

## CDI-II

### Derivadas de Ordem Superior. Extremos

#### 1 Derivadas de Ordem Superior

Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , definida num aberto  $D$ , uma função de classe  $C^1$  e consideremos as respectivas derivadas parciais

$$\frac{\partial f}{\partial x_k} ; k = 1, 2, \dots, n.$$

Note-se que estas derivadas são também funções escalares definidas em  $D$ . Portanto, se forem diferenciáveis podemos considerar as respectivas derivadas parciais.

Assim, teremos as funções

$$\frac{\partial}{\partial x_j} \left( \frac{\partial}{\partial x_k} \right) ; j = 1, 2, \dots, n ; k = 1, 2, \dots, n$$

que são as derivadas parciais de ordem dois (ou de segunda ordem) de  $f$ . Por convenção, serão designadas por

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j} \left( \frac{\partial}{\partial x_k} \right) ; \text{ se } j \neq k,$$

e por

$$\frac{\partial^2 f}{\partial x_k^2} = \frac{\partial}{\partial x_k} \left( \frac{\partial}{\partial x_k} \right) ; \text{ se } j = k.$$

Se as derivadas de ordem dois forem funções diferenciáveis, podemos também considerar as respectivas derivadas parciais

$$\frac{\partial}{\partial x_i} \left( \frac{\partial^2 f}{\partial x_i \partial x_j \partial x_k} \right) ; i = 1, 2, \dots, n ; j = 1, 2, \dots, n ; k = 1, 2, \dots, n,$$

ou seja, as derivadas parciais de ordem três de  $f$ , que serão designadas por

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}.$$

**Exemplo 1.1** Seja  $f(x, y) = xy^2 + yx^3$ . Então, as derivadas parciais de ordem um serão as funções

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= y^2 + 3yx^2 \\ \frac{\partial f}{\partial y}(x, y) &= 2xy + x^3.\end{aligned}$$

As derivadas parciais de ordem dois serão as funções

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2}(x, y) &= \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) (x, y) = 6xy \\ \frac{\partial^2 f}{\partial y^2}(x, y) &= \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) (x, y) = 2x \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) (x, y) = 2y + 3x^2 \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) &= \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) (x, y) = 2y + 3x^2\end{aligned}$$

e algumas de ordem três serão

$$\begin{aligned}\frac{\partial^3 f}{\partial x^3}(x, y) &= \frac{\partial}{\partial x} \left( \frac{\partial^2 f}{\partial x^2} \right) (x, y) = 6y \\ \frac{\partial^3 f}{\partial y \partial x^2}(x, y) &= \frac{\partial}{\partial y} \left( \frac{\partial^2 f}{\partial x^2} \right) (x, y) = 6x \\ \frac{\partial^3 f}{\partial y^3}(x, y) &= \frac{\partial}{\partial y} \left( \frac{\partial^2 f}{\partial y^2} \right) = 0 \\ \frac{\partial^3 f}{\partial x \partial y^2}(x, y) &= \frac{\partial}{\partial x} \left( \frac{\partial^2 f}{\partial y^2} \right) = 2 \\ \frac{\partial^3 f}{\partial x^2 \partial y}(x, y) &= \frac{\partial}{\partial x} \left( \frac{\partial^2 f}{\partial x \partial y} \right) = 6x\end{aligned}$$

Diz-se que uma função  $f$  é de classe  $C^k$  se as derivadas parciais de ordem menor ou igual a  $k$  existirem e forem funções contínuas.

Diz-se que  $f$  é de classe  $C^\infty$  se for de classe  $C^k$  para qualquer  $k \in \mathbb{N}$ .

Note-se que as derivadas parciais de ordem dois da função do exemplo anterior,  $\frac{\partial^2 f}{\partial x \partial y}$  e  $\frac{\partial^2 f}{\partial y \partial x}$ , são iguais.

Esta coincidência não acontece por acaso. De facto temos

**Teorema 1.1 (Schwarz)** *Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^2$  no aberto  $D$ . Então*

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}.$$

A demonstração deste teorema pode ser vista na bibliografia da disciplina (c.f. [2, 3, 1]). É claro que basta considerar o caso em  $\mathbb{R}^2$  e notar que,

$$\frac{\frac{f(a+h,b+k)-f(a,b+k)}{h} - \frac{f(a+h,b)-f(a,b)}{h}}{k} = \frac{\frac{f(a+h,b+k)-f(a+h,b)}{k} - \frac{f(a,b+k)-f(a,b)}{k}}{h}.$$

## 2 Extremos de Funções Escalares

Uma forma bastante conveniente de analisar o comportamento de uma função escalar num ponto é a de restringir a uma linha recta que passe por esse ponto. Foi deste modo que se introduziu a noção de derivada direccional segundo um vector.

Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^1$  no aberto  $D$ .

Consideremos a recta que passa pelo ponto  $a$  e tem a direcção do vector  $h$ , ou seja a linha descrita pela função  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$  definida por

$$\gamma(t) = a + th.$$

Note-se que  $\gamma(0) = a$  e  $\gamma(1) = a + h$ .

Seja  $x = \gamma(t)$  um ponto desta recta e tal que o segmento de recta entre  $a$  e  $x$  esteja contido em  $D$ .

Então, teremos  $f(x) = f(\gamma(t))$  e a função  $f$  passa a ser analisada apenas na recta que passa pelo ponto  $a$  com a direcção do vector  $h$ , recorrendo à função composta

$$\begin{array}{ccccc} \mathbb{R} & \xrightarrow{\gamma} & \mathbb{R}^n & \xrightarrow{f} & \mathbb{R} \\ t & \mapsto & \gamma(t) & \mapsto & f(\gamma(t)) \end{array}$$

que é uma função real de variável real que designaremos por  $g$ , ou seja

$$g(t) = f(\gamma(t)).$$

É claro que  $\gamma$  é de classe  $C^1$  e  $\gamma'(t) = h$ . Portanto,

$$g'(t) = \nabla f(\gamma(t)) \bullet \gamma'(t) = \nabla f(\gamma(t)) \bullet h,$$

e para  $t = 0$ , teremos

$$g'(0) = \nabla f(\gamma(0)) \bullet \gamma'(0),$$

ou seja,

$$g'(0) = \nabla f(a) \bullet h.$$

Sendo  $g$  de classe  $C^1$ , pelo teorema de Lagrange para funções reais de variável real, existirá  $t_0 \in ]0, 1[$  tal que

$$g(1) - g(0) = g'(t_0),$$

ou seja,

$$f(a+h) - f(a) = \nabla f(c) \bullet h$$

em que  $c = \gamma(t_0)$  é um ponto no segmento de recta entre  $a$  e  $a+h$ .

**Teorema 2.1 (Lagrange)** *Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^1$  no aberto  $D$  e sejam  $a$  e  $a+h$  dois pontos em  $D$  tais que o segmento de recta entre eles esteja contido em  $D$ . Então existe um ponto  $c$  nesse segmento de recta tal que*

$$f(a+h) - f(a) = \nabla f(c) \bullet h,$$

com  $c$  distinto de  $a$  e de  $a+h$ .

Seja  $a \in D$  e consideremos uma bola  $B_\epsilon(a) \subset D$  tal que  $\nabla f(x) = 0$  para qualquer ponto  $x \in B_\epsilon(a)$ .

Pelo teorema de Lagrange, teremos  $f(a+h) = f(a)$  para qualquer vector  $h$  tal que  $a+h \in B_\epsilon(a)$ , ou seja, a função  $f$  será constante na bola  $B_\epsilon(a)$ .

Portanto, **uma função de classe  $C^1$  e com gradiente nulo numa bola será constante nessa bola.**

**Definição 2.1 (Ponto Crítico)** *Diz-se que  $a \in D$  é um ponto crítico da função  $f$  se  $\nabla f(a) = 0$ .*

Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^2$  e seja  $a \in D$  um ponto crítico de  $f$ .

Consideremos a recta que passa em  $a$  e com a direcção de um vector  $h \in \mathbb{R}^n$ , ou seja, o conjunto de pontos da forma  $a+th$  com  $t \in \mathbb{R}$ .

Tal como acima, seja  $\gamma(t) = a+th$  e consideremos a função composta

$$\begin{array}{ccccc} \mathbb{R} & \xrightarrow{\gamma} & \mathbb{R}^n & \xrightarrow{f} & \mathbb{R} \\ t & \mapsto & \gamma(t) & \mapsto & f(\gamma(t)). \end{array}$$

Sendo  $a$  um ponto crítico, é claro que  $g'(0) = \nabla f(a) = 0$ .

Pela fórmula de Taylor para funções reais de variável real teremos

$$g(t) - g(0) = g'(0) + \frac{1}{2!}g''(0)t^2 + o(t^2) = \frac{1}{2!}g''(0)t^2 + o(t^2),$$

ou seja,

$$\frac{g(t) - g(0)}{t^2} = \frac{1}{2!}g''(0) + \frac{o(t^2)}{t^2}. \quad (1)$$

Sabendo que  $\lim_{t \rightarrow 0} \frac{o(t^2)}{t^2} = 0$ , para  $t$  suficientemente próximo de zero, a diferença  $g(t) - g(0)$  tem o mesmo sinal da derivada  $g''(0)$ .

Note-se que

$$g'(t) = \nabla f(\gamma(t)) \bullet h = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(\gamma(t))h_k$$

e, portanto,

$$g''(t) = \sum_{k=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(\gamma(t))h_j h_k,$$

ou seja,

$$g''(0) = \sum_{k=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(a)h_j h_k.$$

À matriz com  $n$  linhas e  $n$  colunas cujas entradas são as derivadas parciais de ordem dois, designada pelo símbolo  $D^2 f(a)$ , ou seja,

$$D^2 f(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix}$$

chama-se **matriz Hessiana** de  $f$  no ponto  $a$ .

Assim, a derivada  $g''(0)$  poderá ser apresentada na forma matricial

$$g''(0) = h^T D^2 f(a)h$$

ou na forma vectorial

$$g''(0) = h \bullet D^2 f(a)h.$$

Portanto, da fórmula de Taylor (1), obtemos

$$\frac{f(a + th) - f(a)}{t^2} = \frac{1}{2!}h \bullet D^2 f(a)h + \frac{o(t^2)}{t^2}.$$

Seja  $\lambda \in \mathbb{R}$  um valor próprio da matriz Hessiana  $D^2f(a)$  e  $h \neq 0$  um vector próprio associado a  $\lambda$ , ou seja,

$$D^2f(a)h = \lambda h.$$

Então, teremos

$$g''(0) = h \bullet D^2f(a)h = \lambda h \bullet h = \lambda \|h\|^2$$

e, portanto, o sinal de  $g''(0)$  será o sinal do valor próprio  $\lambda$ .

Portanto, se  $a$  for um ponto crítico de  $f$ , na direcção do vector próprio  $h$  associado ao valor próprio  $\lambda$  da matriz Hessiana  $D^2f(a)$ , teremos

$$\frac{f(a+th) - f(a)}{t^2} = \frac{1}{2!}\lambda \|h\|^2 + \frac{o(t^2)}{t^2}$$

Note-se que, pelo teorema de Schwarz, a matriz Hessiana é simétrica e, por isso, é diagonalizável, os respectivos valores próprios são números reais e os correspondentes vectores próprios constituem uma base ortonormada de  $\mathbb{R}^n$ .

Assim, para classificar os pontos críticos devemos analisar o comportamento da função nas linhas rectas determinadas pelos vectores próprios através dos sinais dos correspondentes valores próprios da matriz Hessiana  $D^2f(a)$ .

A uma linha recta determinada por um vector próprio chamaremos direcção própria ou direcção singular.

Podem ocorrer as situações seguintes.

- a) Os valores próprios de  $D^2f(a)$  são todos positivos:  $a$  é um ponto de mínimo de  $f$ .
- b) Os valores próprios de  $D^2f(a)$  são todos negativos:  $a$  é um ponto de máximo de  $f$ .
- c) A matriz Hessiana  $D^2f(a)$  tem pelo menos um valor próprio positivo e pelo menos um negativo:  $a$  não é um extremo de  $f$ . (Por vezes chamado ponto de sela)
- d) A matriz Hessiana  $D^2f(a)$  tem pelo menos um valor próprio nulo e os restantes têm o mesmo sinal. Neste caso, a função  $f$  deve ser analisada nas direcções próprias associadas aos valores próprios nulos recorrendo às derivadas de ordem superior a dois.

No último caso, esta análise pode não ser conclusiva. Então só um estudo directo do comportamento da função nas vizinhanças de  $a$  poderá esclarecer o problema.

### 3 Exemplos

Nos exemplos seguintes iremos determinar e classificar os pontos críticos de cada uma das funções.

**Exemplo 3.1** Consideremos a função  $f(x, y) = x^2 + y^2$ . É claro que  $f$  é, pelo menos, de classe  $C^2$ .

a) Pontos Críticos:  $\nabla f(x, y) = (0, 0)$ .

$$\nabla f(x, y) = (2x, 2y) = (0, 0) \Leftrightarrow (x, y) = (0, 0).$$

A origem é o único ponto crítico.

b) Classificação do ponto crítico  $(0, 0)$ .

A matriz Hessiana

$$D^2 f(0, 0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0, 0) & \frac{\partial^2 f}{\partial y \partial x}(0, 0) \\ \frac{\partial^2 f}{\partial x \partial y}(0, 0) & \frac{\partial^2 f}{\partial y^2}(0, 0) \end{bmatrix}_{(0,0)} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

apresenta dois valores próprios positivos  $\lambda_1 = \lambda_2 = 2$  e, portanto, o ponto crítico  $(0, 0)$  é um ponto de mínimo de  $f$ .

Note-se que esta análise é desnecessária dado que  $f(x, y) = x^2 + y^2 \geq 0$  e a origem é o único ponto em que  $f$  é nula. Na figura 1 encontra-se o gráfico desta função.

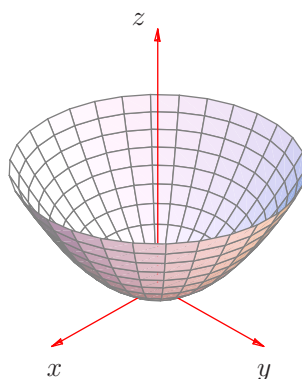


Figura 1: Exemplo de ponto de mínimo:  $f(x, y) = x^2 + y^2$

**Exemplo 3.2** Consideremos a função  $f(x, y) = x^2 - y^2$ . É claro que  $f$  é, pelo menos, de classe  $C^2$ .

a) Pontos Críticos:  $\nabla f(x, y) = (0, 0)$ .

$$\nabla f(x, y) = (2x, -2y) = (0, 0) \Leftrightarrow (x, y) = (0, 0).$$

A origem é o único ponto crítico.

b) Classificação do ponto crítico  $(0, 0)$ .

A matriz Hessiana

$$D^2 f(0, 0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0, 0) & \frac{\partial^2 f}{\partial y \partial x}(0, 0) \\ \frac{\partial^2 f}{\partial x \partial y}(0, 0) & \frac{\partial^2 f}{\partial y^2}(0, 0) \end{bmatrix}_{(0,0)} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

apresenta um valor próprio positivo  $\lambda_1 = 2$  e um valor próprio negativo  $\lambda_2 = -2$  e, portanto, o ponto crítico  $(0, 0)$  não é um extremo de  $f$ .

Neste caso dizemos que é um ponto de sela. Na figura 2 encontra-se o gráfico de  $f$  que ilustra e justifica a designação de ponto de sela.

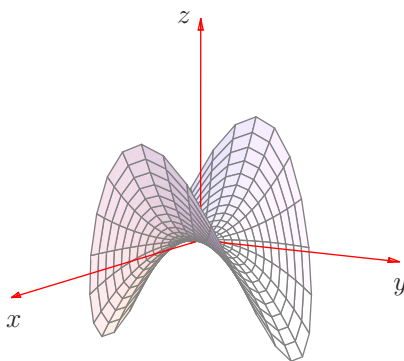


Figura 2: Exemplo de ponto de sela:  $f(x, y) = x^2 - y^2$

Note-se que na direcção em que  $y = 0$  a função apresenta um mínimo e na direcção  $x = 0$  a função apresenta um máximo na origem. Trata-se de um ponto de sela.

**Exemplo 3.3** Consideremos a função  $f(x, y) = (x - y)^2 - x^4 - y^4$ .

a) Pontos Críticos:  $\nabla f(x, y) = (0, 0)$ .

$$\nabla f(x, y) = (2(x - y) - 4x^3, -2(x - y) - 4y^3) = (0, 0) \Leftrightarrow \begin{cases} x - y - 2x^3 = 0 \\ -(x - y) - 2y^3 = 0 \end{cases}$$

ou seja,

$$\begin{cases} x - y - 2x^3 = 0 \\ x^3 + y^3 = 0 \end{cases} \Leftrightarrow \begin{cases} x - y - 2x^3 = 0 \\ y = -x \end{cases}$$

donde se conclui que os pontos críticos são:  $(0, 0)$ ,  $(-1, 1)$ ,  $(1, -1)$ .

Para os classificar recorreremos à matriz Hessiana

$$D^2f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{bmatrix} = \begin{bmatrix} 2 - 12x^2 & -2 \\ -2 & 2 - 12y^2 \end{bmatrix}$$

b) Classificação dos pontos críticos  $(-1, 1)$  e  $(1, -1)$ .

As matrizes Hessianas nestes dois pontos são iguais,

$$D^2f(-1, 1) = D^2f(1, -1) = \begin{bmatrix} -10 & -2 \\ -2 & -10 \end{bmatrix}$$

e apresentam dois valores próprios negativos,  $\lambda_1 = -8$  e  $\lambda_2 = -12$ . Portanto, estes dois pontos são pontos de máximo de  $f$ .

c) Classificação do ponto crítico  $(0, 0)$ .

A matriz Hessiana

$$D^2f(0, 0) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

tem um valor próprio nulo  $\lambda_1 = 0$  e outro positivo  $\lambda_2 = 4$ .

Portanto, na direcção definida pelo vector próprio associado a  $\lambda_2 = 4$ , a função  $f$  tem um mínimo na origem. Isto quer dizer que se a origem for um extremo de  $f$  deverá ser um ponto de mínimo.

Na direcção singular correspondente ao valor próprio nulo  $\lambda_1 = 0$  deveremos passar à análise das derivadas de ordem superior a dois. No entanto, podemos analisar o comportamento de  $f$  directamente em torno da origem.

Note-se que na direcção definida por  $y = x$  temos  $f(x, x) = -2x^4 \leq 0$  e, portanto, a função  $f$  tem um ponto de máximo na origem.

Concluimos assim que a origem não é um extremo de  $f$ .

Na figura 3 encontra-se o gráfico de  $f$  onde se pode constatar a natureza dos pontos críticos.

**Exemplo 3.4** Consideremos a função  $f(x, y) = y^2 - 4x^2y + 3x^4$ .

a) Pontos críticos:  $\nabla f(x, y) = (0, 0)$ .

$$\nabla f(x, y) = (-8xy + 12x^3, 2y - 4x^2) = (0, 0) \Leftrightarrow \begin{cases} x(-2y + 3x^2) = 0 \\ y - 2x^2 = 0 \end{cases}$$

donde se conclui que o único ponto crítico é a origem.

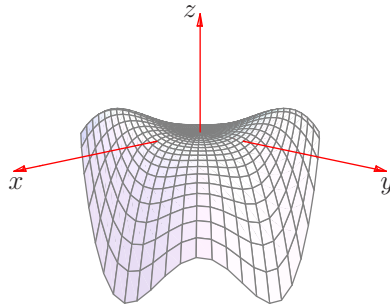


Figura 3: Gráfico da função:  $f(x, y) = (x - y)^2 - x^4 - y^4$

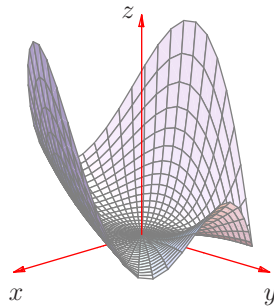


Figura 4: Gráfico da função:  $f(x, y) = (y - x^2)(y - 3x^2)$

b) Classificação do ponto crítico  $(0, 0)$ .

A matriz Hessiana

$$D^2f(0, 0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0, 0) & \frac{\partial^2 f}{\partial y \partial x}(0, 0) \\ \frac{\partial^2 f}{\partial x \partial y}(0, 0) & \frac{\partial^2 f}{\partial y^2}(0, 0) \end{bmatrix} = \begin{bmatrix} -8y + 36x^2 & -8x \\ -8x & 2 \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

tem um valor próprio nulo  $\lambda_1 = 0$  e outro positivo  $\lambda_2 = 2$ . Portanto, se a origem for extremo será um mínimo.

Note-se que a função  $f$  pode ser dada de outra forma

$$f(x, y) = y^2 - 4x^2y + 3x^4 = (y - x^2)(y - 3x^2).$$

Em torno da origem teremos:

- i)  $f(x, y) > 0$  para  $y > 3x^2$  ou para  $y < x^2$ .
- ii)  $f(x, y) < 0$  para  $x^2 < y < 3x^2$ .

Assim, em torno da origem, a função  $f$  toma valores tanto positivos como negativos, ou seja, a origem não é um extremo de  $f$ .

Na figura 4 encontra-se o gráfico de  $f$  onde se pode constatar a natureza da origem como ponto crítico.

\*\*\*

## Referências

- [1] Tom M. Apostol. *Calculus II*. Editorial Reverté, SA, 1977.
- [2] J. Campos Ferreira. *Introdução à Análise em  $R^n$* . AEIST, 1978.
- [3] J. E. Marsden and A. J. Tromba. *Vector Calculus*. W. H. Freeman and Company, 1998.