



Lizet Santa Cruz Calderón

Injetividade de aplicações do R^n

Dissertação de Mestrado
Pós-Graduação em Matemática

Lizet Santa Cruz Calderón

Injetividade de aplicações do R^n

Orientador:
Prof. Dr. Benito Frazão Pires

Coorientador:
Prof. Dr. Ali Messaoudi

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS, LETRAS E CIÊNCIAS EXATAS
CAMPUS DE SÃO JOSÉ DO RIO PRETO

São José do Rio Preto

27 de fevereiro de 2015

Santa Cruz Calderón Lizet.

Injetividade de aplicações do R^n / Lizet Santa Cruz Calderón. São José do Rio Preto, 2015.

45 f.

Orientador: Benito Frazão Pires.

Coorientador: Ali Messaoudi .

Dissertação (mestrado) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas.

1. Matemática. 2. Sistemas dinâmicos diferenciais. 3. Domínios convexos. 4. Matrizes (Matemática) I. Pires, Benito Frazão II. Messaoudi, Ali III. Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas IV. Título.

CDU - 517.938

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Lizet Santa Cruz Calderón

Injetividade de aplicações do R^n

Dissertação apresentada para obtenção do título de Mestre em Matemática, área de Sistemas Dinâmicos, junto ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Benito Frazão Pires
Professor Doutor
Universidade de São Paulo- Riberão Preto
Orientador

Prof. Dr. Francisco Braun
Professor Doutor
Federal de São Carlos- São Carlos

Prof. Dr. Claudio Gomes Pessoa
Professor Titular
UNESP - São José do Rio Preto

São José do Rio Preto, 27 de fevereiro de 2015.

A meu pai, Apolinar.
As minhas irmãs, Tania e Susana.
A meu irmão, Abel.
Ao meu namorado John e amigos.
Dedico.

Agradecimentos

Ao concluir este trabalho, agradeço:

Primeiramente agradeço a Deus, porque ele tem me abençoado em cada coisa que eu fiz, ele me cuida e ajuda para conseguir cada coisa.

À minha mãe que sempre me protegeu a qual se virou meu anjo para me proteger em cada momento difícil.

Ao meu pai por acreditar em mim, pela confiança, pelo amor e conselhos, pela ajuda, pela força que ele deu para conseguir meus objetivos.

A minha família, no especial a meus irmãos Tania, Susana, Abel, pelos momentos felizes e pelos momentos aproveitamos para conversar bastante, ao meu bebezinho Gianpiero, o qual amo. Ao meu orientador Benito Pires, pela dedicação, pela paciência, pela ajuda incondicional, pelo tempo dedicado, pela amizade, pela confiança ao designar a mim este trabalho.

Ao meu amigo Daniel Viera, quem me ajuda demais desde o início do mestrado, quem deu sua amizade, e seu tempo quando mais precisava.

Ao meu namorado John, pela confiança, ajuda e pelo amor e carinho.

Aos meus amigos Gino e Rodiak quem me ajudaram quando cheguei ao Brasil.

A CAPES pelo auxílio financeiro.

‘Se você tentar alguma coisa e falhar, repetir
se você nunca falhar, você deve tentar novamente.
Se você continuar a falhar, é hora de olhar sua estratégia.
Nunca se apegam a uma maneira de fazer as coisas,
aprender a se adaptar e continuar tentando
até chegar ao seu objetivo.’

Farid de Alba

Resumo

Neste trabalho seguimos os testes de B.A. Coomes [6] e Gale e Nikaidô [7], onde estudamos condições que devem ser colocadas sobre o jacobiano $J\mathbf{f}$ de uma aplicação $\mathbf{f} : S \rightarrow \mathbb{R}^n$ para que ela seja injetiva num conjunto convexo $S \subset \mathbb{R}^n$. O primeiro resultado afirma que se $J\mathbf{f}(\mathbf{x})$ for uma matriz positiva-definida para todo $\mathbf{x} \in S$, então $\mathbf{f} : S \rightarrow \mathbb{R}^n$ é injetiva. O segundo resultado diz que se $J\mathbf{f}(\mathbf{x})$ for uma P-matriz para todo $\mathbf{x}$ pertencente a um retângulo fechado n -dimensional $S \subset \mathbb{R}^n$ então $\mathbf{f}$ é injetiva. Um terceiro critério de injetividade envolvendo matrizes de diagonal estritamente dominante também foi incluído no texto.

Palavras-chave: Teorema do Valor Médio de McLeod, Matriz Positiva-definida, Matriz Diagonal dominante, P-matriz.

Abstract

We provide conditions on the Jacobian matrix $J\mathbf{f}$ of a C^1 map $\mathbf{f} : S \rightarrow \mathbb{R}^n$ in order that $\mathbf{f}$ is injective in a convex set $S \subset \mathbb{R}^n$. The first result states that if $J\mathbf{f}(\mathbf{x})$ is a positive-definite matrix for every $x \in S$, then $\mathbf{f}$ is injective. The second result says that if $J\mathbf{f}(\mathbf{x})$ is a P-matrix for every x belonging to a n -dimensional closed rectangle $S \subset \mathbb{R}^n$, then $\mathbf{f}$ is injective. A third injectivity criterium involving diagonally dominant matrices was also included in the text.

Keywords: Mean Value Theorem McLeod, Positive-definite matrix, Diagonally dominant matrix, P-matrix.

Sumário

1	Introdução	p. 10
2	Preliminares	p. 13
2.1	Cálculo diferencial no $\mathbb{R}^n$	p. 13
2.2	O Teorema do Valor Médio de Cauchy	p. 15
2.3	O Teorema do Valor Médio de McLeod	p. 17
3	Injetividade de aplicações com Jacobiano positivo-definido	p. 22
4	Injetividade de aplicações com Jacobiano diagonalmente dominante	p. 28
5	Injetividade de aplicações com Jacobiano do tipo P-matriz	p. 33
6	Exemplos	p. 43
	Referências Bibliográficas	p. 46

Capítulo 1

Introdução

Neste trabalho abordamos condições sobre a matriz jacobiana $J\mathbf{f}$ de uma aplicação $\mathbf{f} : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^1 para que $\mathbf{f}$ seja injetiva num convexo S . No caso $n = 1$, uma das ferramentas importantes para saber se uma aplicação $\mathbf{f}$ é injetiva é uma consequência do Teorema do Valor Médio de Cauchy o qual assegura que se $f : [a, b] \rightarrow \mathbb{R}$ é C^1 e $f'(x) \neq 0$ para todo $x \in [a, b]$, então $\mathbf{f}$ é monótona e, portanto, injetiva.

Em dimensões maiores ($n \geq 2$), a condição de que o determinante jacobiano $|J\mathbf{f}(x)| \neq 0$ em todo ponto $\mathbf{x} \in S$ não assegura a injetividade de $\mathbf{f}$ em S . De fato, considere a aplicação $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $\mathbf{f}(x, y) = (\sqrt{2}e^{x/2} \cos(ye^{-x}), \sqrt{2}e^{x/2} \sin(ye^{-x}))$. O jacobiano de $\mathbf{f}$ no ponto $(x, y) \in \mathbb{R}^2$ é dado por:

$$J\mathbf{f}(x, y) = \begin{pmatrix} \sqrt{2}e^{x/2} \left[\frac{1}{2} \cos(ye^{-x}) + ye^{-x} \sin(ye^{-x}) \right] & -\sqrt{2}e^{-x/2} \sin(ye^{-x}) \\ \sqrt{2}e^{x/2} \left[\frac{1}{2} \sin(ye^{-x}) - ye^{-x} \cos(ye^{-x}) \right] & \sqrt{2}e^{-x/2} \cos(ye^{-x}) \end{pmatrix}.$$

Assim, $|J\mathbf{f}(x, y)| = 1$ para todo $(x, y) \in \mathbb{R}^2$. Entretanto, a aplicação $\mathbf{f}$ não é injetiva pois $\mathbf{f}(x, y + 2k\pi e^x) = \mathbf{f}(x, y)$ para todo $k \in \mathbb{Z}$ e para todo $(x, y) \in \mathbb{R}^2$. Desta forma, é preciso colocar condições mais fortes sobre o jacobiano de $\mathbf{f}$ para assegurar a injetividade.

Em 1964 McLeod generalizou o Teorema do Valor Médio de Cauchy. A versão que vamos utilizar neste trabalho é mais simples que a versão apresentada por ele em [8].

Teorema 2.4 *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então para*

todo $\mathbf{a}, \mathbf{b} \in S$, existem $\mathbf{x}_1, \dots, \mathbf{x}_{n+1} \in S$ e $t_1, \dots, t_{n+1} \geq 0$ com $\sum_{k=1}^{n+1} t_k = 1$ tais que

$$\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = \left(\sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k) \right) (\mathbf{b} - \mathbf{a}).$$

O Teorema de McLeod deu condições a Meisters e Olech [9] para provar o seguinte resultado, que é o primeiro critério de injetividade apresentado neste trabalho.

Corolário 2.7 *Sejam $S \subset \mathbb{R}^n$ um conjunto convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então $\mathbf{f}$ é injetiva se a seguinte condição for satisfeita:*

$$\text{conv} \{ J\mathbf{f}(\mathbf{x}) \cdot (\mathbf{b} - \mathbf{a}) : \mathbf{x} \in [\mathbf{a}, \mathbf{b}] \} \subset \mathbb{R}^n \setminus \{0\}, \quad \forall \mathbf{a}, \mathbf{b} \in S \mid \mathbf{a} \neq \mathbf{b},$$

onde $\text{conv}(A)$ denota a cobertura convexa do conjunto A e $[\mathbf{a}, \mathbf{b}] = \{(1-t)\mathbf{a} + t\mathbf{b} : 0 \leq t \leq 1\}$.

Com base neste resultado de Meisters e Olech deduziremos três critérios sobre a matriz jacobiana de $\mathbf{f}$ que garantem a injetividade de $\mathbf{f}$. O primeiro critério (ver Capítulo 3) assegura a injetividade de $\mathbf{f} : S \rightarrow \mathbb{R}^n$ sob a condição de que $J\mathbf{f}(\mathbf{x})$ é positivo-definido para todo $\mathbf{x} \in S$, onde $S \subset \mathbb{R}^n$ é um convexo (Teorema 3.1). Este resultado é corolário de um resultado mais geral (Teorema 3.2) em que é suficiente que a parte simétrica de $J\mathbf{f}(\mathbf{x})$ seja positiva-definida para todo $\mathbf{x} \in S$. Pelo Critério de Sylvester (Proposição 3.1), obtemos que $\mathbf{f}$ será injetiva se $J\mathbf{f}(\mathbf{x})$ for simétrica e todos os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ forem positivos para todo $\mathbf{x} \in S$. Em 1953, P. A. Samuelson [13], renomado economista laureado com o prêmio Nobel, conjecturou que se os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ não se anularem então $\mathbf{f}$ é injetiva. Em 1965, Gale e Nikaidô [7] encontraram um contraexemplo em $\mathbb{R}^2$: uma aplicação $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ não-injetiva com todos os menores principais líderes positivos (ver Exemplo 6.3). Desta forma, a hipótese de simetria no Teorema 3.1 é importante. Uma versão do Teorema 3.1 sem a hipótese de simetria do jacobiano foi dada por Gale e Nikaidô em [7]. Trata-se do Teorema 5.1 dado no Capítulo 5. Este teorema substitui a condição de que $J\mathbf{f}(\mathbf{x})$ é simétrico e tem todos os menores principais líderes positivos pela condição de que $J\mathbf{f}(\mathbf{x})$ tem todos os menores principais (líderes ou não) positivos. Assim, a hipótese de simetria é substituída por uma condição mais forte sobre os menores principais. Entretanto, o resultado obtido só vale quando S é um n -retângulo fechado. No Capítulo 3, apresentamos um critério de injetividade baseado na hipótese de que o jacobiano de $\mathbf{f}$ é de diagonal estritamente dominante. A relação do

Capítulo 4 com os outros capítulos é estabelecida pelo Lema 5.5, isto é, que toda matriz de diagonal positiva e estritamente dominante é P-matriz.

Capítulo 2

Preliminares

2.1 Cálculo diferencial no $\mathbb{R}^n$

Vamos denotar por $\mathbb{R}^n$ o espaço euclidiano n -dimensional dotado da norma

$$|\mathbf{x}| = |(x_1, \dots, x_n)| = \sqrt{x_1^2 + \dots + x_n^2}$$

e por $\mathcal{M}_{n \times n}$ o espaço das matrizes quadradas de ordem n . Dadas uma ênupla $\mathbf{x} \in \mathbb{R}^n$ e uma matriz $A = (a_{ij}) \in \mathcal{M}_{n \times n}$, definimos a ênupla

$$A \cdot \mathbf{x} = \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{nj}x_j \right).$$

O espaço $\mathcal{M}_{n \times n}$ pode ser munido da norma

$$\|A\| = \sup_{|\mathbf{x}|=1} |A \cdot \mathbf{x}|.$$

É fácil provar que para quaisquer $\mathbf{x} \in \mathbb{R}^n$ e $A, B \in \mathcal{M}_{n \times n}$,

$$|A \cdot \mathbf{x}| \leq \|A\| \cdot |\mathbf{x}| \quad \text{e} \quad \|AB\| \leq \|A\| \cdot \|B\|.$$

Seja $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ a projeção na i -ésima coordenada, isto é, a aplicação definida por $\pi_i(x_1, \dots, x_n) = x_i$. Dados um conjunto $S \subset \mathbb{R}^n$ e uma aplicação $\mathbf{f} : S \rightarrow \mathbb{R}^n$, definimos $f_i = \pi_i \circ \mathbf{f}$ e escrevemos $\mathbf{f} = (f_1, \dots, f_n)$. Dizemos que $\mathbf{f}$ é C^1 se existir um aberto $S \subset U \subset \mathbb{R}^n$ tal que as derivadas parciais $\frac{\partial f_i}{\partial x_j}$, $1 \leq i, j \leq n$, existem e são contínuas em U .

Nesse caso, definimos a *matriz jacobiana de $\mathbf{f}$ em $\mathbf{x} \in S$* como sendo a matriz

$$J\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \frac{\partial f_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}.$$

Como $\mathbf{f}$ é C^1 , as entradas de $J\mathbf{f}(\mathbf{x})$ são funções contínuas das variáveis $x_1, \dots, x_n$.

No caso em que $S \subset \mathbb{R}$, a matriz jacobina de $\mathbf{f}$ em $t \in \mathbb{R}$ é a ênupla:

$$\mathbf{f}'(t) = J\mathbf{f}(t) = \left(f'_1(t), \dots, f'_n(t) \right),$$

onde $f'_i(t) = \frac{df_i}{dt}(t)$. Além disso, se $a, b \in \mathbb{R}$ e $a \leq b$, então

$$\int_a^b \mathbf{f}(t) dt = \left(\int_a^b f_1(t) dt, \int_a^b f_2(t) dt, \dots, \int_a^b f_n(t) dt \right). \quad (2.1)$$

Teorema 2.1 (Teorema Fundamental do Cálculo). *Seja $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então para todo $a \leq b$,*

$$\int_a^b \mathbf{f}'(t) dt = \mathbf{f}(b) - \mathbf{f}(a).$$

Prova. Segue imediatamente do Teorema Fundamental do Cálculo de uma variável e da definição dada pela equação (2.1). $\square$

Lema 2.1. *Seja $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ uma aplicação contínua, então*

$$\left| \int_a^b \mathbf{f}(t) dt \right| \leq \int_a^b |\mathbf{f}(t)| dt. \quad (2.2)$$

Prova. Para cada $n \geq 1$, seja $a = x_0 < x_1 < \cdots < x_n = b$ uma partição de $[a, b]$ formada por $n + 1$ pontos igualmente espaçados. Pela definição de integral em termos de somas de Riemann, tem-se que

$$\int_a^b \mathbf{f}(t) dt = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbf{f}(x_i) \frac{b-a}{n}.$$

Assim,

$$\left| \int_a^b \mathbf{f}(t) dt \right| = \lim_{n \rightarrow \infty} \left| \sum_{i=1}^n \mathbf{f}(x_i) \frac{b-a}{n} \right| \leq \lim_{n \rightarrow \infty} \sum_{i=1}^n |\mathbf{f}(x_i)| \frac{b-a}{n} = \int_a^b |\mathbf{f}(t)| dt.$$

□

2.2 O Teorema do Valor Médio de Cauchy

O Teorema do Valor Médio tal qual apresentado nos livros de Cálculo Diferencial é devido a A. L. Cauchy. Uma versão é a seguinte:

Teorema 2.2 (Teorema do Valor Médio de Cauchy). *Se $f : [a, b] \rightarrow \mathbb{R}$ é uma função C^1 , então existe um ponto $c \in (a, b)$ tal que $f(b) - f(a) = f'(c)(b - a)$.*

Para generalizar o Teorema 2.2 para dimensões maiores, precisamos introduzir o conceito de subconjunto convexo de $\mathbb{R}^n$. Dados dois pontos $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, definimos

$$[\mathbf{a}, \mathbf{b}] = \{(1-t)\mathbf{a} + t\mathbf{b} \mid 0 \leq t \leq 1\}.$$

Dizemos que $S \subset \mathbb{R}^n$ é *convexo* se $[\mathbf{x}, \mathbf{y}] \subset S$ para todo $\mathbf{x}, \mathbf{y} \in S$. De maneira análoga define-se subconjunto convexo de $\mathcal{M}_{n \times n}$. Apresentamos abaixo uma generalização natural do Teorema 2.2. A prova deste teorema será dada mais adiante.

Teorema 2.3 (Teorema do Valor Médio no $\mathbb{R}^n$). *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então para cada $\mathbf{a}, \mathbf{b} \in S$, existe $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ tal que*

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leq \|J\mathbf{f}(\mathbf{c})\| \cdot |\mathbf{b} - \mathbf{a}|. \quad (2.3)$$

A forma como foi enunciado o Teorema 2.3 é ligeiramente mais forte do que aquela comumente encontrada na literatura, onde no lugar de $\|J\mathbf{f}(\mathbf{c})\|$ aparece $\sup_{\mathbf{x} \in [\mathbf{a}, \mathbf{b}]} \|J\mathbf{f}(\mathbf{x})\|$. Observa-se também que a afirmação do Teorema 2.3 é mais fraca do que aquela do Teorema 2.2. O próximo exemplo mostra que não é possível substituir a desigualdade em (2.3) por uma igualdade.

Exemplo 2.1. *Seja $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a aplicação definida por $\mathbf{f}(x_1, x_2) = (\sin x_1, \cos x_1)$. Então,*

$$J\mathbf{f}(c_1, c_2) = \begin{pmatrix} \cos c_1 & 0 \\ -\sin c_1 & 0 \end{pmatrix}.$$

Sejam $\mathbf{a} = (2\pi, 2\pi)$, $\mathbf{b} = (0, 0)$ e $\mathbf{x} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Como $|\mathbf{x}| = 1$, temos que

$$\|J\mathbf{f}(\mathbf{c})\| \geq |J\mathbf{f}(\mathbf{c}) \cdot \mathbf{x}| = \frac{1}{\sqrt{2}}|\cos c_1, -\sin c_1| = \frac{1}{\sqrt{2}}, \quad \forall \mathbf{c} = (c_1, c_2) \in \mathbb{R}^2.$$

Portanto, para todo $\mathbf{c} \in \mathbb{R}^2$, temos que

$$|f(\mathbf{b}) - f(\mathbf{a})| = 0 < 2\pi = \frac{1}{\sqrt{2}}2\sqrt{2}\pi = \frac{1}{\sqrt{2}}|(2\pi, 2\pi)| = \frac{1}{\sqrt{2}}|\mathbf{b} - \mathbf{a}| \leq \|J\mathbf{f}(\mathbf{c})\| \cdot |(\mathbf{b} - \mathbf{a})|.$$

Isto mostra que não existe $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ tal que valha a igualdade em (2.3).

Para demonstrar o Teorema 2.3, vamos precisar do seguinte lema. Trata-se de uma versão do Teorema 2.3 para caminhos C^1 , isto é, para aplicações suaves do intervalo no $\mathbb{R}^n$.

Lema 2.2 (Rudin [14], p. 113). *Seja $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ uma aplicação C^1 definida num intervalo $[a, b] \subset \mathbb{R}$. Então existe $c \in [a, b]$ tal que*

$$|\mathbf{f}(b) - \mathbf{f}(a)| \leq |\mathbf{f}'(c)| \cdot |b - a|.$$

Prova. Seja $\mathbf{v} = \mathbf{f}(b) - \mathbf{f}(a)$. Precisamos provar que $|\mathbf{v}| \leq |\mathbf{f}'(c)| \cdot |b - a|$. Podemos assumir que $|\mathbf{v}| > 0$. Seja $\phi : [a, b] \rightarrow \mathbb{R}$ a função definida por

$$\phi(t) = \langle \mathbf{f}(t), \mathbf{v} \rangle, \quad t \in [a, b],$$

onde $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ denota o produto escalar entre $\mathbf{x}$ e $\mathbf{y}$. Segue das propriedades do produto escalar que a função ϕ é C^1 . Pelo Teorema 2.2, existe $c \in [a, b]$ tal que

$$\phi(b) - \phi(a) = \phi'(c)(b - a) = \langle \mathbf{f}'(c), \mathbf{v} \rangle(b - a). \quad (2.4)$$

Pela definição de ϕ , temos que

$$\phi(b) - \phi(a) = \langle \mathbf{f}(b), \mathbf{v} \rangle - \langle \mathbf{f}(a), \mathbf{v} \rangle = \langle \mathbf{f}(b) - \mathbf{f}(a), \mathbf{v} \rangle = |\mathbf{v}|^2. \quad (2.5)$$

Das equações (2.4) e (2.5) e da Desigualdade de Cauchy Schwarz segue que

$$|\mathbf{v}|^2 = \langle \mathbf{f}'(c), \mathbf{v} \rangle(b - a) \leq |\mathbf{f}'(c)| \cdot |\mathbf{v}|(b - a).$$

Dividindo por $|\mathbf{v}|$ obtemos a desigualdade que precisávamos provar. □

Prova do Teorema 2.3. Seja $\gamma : [0, 1] \rightarrow S$ definida por $\gamma(t) = (1 - t)\mathbf{a} + t\mathbf{b}$. Observe

que $\mathbf{f} \circ \gamma : [0, 1] \rightarrow \mathbb{R}^n$ é uma aplicação C^1 . Além disso,

$$|(\mathbf{f} \circ \gamma)'(t)| = |J\mathbf{f}(\gamma(t)) \gamma'(t)| \leq \|J\mathbf{f}(\gamma(t))\| \cdot |\mathbf{b} - \mathbf{a}|, \quad \forall t \in [0, 1]. \quad (2.6)$$

Pelo Teorema 2.2 e pela equação (2.6), existe $c \in [0, 1]$ tal que

$$|\mathbf{f} \circ \gamma(1) - \mathbf{f} \circ \gamma(0)| \leq |(\mathbf{f} \circ \gamma)'(c)| \cdot |1 - 0| \leq \|J\mathbf{f}(\gamma(c))\| \cdot |\mathbf{b} - \mathbf{a}|.$$

Definindo $\mathbf{c} = \gamma(c)$, obtemos

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| = |\mathbf{f} \circ \gamma(1) - \mathbf{f} \circ \gamma(0)| \leq \|J\mathbf{f}(\mathbf{c})\| \cdot |\mathbf{b} - \mathbf{a}|.$$

Note que, por definição, $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$. ■

2.3 O Teorema do Valor Médio de McLeod

Nesta seção apresentamos uma outra generalização do Teorema do Valor Médio de Cauchy denominada Teorema do Valor Médio de McLeod [8], cujo enunciado é o seguinte.

Teorema 2.4 (Teorema do Valor Médio de McLeod). *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então para todo $\mathbf{a}, \mathbf{b} \in S$, existem $\mathbf{x}_1, \dots, \mathbf{x}_{n+1} \in S$ e $t_1, \dots, t_{n+1} \geq 0$ com $\sum_{k=1}^{n+1} t_k = 1$ tais que*

$$\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = \left(\sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k) \right) (\mathbf{b} - \mathbf{a}).$$

Note que as hipóteses do Teorema 2.4 são as mesmas do Teorema 2.3. Uma diferença fundamental entre os dois teoremas é que o Teorema 2.4 preserva a igualdade presente no Teorema do Valor Médio de Cauchy. Isto torna o Teorema 2.4 útil para a obtenção de um critério de injetividade através da suposição de que $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$.

Para provar o Teorema 2.4, vamos precisar de alguns conceitos de análise convexa. A prova apresentada aqui baseia-se num Teorema de Choquet (ver [5, p. 38]) e segue o roteiro descrito em [10, p. 249]. Uma abordagem histórica deste teorema é dada em [15].

Seja X um espaço métrico. É imediato que $\emptyset$ e X são convexos. Além disso, interseção arbitrária de convexos é convexo. A *cobertura convexa* (ou *envelope convexo*) de um subconjunto S de X é definida como a interseção de todos os subconjuntos convexos de X que contêm S , sendo denotada por $\text{conv}(S)$. Assim, a cobertura convexa é o menor convexo (segundo a inclusão de conjuntos) que contém o conjunto S . No caso em que S já é convexo

tem-se $\text{conv}(S) = S$.

Dizemos que $\mathbf{x} \in X$ é *combinação convexa* de $\mathbf{x}_1, \dots, \mathbf{x}_m \in S$ se existirem $t_1, \dots, t_m \geq 0$ com $t_1 + \dots + t_m = 1$ tais que $\mathbf{x} = t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m$.

Proposição 2.1. *Um conjunto $S \subset X$ é convexo se e somente se S consiste de todas as combinações convexas de seus elementos.*

Prova. Suponha que S seja convexo e $\mathbf{x} \in S$. Vamos provar por indução em m que se $\mathbf{x}$ é qualquer combinação convexa de $\mathbf{x}_1, \dots, \mathbf{x}_m \in S$ então $\mathbf{x} \in S$. O caso $m = 1$ é trivial. Assuma que o resultado seja válido para um inteiro $m \geq 1$. Vamos provar que ele também vale para $m + 1$. Seja $\mathbf{x}$ uma combinação convexa de $\mathbf{x}_1, \dots, \mathbf{x}_{m+1} \in S$, isto é, suponha que existam $t_1, \dots, t_{m+1} \geq 0$ com $\sum_{k=1}^{m+1} t_k = 1$ tais que $\mathbf{x} = \sum_{k=1}^{m+1} t_k\mathbf{x}_k$. Se $t_{m+1} = 1$, então $t_k = 0$ para todo $1 \leq k \leq m$. Neste caso, $\mathbf{x} = \mathbf{x}_{m+1} \in S$. Assim podemos assumir que $t_{m+1} < 1$ e definir $s_j = \frac{t_j}{1-t_{m+1}}$, $1 \leq j \leq m$. Note que, $s_k \geq 0$ para todo $1 \leq k \leq m$. Além disso, $\sum_{k=1}^m s_k = 1$. Pela hipótese de indução,

$$\mathbf{y} = \sum_{j=1}^m s_j\mathbf{x}_j \in S.$$

Como S é convexo, temos que

$$\sum_{k=1}^{m+1} t_k\mathbf{x}_k = (1 - t_{m+1})\mathbf{y} + t_{m+1}\mathbf{x}_{m+1} \in S.$$

Assim, a afirmação vale também para $m + 1$, o que prova que S é combinação convexa de seus elementos.

Reciprocamente, assuma que S é combinação convexa de seus elementos. Em particular, qualquer combinação convexa de dois elementos de S é um elemento de S , o que mostra que S é convexo. $\square$

Corolário 2.1. *A cobertura convexa de um conjunto S consiste de todas as combinações convexas de elementos de S , isto é:*

$$\text{conv}(S) = \left\{ \sum_{k=1}^m t_k\mathbf{x}_k \mid m \geq 1, \mathbf{x}_1, \dots, \mathbf{x}_m \in S, t_1, \dots, t_m \geq 0, \sum_{k=1}^m t_k = 1 \right\}. \quad (2.7)$$

Prova. Seja R o conjunto definido pelo lado direito da equação (2.7). Pela Proposição 2.1, R é um convexo que contém S . Portanto, $\text{conv}(S) \subset R$, pois é combinação convexa de seus elementos $\mathbf{x}_1, \dots, \mathbf{x}_m$. Por outro lado, pela Proposição 2.1, qualquer convexo que

contém S , contém também as combinações convexas de S , isto é, contém R . Assim, $R \subset \text{conv}(S)$, o que conclui a demonstração. $\square$

Teorema 2.5 (Carathéodory). *Seja $S \subset \mathbb{R}^n$. Todo elemento de $\text{conv}(S)$ pode ser escrito como combinação convexa de até $n + 1$ elementos de S .*

Prova. Pelo Corolário 2.1, todo $\mathbf{x} \in \text{conv}(S)$ se escreve da forma

$$\mathbf{x} = t_1 \mathbf{x}_1 + \cdots + t_m \mathbf{x}_m, \quad (2.8)$$

onde $t_1, \dots, t_m > 0$, $t_1 + \dots + t_m = 1$ e $\mathbf{x}_1, \dots, \mathbf{x}_m \in S$. Suponha que $m > n + 1$, então $m - 1 > n$. Assim, os vetores $\mathbf{x}_1 - \mathbf{x}_m, \dots, \mathbf{x}_{m-1} - \mathbf{x}_m$ são linearmente dependentes, portanto existem constantes $c_1, \dots, c_{m-1}$ não todas nulas tais que

$$c_1(\mathbf{x}_1 - \mathbf{x}_m) + \cdots + c_{m-1}(\mathbf{x}_{m-1} - \mathbf{x}_m) = 0.$$

Seja $c_m = -(c_1 + \dots + c_{m-1})$. Então

$$\sum_{i=1}^m c_i \mathbf{x}_i = 0, \quad \sum_{i=1}^m c_i = 0. \quad (2.9)$$

Como $t_i > 0$, $\forall i$, e $c_i > 0$ para algum i , podemos definir

$$\rho := \max \{ \rho^* \geq 0 : t_i - \rho^* c_i \geq 0, i = 1, \dots, m \} = \min_{c_i > 0} \left\{ \frac{t_i}{c_i} \right\} \geq 0, \quad \text{com } 1 \leq k \leq m$$

Segue de (2.8) e (2.9) que as seguintes propriedades valem para $c'_i := t_i - \rho c_i$, $1 \leq i \leq m$:

- (i) $c'_k = 0$ para algum $1 \leq k \leq m$;
- (ii) $\sum_{i \neq k} c'_i = 1$;
- (iii) $\mathbf{x} = \sum_{i \neq k}^m c'_i \mathbf{x}_i$.

A propriedades (i), (ii) e (iii) implicam que $\mathbf{x}$ é combinação convexa dos $m - 1$ vetores $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_m$ de S . Assim, ou $m - 1 = n + 1$ ou então o mesmo raciocínio mostra que $\mathbf{x}$ é combinação de $m - 2$ elementos de S . Procedendo por indução, concluímos que $\mathbf{x}$ é combinação de no máximo $n + 1$ elementos de S . $\square$

Denotamos por $\overline{\text{conv}(S)}$ o fecho topológico de $\text{conv}(S)$ em X .

Lema 2.3. *Se $S \subset \mathbb{R}^n$ é compacto, então $\overline{\text{conv}(S)} = \text{conv}(S)$.*

Prova. Seja $\mathbf{x} \in \overline{\text{conv } S}$, então existe uma sequência $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$ em $\text{conv } S$ tal que $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)}$. Pelo Teorema 2.5, para cada $k \geq 1$, existem números $t_1^{(k)}, \dots, t_{n+1}^{(k)} \geq 0$, com $t_1^{(k)} + \dots + t_{n+1}^{(k)} = 1$ e vetores $\mathbf{x}_1^{(k)}, \dots, \mathbf{x}_{n+1}^{(k)}$ de S tais que

$$\mathbf{x}^{(k)} = t_1^{(k)} \mathbf{x}_1^{(k)} + \dots + t_{n+1}^{(k)} \mathbf{x}_{n+1}^{(k)}. \quad (2.10)$$

Como para cada $1 \leq i \leq n+1$, as sequências $\{t_i^{(k)}\}_{k=1}^\infty$ e $\{\mathbf{x}_i^{(k)}\}_{k=1}^\infty$ pertencem a espaços compactos pois $t_i^{(k)} \in [0, 1]$, podemos assumir, tomando uma subsequência se necessário, que elas convergem para $t_i \in \mathbb{R}$ e $\mathbf{x}_i \in S$, respectivamente. É fácil ver que $t_i \geq 0$, $\forall i$, e que $t_1 + \dots + t_{n+1} = 1$. Passando o limite em (2.10), resulta que

$$\mathbf{x} = t_1 \mathbf{x}_1 + \dots + t_{n+1} \mathbf{x}_{n+1},$$

isto é, $\mathbf{x} \in \text{conv}(S)$. □

Teorema 2.6 (Choquet). *Seja $\mathbf{g} : [0, 1] \rightarrow \mathbb{R}^n$ uma função contínua. Então*

$$\int_0^1 \mathbf{g}(t) dt \in \text{conv}(\mathbf{g}([0, 1])).$$

Prova. Para cada número natural k , sejam $0 = t_0 < t_1 < \dots < t_k = 1$ definidos por $t_i = \frac{i}{k}$, $0 \leq i \leq k$. Seja $\mathbf{g}^{(k)} : [0, 1] \rightarrow \mathbb{R}^n$ a aplicação definida por $\mathbf{g}^{(k)}(t) = \mathbf{g}(t_{i-1})$ se $t \in [t_{i-1}, t_i)$, $1 \leq i \leq k$, e $\mathbf{g}^{(k)}(1) = \mathbf{g}(1)$. Como a função $\mathbf{g}$ é uniformemente contínua, a sequência de aplicações $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots$ converge uniformemente para a aplicação $\mathbf{g}$. Isto garante que $\int_0^1 \mathbf{g}(t) dt = \lim_{k \rightarrow \infty} \int_0^1 \mathbf{g}^{(k)}(t) dt$. Além disso, temos que

$$\int_0^1 \mathbf{g}^{(k)}(t) dt = \sum_{i=1}^k (t_i - t_{i-1}) \mathbf{g}(t_{i-1}). \quad (2.11)$$

Observe que $c_i = t_i - t_{i-1}$, $1 \leq i \leq k$, satisfaz $c_i \geq 0$ e $\sum_{i=1}^k c_i = 1$. Desta forma, a integral que aparece em (2.11) é uma combinação convexa de elementos de $\mathbf{g}([0, 1])$ e, portanto, pertence à $\text{conv}(\mathbf{g}([0, 1]))$. Fazendo $k \rightarrow \infty$, obtemos que $\int_0^1 \mathbf{g}(t) dt$ pertence à $\overline{\text{conv}(\mathbf{g}([0, 1]))}$. O conjunto $\mathbf{g}([0, 1])$ é imagem do compacto $[0, 1]$ por uma aplicação contínua, portanto também é compacto. Pelo Lema 2.3, $\overline{\text{conv}(\mathbf{g}([0, 1]))} = \text{conv}(\mathbf{g}([0, 1]))$, o que conclui a prova. □

Prova do Teorema do Valor Médio de McLeod. Sejam $\mathbf{a}, \mathbf{b} \in S$ e $\gamma : [0, 1] \rightarrow S$ definida por $\gamma(t) = (1-t)\mathbf{a} + t\mathbf{b}$. Pelo Teorema 2.1 e pelo Teorema 2.6 com $\mathbf{g} = (\mathbf{f} \circ \gamma)'$,

obtemos

$$\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = \mathbf{f} \circ \gamma(1) - \mathbf{f} \circ \gamma(0) = \int_0^1 (\mathbf{f} \circ \gamma)'(t) dt \in \text{conv}((\mathbf{f} \circ \gamma)'([0, 1])). \quad (2.12)$$

Além disso, $(\mathbf{f} \circ \gamma)'(t) = J\mathbf{f}(\gamma(t))(\mathbf{b} - \mathbf{a})$ para todo $t \in [0, 1]$. Consequentemente,

$$\text{conv}((\mathbf{f} \circ \gamma)'([0, 1])) \subseteq \text{conv}\{J\mathbf{f}(\mathbf{x}) \cdot (\mathbf{b} - \mathbf{a}) : \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\}. \quad (2.13)$$

Pelo Teorema 2.5 e pelas equações (2.12) e (2.13), existem $\mathbf{x}_1, \dots, \mathbf{x}_{n+1} \in S$, $t_1, \dots, t_{n+1} \geq 0$ com $\sum_{k=1}^{n+1} t_k = 1$ tais que $\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = \sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k) \cdot (\mathbf{b} - \mathbf{a})$. ■

Corolário 2.7 (Meisters). *Sejam $S \subset \mathbb{R}^n$ um conjunto convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então $\mathbf{f}$ é injetiva se a seguinte condição for satisfeita:*

$$\text{conv}\{J\mathbf{f}(\mathbf{x}) \cdot (\mathbf{b} - \mathbf{a}) : \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\} \subset \mathbb{R}^n \setminus \{0\}, \quad \forall \mathbf{a}, \mathbf{b} \in S \mid \mathbf{a} \neq \mathbf{b}.$$

Prova. Suponha que $\mathbf{f}$ não seja injetiva, então existem $\mathbf{a}, \mathbf{b} \in S$ com $\mathbf{b} \neq \mathbf{a}$ tais que $\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = 0$. Segue de (2.12) e (2.13) da prova do Teorema do Valor Médio de McLeod que $0 \in \text{conv}\{J\mathbf{f}(\mathbf{x}) \cdot (\mathbf{b} - \mathbf{a}) : \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\}$. □

Corolário 2.8 (Critério geral de injetividade). *Sejam $S \subset \mathbb{R}^n$ um conjunto convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Então $\mathbf{f}$ é injetiva se a cobertura convexa de $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ for formada por matrizes não-singulares.*

Prova. Suponha que $\mathbf{f}$ não seja injetiva, então existem $\mathbf{a}, \mathbf{b} \in S$ com $\mathbf{a} \neq \mathbf{b}$ tais que $\mathbf{f}(\mathbf{a}) = \mathbf{f}(\mathbf{b})$. Pelo Teorema do Valor Médio de McLeod, existem $t_1, \dots, t_{n+1} \geq 0$ com $t_1 + \dots + t_{n+1} = 1$ e $\mathbf{x}_1, \dots, \mathbf{x}_{n+1} \in S$ tais que

$$0 = \mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}) = \sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k) \cdot (\mathbf{b} - \mathbf{a}) = \left(\sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k) \right) \cdot (\mathbf{b} - \mathbf{a}).$$

Assim, a hipótese de não-injetividade implica que a matriz $\sum_{k=1}^{n+1} t_k J\mathbf{f}(\mathbf{x}_k)$, que pertence à cobertura convexa de $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$, é singular. □

Capítulo 3

Injetividade de aplicações com Jacobiano positivo-definido

Denote por $\mathbf{x} = (x_1, \dots, x_n)$ uma ênupla qualquer do $\mathbb{R}^n$. Dizemos que uma matriz simétrica $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ é *positiva-definida* se a forma quadrática $\langle A \cdot \mathbf{x}, \mathbf{x} \rangle$ for positiva-definida, isto é, se $\sum_{i,j=1}^n a_{ij} x_i x_j > 0$ para todo $\mathbf{x} \neq \mathbf{0}$.

O principal resultado desta seção é o seguinte critério de injetividade.

Teorema 3.1. *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Se para todo $\mathbf{x} \in S$, $J\mathbf{f}(\mathbf{x})$ é positivo-definido, então $\mathbf{f}$ é injetiva. Equivalentemente, $\mathbf{f}$ é injetiva se para todo $\mathbf{x} \in S$, $J\mathbf{f}(\mathbf{x})$ for simétrica e os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ forem positivos, isto é,*

$$\frac{\partial f_1}{\partial x_1}(\mathbf{x}) > 0, \quad \begin{vmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) \end{vmatrix} > 0, \quad \dots, \quad \begin{vmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{vmatrix} > 0.$$

O Teorema 3.1 o qual será provado mais adiante, é, na verdade, corolário de um resultado um pouco mais geral, onde no lugar de $J\mathbf{f}(\mathbf{x})$ aparece a parte simétrica de $J\mathbf{f}(\mathbf{x})$.

Teorema 3.2. *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Se para todo $\mathbf{x} \in S$, a parte simétrica de $J\mathbf{f}(\mathbf{x})$ é positiva-definida, então $\mathbf{f}$ é injetiva.*

Dizemos que uma matriz $A \in \mathcal{M}_{n \times n}$ é *positiva-quase-definida* se sua parte simétrica $\frac{A + A^T}{2}$ for positiva-definida. Assim, o Teorema 3.2 afirma que se $J\mathbf{f}(\mathbf{x})$ for positivo-quase-definido para todo $\mathbf{x} \in S$, então $\mathbf{f}$ é injetiva. Note que enquanto a definição de matriz

positiva-definida requer simetria, a definição de matriz positiva-quase-definida dispensa tal hipótese. No Capítulo 5 vamos ver uma classe de matrizes mais abrangente, chamadas P-matrizes, que inclui em particular as matrizes positiva-quase-definidas.

Para provar os resultados anteriores, vamos precisar de alguns lemas preparatórios que passamos a enunciar.

Lema 3.1. *Se A é uma matriz positiva-definida, então os autovalores de A são números reais positivos.*

Prova. Seja $(\lambda, \mathbf{x})$ um par autovalor-autovetor de A . Como A é positiva-definida e $\mathbf{x} \neq 0$, temos que $\lambda \in \mathbb{R}$ (pois A é simétrica) e $\lambda|\mathbf{x}|^2 = \langle \lambda \mathbf{x}, \mathbf{x} \rangle = \langle A \cdot \mathbf{x}, \mathbf{x} \rangle > 0$, portanto $\lambda > 0$. $\square$

Dizemos que uma matriz quadrada B é *submatriz principal* de $A \in \mathcal{M}_{n \times n}$ se B for obtida eliminando r -linhas e as correspondentes r -colunas de A , onde $0 \leq r \leq n - 1$. O determinante de B é chamado *menor principal* de A . A *submatriz principal líder* de A de ordem k é a submatriz obtida eliminando as $n - k$ últimas linhas e as $n - k$ últimas colunas de A , isto é,

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}, \quad 1 \leq k \leq n. \quad (3.1)$$

O determinante de A_k é chamado *menor principal líder* de A de ordem k .

Uma caracterização muito útil de matriz positiva-definida é dada pelo seguinte resultado.

Proposição 3.1. *Uma matriz simétrica $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ é positiva-definida se e somente se todos os menores principais líderes de A forem positivos, isto é, se os seguintes determinantes forem positivos:*

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \dots, \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} > 0. \quad (3.2)$$

Prova. Suponha que A seja positiva-definida, então $\sum_{i,j=1}^n a_{ij}x_i x_j > 0$ para todo $\mathbf{x} = (x_1, \dots, x_n) \neq \mathbf{0}$. Em particular, tomando $\mathbf{y} = (y_1, \dots, y_k, 0, \dots, 0) \neq \mathbf{0}$, obtemos que $\langle A_k \cdot \mathbf{y}, \mathbf{y} \rangle = \sum_{i,j=1}^k a_{ij}y_i y_j = \sum_{i,j=1}^n a_{ij}y_i y_j > 0$. Isto mostra que a matriz simétrica A_k é positiva-definida para todo $1 \leq k \leq n$. Pelo Lema 3.1, todos os autovalores de A_k são positivos, logo o determinante de A_k é positivo.

Vamos provar agora a recíproca. Suponha que os determinantes $|A_1|, |A_2|, \dots, |A_n|$ sejam todos positivos, isto é, suponha que (3.2) seja satisfeita. Precisamos provar que

$$\langle A \cdot \mathbf{x}, \mathbf{x} \rangle = \sum_{i,j=1}^n a_{ij}x_i x_j > 0 \quad \text{para todo } \mathbf{x} = (x_1, \dots, x_n) \neq \mathbf{0}.$$

Denote por $\mathbf{e}_1, \dots, \mathbf{e}_n$ os vetores

$$\mathbf{e}_1 = (1, 0, \dots, 0, 0), \mathbf{e}_2 = (0, 1, \dots, 0, 0), \quad \dots, \quad \mathbf{e}_n = (0, 0, \dots, 0, 1).$$

Afirmamos que existe uma única coleção de vetores $\mathbf{v}_k = (\alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kk}, 0, \dots, 0) \in \mathbb{R}^n$, $1 \leq k \leq n$, tais que

$$\langle A \cdot \mathbf{v}_k, \mathbf{e}_i \rangle = 0 \quad \text{para todo } 1 \leq i \leq k-1, \quad \text{e} \quad \langle A \cdot \mathbf{v}_k, \mathbf{e}_k \rangle = 1. \quad (3.3)$$

De fato, substituindo $\mathbf{v}_k = \alpha_{k1}\mathbf{e}_1 + \dots + \alpha_{kk}\mathbf{e}_k$ em (3.3) e usando a linearidade de A junto com a bilinearidade do produto interno, obtemos:

$$\begin{aligned} \alpha_{k1}\langle A \cdot \mathbf{e}_1, \mathbf{e}_1 \rangle &+ \alpha_{k2}\langle A \cdot \mathbf{e}_2, \mathbf{e}_1 \rangle + \dots + \alpha_{kk}\langle A \cdot \mathbf{e}_k, \mathbf{e}_1 \rangle = 0 \\ \alpha_{k1}\langle A \cdot \mathbf{e}_1, \mathbf{e}_2 \rangle &+ \alpha_{k2}\langle A \cdot \mathbf{e}_2, \mathbf{e}_2 \rangle + \dots + \alpha_{kk}\langle A \cdot \mathbf{e}_k, \mathbf{e}_2 \rangle = 0 \\ &\vdots \\ \alpha_{k1}\langle A \cdot \mathbf{e}_1, \mathbf{e}_{k-1} \rangle &+ \alpha_{k2}\langle A \cdot \mathbf{e}_2, \mathbf{e}_{k-1} \rangle + \dots + \alpha_{kk}\langle A \cdot \mathbf{e}_k, \mathbf{e}_{k-1} \rangle = 0 \\ \alpha_{k1}\langle A \cdot \mathbf{e}_1, \mathbf{e}_k \rangle &+ \alpha_{k2}\langle A \cdot \mathbf{e}_2, \mathbf{e}_k \rangle + \dots + \alpha_{kk}\langle A \cdot \mathbf{e}_k, \mathbf{e}_k \rangle = 1 \end{aligned}$$

Como $\langle A \cdot \mathbf{e}_i, \mathbf{e}_j \rangle = a_{ij}$, temos que o sistema de equações anterior é equivalente ao seguinte:

$$A_k \begin{pmatrix} \alpha_{k1} \\ \alpha_{k2} \\ \vdots \\ \alpha_{kk} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \quad (3.4)$$

onde A_k é dada por (3.1). Como, por hipótese, o determinante de A_k é positivo, temos que o sistema (3.4) é compatível e determinado. Vamos voltar agora para o problema original. Vamos mostrar que a forma quadrática $\langle A \cdot \mathbf{x}, \mathbf{x} \rangle$ é positiva-definida. Isto segue das seguin-

tes afirmações:

Afirmção (a). $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ é base do $\mathbb{R}^n$.

De fato, resolvendo o sistema (3.4) pela regra de Cramer vemos que na variável α_{kk} resulta que

$$\alpha_{kk} = \frac{|A_{k-1}|}{|A_k|} > 0,$$

onde $A_0 = 1$. Isto mostra em particular que os vetores $\mathbf{v}_k = (\alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kk}, 0, \dots, 0)$, $1 \leq k \leq n$, formam uma base do $\mathbb{R}^n$.

Afirmção (b). $\langle A \cdot \mathbf{v}_k, \mathbf{v}_j \rangle = 0$ para todo $k \neq j$.

Como A é simétrica, temos que $\langle A \cdot \mathbf{v}_k, \mathbf{v}_j \rangle = \langle A \cdot \mathbf{v}_j, \mathbf{v}_k \rangle$, assim podemos assumir que $1 \leq j \leq k-1 \leq n$. Desta forma, como $\mathbf{v}_j = \sum_{i=1}^j \alpha_{ji} \mathbf{e}_i$, segue de (3.3) que

$$\langle A \cdot \mathbf{v}_k, \mathbf{v}_j \rangle = \sum_{i=1}^j \alpha_{ji} \langle A \cdot \mathbf{v}_k, \mathbf{e}_i \rangle = 0.$$

Afirmção (c). $\langle A \cdot \mathbf{v}_k, \mathbf{v}_k \rangle = \alpha_{kk}$ para todo $1 \leq k \leq n$.

Temos que $\mathbf{v}_k = \sum_{i=1}^{k-1} \alpha_{ki} \mathbf{e}_i + \alpha_{kk} \mathbf{e}_k$. Segue de (3.3) que

$$\langle A \cdot \mathbf{v}_k, \mathbf{v}_k \rangle = \sum_{i=1}^{k-1} \alpha_{ki} \langle A \cdot \mathbf{v}_k, \mathbf{e}_i \rangle + \alpha_{kk} \langle A \cdot \mathbf{v}_k, \mathbf{e}_k \rangle = \alpha_{kk}.$$

Para concluir a prova, considere um vetor qualquer $\mathbf{x} \in \mathbb{R}^n$. Pela Afirmção (a), podemos escrever $\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$. Segue das afirmações (b) e (c) que

$$\langle A \cdot \mathbf{x}, \mathbf{x} \rangle = \sum_{i,j=1}^n \langle A \cdot \mathbf{v}_i, \mathbf{v}_j \rangle c_i c_j = \alpha_{11} c_1^2 + \dots + \alpha_{nn} c_n^2 \geq 0.$$

Além disso, $\langle A \cdot \mathbf{x}, \mathbf{x} \rangle = 0$ se e somente se $c_1 = \dots = c_n = 0$, isto é, se e somente se $\mathbf{x} = \mathbf{0}$. Isto prova que A é positiva-definida. $\square$

A parte simétrica de uma matriz $A \in \mathcal{M}_{n \times n}$ é a matriz simétrica definida por $B =$

$\frac{A + A^T}{2}$, onde A^T denota a transposta de A . É claro que uma matriz A é simétrica se e somente se ela for igual a sua parte simétrica.

Lema 3.2. *Uma matriz $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ e sua parte simétrica $B = \frac{A + A^T}{2}$ determinam a mesma forma quadrática, isto é,*

$$\langle A \cdot \mathbf{x}, \mathbf{x} \rangle = \left\langle \frac{A + A^T}{2} \cdot \mathbf{x}, \mathbf{x} \right\rangle \quad \text{para todo } \mathbf{x} \in \mathbb{R}^n.$$

Prova. Primeiro observe que

$$\langle A^T \cdot \mathbf{x}, \mathbf{x} \rangle = \sum_{i,j=1}^n a_{ji} x_i x_j = \sum_{i,j=1}^n a_{ij} x_i x_j = \langle A \cdot \mathbf{x}, \mathbf{x} \rangle$$

Logo

$$\left\langle \frac{A + A^T}{2} \cdot \mathbf{x}, \mathbf{x} \right\rangle = \frac{1}{2} \langle A \cdot \mathbf{x}, \mathbf{x} \rangle + \frac{1}{2} \langle A^T \cdot \mathbf{x}, \mathbf{x} \rangle = \langle A \cdot \mathbf{x}, \mathbf{x} \rangle.$$

□

Teorema 3.2. *Sejam $S \subset \mathbb{R}^n$ um convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Se para todo $\mathbf{x} \in S$, a parte simétrica de $J\mathbf{f}(\mathbf{x})$ é positiva-definida, então $\mathbf{f}$ é injetiva.*

Prova. Pelo Critério geral de injetividade (ver Corolário 2.8), é suficiente mostrar que toda matriz A pertencente à cobertura convexa de $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ é não-singular. Pelo Teorema de Carathéodory (ver Teorema 2.5), existem $t_1, \dots, t_{n^2+1} \geq 0$ com $t_1 + \dots + t_{n^2+1} = 1$ e $\mathbf{x}_1, \dots, \mathbf{x}_{n^2+1} \in S$ tais que $A = \sum_{k=1}^{n^2+1} t_k J\mathbf{f}(\mathbf{x}_k)$. Então, para todo $\mathbf{v} = (v_1, \dots, v_n) \neq \mathbf{0}$, temos que:

$$\begin{aligned} \mathbf{v} \left(\frac{A + A^T}{2} \right) \mathbf{v}^T &= \mathbf{v} \left(\frac{\sum_{k=1}^{n^2+1} t_k J\mathbf{f}(\mathbf{x}_k) + \sum_{k=1}^{n^2+1} t_k J\mathbf{f}(\mathbf{x}_k)^T}{2} \right) \mathbf{v}^T \\ &= \sum_{k=1}^{n^2+1} t_k \mathbf{v} \left(\frac{J\mathbf{f}(\mathbf{x}_k) + J\mathbf{f}(\mathbf{x}_k)^T}{2} \right) \mathbf{v}^T > 0, \end{aligned}$$

pois por hipótese a matriz $\frac{J\mathbf{f}(\mathbf{x}_k) + J\mathbf{f}(\mathbf{x}_k)^T}{2}$, sendo a parte simétrica de $J\mathbf{f}(\mathbf{x}_k)$, é positiva-definida. Isto mostra que a parte simétrica de A é positiva-definida. Falta mostrar que A é não-singular. Seja $(\lambda, \mathbf{v})$ um par autovalor-autovetor de A . Então, pelo Lema 3.2,

$$0 < \left\langle \frac{A + A^T}{2} \cdot \mathbf{v}, \mathbf{v} \right\rangle = \langle A \cdot \mathbf{v}, \mathbf{v} \rangle = \lambda |\mathbf{v}|^2.$$

Assim, os autovalores reais de A são positivos. Os autovalores complexos de A aparecem

em pares conjugados, portanto seu produto é positivo. O determinante de A , sendo produto dos autovalores de A , é positivo e assim, A é não-singular.

□

Como já foi observado, o Teorema 3.1 é um caso particular do Teorema 3.2.

Capítulo 4

Injetividade de aplicações com Jacobiano diagonalmente dominante

Neste capítulo vamos apresentar outro critério para que uma aplicação $\mathbf{f} : S \rightarrow \mathbb{R}^n$ seja injetiva, onde $S \subset \mathbb{R}^n$ é convexo. Em vez de exigir que a parte simétrica de $J\mathbf{f}(x)$ seja positiva-definida, vamos exigir que $J\mathbf{f}(\mathbf{x})$ tenha uma diagonal que domina as entradas fora da diagonal. Naturalmente, existem matrizes de diagonal estritamente dominante cuja parte simétrica não é positiva-definida.

Dizemos que $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) > 0$ se $\alpha_i > 0$ para todo $1 \leq i \leq n$.

Definição 4.1. Dizemos que $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ é de diagonal estritamente dominante por linhas com respeito a $\boldsymbol{\alpha} > 0$ se

$$\alpha_i |a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |a_{ij}| \quad \text{para todo } i \in \{1, \dots, n\}. \quad (4.1)$$

Dizemos que A é de diagonal estritamente dominante por colunas com respeito a $\boldsymbol{\alpha} > 0$ se

$$\alpha_j |a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^n \alpha_i |a_{ij}| \quad \text{para todo } j \in \{1, \dots, n\}. \quad (4.2)$$

Dizemos que A é de diagonal estritamente dominante se (4.1) ou (4.2) vale para algum $\boldsymbol{\alpha} > 0$.

Exercício 4.1. Considere as matrizes

$$A = \begin{pmatrix} -4 & 4 \\ 3 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} -4 & 4 \\ 5 & 3 \end{pmatrix}.$$

A matriz A é de diagonal estritamente dominante por linhas com respeito a $\alpha = (\frac{5}{4}, 1)$. Ela também é de diagonal estritamente dominante por colunas com respeito a $\alpha = (1, 1)$. A matriz B não é estritamente dominante por linhas. De fato, não existe $\alpha = (\alpha_1, \alpha_2) > 0$ tal que $4\alpha_1 > 4\alpha_2$ e $3\alpha_2 > 5\alpha_1$.

Vamos precisar do seguinte resultado sobre a distribuição espacial dos autovalores de uma matriz A .

Teorema 4.2 (Teorema de Gerschgorin). *Seja $A = (a_{ij})$ uma matriz quadrada de ordem $n \geq 2$ e seja*

$$\Lambda_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad 1 \leq i \leq n.$$

Se λ é autovalor real de A , então $|\lambda - a_{rr}| \leq \Lambda_r$ para algum $1 \leq r \leq n$.

Prova. Seja $(\lambda, \mathbf{x})$ um par autovalor-autovetor da matriz A , onde $\mathbf{x} = (x_1, \dots, x_n)$ é normalizado de tal forma que a máxima componente de $\mathbf{x}$ em valor absoluto é igual a 1. Como $A \cdot \mathbf{x} = \lambda \mathbf{x}$, temos que

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Logo,

$$\sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j = (\lambda - a_{ii})x_i, \quad 1 \leq i \leq n.$$

Em particular, se $|x_r| = 1$ então

$$|\lambda - a_{rr}| = |(\lambda - a_{rr})x_r| \leq \sum_{\substack{j=1 \\ j \neq r}}^n |a_{rj}x_j| = \Lambda_r.$$

□

Lema 4.1. *Se A é de diagonal estritamente dominante então A é não-singular.*

Prova. Suponha primeiramente que $A = (a_{ij})$ seja de diagonal estritamente dominante por linhas com relação a $\alpha = (\alpha_1, \dots, \alpha_n) > 0$. Seja

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |a_{ij}|, \quad 1 \leq i \leq n.$$

Então, de (4.1),

$$\alpha_i |a_{ii}| > r_i \quad \text{para todo } i \in \{1, \dots, n\}. \quad (4.3)$$

Considere a matriz

$$B = (b_{ij}) = \begin{pmatrix} \alpha_1 a_{11} & \cdots & \alpha_n a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_1 a_{n1} & \cdots & \alpha_n a_{nn} \end{pmatrix}.$$

Pelo Teorema 4.2, se λ é autovalor de B então

$$|\lambda - b_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |b_{ij}| \quad \text{para algum } i.$$

Isto é,

$$|\lambda - \alpha_i a_{ii}| \leq r_i \quad \text{para algum } i. \quad (4.4)$$

Segue de (4.3) e (4.4) que B não tem autovalor igual a zero e portanto, é não-singular. Como, $\det(B) = \alpha_1 \cdots \alpha_n \det A$, temos que A também é não-singular.

O caso em que A é de diagonal estritamente dominante por colunas reduz-se ao caso anterior, pois nesse caso a transposta de A é de diagonal estritamente dominante por linhas. $\square$

Seja $\mathcal{F} \subset \mathcal{M}_{n \times n}$ uma família de matrizes quadradas de ordem n . Dizemos que $\mathcal{F}$ é *uniformemente de diagonal estritamente dominante por linhas* (respectivamente, por colunas) se cada matriz $M \in \mathcal{F}$ for de diagonal estritamente dominante por linhas (respectivamente, por colunas) com relação ao mesmo $\alpha > 0$.

Teorema 4.3. *Sejam $S \subset \mathbb{R}^n$ um conjunto convexo e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Suponha que a família $\{J\mathbf{f}(x) : \mathbf{x} \in S\}$ seja uniformemente de diagonal estritamente dominante, então $\mathbf{f}$ é injetiva.*

Prova. Suponha que a família $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ seja uniformemente de diagonal estritamente dominante por linhas com respeito a $\alpha = (\alpha_1, \dots, \alpha_n) > 0$. Seja $a_{ij}(\mathbf{x}) = \frac{\partial f_i}{\partial x_j}(\mathbf{x})$, então $J\mathbf{f}(\mathbf{x}) = (a_{ij}(\mathbf{x}))$. Como $J\mathbf{f}(\mathbf{x})$ é de diagonal estritamente dominante por

linhas, temos que $a_{ii}(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in S$ e $1 \leq i \leq n$. Além disso, como S é conexo e $\mathbf{x} \mapsto a_{ii}(\mathbf{x})$ é contínua, temos que $a_{ii}(\mathbf{x}) > 0$ para todo $\mathbf{x} \in S$ ou $a_{ii}(\mathbf{x}) < 0$ para todo $\mathbf{x} \in S$. Em particular, se $\mathbf{x}_1, \dots, \mathbf{x}_{n+1} \in S$ e $t_1, \dots, t_{n+1} \geq 0$ então

$$\left| \sum_{k=1}^{n+1} t_k a_{ii}(\mathbf{x}_k) \right| = \sum_{k=1}^{n+1} t_k |a_{ii}(\mathbf{x}_k)| \quad \text{para todo } i \in \{1, \dots, n\}. \quad (4.5)$$

Pelo Corolário 2.8 e pelo Lema 4.1, é suficiente mostrar que toda matriz A pertencente à cobertura convexa de $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ é de diagonal estritamente dominante. Pelo Teorema 2.5, podemos supor que $A = (b_{ij}) = \sum_{k=1}^{n^2+1} t_k J\mathbf{f}(\mathbf{x}_k)$, onde $t_1, \dots, t_{n^2+1} \geq 0$, $t_1 + \dots + t_{n^2+1} = 1$ e $\mathbf{x}_1, \dots, \mathbf{x}_{n^2+1} \in S$. Pela hipótese inicial sobre a família $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ temos que

$$\alpha_i |a_{ii}(\mathbf{x})| > \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |a_{ij}(\mathbf{x})| \quad \text{para todo } i \in \{1, \dots, n\} \quad \text{e } \mathbf{x} \in S. \quad (4.6)$$

De (4.5) e (4.6), obtemos:

$$\begin{aligned} \alpha_i |b_{ii}| &= \alpha_i \left| \sum_{k=1}^{n^2+1} t_k a_{ii}(\mathbf{x}_k) \right| = \sum_{k=1}^{n^2+1} \alpha_i t_k |a_{ii}(\mathbf{x}_k)| > \sum_{k=1}^{n^2+1} t_k \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |a_{ij}(\mathbf{x}_k)| \\ &= \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j \sum_{k=1}^{n^2+1} t_k |a_{ij}(\mathbf{x}_k)| \geq \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j \left| \sum_{k=1}^{n^2+1} t_k a_{ij}(\mathbf{x}_k) \right| \\ &= \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |b_{ij}|. \end{aligned}$$

Isto mostra que A é de diagonal estritamente dominante, o que conclui a prova. O caso em que a família $\{J\mathbf{f}(\mathbf{x}) : \mathbf{x} \in S\}$ é uniformemente de diagonal estritamente dominante por colunas é análogo. $\square$

Exercício 4.4. Seja $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $\mathbf{f}(x, y) = (x^3 + 2x + 5y, 3x - 16y + \sin x)$. O domínio de $\mathbf{f}$ é convexo, $\mathbf{f}$ é C^1 e

$$J\mathbf{f}(x, y) = \begin{pmatrix} 3x^2 + 2 & 5 \\ 3 + \cos x & -16 \end{pmatrix}.$$

Note que $J\mathbf{f}(\mathbf{x})$ é de diagonal estritamente dominante por linhas com relação a $\boldsymbol{\alpha} = (3, 1)$ para todo $\mathbf{x} \in \mathbb{R}^2$. Pelo Teorema 4.3, $\mathbf{f}$ é injetiva.

Observe também que a parte simétrica de $J\mathbf{f}(x, y)$ é dada por

$$B = \begin{pmatrix} 3x^2 + 2 & 4 + \frac{1}{2}\cos x \\ 4 + \frac{1}{2}\cos x & -16 \end{pmatrix}.$$

Note que $\det B = -48 - 48x^2 - 4\cos x - \frac{1}{4}\cos^2 x < 0$. Isto mostra que a parte simétrica de $J\mathbf{f}(\mathbf{x})$ não é positiva-definida, o que torna impossível utilizar o procedimento descrito no Capítulo anterior.

Capítulo 5

Injetividade de aplicações com Jacobiano do tipo P -matriz

O critério de injetividade (Teorema 3.2) dado no Capítulo 3 exige que o jacobiano $J\mathbf{f}(\mathbf{x})$ da aplicação $\mathbf{f} : S \rightarrow \mathbb{R}^n$ seja positivo-quase-definido para todo $\mathbf{x} \in S$, isto é, que a parte simétrica de $J\mathbf{f}(\mathbf{x})$ seja positiva-definida para todo ponto $\mathbf{x}$ de um convexo S . Um corolário deste resultado afirma que se, para cada $\mathbf{x} \in S$, $J\mathbf{f}(\mathbf{x})$ for simétrica e todos os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ forem positivos, então $\mathbf{f}$ é injetiva (Teorema 3.1).

Nesta seção daremos condições suficientes para que uma aplicação $\mathbf{f} : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ seja injetiva sem exigir simetria de $J\mathbf{f}(\mathbf{x})$. A retirada da hipótese de simetria pede uma condição mais forte sobre os menores principais da matriz $J\mathbf{f}(\mathbf{x})$. Mais especificamente, a condição de que os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ sejam positivos será substituída pela condição de que todos os menores principais de $J\mathbf{f}(\mathbf{x})$, líderes ou não, sejam positivos. Para enunciar o resultado principal deste capítulo, vamos precisar das seguintes definições.

Dizemos que uma matriz $A \in \mathcal{M}_{n \times n}$ é uma P -matriz se todos os menores principais de A , líderes ou não, forem positivos. Note que uma P -matriz não precisa ser simétrica. Dizemos que $S \subset \mathbb{R}^n$ é um n -retângulo se S for o produto cartesiano de n subintervalos $I_1, \dots, I_n$ não-vazios de $\mathbb{R}$, isto é, $S = I_1 \times \dots \times I_n$. Se cada intervalo I_i for compacto (respectivamente, fechado), então dizemos que S é um n -retângulo compacto (respectivamente, um n -retângulo fechado).

O resultado principal deste capítulo é o seguinte.

Teorema 5.1 (Gale-Nikaidô). *Sejam $S \subset \mathbb{R}^n$ um n -retângulo fechado e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Se para todo $\mathbf{x} \in S$, $J\mathbf{f}(\mathbf{x})$ é uma P -matriz, então $\mathbf{f}$ é injetiva.*

O Corolário 5.1 afirma que toda matriz positiva-quase-definida é também uma P -matriz,

portanto as hipóteses sobre $J\mathbf{f}(x)$ no Teorema 5.1 incluem todos os casos cobertos pelos teoremas 3.1 e 3.2. Poderia se pensar então que o Teorema 5.1 é mais geral do que os resultados apresentados no Capítulo 3. Isto não é verdade porque enquanto no Teorema 3.2 o conjunto S pode ser qualquer convexo, no Teorema 5.1 S precisa ser um n -retângulo. A prova do Teorema 5.1 apresentada nesta seção segue a abordagem original dos autores do resultado (ver [7]). Vamos precisar de alguns lemas preparatórios para a demonstração.

Lema 5.1. *Se A é uma P -matriz, então A^T é uma P -matriz.*

Prova. Seja B uma submatriz principal de A^T , então B^T é submatriz principal de A e, portanto, $\det(B) = \det(B^T) > 0$. Logo A^T é uma P -matriz. $\square$

Lema 5.2. *Toda submatriz principal de uma P -matriz é uma P -matriz.*

Prova. Seja B uma submatriz principal de A . Seja B' uma submatriz principal de B . Como B' também é uma submatriz principal de A , segue que $\det B' > 0$, logo B é P -matriz. $\square$

Sejam $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ e A_{ij} a submatriz de A obtida pela eliminação da i -ésima linha e da j -ésima coluna de A . Chamamos *cofator* de a_{ij} o termo $(-1)^{i+j} \det A_{ij}$, denotado por a_{ij}^c . A *matriz dos cofatores* de A é a matriz $\text{cof } A = (a_{ij}^c)$. A *matriz adjunta* de A é a transposta da matriz de cofatores de A , isto é, $\text{adj } A = [\text{cof } A]^T$. Para calcular a inversa de uma matriz não-singular A , simplesmente dividimos a matriz adjunta de A pelo determinante de A , isto é,

$$A^{-1} = \frac{\text{adj } A}{\det A}.$$

Lema 5.3. *Seja A uma P -matriz, então A é não-singular e as entradas da diagonal principal de A^{-1} são positivas.*

Prova. Como A é P -matriz, A é não-singular. Seja $A^{-1} = (a_{ij}^*)$ a inversa de A , então

$$a_{ii}^* = \frac{(-1)^{2i} \det A_{ii}}{\det A} = \frac{\det A_{ii}}{\det A}, \quad 1 \leq i \leq n.$$

Como A e A_{ii} são submatrizes principais de A , temos que $\det A > 0$ e $\det A_{ii} > 0$ para todo $1 \leq i \leq n$. Assim, $a_{ii}^* > 0$. $\square$

Seja $\mathbf{x} = (x_1, \dots, x_n)$. Dizemos que $\mathbf{x} > \mathbf{0}$ (respectivamente, $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} < \mathbf{0}$, $\mathbf{x} \leq \mathbf{0}$) se $x_i > 0$ (respectivamente, $x_i \geq 0$, $x_i < 0$, $x_i \leq 0$) para todo $1 \leq i \leq n$.

Proposição 5.1. *Se A é uma P-matriz, então as desigualdades*

$$A \cdot \mathbf{x} \leq \mathbf{0} \quad \text{e} \quad \mathbf{x} \geq \mathbf{0} \quad (5.1)$$

admitem somente a solução trivial $\mathbf{x} = \mathbf{0}$.

Prova. Seja $A \in \mathcal{M}_{n \times n}$. Vamos proceder por indução sobre a ordem. O caso $n = 1$ é trivial. Suponha que a proposição seja válida para n . Provaremos que ela vale para $n + 1$. Sejam A uma P-matriz quadrada de ordem $n + 1$ e $\mathbf{x} \in \mathbb{R}^{n+1}$ tal que $\mathbf{x} \geq \mathbf{0}$ e $A\mathbf{x} \leq \mathbf{0}$. Pelo Lema 5.3, A é invertível e as entradas da diagonal de $A^{-1} = (a_{ij}^*)$ são positivas, ou seja, $a_{ii}^* > 0$ para todo $1 \leq i \leq n + 1$. Denote por $\mathbf{b}_j$ a j -ésima coluna da matriz A^{-1} , assim $A^{-1} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_{n+1}]$. Defina

$$\beta = \min \left\{ \frac{x_i}{a_{i1}^*} : a_{i1}^* > 0, 1 \leq i \leq n + 1 \right\}.$$

É claro que $\beta \geq 0$. Além disso, existe $k \in \{1, \dots, n + 1\}$ tal que $\beta = \frac{x_k}{a_{k1}^*}$. Defina o vetor $\mathbf{v} = \mathbf{x} - \beta \mathbf{b}_1$. Segue da definição de β que $\mathbf{v} \geq \mathbf{0}$ e $v_k = x_k - \beta a_{k1}^* = 0$, onde $\mathbf{v} = (v_1, \dots, v_{n+1})$. Como A é invertível, temos que

$$A[\mathbf{b}_1 \ \cdots \ \mathbf{b}_{n+1}] = AA^{-1} = [\mathbf{e}_1 \ \cdots \ \mathbf{e}_{n+1}] = I,$$

onde I denota a matriz identidade. Em particular, $A \cdot \mathbf{b}_1 = \mathbf{e}_1$, onde $\mathbf{e}_1^T = (1 \ 0 \ \cdots \ 0)$. Isto combinado com o fato que $A \cdot \mathbf{x} \leq \mathbf{0}$, e $\beta > 0$, implica que

$$A \cdot \mathbf{v} = A \cdot \mathbf{x} - \beta A \cdot \mathbf{b}_1 = A \cdot \mathbf{x} - \beta \mathbf{e}_1 \leq \mathbf{0}. \quad (5.2)$$

Seja B a submatriz principal de A obtida pela eliminação da k -ésima linha e da k -ésima coluna de A . Seja $\mathbf{v}^*$ o vetor de n coordenadas obtido pela eliminação da k -ésima coordenada ($v_k = 0$) do vetor $\mathbf{v}$. É claro que $B \cdot \mathbf{v}^* \leq \mathbf{0}$, e $\mathbf{v}^* \geq \mathbf{0}$. Do Lema 5.1, B é P-matriz de ordem n . Pela hipótese de indução, temos que $\mathbf{v}^* = \mathbf{0}$. Como $v_k = 0$, temos que $\mathbf{v} = \mathbf{0}$. Segue da equação (5.2) que $A \cdot \mathbf{x} = \beta \mathbf{e}_1$, portanto, $A \cdot \mathbf{x} \geq \mathbf{0}$. Como $A \cdot \mathbf{x} \leq \mathbf{0}$, temos que $A \cdot \mathbf{x} = \mathbf{0}$. Finalmente, como A é não-singular, então $\mathbf{x} = \mathbf{0}$. $\square$

Definição 5.1. *Sejam $A \in \mathcal{M}_{n \times n}$ e $\mathbf{x} = (x_1, \dots, x_n)$. Dizemos que A inverte o sinal de $\mathbf{x}$ se $(A \cdot \mathbf{x})_i x_i \leq 0$ para todo $i \in \{1, \dots, n\}$.*

Proposição 5.2. *A é uma P-matriz se e somente se A não inverte o sinal de nenhum vetor não-nulo.*

Prova. Primeiro provaremos que a condição que não inverte sinal é necessária. Suponha que A seja uma P-matriz que inverta o sinal do vetor $\mathbf{x} = (x_1, \dots, x_n) \neq 0$, isto é, que $(A \cdot \mathbf{x})_i x_i \leq 0$ para todo $i \in \{1, \dots, n\}$. Precisamos considerar dois casos.

Primeiro caso: $\mathbf{x} = (x_1, \dots, x_n) \geq \mathbf{0}$. Defina o conjunto $F = \{i : x_i > 0\}$. É claro que $F \neq \emptyset$, pois $\mathbf{x} \geq \mathbf{0}$ e $\mathbf{x} \neq \mathbf{0}$. Seja $\hat{A}$ a submatriz principal de A obtida pela eliminação de toda k -ésima linha e k -ésima coluna de A tal que $k \in \{1, \dots, n\} \setminus F$. Segue de $F \neq \emptyset$ que $\hat{A}$ tem ordem estritamente menor que n . Seja $\hat{\mathbf{x}} > \mathbf{0}$ o vetor obtido de $\mathbf{x}$ pela eliminação de todas as componentes nulas de $\mathbf{x}$, isto é, de todas as componentes x_k tais que $k \in \{1, \dots, n\} \setminus F$. Pelo Lema 5.2, como $\hat{A}$ é submatriz principal da P-matriz A , temos que $\hat{A}$ é P-matriz que inverte o sinal de $\hat{\mathbf{x}}$, isto é, $(\hat{A} \cdot \hat{\mathbf{x}})_i x_i \leq 0$ para todo $i \in \{1, \dots, n\} \setminus F$. Isto juntamente com $\hat{\mathbf{x}} > \mathbf{0}$ implica que $(\hat{A} \cdot \hat{\mathbf{x}})_i \leq 0$ para todo $i \in \{1, \dots, n\} \setminus F$. Pela Proposição 5.1, $\hat{\mathbf{x}} = \mathbf{0}$, logo $\mathbf{x} = \mathbf{0}$.

Segundo caso: $\mathbf{x} = (x_1, \dots, x_n)$ tem pelo menos uma componente $x_k < 0$. Seja G o conjunto não-vazio definido por $G = \{i : x_i < 0\}$. Seja D a matriz definida por

$$\begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix},$$

onde $d_i = -1$ se $i \in G$, e $d_i = 1$ se $i \notin G$. Em outras palavras, $d_i = -1$ se $x_i < 0$, e $d_i = 1$ se $x_i \geq 0$. Isto implica que

$$(D \cdot \mathbf{x})_i = d_i x_i \geq 0, \quad \forall i \in \{1, \dots, n\}.$$

Note que D é uma involução, isto é, $D^2 = I$. A matriz $B = DAD$ é uma P-matriz pois a matriz B foi obtida de A multiplicando um conjunto de linhas e o correspondente conjunto de colunas de A por -1 . Afirmamos que B inverte o sinal de $\mathbf{y} = D \cdot \mathbf{x} \geq \mathbf{0}$. De fato,

$$B \cdot \mathbf{y} = DADD \cdot \mathbf{x} = DAD^2 \cdot \mathbf{x} = DA \cdot \mathbf{x}. \quad (5.3)$$

Portanto,

$$(B \cdot \mathbf{y})_i y_i = (DA \cdot \mathbf{x})_i (D \cdot \mathbf{x})_i = d_i (A \cdot \mathbf{x})_i d_i x_i = d_i^2 (A \cdot \mathbf{x})_i x_i = (A \cdot \mathbf{x})_i x_i \leq 0.$$

Pelo primeiro caso, temos que $\mathbf{y} = \mathbf{0}$, logo $\mathbf{x} = \mathbf{0}$.

Para provar que a condição é suficiente, suponha que A seja uma P-matriz que não

inverte o sinal de nenhum vetor $\mathbf{x} = (x_1, \dots, x_n)$ exceto $\mathbf{x} = \mathbf{0}$. Isto quer dizer que se $(A\mathbf{x})_i x_i \leq 0$ para todo $1 \leq i \leq n$, então $\mathbf{x} = \mathbf{0}$.

Seja B qualquer submatriz principal de A . Afirmamos que $\det B > 0$. De fato, suponha que $\det B \leq 0$, então B tem um autovalor negativo λ correspondente a um autovetor $\mathbf{x} \neq \mathbf{0}$ (pois o determinante é o produto dos autovalores). Portanto,

$$(B \cdot \mathbf{x})_i x_i = (\lambda \mathbf{x})_i x_i = \lambda (x_i)^2 \leq 0, \quad \forall i \in \{1, \dots, n\}.$$

ou seja B inverte o sinal do vetor $\mathbf{x}$, o que implica que A inverte o sinal do vetor $\hat{\mathbf{x}}$. Daí segue que $\hat{\mathbf{x}} = \mathbf{0}$, então $\mathbf{x} = \mathbf{0}$, contradizendo o fato de $\mathbf{x}$ ser um autovetor. Assim, toda submatriz principal de A tem determinante positivo. Nestas condições, A é P-matriz. $\square$

Corolário 5.1. *Toda matriz $A \in \mathcal{M}_{n \times n}$ positiva-quase-definida é P-matriz, mas nem toda P-matriz é positiva-quase-definida.*

Prova. Seja $A \in \mathcal{M}_{n \times n}$ uma matriz positiva-quase-definida. Suponha por absurdo que A não seja P-matriz. Pela Proposição 5.2, existe $\mathbf{x} = (x_1, \dots, x_n) \neq \mathbf{0}$ tal que $(A\mathbf{x})_i x_i \leq 0$ para todo $1 \leq i \leq n$. Assim, pelo Lema 3.2, temos que

$$\left\langle \frac{A + A^T}{2} \cdot \mathbf{x}, \mathbf{x} \right\rangle = \langle A \cdot \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^n (A\mathbf{x})_i x_i \leq 0,$$

o que contradiz o fato de A ser positiva-quase-definida. Portanto, A é P-matriz. Para mostrar que a recíproca não vale considere a P-matriz

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

Observe que para todo $\mathbf{x} = (x_1, x_2)$ satisfazendo $x_1 = -x_2$,

$$\left\langle \frac{A + A^T}{2} \cdot \mathbf{x}, \mathbf{x} \right\rangle = (x_1)^2 + 2x_1x_2 + (x_2)^2 = (x_2)^2 - 2(x_2)^2 + (x_2)^2 = 0.$$

Portanto, A não é positiva-quase-definida. $\square$

Lema 5.4. *Uma matriz A é P-matriz se e somente se a sua inversa A^{-1} é P-matriz.*

Prova. Seja A uma P-matriz e A^{-1} a sua inversa. Seja $\mathbf{x} = (x_1, \dots, x_n)$ um vetor tal que $(A^{-1}\mathbf{x})_i x_i \leq 0$ para todo $i \in \{1, \dots, n\}$. Defina $\mathbf{y} = A^{-1}\mathbf{x}$. Então

$$(A\mathbf{y})_i y_i = (A^{-1}\mathbf{x})_i x_i \leq 0.$$

Pela Proposição 5.2, como A é P-Matriz, temos que $\mathbf{y} = \mathbf{0}$, logo $\mathbf{x} = \mathbf{0}$. Pela Proposição 5.2 novamente, temos que A^{-1} é P-matriz. $\square$

Lema 5.5. *Toda matriz de diagonal positiva e estritamente dominante é P-matriz.*

Prova. Seja $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ uma matriz de diagonal positiva ($a_{ii} > 0$) e estritamente dominante. Podemos assumir que A é de diagonal estritamente dominante por linhas (ver Definição 4.1), assim existe $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) > \mathbf{0}$ tal que

$$\alpha_i a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^n \alpha_j |a_{ij}| \quad \text{para todo } i \in \{1, \dots, n\}. \quad (5.4)$$

Primeiro provaremos que $\det A \neq 0$. Suponha que $\det A = 0$, isto é, que exista $\mathbf{x} = (x_1, \dots, x_n) \neq \mathbf{0}$, tal que $A\mathbf{x} = \mathbf{0}$. Isto é equivalente a

$$a_{ii}x_i = - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \quad \text{para todo } i \in \{1, \dots, n\}. \quad (5.5)$$

Defina

$$\rho = \max_i \frac{|x_i|}{\alpha_i} = \frac{|x_k|}{\alpha_k} \quad \text{para algum } k \in \{1, \dots, n\}. \quad (5.6)$$

O fato que $\mathbf{x} \neq \mathbf{0}$ implica que $\rho > 0$. Segue de (5.4), (5.5) e (5.6) que

$$a_{kk}|x_k| = |a_{kk}x_k| = \left| \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj}x_j \right| \leq \rho \sum_{\substack{j=1 \\ j \neq k}}^n \alpha_j |a_{kj}| < \rho \alpha_k a_{kk}.$$

Logo, como $a_{kk} > 0$, temos que $|x_k| < \rho \alpha_k$, isto é, $\rho < \rho$, o que é uma contradição. Isto mostra que $\det A \neq 0$. Vamos provar agora que, de fato, $\det A > 0$.

Segue do fato de A ter diagonal positiva estritamente dominante que a matriz $A + cI$ tem diagonal positiva e estritamente dominante para todo $c \geq 0$, onde I é a matriz identidade. Além disso,

$$\det(A + cI) = \det A + c^n + \phi(c),$$

onde ϕ é um polinômio em c de grau no máximo $n - 1$. Assim,

$$\lim_{c \rightarrow \infty} \det(A + cI) = \lim_{c \rightarrow \infty} c^n \left(\frac{\det A}{c^n} + 1 + \frac{\phi(c)}{c^n} \right) = \infty. \quad (5.7)$$

A função $c \in [0, \infty) \mapsto \det(A + cI)$ é contínua. Da primeira parte da prova, o fato de $A + cI$ ter diagonal positiva estritamente dominante garante que $\det(A + cI) \neq 0$ para todo $c \geq 0$. Isto juntamente com (5.7) implica que $\det(A + cI) > 0$ para todo $c \geq 0$. Em particular,

$\det A > 0$. Assim, toda matriz que tem diagonal positiva estritamente dominante tem determinante positivo.

Para concluir a prova, seja $\hat{A}$ submatriz principal de uma matriz A com diagonal positiva estritamente dominante. Então $\hat{A}$ também tem diagonal positiva estritamente dominante, logo $\det \hat{A} > 0$. Isto mostra que A é P-matriz. $\square$

Corolário 5.2. *Se A é uma P-matriz então existe um número $\lambda > 0$ tal que para todo vetor $\mathbf{x} \geq \mathbf{0}$ com norma $|\mathbf{x}| = 1$ alguma componente de $A\mathbf{x}$ é maior ou igual a λ .*

Prova. Dado $\mathbf{x} \in \mathbb{R}^n$, denote por $M(\mathbf{x}) = \max_i (A\mathbf{x})_i$ a máxima componente de $A\mathbf{x}$. Como $B = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ e } |\mathbf{x}| = 1\}$ é compacto e $\mathbf{x} \mapsto M(\mathbf{x})$ é contínua, a função M atinge o seu valor mínimo λ no conjunto B . Isto é,

$$\exists \lambda \in \mathbb{R} \mid \mathbf{x} \geq \mathbf{0}, |\mathbf{x}| = 1 \implies \max_i (A\mathbf{x})_i \geq \lambda.$$

Afirmamos que $\lambda > 0$. De fato, se $\lambda \leq 0$ então $A\mathbf{x} \leq \mathbf{0}$. Neste caso, pelo Teorema 5.1, $\mathbf{x} = \mathbf{0}$, o que contradiz $|\mathbf{x}| = 1$. $\square$

Proposição 5.3. *Seja A uma P-matriz, então as desigualdades*

$$A\mathbf{x} > \mathbf{0} \quad \text{e} \quad \mathbf{x} > \mathbf{0}$$

admitem pelo menos uma solução.

Prova. Ver [12, Theorem 2, p. 9]. $\square$

Proposição 5.4. *Sejam $S \subset \mathbb{R}^n$ um retângulo compacto n -dimensional e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 cujo jacobiano $J\mathbf{f}(\mathbf{x})$ é uma P-matriz para todo $\mathbf{x} \in S$. Então para cada $\mathbf{x}, \mathbf{a} \in S$, as desigualdades*

$$\mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{a}) \quad \text{e} \quad \mathbf{x} \geq \mathbf{a} \tag{5.8}$$

têm somente a solução $\mathbf{x} = \mathbf{a}$.

Prova. Procedamos por indução sobre n . Suponha que $n = 1$. Então a condição de que $Jf(x)$ é uma P-matriz implica que $Jf(x) = f'(x) > 0$ para todo $x \in S$, onde $S \subset \mathbb{R}$ é um intervalo. Então f é estritamente crescente. Assim, $x \geq a$ implica $f(x) \geq f(a)$. Adicionando a hipótese de que $f(x) \leq f(a)$, resulta $f(x) = f(a)$, isto é, $x = a$.

Seja X o conjunto solução de (5.8), isto é, suponha que

$$X = \{\mathbf{x} \in S \mid \mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{a}) \text{ e } \mathbf{x} \geq \mathbf{a}\}.$$

Claramente, X é um subconjunto fechado de S e, portanto, um compacto. Além disso, $\mathbf{a} \in X$. Para concluir a prova, basta provar que o conjunto $\widehat{X} = X \setminus \{\mathbf{a}\}$ é vazio.

Afirmamos que $\widehat{X}$ é compacto. Para provar isto, vamos mostrar que $\mathbf{a}$ é um ponto isolado de X . Como $\mathbf{f}$ é C^1 e, portanto, diferenciável, temos que

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \left(\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})}{|\mathbf{x} - \mathbf{a}|} - J\mathbf{f}(\mathbf{a}) \cdot \frac{\mathbf{x} - \mathbf{a}}{|\mathbf{x} - \mathbf{a}|} \right) = \mathbf{0}. \quad (5.9)$$

Como $J\mathbf{f}(\mathbf{a})$ é uma P-matriz, pela Proposição 5.2 existe $\lambda > 0$ tal que para todo $\mathbf{x} > \mathbf{a}$ alguma componente de $J\mathbf{f}(\mathbf{a}) \cdot \frac{\mathbf{x} - \mathbf{a}}{|\mathbf{x} - \mathbf{a}|}$ é maior ou igual a λ . Assim, de (5.9), existe uma vizinhança $U_{\mathbf{a}} \subset S$ de $\mathbf{a}$ tal que para todo $\mathbf{x} \in U_{\mathbf{a}} \setminus \{\mathbf{a}\}$, alguma componente de $\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})}{|\mathbf{x} - \mathbf{a}|}$ é positiva. Em particular, $U_{\mathbf{a}} \cap X = \{\mathbf{a}\}$, o que mostra que $\mathbf{a}$ é um ponto isolado de X . Assim, $\widehat{X} \setminus \{\mathbf{a}\}$ é fechado e, portanto, compacto.

Podemos assumir que $\widehat{X}$ é um compacto não-vazio, caso contrário a prova estaria terminada. Seja $\bar{\mathbf{x}}$ um ponto de mínimo global da função contínua $\theta(x_1, \dots, x_n) = x_1 + \dots + x_n$. Por definição, não existe nenhum ponto $\mathbf{y} \in \widehat{X}$ tal que $\mathbf{y} < \bar{\mathbf{x}}$. Como $\bar{\mathbf{x}} \in \widehat{X}$, temos que $\bar{\mathbf{x}} \geq \mathbf{a}$. Assim, há dois casos a considerar.

Primeiro caso: $\bar{\mathbf{x}} > \mathbf{a}$. Como $J\mathbf{f}(\bar{\mathbf{x}})$ é uma P-matriz, pela Proposição 5.3 existe um vetor $\mathbf{u} < \mathbf{0}$ tal que $J\mathbf{f}(\bar{\mathbf{x}})\mathbf{u} < 0$. Como $\bar{\mathbf{x}} > \mathbf{a}$, podemos escolher $\lambda > 0$ suficiente pequeno tal que $\mathbf{x}(\lambda) = \bar{\mathbf{x}} + \lambda\mathbf{u} > \mathbf{a}$. Em particular, $\mathbf{a} < \mathbf{x}(\lambda) < \bar{\mathbf{x}}$. Como S é um n -retângulo que contém $\mathbf{x}$ e $\mathbf{a}$, temos que $\mathbf{x}(\lambda)$ pertence à S para todo λ suficientemente pequeno. Além disso, como $\mathbf{f}$ é C^1 , temos que $\mathbf{f}$ é diferenciável e

$$\mathbf{f}(\mathbf{x}(\lambda)) = \mathbf{f}(\bar{\mathbf{x}}) + \lambda J\mathbf{f}(\bar{\mathbf{x}}) \cdot \mathbf{u} + o(\lambda|\mathbf{u}|).$$

Desta forma,

$$\frac{\mathbf{f}(\mathbf{x}(\lambda)) - \mathbf{f}(\bar{\mathbf{x}})}{\lambda|\mathbf{u}|} - J\mathbf{f}(\bar{\mathbf{x}}) \cdot \frac{\mathbf{u}}{|\mathbf{u}|}$$

pode ser tornado arbitrariamente pequeno escolhendo valores positivos de λ ainda menores. Assim, para λ arbitrariamente pequeno, $\mathbf{f}(\mathbf{x}(\lambda)) - \mathbf{f}(\bar{\mathbf{x}}) < \mathbf{0}$, isto é, $\mathbf{f}(\mathbf{x}(\lambda)) < \mathbf{f}(\bar{\mathbf{x}}) \leq \mathbf{f}(\mathbf{a})$. Isto é, $\mathbf{x}(\lambda) \in \widehat{X}$, contrariando a minimalidade de $\bar{\mathbf{x}}$. Esta contradição mostra que $\widehat{X} = \emptyset$.

Segundo caso: alguma componente de $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)$ é igual à correspondente com-

ponente de $\mathbf{a} = (a_1, \dots, a_n)$. Rearranjando as equações e variáveis, podemos assumir que $\bar{x}_1 = a_1$. Então defina uma nova aplicação $\widehat{\mathbf{f}} : \widehat{S} \rightarrow \mathbb{R}^{n-1}$ através da fórmula

$$\widehat{\mathbf{f}}(x_2, \dots, x_n) = (f_2(a_1, x_2, \dots, x_n), \dots, f_n(a_1, x_2, \dots, x_n)),$$

onde $\widehat{S} = \{(x_2, \dots, x_n) : (a_1, x_2, \dots, x_n) \in S\}$. Observe que $\widehat{S}$ é uma seção transversal do n -retângulo compacto S e, portanto, é um $(n-1)$ -retângulo compacto. A matriz jacobiana desta nova aplicação é

$$J\widehat{\mathbf{f}}(x_2, \dots, x_n) = \begin{pmatrix} \frac{\partial f_2}{\partial x_2}(\mathbf{z}) & \frac{\partial f_2}{\partial x_3}(\mathbf{z}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{z}) \\ \frac{\partial f_3}{\partial x_2}(\mathbf{z}) & \frac{\partial f_3}{\partial x_3}(\mathbf{z}) & \dots & \frac{\partial f_3}{\partial x_n}(\mathbf{z}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_2}(\mathbf{z}) & \frac{\partial f_n}{\partial x_3}(\mathbf{z}) & \dots & \frac{\partial f_n}{\partial x_n}(\mathbf{z}) \end{pmatrix},$$

onde $\mathbf{z} = (a_1, x_2, \dots, x_n) \in S$. Assim, para todo $(x_2, \dots, x_n) \in \widehat{S}$, $J\widehat{\mathbf{f}}(x_2, \dots, x_n)$ é submatriz principal de $J\mathbf{f}(\mathbf{z})$ obtida eliminando a primeira linha e a primeira coluna de $J\mathbf{f}(\mathbf{z})$. Pelo Lema 5.2, $J\widehat{\mathbf{f}}(x_2, \dots, x_n)$ é P-matriz para todo $(x_2, \dots, x_n) \in \widehat{S}$. Por hipótese, $\bar{x} \geq \mathbf{a}$ e $\mathbf{f}(\bar{x}) \leq \mathbf{f}(\mathbf{a})$, portanto $(\bar{x}_2, \dots, \bar{x}_n) \geq (a_2, \dots, a_n)$ e $\widehat{\mathbf{f}}(\bar{x}_2, \dots, \bar{x}_n) \leq \widehat{\mathbf{f}}(a_2, \dots, a_n)$. Desta forma, o ponto $(\bar{x}_2, \dots, \bar{x}_n)$ pertencente ao $(n-1)$ -retângulo compacto $\widehat{S} \subset \mathbb{R}^{n-1}$ satisfaz (5.8) com relação à função $\widehat{\mathbf{f}} : \widehat{S} \rightarrow \mathbb{R}^{n-1}$. Pela hipótese de indução, $(\bar{x}_2, \dots, \bar{x}_n) = (a_2, \dots, a_n)$. Como $a_1 = \bar{x}_1$, segue que $\bar{\mathbf{x}} = \mathbf{a}$, o que contradiz o fato de que $\bar{\mathbf{x}} \in \widehat{X}$. Logo $\bar{\mathbf{x}} = \emptyset$.

□

Teorema 5.1 (Gale-Nikaidô) *Sejam $S \subset \mathbb{R}^n$ um n -retângulo fechado e $\mathbf{f} : S \rightarrow \mathbb{R}^n$ uma aplicação C^1 . Se para todo $\mathbf{x} \in S$, $J\mathbf{f}(\mathbf{x})$ é uma P-matriz, então $\mathbf{f}$ é injetiva.*

Prova. Como todo n -retângulo fechado é união enumerável de n -retângulos compactos encaixados, podemos assumir que S é um n -retângulo compacto.

Sejam $\mathbf{a}, \mathbf{b} \in S$ tais que $\mathbf{f}(\mathbf{a}) = \mathbf{f}(\mathbf{b})$. Provaremos que $\mathbf{a} = \mathbf{b}$. Reordenando as variáveis se necessário, podemos supor que $\mathbf{a} = (a_1, \dots, a_n)$ e $\mathbf{b} = (b_1, \dots, b_n)$ satisfazem

$$a_i \leq b_i \ (i \leq k) \quad \text{e} \quad a_i > b_i \ (i > k),$$

para algum $k \in \{1, \dots, n\}$. Se $k = n$, então $\mathbf{f}(\mathbf{a}) = \mathbf{f}(\mathbf{b})$ e $\mathbf{a} \leq \mathbf{b}$. Neste caso, as condições

da Proposição 5.4 são satisfeitas e então $\mathbf{a} = \mathbf{b}$. O caso $k = 0$ é análogo. Se $0 < k < n$, defina a aplicação $D : \mathbb{R}^n \rightarrow \mathbb{R}^n$ por

$$D(x_1, \dots, x_n) = (x_1, \dots, x_k, -x_{k+1}, \dots, -x_n).$$

Note que D é injetiva. Além disso, D é uma involução, isto é, $D^2 = I$, ou seja, $D^{-1} = D$. Observe também que $D(S)$ é novamente um n -retângulo fechado. Sejam $\mathbf{a}^* = D(\mathbf{a})$ e $\mathbf{b}^* = D(\mathbf{b})$. Seja $\mathbf{g} : D(S) \rightarrow \mathbb{R}^n$ a aplicação definida por $\mathbf{g} = D \circ \mathbf{f} \circ D = D \circ \mathbf{f} \circ D^{-1}$. Como D é uma transformação linear, temos que o Jacobiano de g no ponto $\mathbf{x}$ é $J\mathbf{g}(\mathbf{x}) = DJ\mathbf{f}(\mathbf{x})D$ para todo $\mathbf{x} \in S$. Em outras palavras, a matriz $J\mathbf{g}(\mathbf{x})$ é uma P-matriz pois ela foi obtida de $J\mathbf{f}(\mathbf{x})$ multiplicando um conjunto de linhas e o correspondente conjunto de colunas de D por -1 . Pela definição de $\mathbf{g}$ temos que $\mathbf{g}(\mathbf{a}^*) = \mathbf{g}(\mathbf{b}^*)$ e $\mathbf{a}^* \leq \mathbf{b}^*$. Assim, pela Proposição 5.4, $\mathbf{a}^* = \mathbf{b}^*$, o que implica em $\mathbf{a} = \mathbf{b}$. $\square$

Capítulo 6

Exemplos

O propósito deste capítulo é mostrar como os resultados dos capítulos anteriores podem ser utilizados para garantir a injetividade de uma aplicação C^1 $\mathbf{f} : S \rightarrow \mathbb{R}^n$ em um subconjunto convexo $S \subset \mathbb{R}^n$.

Exemplo 6.1 ([12, p. 22]). *Seja $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a função definida por*

$$f(x, y, z) = (x^2 + z^2, x^2 + y^2, y^2 + z^2).$$

Então $\mathbf{f}$ é injetiva em $\mathbb{R}_+^3 = \{(x, y, z) : x > 0, y > 0 \text{ e } z > 0\}$ mas não é injetiva em $\mathbb{R}^3$.

Prova. A igualdade $\mathbf{f}(1, 1, 1) = \mathbf{f}(-1, -1, -1)$ implica que $\mathbf{f}$ não é injetiva em $\mathbb{R}^3$. Seja

$$A_k = \left[\frac{1}{k}, k\right] \times \left[\frac{1}{k}, k\right] \times \left[\frac{1}{k}, k\right].$$

Então A_k é um 3-retângulo compacto. Além disso, $A_1 \subset A_2 \subset \dots$ é uma sequência encaixada de retângulos compactos tais que $\bigcup_{k=1}^{\infty} A_k = \mathbb{R}_+^3$. Assim, para provar que $\mathbf{f}$ é injetiva em $\mathbb{R}_+^3$ é suficiente provar que $\mathbf{f}$ é injetiva em cada A_k . Isto pode ser feito estudando o jacobiano de $\mathbf{f}$ em cada ponto de A_k . Para cada $(x, y, z) \in A_k$, temos que

$$J\mathbf{f}(x, y, z) = \begin{pmatrix} 2x & 0 & 2z \\ 2x & 2y & 0 \\ 0 & 2y & 2z \end{pmatrix}.$$

Os menores principais de $J\mathbf{f}(x, y, z)$ são:

$$2x, 2y, 2z, \begin{vmatrix} 2x & 0 \\ 2x & 2y \end{vmatrix} = 4xy, \begin{vmatrix} 2x & 2z \\ 0 & 2z \end{vmatrix} = 4xz, \begin{vmatrix} 2y & 0 \\ 2y & 2z \end{vmatrix} = 4yz, \begin{vmatrix} 2x & 0 & 2x, \\ 2x & 2y & 0 \\ 0 & 2y & 2z \end{vmatrix} = 16xyz.$$

Como todos os menores principais são positivos em $A_k \subset \mathbb{R}_3^+$, temos que $J\mathbf{f}(x, y, z)$ é P-matriz. Pelo Teorema 5.1, $\mathbf{f}$ é injetiva em A_k e, portanto, em $\mathbb{R}_3^+$.

Observe que as conclusões alcançadas acima não poderiam ser obtidas através dos resultados do Capítulo 3. De fato, a parte simétrica do jacobiano de $\mathbf{f}$ é dada por:

$$\frac{J\mathbf{f}(x, y, z) + [J\mathbf{f}(x, y, z)]^T}{2} = \begin{pmatrix} 2x & x & z \\ x & 2y & y \\ z & y & 2y \end{pmatrix}.$$

Note que o menor principal líder

$$\begin{vmatrix} 2x & x \\ x & 2y \end{vmatrix} = 4xy - x^2$$

é negativo no ponto $P = (4, 1/2, 1) \in \mathbb{R}_3^+$. Portanto, $J\mathbf{f}(4, 1/2, 1)$ não é positiva-quase-definida, o que inviabiliza o uso do Teorema 3.2. $\square$

Exemplo 6.2 ([12, p. 26]). *Seja $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a função definida por*

$$f(x, y) = \left(\frac{x^3}{3} - \frac{x}{2} - y, x + y \right),$$

então $\mathbf{f}$ é injetiva.

Prova. O jacobiano de $\mathbf{f}$ é dado por

$$J\mathbf{f}(x, y) = \begin{pmatrix} x^2 - \frac{1}{2} & -1 \\ 1 & 1 \end{pmatrix}.$$

Claramente, $J\mathbf{f}(x, y)$ não é uma P-matriz quando $x^2 < \frac{1}{2}$. Seja

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \text{então} \quad LJ\mathbf{f}(x, y) = \begin{pmatrix} x^2 + \frac{1}{2} & 0 \\ 1 & 1 \end{pmatrix}.$$

A matriz $LJ\mathbf{f}(x, y)$ é claramente uma P-matriz para todo $(x, y) \in \mathbb{R}^2$, portanto $L\mathbf{f}$ é injetiva pelo Teorema 5.1. Logo, como L é um isomorfismo, temos que $\mathbf{f}$ também é injetiva. $\square$

Exemplo 6.3 ([7, p. 82]). Seja $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a função definida por

$$\mathbf{f}(x, y) = (e^{2x} - y^2 + 3, 4e^{2x}y - y^3).$$

Os menores principais líderes de $J\mathbf{f}(\mathbf{x})$ são positivos para todo $\mathbf{x} \in [-2, 2] \times [-2, 2]$ mas a função não é injetiva.

Prova. Como $f(0, 2) = (0, 0) = f(0, -2)$, temos que $\mathbf{f}$ não é injetiva. O jacobiano de $\mathbf{f}$ é dado por

$$J\mathbf{f}(x, y) = \begin{pmatrix} 2e^{2x} & -2y \\ 8e^{2x}y & 4e^{2x} - 3y^2 \end{pmatrix}.$$

Os menores principais líderes de $J\mathbf{f}(x, y)$ são

$$2e^{2x} > 0, \quad |J\mathbf{f}(x, y)| = 2e^{2x}(4e^{2x} + 5y^2) > 0.$$

Entretanto, o Teorema 3.1 não se aplica porque $J\mathbf{f}(x, y)$ não é simétrica. $\square$

Exemplo 6.4 ([1, p. 2]). A aplicação $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$\mathbf{f}(x, y) = (x + x^3y^2, y + x^2y^3)$$

é injetiva.

Prova. De fato, o jacobiano de $\mathbf{f}$ é

$$J\mathbf{f}(x, y) = \begin{pmatrix} 1 + 3x^2y^2 & 2x^3y \\ 2xy^3 & 1 + 3x^2y^2 \end{pmatrix}.$$

Os menores principais de $J\mathbf{f}(x, y)$ são:

$$1 + 3x^2y^2 > 0, \quad 1 + 3x^2y^2 > 0, \quad |J\mathbf{f}(x, y)| = 1 + 6x^2y^2 + 5x^4y^4 > 0,$$

portanto $\mathbf{f}(x, y)$ é uma P-matriz para todo $(x, y) \in \mathbb{R}^2$. Pelo Teorema 5.1, $\mathbf{f}$ é injetiva. $\square$

Referências Bibliográficas

- [1] L. A. Campbell, *Remarks on the real Jacobian Conjecture and Samuelson Maps*, App. Math. Lett. Vol. 10, No. 4, 1–3, 1997.
- [2] L. A. Campbell, *Decomposing Samuelson Maps*, Linear Algebra and Its Applications Vol. 187, 227–238, 1993.
- [3] L. A. Campbell, *Samuelson Maps - Recent Results*, Mathl. Comput. Modelling **17**(12) 3–8, 1993.
- [4] L.A. Campbell, *Rational Samuelson maps are univalent*, Journal of Pure and Applied Algebra **92**(3), 227–240, 1994.
- [5] M. Choquet, *Equations différentielles*, Centre de Documentation Universitaires, Paris, 1961.
- [6] B. A. Coomes, *On conditions sufficient for injectivity of maps*, IMA Preprint Series 544, University of Minnesota, 1989.
- [7] D. Gale e H. Nikaidô, *The Jacobian Matrix and Global Univalence of Mappings*, Math. Annalen **159**, 811–93, 1965.
- [8] R. McLeod, *Mean value theorem for vector valued functions*, Proc. Edinburgh Math. Soc. (2) 14 (1965), 197–209.
- [9] G. Meisters e C. Olech, *A Jacobian condition for injectivity of differentiable plane maps*, Ann. Polon. Math. **51**, 249–254, 1990.
- [10] P. Morales, *Mean value theorem for the m -integral of Dinculeanu*, Canad. Math. Bull. **15**(2), 1972.
- [11] S. Nickel, *Convex analysis*, University of Kaiserslautern, 1998.
- [12] T. Parthasarathy, *On Global univalence theorems*, Springer, 1983.
- [13] P. A. Samuelson, *Prices of factors and goods in general equilibrium*, Review of Economic Studies **21**, 1–20, 1953.
- [14] W. Rudin, *Principles of mathematical analysis*, third edition, McGraw-Hill, 1976.
- [15] R. Výborný, *Mean value theorems and a Taylor theorem for vector valued functions*, Bull. Austral. Math. Soc. **24** (1981), 69–77.

Autorizo a reprodução xerográfica para fins de pesquisa.

São José do Rio Preto, 27 de Janeiro de 2015.

Lizet Santa Cruz Calderón