{1}{2}} - 0,05\right) \tag{71}$$

Agora podemos analisar o diagrama de fase no espaço (s, c).

Primeiramente analisaremos o locus de pontos onde $\dot{c} = 0 \quad e \quad \dot{s} = 0$.

$$\dot{s} = 0 \rightarrow c = 2s^{0,5}$$

Note que $c = 2s^{0,5}$é uma função côncava e crescente em s. Essa curva sai da origem, e possui uma inclinação infinitamente grande. (se $s = 0$ então $c = 0$).

Já a equação $\dot{c} = 0$ nos indica que se $c = 0$, $s = 400$. Esses locus críticos dividem nosso quadrante positivo em quatro regiões ou iso setores.

Na região I $c < 2s^{0,5}$ então $\dot{s} > 0$. Como $c > 0$ e $s < 400$ isso implica que $\dot{c} > 0$. na região II ainda temos $c < 2s^{0,5}$; desse modo, $\dot{s} > 0$. e $s > 400$ então $\dot{c} < 0$. A região III está acima da curva $c = 2s^{0,5}$ então $\dot{s} < 0$, $\dot{s} < 400$ e $\dot{c} < 0$.

Na região IV $\dot{s} < 0$ e $\dot{c} > 0$.

Referências

CHIANG, A.; WAINWRIGHT, K. Matemática para economistas. 4a ed. Elsevier, 2005.

CYSNE; R. MOREIRA, H. Curso de matemática para economistas. 2 ed. São Paulo, Altas, 2000.

DADKHAH, K. Foundations of Mathematical and Computational Economics, 2ed. Springer, 2011.

GUIDORIZZI, H. Curso de Calculo, vol. 1. LTC, 2001.

LEONARD, D.; VAN LONG, N. Optimal Control Theory and Static Optimization in Economics. Cambridge University Press, 1992.

PONTRYAGIN, L. S.; V. G. BOLTYANSKII; R. V. GAMKRELIDZE; E. F. MISHCHENKO. The Mathematical Theory of Optimal Processes. New York, Wiley, 1962.

SALA-I-MARTIN, X. Apuntes de Crecimiento Económico. Barcelona: Antoni Bosch (Ed.), 2000.

SIMON, C.; BLUME, L. Matemática para economistas. Porto Alegre: Bookman, 2004.

SUNDARAM, R. K. A First Course in Optimization Theory, 1996.