Universidade Federal de Sergipe

Centro de Ciências Exatas e Tecnologia

Departamento de Matemática

Pós-Graduação em Matemática

Otimização: estudo de máximos e mínimos de funções que definem problemas cotidianos

São Cristóvão – SE
Agosto de 2018

Universidade Federal de Sergipe

Centro de Ciências Exatas e Tecnologia

Departamento de Matemática

Pós-Graduação em Matemática

Otimização: estudo de máximos e mínimos de funções que definem problemas cotidianos

por

THED FREITAS FERREIRA

sob a orientação do

Prof. Dr. Gerson Cruz Araujo

São Cristóvão – SE
Agosto de 2018

Otimização: estudo de máximos e mínimos de funções que definem problemas cotidianos

por
THED FREITAS FERREIRA

Dissertação apresentada ao Corpo Docente da Pós-Graduação em Matemática da Universidade Federal de Sergipe como requisito parcial para a obtenção do título de Mestrado profissional em Matemática.

Área de Concentração: Análise

Aprovada em 02 de Agosto de 2018.

Banca Examinadora:

Prof. Dr. Gerson Cruz Araujo – UFS
(Orientador)

Prof. Dr. Naldisson dos Santos - UFS
(Examinador Interno)

Prof. Dr. Paulo de Souza Rabelo - UFS
(Examinador Externo)

Mas hoje eu sei, que só
através do amor,
o Homem pode se encontrar,
com a perfeição dos sábios...

Agradecimentos

Dedico a Deus todos os méritos da produção desse trabalho e a ele agradeço por ter me dado força de sempre buscar o melhor.

Dedico este trabalho aos meus pais Pedro e Norailde com todo meu amor e gratidão, vocês que me deram a vida e me ensinaram a vivê-la com dignidade. Espero ser merecedor de todos os esforços a mim dedicados, levarei vocês como exemplo de vida, os verdadeiros chefes de uma família exemplar formadas por meus irmãos Adailton e Filipe, cunhadas, que de forma direta ou indireta fazem parte dessa conquista, amo muito vocês.

Dedico a minha filha Thayla Monique o maior presente que eu poderia ganhar de Deus e a minha esposa Pricilla Carla pessoa guerreira, companheira, que suportou e me apoiou em todas as etapas desse curso. Agradeço a Deus por ter colocado você em meu caminho, sou muito grato por tudo Pricilla, amo você.

Ao término desse trabalho, deixo aqui meus sinceros agradecimentos.

Agradeço ao meu orientador o Professor Dr Gerson Cruz por ter aceitado de imediato caminhar ao meu lado neste projeto e pelos ensinamentos a mim repassados.

Agradeço aos professores do PROFMAT-UFS pelo empenho e dedicação.

Aos meus amigos queridos de PROFMAT-UFS, obrigado por sempre lutarmos juntos para que todos do curso chegassem até o final, sou fã de vocês.

Agradeço ao Professor Dr Naldisson do Santos e ao Professor Dr Paulo de Souza Rabelo por terem aceitado o convite para fazerem parte da minha banca e defesa.

Resumo

No presente trabalho, o objetivo é apresentar um estudo sobre os métodos de otimização e aplicá-los na solução de problemas cotidianos. Em matemática, otimização refere-se ao estudo de problemas em que se deseja maximizar ou minimizar uma determinada função através da escolha sistemática dos valores de variáveis dentro de um conjunto viável. Podemos analisar problemas de otimização por diferenciabilidade com e sem restrições, introduzindo o método dos multiplicadores de Lagrange. As resoluções apresentadas, baseam-se numa pequena fundamentação teórica, têm a preocupação de abranger diferentes abordagens e proporcionar o relacionamento de conceitos. Um dos métodos de se determinar os máximos e mínimos de funções é utilizando-se o cálculo em várias variáveis, o qual será abordado nesse trabalho. Possíveis situações relacionadas ao cotidiano são apresentadas para que o processo de otimização possa ser abordado por alunos do Ensino Médio.

Palavras-chave: Otimização; Máximos e Mínimos de funções; Multiplicadores de Lagrange.

Abstract

In the present work, the objective is to present a study on the methods of optimization and to apply them in the solution of daily problems. In mathematics, optimization refers to the study of problems in which one wants to maximize or minimize a given function by systematically choosing the values of variables within a viable set. We can analyze optimization problems by differentiability with and without constraints, introducing the Lagrange multiplier method. The resolutions presented, based on a small theoretical basis, are concerned to cover different approaches and provide the relationship of concepts. One of the methods to determine the maximum and minimum of functions is to use the calculation in several variables, which will be approached in this work. Possible situations related to everyday life are presented so that the optimization process can be approached by high school students.

Keywords: Optimization; Maximum and minimum functions; Lagrange multipliers.

Sumário

Introdução

Os problemas de otimização consistem em determinar os valores extremos de uma função, isto é, o maior ou menor valor que uma função pode assumir em um dado intervalo. Estes problemas são comuns em nossa vida diária e aparecem, por exemplo, quando procuramos determinar o nível de produção mais econômico de uma fábrica, as dimensões de embalagens de produtos que maximizam a capacidade das mesmas, etc. Desde a Grécia Antiga, no século III a.C., os estudiosos gregos já se interessavam em saber o estado "ótimo"de um determinado fenômeno, como por exemplo, intuitivamente sabiam que de todas as curvas com igual perímetro, as que envolviam maior área era o círculo. Assim como este, inumeros problemas eram resolvidos utilizando processos engenhosos, não havendo uma forma sistemática de os solucionar. Só no século XVII, Pierre de Fermat (1607 - 1665) desenvolveu o primeiro método geral para a determinação de máximos e mínimos. No entanto, este método era um procedimento algorítmico desprovido de qualquer fundamentação demonstrativa. A generalização da resolução deste tipo de problema aparece com os trabalhos de Isaac Newton (1643 - 1727) e Gottfried Leibniz (1646 - 1716) no desenvolvimento do cáculo diferencial integral.

A Otimização, com suas múltiplas subáreas, é um rico campo aberto ao desenvolvimento da Matemática Contemporânea. Neste sentido, a essência da otimização é melhorar algo em um conjunto de alternativas disponíveis. Este algo tem uma representação matemática que recebe o nome de função objetivo, porém em diversos momentos nos deparamos com situações limitadoras, dessa forma criam-se novas funções, denominadas funções restrições.

Existe vários problemas clássicos de otimização com restrições nas áreas de engenharia e matemática aplicada que vêm sendo estudados utilizando-se diversas técnicas encontradas na literatura em que um dos objetivos é minimizar o custo computacional necessário para a localização da solução ótima.

À medida que o número de funções e o número de variáveis aumentam, a dificuldade em se determinar o conjunto de soluções ótimas também aumentam. É neste contexto que surge a necessidade de desenvolver técnicas matemáticas que refinem o processo de otimização.

O método que utilizamos no decorrer desse trabalho é baseado no cálculo diferencial com várias variáveis. Dessa maneira, começamos nosso projeto com um estudo dos conceitos preliminares utilizados no cálculo, para depois estudarmos os métodos de otimização.

A fim de atingir os objetivos propostos, o trabalho será realizado da seguinte forma:

No primeiro capítulo, apresentamos um prefácio a Topologia do $\mathbb{R}^n$. Iniciamos com conceitos de produto interno e norma, bolas e conjuntos limitados, sequência convergentes e de Cauchy, e terminamos com conjuntos abertos e conjuntos fechados.

No segundo capítulo, é apresentado todo o embasamento teórico necessário para o entendimento do conceito de otimização. Estudos sobre funções de várias variáveis, envolvendo gráficos, limites, continuidade e diferenciabilidade, Teorema da Função Inversa e da Função Implícita. Derivadas parciais de ordem superior foram abordados para o auxílio da compreensão dos problemas de otimização.

No terceiro capítulo, simultaneamente a apresentação de diferentes teoremas e definições fundamentais, analisaremos problemas de otimização sem restrições e, posteriormente, problemas condicionados à uma diversidade de restrições. Aplicação do método de multiplicadores de Lagrange complementam o capítulo.

Por fim, no capítulo 4, abordaremos alguns problemas envolvendo máximos e mínimos de funções quadráticas e, a desigualdade das médias, que possívelmente poderiam ser trabalhados em sala de aula no ensino médio. A proposta dos problemas desse capítulo, tem o intuito de introduzir nas aulas de matemáticas o contéudo de otimização visto apenas no ensino superior.

Capítulo 1

Cálculo de funções em várias variáveis

Neste capítulo serão introduzidos alguns conceitos importantes do cálculo diferencial de funções de variáveis reais. Uma função de várias variáveis reais é uma regra que descreve como uma quantidade é determinada por outras quantidades, de modo único. Através das funções de várias variáveis poderemos idealizar uma grande quantidade de fenômenos dos mais diversos ramos da ciência.

1.1 Funções de várias variáveis

Definição 1.1. *Uma função real ou escalar f de n váriaveis associa a cada $n-$upla $x = (x_1, x_2, ..., x_n) \in X \subset \mathbb{R}^n$ um único número real $w = f(x_1, x_2, ..., x_n)$. O subconjunto X de $\mathbb{R}^n$ é chamado domínio da função f. Podemos denotar por*

$$f : X \subset \mathbb{R}^n \rightarrow \mathbb{R} \tag{1.1}$$

$$(x_1, ..., x_n) \mapsto w = f(x_1, ..., x_n) \tag{1.2}$$

Exemplo 1.1. *Consideremos a função,*

$$\begin{aligned} f : X \subset \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto f(x, y) = \frac{1}{x - y}. \end{aligned}$$

Temos que, $X = \{(x, y) \in \mathbb{R}^2 \mid x \neq y\}$. Ilustremos abaixo tal domínio:

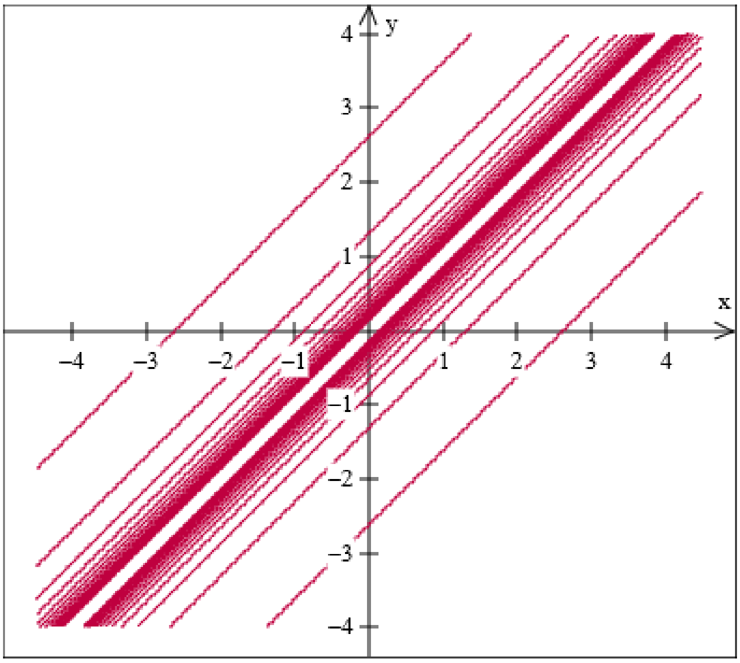

Figura 1.1: Gráfico do Domínio da função $f(x, y) = \dfrac{1}{x - y}$

1.2 Gráficos de uma função em várias variáveis

Definição 1.2. *Seja $f : X \subset \mathbb{R}^n \to \mathbb{R}$ uma função de n variáveis. Definimos o gráfico de f, denotado por G_f como subconjunto de R^{n+1}, onde $(x_1, ..., x_n) \in X$, ou seja,*

$$G_f := \{(x_1, ..., x_n, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_{n+1} = f(x_1, ..., x_n); (x_1, ..., x_n) \in X\} \tag{1.3}$$

observemos que, no caso $n = 2$, $G_f := \{(x, y, z) \in \mathbb{R}^3; z = f(x, y); (x, y) \in X\}$ é uma superfície em $\mathbb{R}^3$.

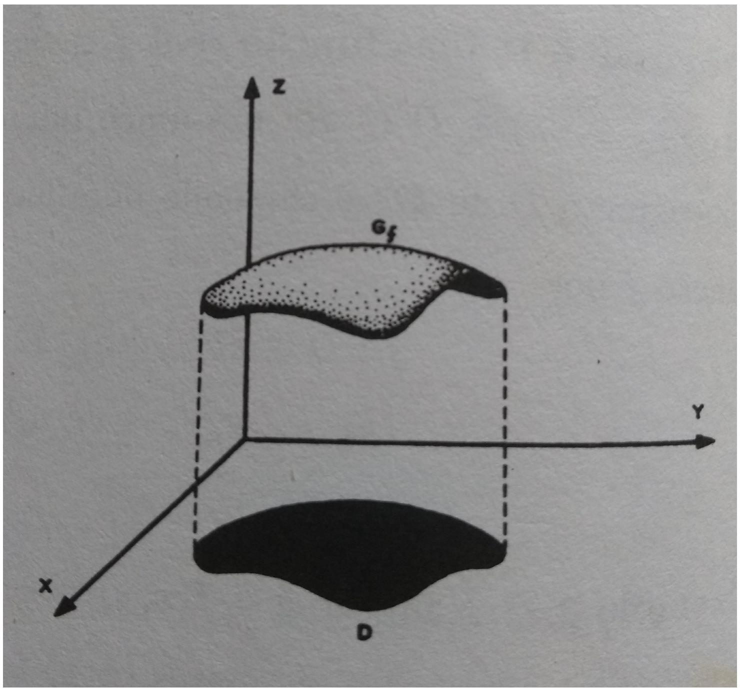

Figura 1.2: Gráfico de uma função definido em um domínio bidimensional

Exemplo 1.2. Determine o domínio e esboce o gráfico da função

$$f(x,y) = \sqrt{1 - x^2 - y^2}$$

Solução:

O domínio de f é o conjunto de todos os pares $(x,y) \in \mathbb{R}^2$ para os quais $1 - x^2 - y^2 \geq 0$, ou seja, o domínio X de f é o disco circular $x^2 + y^2 \leq 1$, de raio 1 e centro na origem.

Um ponto (x,y,z) pertence ao gráfico de f se, e somente se, $(x,y) \in X$ e $z = f(x,y)$, isto é, $z = \sqrt{1 - x^2 - y^2}$. A condição $z = \sqrt{1 - x^2 - y^2}$ é equivalente às duas condições, a saber, $z \geq 0$ e $x^2 + y^2 + z^2 = 1$. Deste modo, o gráfico de f consiste da porção da esfera $x^2 + y^2 + z^2 = 1$ acima do plano xy, conforme a figura 1.3:

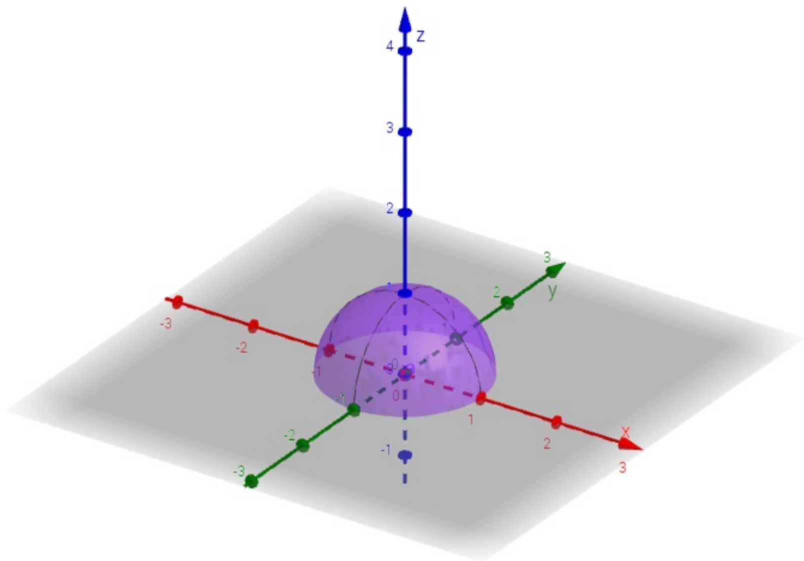

Figura 1.3: Gráfico da função $z = \sqrt{1 - x^2 - y^2}$

Exemplo 1.3. *A temperatura em um ponto* (x, y) *de uma placa de metal plana é definida pela função* $T(x, y) = 9x^2 + 4y^2$ *graus Celsius.*

a) *Encontre a temperatura no ponto* $(1, 2)$.

b) *Determine a equação da curva ao longo da qual a tempertura tem um valor constante igual a 36 graus Celsius.*

c) *Esboce a curva do item (b)*

Solução:

a) $T(1, 2) = 9 + 16 = 25$ *graus Celsius.*

b) *A curva tem a equação* $T(x, y) = 9x^2 + 4y^2 = 36 \Rightarrow \dfrac{x^2}{4} + \dfrac{y^2}{9} = 1.$

c) *A curva de equação* $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$ *é a elipse. Veja a figura 1.4:*

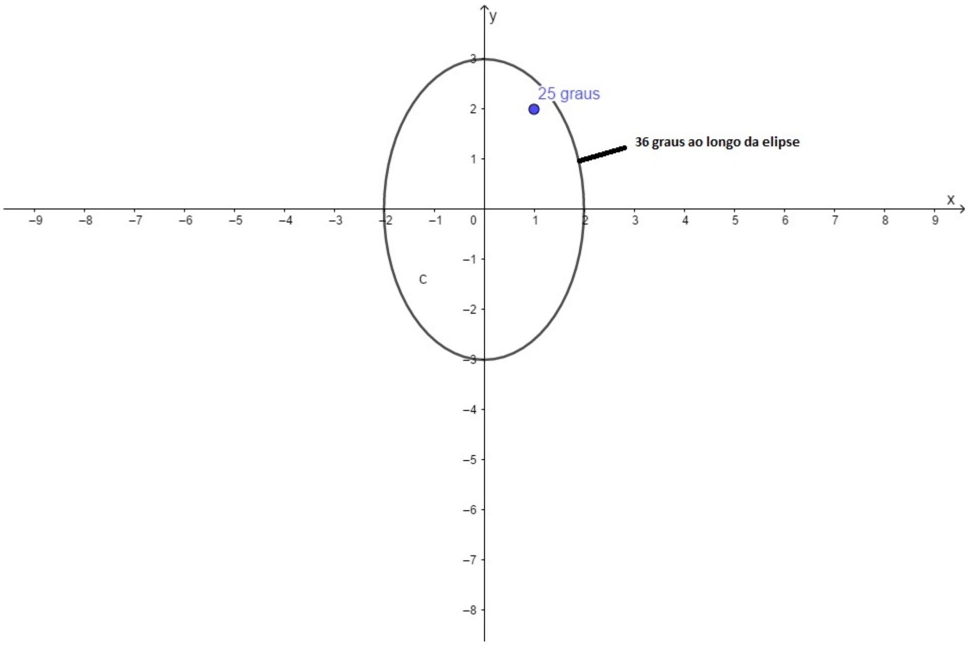

Figura 1.4: Gráfico da Elipse $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$

Em geral, os gráficos nos fornecem uma maneira de visualizarmos funções de várias variáveis. Uma outra maneira de visualizarmos tais funções é desenhar as suas curvas de nível, as quais serão comentadas na observação a seguir:

Observação 1.1. *Seja f uma função de duas variáveis e k um número real. O conjunto dos pontos (x, y) no domínio de f para os quais $f(x, y) = k$ é chamado de uma **Curva de Nível** de f. Esta contém os pontos do domínio de f para os quais o gráfico de f tem altura k. Ao esboçarmos a curva de nível no plano xy, devemos associar a mesma o seu correspondente valor de k.*

*Para diferentes valores para a constante k, obtemos um conjunto de curvas de nível. Este conjunto de curvas é chamado **Mapa de Contorno de Superfície** S.*

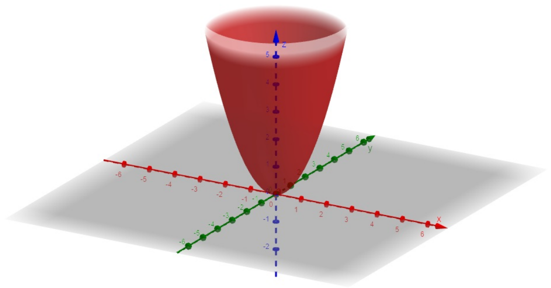

Figura 1.5: O gráfico da função $z = x^2 + y^2$

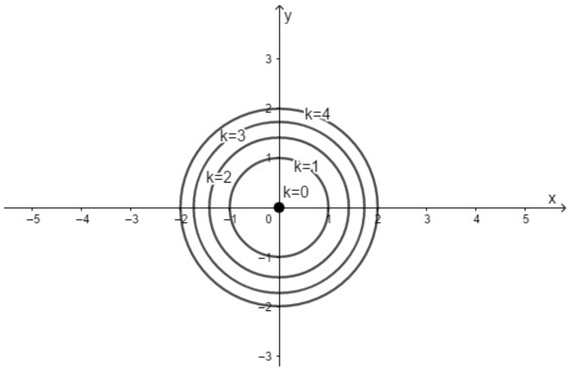

Figura 1.6: As curvas de níveis para valores de k, $f(x, y) = x^2 + y^2 = k$

Observação 1.2. *Referente às funções de três variáveis $w = f(x, y, z)$, não podemos visualizar seu gráfico. Mas, podemos considerar as superfícies de equação $f(x, y, z) = k$, quando k varia no conjunto imagem de f. Estas superfícies são chamadas **superfícies de níveis para f**.*

Exemplo 1.4. *Descreva as superfícies de nível da função $f(x, y, z) = x^2 + y^2 - z^2$.*

Solução: *As superfícies de nível possui equação*

$$x^2 + y^2 - z^2 = k$$

Vamos analizar os seguintes casos para k:

(i) *Se $k > 0$, a superfície de nível é o hiperbolóide de uma folha de equação $H : \dfrac{x^2}{k} + \dfrac{y^2}{k} - \dfrac{z^2}{k} = 1$.*

Observe a figura 1.7:

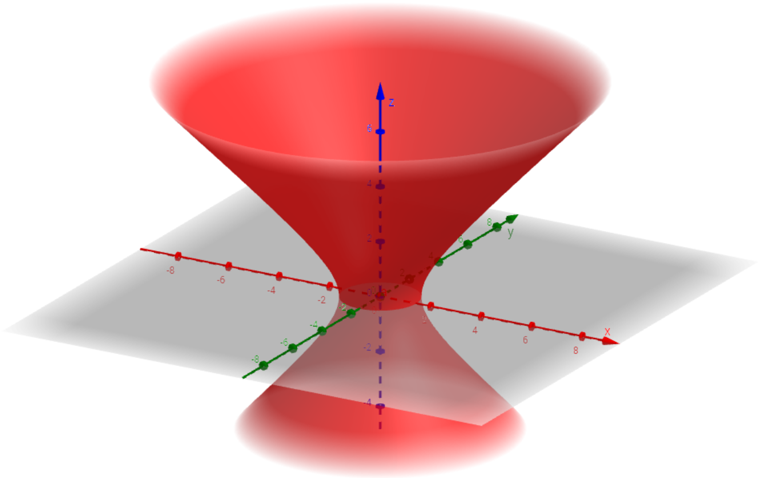

Figura 1.7: Hiperbolóide de uma folha de equação $\dfrac{x^2}{k}+\dfrac{y^2}{k}-\dfrac{z^2}{k}=1$

(ii) *Se $k<0$, a superfície de nível é o hiperbolóide de duas folhas de equação $H:-\dfrac{x^2}{|k|}-\dfrac{y^2}{|k|}+\dfrac{z^2}{|k|}=1$.*

Observe a figura 1.8:

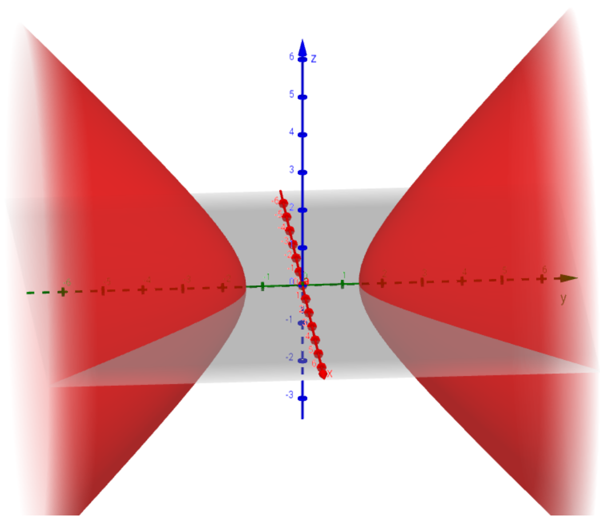

Figura 1.8: Hiperbolóide de duas folhas de equação $-\dfrac{x^2}{|-k|}+\dfrac{y^2}{|-k|}-\dfrac{z^2}{|-k|}=1$

(iii) *Se $k=0$, a superfície de nível é um cone circular de equação $x^2+y^2-z^2=0$.*

Observe a figura 1.9

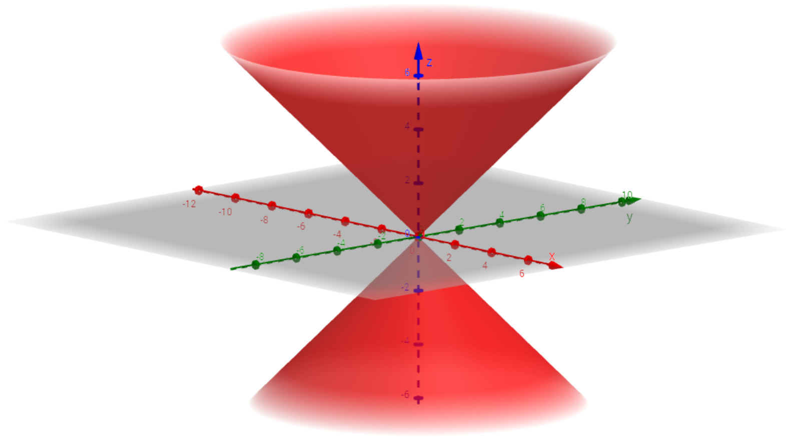

Figura 1.9: Cone circular de equação $x^2 + y^2 - z^2 = 0$

1.3 Limites de funções em várias variáveis

Trataremos sobre alguns resultados importantes acerca de limites de funções cujo domínio é qualquer conjunto constituído de números reais .

Definição 1.3. *Sejam* $f : X \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}^n$ *e* $a \in X$*'. Dizemos que o limite de* f*, quando* x *tende para* a *será* $l \in \mathbb{R}^n$*, se dado* $\epsilon > 0$*, existe* $\delta > 0$ *tal que*

para todo $x \in X$ *com* $0 < \|x - a\| < \delta$ *tem-se que* $\|f(x) - l\| < \epsilon$

Denotamos $\lim\limits_{x \longrightarrow a} f(x) = f(a)$.

Vamos exibir dois exemplos, o qual utilizaremos a definição formal de limite 1.3.

Exemplo 1.5. *Mostre que:*

a) $\lim\limits_{(x,y)\to(0,0)} \dfrac{3x^2y}{x^2 + y^2} = 0.$

b) $\lim\limits_{(x,y)\to(0,0)} \dfrac{5}{x^2 - 3y^2} = \infty$

Solução:

a) *De fato, dado* $\varepsilon > 0$*, tomemos* $\delta = \dfrac{\varepsilon}{3} > 0$*. Deste modo,*

$$\|(x,y) - (0,0)\| = \|(x,y)\| = \sqrt{x^2+y^2} < \delta \Leftrightarrow x^2 + y^2 < \delta^2. \tag{1.4}$$

Como

$$\frac{x^2}{x^2+y^2} \leq 1, \tag{1.5}$$

Multiplicando ambos os lados da equação 1.5 por $3|y|$*, obtemos*

$$\frac{3x^2|y|}{x^2+y^2} \leq 3|y| \leq 3\sqrt{x^2+y^2}. \tag{1.6}$$

Assim,

$$\left|\frac{3x^2y}{x^2+y^2}\right| \quad \leq \quad 3\delta = \varepsilon. \tag{1.7}$$

Logo,

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = 0.$$

b) *De fato, dado* $A > 0$*, tomemos* $\delta = \sqrt{\dfrac{5}{A}} > 0$ *no qual,*

$$\|(x,y) - (0,0)\| = \|(x,y)\| = \sqrt{x^2+y^2} < \delta \Leftrightarrow x^2 + y^2 < \delta^2. \tag{1.8}$$

Desde que:

$$x^2 - 3y^2 = x^2 + y^2 - 4y^2 \leqslant x^2 + y^2, \tag{1.9}$$

segue que

$$\begin{aligned} f(x,y) = \frac{5}{x^2-3y^2} \quad &\geq \quad \frac{5}{x^2+y^2} \\ &> \quad \frac{5}{\delta^2} \\ &= \quad A. \end{aligned}$$

Assim para $\|(x,y)\| < \delta$*, temos* $f(x,y) > A$*.*

Logo,

$$\lim_{(x,y)\to(0,0)} f(x,y) = \infty.$$

1.4 Continuidade de funções em várias variáveis

Topologicamente e de suma importância termos noção sobre continuidade, já que ao definirmos *contração* aparecerá a ideia de continuidade uniforme e ambas tem uma relação importante.

Definição 1.4. *Sejam $X \subseteq \mathbb{R}^m, y \in X$ e $f : X \to \mathbb{R}^n$ uma função. Diremos que f é contínua em $y \in X$, se dado $\epsilon > 0$, existe $\delta = \delta(\epsilon, y) > 0$ tal que*

para todo $x \in B_\delta(y) \cap X$ tem-se que $f(x) \in B_\epsilon(f(y))$.

Observação 1.3. *Veja que se $x \in X \cap X'$, então f é contínua em y se, e somente se, $\lim\limits_{x \longrightarrow y} f(x) = f(y)$.*

Exemplo 1.6. *Todo função polinômial $p : \mathbb{R} \to \mathbb{R}$ é uma função contínua, pois $\lim\limits_{x \longrightarrow y} p(x) = p(y)$, para todo $y \in \mathbb{R}$.*

Exemplo 1.7. *As funções $f : \mathbb{R}^2 \to \mathbb{R}^2$, definidas por $f(x, y) = (x sen(y), ye^x)$ e $g : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $g(x, y, z) = (x^2 + y, z^2)$, são funções contínuas.*

Exemplo 1.8. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma aplicação definida por:*

$$f(x) = \begin{cases} sen\left(\dfrac{1}{x}\right) & , se \;\; x \neq 0; \\ 0 & , se \;\; x = 0. \end{cases}$$

Observe que $0 \in \mathbb{R} \cap \mathbb{R}' = \mathbb{R}$. Temos ainda que, $\lim\limits_{x \longrightarrow 0} sen\left(\dfrac{1}{x}\right)$ não existe. Daí, segue que $\lim\limits_{x \longrightarrow 0} sen\left(\dfrac{1}{x}\right) \neq 0 = f(0)$. Logo temos que $\lim\limits_{x \longrightarrow 0} sen\left(\dfrac{1}{x}\right) \neq f(0)$. Ou seja, f não é contínua em $x = 0$.

Agora, veremos uma outra maneira de definir uma função ser contínua em um ponto.

Teorema 1.4.1. *Seja $X \subseteq \mathbb{R}^m, y \in X$. Então $f : X \to \mathbb{R}^n$ é contínua em $y \in X$ se, e somente se, toda sequência $(x_m) \subseteq X$ com $\lim\limits_{m \longrightarrow \infty} x_m = y$ satisfaz $\lim\limits_{m \longrightarrow \infty} f(x_m) = f(y)$.*

Demonstração. Suponha que f seja contínua em $y \in X$. Seja $(x_m) \subseteq X$ tal que $\lim x_m = y$. Assim, dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$\text{para todo } x \in X \text{ com } \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon.$$

Como $\lim x_m = y$, então existe $n_0 \in \mathbb{N}$ tal que,

$$\|x_m - y\| < \delta, \ para\ todo\ m \geq n_0.$$

Logo,

$$\|f(x_m) - f(y)\| < \epsilon, \ para\ todo\ m \geq n_0.$$

Ou seja, $\lim\limits_{m \longrightarrow \infty} f(x_m) = f(y)$.

Reciprocamente, considere que f não seja contínua em y, então existe $\epsilon > 0$ tal que para todo $\delta > 0$, podemos encontrar $x_\delta \in X$ com $\|x_\delta - y\| < \delta$ e $\|f(x_\delta) - f(y)\| \geq \epsilon$. Assim existe $(x_m) \subseteq X$ com

$$\|x_m - y\| < \frac{1}{m} \ e\ \|f(x_m) - f(y)\| \geq \epsilon.$$

Logo, $\lim\limits_{m \longrightarrow \infty} \|x_m - y\| = 0$, porém, $\lim\limits_{m \longrightarrow \infty} f(x_m) \neq f(y)$. □

1.4.1 Continuidade uniforme

Como foi citado anteriormente, necessitamos compreender um pouco do conceito que estabelece o que é uma função uniformemente contínua, pois, relacionaremos tal definição com o estudo de *contrações.*

Definição 1.5. *Seja $f : X \to \mathbb{R}^n$ uma função, onde $X \subseteq \mathbb{R}^m$. Dizemos que f é uniformemente contínua, se dado $\epsilon > 0$, existe $\delta = \delta(\epsilon) > 0$ tal que, para todo $x, y \in X$ com*

$$\|x - y\| < \delta, \ tem - se\ \|f(x) - f(y)\| < \epsilon.$$

Exemplo 1.9. *Seja $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = x + y$. Então f é uniformemente contínua. De fato, dado $\epsilon > 0$, observe que, $\forall\ (x, y), (a, b) \in \mathbb{R}^2$, temos as seguintes equivalências:*

$$\|(x, y) - (a, b)\|_s < \delta \Leftrightarrow \|(x - a, y - b)\|_s < \delta \Leftrightarrow |x - a| + |y - b| < \delta = \epsilon,$$

onde $\|.\|_s$ é a norma da soma. Com isso, concluímos que

$$|f(x, y) - f(a, b)| = |(x + y) - (a + b)| = |(x - a) + (y - b)| < \epsilon,$$

ou seja, f é uniformemente contínua.

Vejamos outra maneira de definir continuidade uniforme.

Teorema 1.4.2. *Seja $X \subseteq \mathbb{R}^m$. Então $f : X \to \mathbb{R}^n$ é uniformemente contínua se, e somente se, para todo $(x_m), (y_m) \subseteq X$ com $\lim\limits_{m \longrightarrow \infty} \|x_m - y_m\| = 0$, tem-se que $\lim\limits_{m \longrightarrow \infty} \|f(x_m) - f(y_m)\| = 0$.*

Demonstração. Suponha que f é uniformemente contínua. Sejam para todo $(x_m), (y_m) \subseteq X$ tais que $\lim\limits_{m \longrightarrow \infty} \|x_m - y_m\| = 0$. Então, dado $\epsilon > 0$,existe $\delta > 0$ tal que

$$\|x - y\| < \delta \text{ com } x, y \in X, \text{ tem-se que } \|f(x) - f(y)\| < \epsilon.$$

Mas, $\lim\limits_{m \longrightarrow \infty} \|x_m - y_m\| = 0$. Logo, para $\delta > 0$ existe $n_0 \in \mathbb{N}$ tal que,

$$\|x_m - y_m\| < \delta, \ \textit{para todo } m \geq n_0.$$

Daí, $\|f(x_m) - f(y)\| < \epsilon$, para todo $m \geq n_0$. Ou seja, $\lim\limits_{m \longrightarrow \infty} \|f(x_m) - f(y_m)\| = 0$.

Agora, suponhamos válida a recíproca estabelecida no teorema. Se f não fosse uniformemente contínua, então existe $\epsilon > 0$ com a seguinte propriedade: Para todo $m \in \mathbb{N}$ poderíamos achar pontos x_m, y_m em X tais que,

$$\|x_m - y_m\| < \frac{1}{m} \ e \ \|f(x_m) - f(y)\| \geq \epsilon.$$

Então teríamos

$$\lim_{m \longrightarrow \infty} \|x_m - y_m\| = 0,$$

mas $\lim\limits_{m \longrightarrow \infty} |f(x_m) - f(y_m)\| \neq 0$. Esta condição conclui a prova do teorema.

$\square$

Exemplo 1.10. *A função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^2$, não é uniformemente contínua. Com efeito, tomando $x_m = m + \dfrac{1}{m}$ e $y_m = m$ temos que,*

$$\lim_{m \longrightarrow \infty} (x_m - y_m) = \lim_{m \longrightarrow \infty} \left(m + \frac{1}{m} - m \right) = \lim_{m \longrightarrow \infty} \frac{1}{m} = 0,$$

mas,

$$f(x_m) - f(y_m) = m^2 + 2 + \frac{1}{m^2} - m^2 = 2 + \frac{1}{m^2} > 2.$$

Logo, não se tem $\lim\limits_{m \longrightarrow \infty} [f(x_m) - f(y_m)] = 0$.

Exemplo 1.11. *Seja $p : \mathbb{R}^2 \to \mathbb{R}$ dada por $p(x, y) = x.y$. Temos que p não é uniformemente contínua. Com efeito, sejam $x_m = (m, 0)\,, y_m = \left(m, \dfrac{1}{m} \right) \subseteq \mathbb{R}^2$, então*

$$\lim_{m\longrightarrow\infty}\|x_m-y_m\| = \lim_{m\longrightarrow\infty}\left\|(m,0)-\left(m,\frac{1}{m}\right)\right\| = \lim_{m\longrightarrow\infty}\left\|\left(0,-\frac{1}{m}\right)\right\| = \lim_{m\longrightarrow\infty}\frac{1}{m}=0.$$

Mas,

$$\lim_{m\longrightarrow\infty}|p(x_m)-p(y_m)| = \lim_{m\longrightarrow\infty}\left|p(m,0)-p\left(m,\frac{1}{m}\right)\right| = 1\neq 0.$$

Logo, pelo Teorema anterior, p não é uniformemente contínua.

Observação 1.4. *Pela definição, deduzimos que toda função uniformemente contínua é contínua. Porém, a recíproca não é verdadeira. No exemplo acima, vimos que p não é uniformemente continua, mas p é claramente contínua.*

1.5 Diferenciabilidade de funções em várias varáveis

Nesta seção vamos compreender o significado de diferenciabilidade para uma função de várias variáveis, bem como as consequências da diferenciabilidade.

1.5.1 Derivadas parciais

Nesta subseção, trataremos das derivadas parciais que estão inclusas na definição de matriz jacobiana e, consequentemente, em diferenciabilidade. Esta última tem papel imprescindível na equivalência entre os Teoremas da Função Inversa e Função Implícita.

Definição 1.6. *Seja $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ uma função definida no aberto X. Dizemos que a i-ésima derivada parcial de f em $a=(a_1,a_2,\cdots,a_n)$ existe se o limite*

$$\lim_{t\longrightarrow 0}\frac{f(a+te_i)-f(a)}{t},$$

existe, $i=1,...,n$.

Neste caso, escrevemos $\dfrac{\partial f}{\partial x_i}(a) = \lim\limits_{t\longrightarrow 0}\dfrac{f(a+te_i)-f(a)}{t}$, para a derivada de f em relação a i-ésima variável.

Outras notações: $\frac{\partial f}{\partial x_i}(a) = \partial_i f(a) = D_i f(a) = f_{x_i}(a)$.

Interpretação geométrica

A interpretação geométrica das derivadas parciais de uma função de duas variáveis é análoga aquela para função de uma variável real, vista em cursos iniciais de cálculo.

Para visualisar o significado do conceito de derivada parcial, considere as figuras 1.10 e 1.11

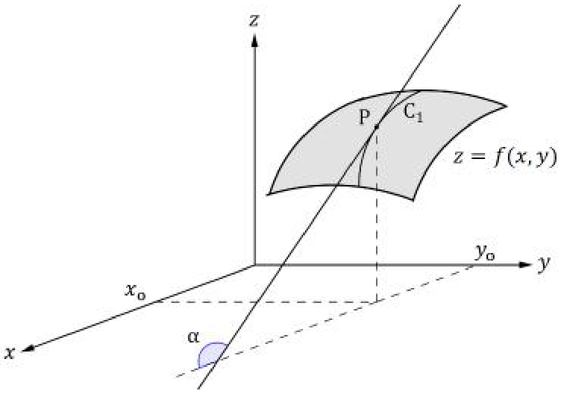

Figura 1.10: interpretação geomérica das derivadas parciais de $f(x, y)$

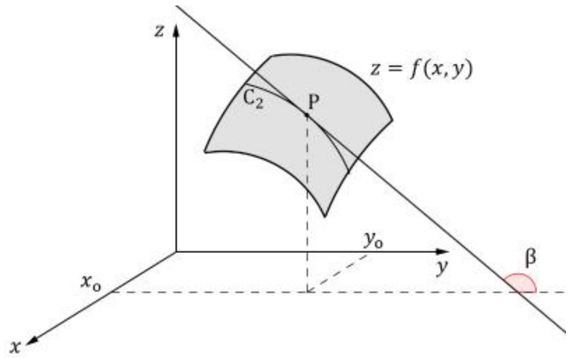

Figura 1.11: interpretação geomérica das derivadas parciais de $f(x, y)$

Seja (x_0, y_0) um ponto no plano xy e $P_0 = (x_0, y_0, f(x_0, y_0))$ o ponto correspondente na superfície pelo plano $y = y_0$. A interseção é uma curva de equação $z = f(x, y_0)$ neste plano. O número $\dfrac{\partial f}{\partial x}(x_0, y_0)$ é o coeficiente angular da reta tangente a esta curva quando $x = x_0$, assim:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \tan \alpha.$$

Analogamente, a interseção $z = (x, y)$ com plano $x = x_0$ é a curva $z = f(x_0, y)$ neste plano. O número $\dfrac{\partial f}{\partial y}(x_0, y_0)$ é o coeficiente angular da reta tangente a esta curva quando $y = y_0$, assim:

$$\frac{\partial f}{\partial y}(x_0, y_0) = \tan \beta.$$

Exemplo 1.12. *Seja* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *dada por* $f(x, y) = \dfrac{xy}{x^2 + y^2}$ *se* $(x, y) \neq (0, 0)$ *e* $f(0, 0) = 0$.

Então,

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \longrightarrow 0} \frac{f((0, 0) + t(1, 0)) - f(0, 0)}{t} = \lim_{t \longrightarrow 0} \frac{f(t, 0)}{t} = \lim_{t \longrightarrow 0} \frac{t.0}{t^2 + 0^2} = 0$$

Analogamente, $\frac{\partial f}{\partial y}(0, 0) = 0$.

Observação 1.5. *É importante ressaltar que, se as derivadas parciais existirem não garante que a função seja contínua, olhemos para o exemplo anterior, as derivadas parciais existem mas f é descontínua em (0,0). De fato, existe* $\epsilon = \frac{1}{2} > 0$ *tal que,*

para todo $\delta > 0$*, podemos encontrar,* $(x_\delta, y_\delta) \in \mathbb{R}^2$ *com* $\|(x_\delta, y_\delta)\| < \delta$*.*

Porém,

$$\|f(x_\delta, y_\delta)\| = \epsilon = \frac{1}{2}$$

Exemplo 1.13. *Seja* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *dada por* $f(x,y) = (x\cos y, e^x sen\ y)$*. Temos que f é diferenciável em* $(x,y) \in \mathbb{R}^2$*. Veja que*

$$f_1(x,y) = x\cos y \text{ e } f_2(x,y) = e^x sen\ y$$

Logo,

$$\begin{aligned}
\frac{\partial f_1}{\partial x}(x,y) &= \lim_{t\longrightarrow 0} \frac{f((x,y)+t(1,0)) - f(x,y)}{t} \\
&= \lim_{t\longrightarrow 0} \frac{f_1((x+t,y)) - f_1(x,y)}{t} \\
&= \lim_{t\longrightarrow 0} \frac{(x+t)\cos y - x\cos y}{t} \\
&= \lim_{t\longrightarrow 0} \frac{t\cos y}{t} \\
&= \cos y.
\end{aligned}$$

Seguindo o mesmo processo, obtemos

$$\frac{\partial f_1}{\partial x}(x,y) = \cos y,\ \frac{\partial f_1}{\partial y}(x,y) = \ x sen\ y,\ \frac{\partial f_2}{\partial x}(x,y) = e^x sen\ y,\ \frac{\partial f_2}{\partial y}(x,y) = e^x\cos y.$$

1.5.2 Derivadas direcionais

Aqui, falaremos sobre a derivada direcional e mostraremos um exemplo sobre tal.

Definição 1.7. *Seja* $X \subseteq \mathbb{R}^n$ *aberto. Sejam* $f : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *uma função real e* $v \in \mathbb{R}^n$*. Chamamos e denotamos o limite*

$$\frac{\partial f}{\partial v}(a) = \lim_{t\longrightarrow 0} \frac{f(a+tv) - f(a)}{t}.$$

de derivada direcional de f em $a \in X$*, caso este limite exista. Caso contrário, diremos que a derivada direcional de f não existe no ponto a.*

Exemplo 1.14. *Seja* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definida por*

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^4 + y^2} \text{ se } (0,0) \neq (x,y) \\ 0 \text{ se } x = y. \end{cases}$$

Logo, para $v = (x,y)$ *, obtemos*

$$\begin{aligned} \frac{\partial f}{\partial v}(0,0) &= \lim_{t \longrightarrow 0} \frac{f((0,0) + t(x,y)) - f(0,0)}{t} \\ &= \lim_{t \longrightarrow 0} \frac{f(tx,ty)}{t} = \lim_{t \longrightarrow 0} \frac{t^3 x^3 t y}{t(t^4 x^4 + t^2 y^2)} \\ &= \lim_{t \longrightarrow 0} \frac{t^4 x^3 y}{t^3(t^2 x^4 + y^2)} = \lim_{t \longrightarrow 0} \frac{t x^3 y}{(t^2 x^4 + y^2)} = 0. \end{aligned}$$

Portanto, $\frac{\partial f}{\partial v}(0,0) = 0.$

1.5.3 Diferenciabilidade

Será feita a definição de diferenciabilidade e apresentado exemplos que reforçam este conceito.

Definição 1.8. *Seja* $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *uma função real definida no aberto* $X \subseteq \mathbb{R}^n$*. Dizemos que* f *é diferenciável em* $a \in X$ *se :*

1. $\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \cdots, \frac{\partial f}{\partial x_n}(a)$ *existem*

2. *Para todo,* $v = (v_1, v_2, \cdots, v_n)$ *tal que* $a + v \in X$ *tem -se,*

$$f(a+v) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot v_i + r(v),$$

onde $\lim\limits_{v \longrightarrow 0} \dfrac{r(v)}{\|v\|} = 0.$
Dizemos que f *é diferenciável em* X*, se* f *é diferenciável em todo* $a \in X$*.*

Exemplo 1.15. *Seja* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *definida por* $f(x,y) = 2x + 3y$*. Dizemos que* f *é diferenciável em* $(-1,1)$*, pois*

1. $\frac{\partial f}{\partial x}(-1,1) = 2$ *e* $\frac{\partial f}{\partial y}(-1,1) = 3$*;*

2. *Note que, para* $r(x,y) = \frac{\partial f}{\partial x}(-1,1) \cdot x + \frac{\partial f}{\partial y}(-1,1) \cdot y$ *obtemos*

$$r(x,y) = 2x + 3y - (2-3) - 2(x-1) - 3(y+1) = 2x + 3y + 1 - 2x + 2 - 3y - 3 = 0$$

Portanto, $\lim\limits_{(x,y)\longrightarrow 0} \frac{r(x,y)}{\|(x,y)\|} = 0$. *Assim, pela definição acima temos que f é diferenciável em $(-1,1)$.*

Definição 1.9. *Seja $X \subseteq \mathbb{R}^m$ um conjunto aberto. Seja $f : X \longrightarrow \mathbb{R}^n$ uma função. Dizemos que f é diferenciável em $a \in X$, se f_i é diferenciável em $a \in X$, para todo $i = 1, 2, ..., n$, onde, $f = (f_1, \cdots, f_n)$.*

Os conceitos de diferenciabilidade e derivabilidade em $\mathbb{R}$ coincidem. Vejamos a prova deste fato na seguinte proposição.

Proposição 1.1. *Seja $f : \mathbb{R} \longrightarrow \mathbb{R}$ é diferenciável em $a \in \mathbb{R}$ se, e somente se, f é derivável em a $\in \mathbb{R}$.*

Demonstração. Como, f é diferenciável em $a \in \mathbb{R} \Leftrightarrow f'(a)$ existe $\Leftrightarrow \lim\limits_{v\longrightarrow 0} \dfrac{f(a+v)-f(a)}{v} = f'(a) \Leftrightarrow$ $\lim\limits_{v\longrightarrow 0} \dfrac{f(a+v)-f(a)-f'(a)}{v} \Leftrightarrow \lim\limits_{v\longrightarrow 0} \dfrac{|f(a+v)-f(a)-f'(a)|}{|v|} = 0 \Leftrightarrow f$ é diferenciável em a $\in \mathbb{R}$.

□

Exemplo 1.16. *Considere a função $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $f(x) = \sqrt{x}$. Veja que f não é diferenciável no ponto 0, pois, f não é derivável em 0. Com efeito,*

$$\lim_{v\longrightarrow 0} \frac{f(x)-f(0)}{x-0} = \lim_{v\longrightarrow 0} \frac{\sqrt{x}-0}{x} = \lim_{v\longrightarrow 0} \frac{1}{\sqrt{x}}.$$

Este limite não existe.

Proposição 1.2. *Seja $f; X \subseteq \mathbb{R}^n$ diferenciável em $a \in X$, então f é contínua em a.*

Demonstração. Se f for diferenciável em $a \in$ X, então $\lim\limits_{v\longrightarrow 0} \dfrac{r(v)}{\|v\|} = 0$. Daí,

$$\lim_{v\longrightarrow 0} r(v) = \lim_{v\longrightarrow 0} \frac{r(v)}{\|v\|} \cdot \|v\| = 0$$

pois, $\lim\limits_{v\longrightarrow 0} \|v\| = 0$.

Logo,

$$\lim_{v\longrightarrow 0}[f(a+v)-f(a)] = \lim_{v\longrightarrow 0}\left[\sum_{i=1}^{n} \frac{\partial f}{\partial x_i(a)}\right] \cdot (v_i + r(v)) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i(a)} \cdot \Big(\lim_{v\longrightarrow 0} v_i + \lim_{v\longrightarrow 0} r(v)\Big) = 0,$$

onde $v = (v_1, \cdots, v_n)$, ou seja, f é contínua em a.

□

1.5.4 Matriz jacobiana e vetor gradiente

Nesta subseção, definiremos matriz jacobiana, vetor gradiente e ainda mostraremos alguns exemplos para que haja uma melhor compreensão acerca das definições.

Definição 1.10. *Seja $X \subseteq \mathbb{R}^m$ aberto. Seja $f : X \longrightarrow \mathbb{R}^n$ diferenciável em $a \in X$. A matriz*

$$Jf(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right) \in M_{n\times m}(\mathbb{R}),$$

é chamada matriz jacobiana de f no ponto a.

Exemplo 1.17. *Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $f(x,y) = (x\cos y, e^x sen\ y)$. Temos que f é diferenciável em $(x,y) \in \mathbb{R}^2$. Veja que*

$$f_1(x,y) = x\cos y, f_2 = e^x sen\ y.$$

Logo,

$$\frac{\partial f_1}{\partial x}(x,y) = \cos y, \frac{\partial f_1}{\partial y}(x,y) = -xsen\ y, \frac{\partial f_2}{\partial x}(x,y) = e^x sen\ y, \frac{\partial f_2}{\partial y}(x,y) = e^x \cos y$$

Daí, a matriz jacobiana de f em (x,y) é dada por:

$$Jf(x,y) = \begin{bmatrix} \cos y & -xsen\ y \\ e^x sen\ y & e^x \cos y \end{bmatrix}.$$

Definição 1.11. *Seja $X \subseteq \mathbb{R}^m$ aberto. Seja $f : X \longrightarrow \mathbb{R}$ uma função diferenciável em $a \in X$. O vetor definido por*

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_m}(a)\right)$$

é chamado vetor gradiente de f em a.

Exemplo 1.18. *Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x,y) = xe^y$, então*

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = (e^y, xe^y).$$

1.5.5 Derivada como aplicação linear

Os principais resultados deste trabalho fazem referência ao conceito de derivada entre espaços reais de dimensão maior ou igual a um. Com isso, precisamos estabelecer, precisamente, como estender a conhecida derivada vista em Análise na Reta. Esta nova teoria está associada à definição de aplicação linear em Álgebra linear.

Definição 1.12. *Seja* $X \subseteq \mathbb{R}^m$ *diferenciável em* $a \in X$. *A derivada de* f *em* $a \in X$ *é transformação linear* $f'(a) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ *tal que*

$$[f'(a] = Jf(a),$$

matriz de $f'(a)$ *em relação as bases canônicas de* $\mathbb{R}^m$ *e* $\mathbb{R}^n$.

Exemplo 1.19. *Vimos que para* $f(x,y) = (x\cos y, e^y sen\ y,$ *a matriz jacobiana no ponto (x,y) é dada por,*

$$Jf(x,y) = \begin{pmatrix} \cos y & -x sen\ y \\ e^x sen\ y & e^x \cos y \end{pmatrix}.$$

Daí, $f'(x,y)$*:* $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *dada por,*

$$\begin{aligned} f'(x,y)\cdot(a,b) &= aJf(x,y)\cdot e_1 + bJf(x,y)\cdot e_2 \\ &= a(\cos y, e^x sen\ y) + b(-x sen\ y, e^x \cos y) \\ &= (a\cos y - xb sen\ y, ae^x sen\ y + be^x \cos y). \end{aligned}$$

Por fim,

$$f'(x,y)\cdot(a,b) = (a\cos y - xb\sin y, ae^x \sin y + be^x \cos y)$$

Vejamos uma maneira de provar que a derivada é única.

Teorema 1.5.1. *Seja* $f : X \longrightarrow \mathbb{R}^m$ *uma aplicação diferenciável em* $a \in X$ *definida no aberto* $X \subseteq \mathbb{R}^m$. *Se uma transformação linear* $T : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ *é tal que para* a*,* $a+v \in X$ *tem-se:*

$$f(a+v) - f(a) = T\cdot v + r(v),$$

com $\lim\limits_{v\longrightarrow 0} \frac{r(v)}{\|v\|} = 0$. *Então* $T = f'(a)$.

Demonstração. De fato, tomando tv no lugar de v segue que,

$$\frac{f(a+tv) - f(a)}{t} = T\cdot v \pm \frac{r(tv)}{\|tv\|}\cdot\|tv\|.$$

Logo,

$$T\cdot v = \lim_{v\longrightarrow 0} \frac{f(a+tv) - f(a)}{t} = \frac{\partial f}{\partial v}(a) = f'(a)\cdot v.$$

Portanto, segue que T $= f'(a)$.

□

Exemplo 1.20. *Seja $f : M_n(\mathbb{R}) \longrightarrow M_n(\mathbb{R})$ uma aplicação diferenciável definida por $f(X) = X^2$ onde $M_n(\mathbb{R})$ é o conjunto das matrizes reais $n \times n$. Vamos utilizar o teorema acima para mostrar que $Df(X) \cdot H = XH + HX$. Com efeito, temos que*

$$\begin{aligned} f(X+H) &= f(X) + XH + HX + r(H) \\ (X+H)^2 &= X^2 + XH + HX + r(H) \\ X^2 + XH + HX + H^2 &= X^2 + XH + HX + r(H) \\ H^2 &= r(H). \end{aligned}$$

Passando a norma em ambos os lados da igualdade temos que,

$$\|r(H)\| = \|H^2\|$$

Consequentemente,

$$\frac{\|r(H)\|}{\|H\|} = \frac{\|H^2\|}{\|H\|} = \|H\|.$$

Portanto,

$$\lim_{H \longrightarrow 0} \frac{\|r(H)\|}{\|H\|} = \lim_{H \longrightarrow 0} \|H\| = 0,$$

ou seja, $Df(X) \cdot H = XH + HX$.

Exemplo 1.21. *Seja $f : X \longrightarrow \mathbb{R}^m$ e $g : X \longrightarrow \mathbb{R}^p$ diferenciáveis em $z \in X \subseteq \mathbb{R}^n$ aberto. Defina $(f,g) : X \longrightarrow \mathbb{R}^m \times \mathbb{R}^p$ por $(f,g)(x) = (f(x), g(x))$.Então (f,g) é diferenciável em $a \in X$ e*

$$(f,g)'(a) \cdot v = (f'(a) \cdot v, g'(a) \cdot v).$$

De fato,

$$\begin{aligned} r(v) = &\ (f(a+v), g(a+v)) - (f(a), g(a)) - (f'(a) \cdot v, g'(a) \cdot v) \\ = &\ (f(a+v) - f(a) - f'(a) \cdot v, g(a+v) - g(a) - g'(a) \cdot v) \end{aligned}$$

Como f e g são diferenciáveis em a então

$$r_1(v) = f(a+v) - f(a) - f'(a) \cdot v \text{ e } r_2(v) = g(a+v) - g(a) - g'(a) \cdot v.$$

satisfazem,

$$\lim_{v \longrightarrow 0} \frac{r_1(v)}{\|v\|} = \lim_{v \longrightarrow 0} \frac{r_2(v)}{\|v\|} = 0.$$

Daí, $r(v) = (r_1(v), r_2(v))$. Com isso,

$$\lim_{v \longrightarrow 0} \frac{r(v)}{\|v\|} = \left(\lim_{v \longrightarrow 0} \frac{r_1(v)}{\|v\|}, \lim_{v \longrightarrow 0} \frac{r_2(v)}{\|v\|} \right) = (0,0).$$

Ou seja, (f,g) é diferenciável e $a \in X$ e $(f,g)'(a) \cdot v = (f'(a) \cdot v, g'(a) \cdot v)$, para todo $v \in \mathbb{R}^n$.

1.5.6 Teorema do valor médio

Este teorema nos ajudará na demonstração do Teorema da função Inversa. Por isso, faremos sua demonstração.

Teorema 1.5.2. *Dada $f : X \longrightarrow \mathbb{R}$ diferenciável no aberto $X \subseteq \mathbb{R}^n$, se o segmento de reta $[a, a+v]$ estiver contido em X, então existe $\theta \in (0,1)$ tal que*

$$f(a+v) - f(a) = \frac{\partial f}{\partial v}(a+\theta v) = \langle \nabla f(a+\theta v), v\rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+\theta v) \cdot v_i$$

onde $v = (v_1, \cdot, v_n)$.

Demonstração. Seja $\lambda : [0,1] \longrightarrow X$ dado por $\lambda(t) = a + tv$ e considere $f \circ \lambda : [0,1] \longrightarrow \mathbb{R}$. Pelo Teorema do Valor Médio, temos que existe $\theta \in (0,1)$, tal que,

$$f(a+v) - f(a) = f(\lambda(1)) - f(\lambda(0)) = (f \circ \lambda)'(\theta)(1-0) = (f \circ \lambda)'(\theta)$$

Portanto,

$$\begin{aligned} f(a+v) - f(a) &= \lim_{t \longrightarrow 0} \frac{f \circ \lambda(\theta + t) - (f \circ \lambda)(\theta)}{t} = \lim_{t \longrightarrow 0} \frac{f(a + (\theta + t)v) - f(a + \theta v)}{t} \\ &= \frac{\partial f}{\partial v}(a+\theta v) = \langle \nabla f(a+\theta v), v\rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+\theta v) \cdot v_i. \end{aligned}$$

□

1.5.7 Desigualdade do valor médio

O seguinte resultado é útil para a demonstração do Teorema da Função Inversa e faz associação à ideia de contração.

Definição 1.13. *Seja $f : [a,b] \longrightarrow \mathbb{R}^n$ uma aplicação contínua em $[a,b]$ e derivável em $t \in (a,b)$. Considere que $\|f'(t)\| \leq c$, para todo $t \in (a,b)$, onde c é constante. Então,*

$$\|f(b) - f(a)\| \leqslant c(b-a)$$

.

Demonstração. Defina $g : [a,b] \longrightarrow \mathbb{R}$ por $g(t) = \langle f(t), f(b) - f(a)\rangle$. Daí,

$$g'(t) = \langle f'(t), f(b) - f(a)\rangle, \ \textit{para todo}\ t \in (a,b).$$

Note que g é derivável, pois f o é. Daí, existe $\theta \in (a,b)$ tal que,

$$g(b) - g(a) = g'(\theta)(b-a)$$

Consequentemente, $\langle f(b), f(b) - f(a)\rangle - \langle f(a), f(b) - f(a)\rangle = [\langle f'(\theta), f(b) - f(a)\rangle](b-a)$.

Logo,

$$\|f(b) - f(a)\|^2 \leq \|f'(\Theta)\| \|f(b) - f(a)\|(b-a)$$

e daí, $\|f(b) - f(a)\| \leqslant c(b-a)$. □

Teorema 1.5.3. ***(Desigualdade do Valor Médio)****. Sejam $X \subseteq \mathbb{R}^m$ aberto e $f : X \longrightarrow \mathbb{R}^m$ uma função diferenciável tal que*

$$\|f'(x)\| \leqslant M, \ \textit{para todo } x \in X.$$

Então, se o segmento $[a, a+v] \subseteq X$, temos que

$$\|f(a+v) - f(a)\| \leqslant M\|v\|.$$

Demonstração. Seja $\lambda : [0,1] \longrightarrow \mathbb{R}^m$ uma função contínua em $[0,1]$ dada por:

$$\lambda(t) = f(a+tv), \ \textit{para todo } t \in [0,1].$$

Assim sendo,

$$\|f(a+v) - f(a)\| = \|\lambda(1) - \lambda(0)\|$$

e $\lambda'(t) = Jf(a+tv) \cdot v$.

Daí, usando o Teorema anterior,

$$\begin{aligned}
\|f(a+v) - f(a)\| &\leqslant \|\lambda'(t)\|(1-0) \\
&\leqslant \|\lambda'(a+tv) \cdot v\| \\
&\leqslant \|\lambda'(a+tv)\| \|v\| \\
&\leqslant M\|v\|
\end{aligned}$$

□

Corolário 1.1. *Sejam X aberto e convexo do $\mathbb{R}^n$ e $f : X \longrightarrow \mathbb{R}^n$ uma aplicação diferenciável tal que*

$$\|f'(x)\| \leqslant M\|b-a\|, \forall a, b \in X$$

Demonstração. Seja a, b $\in X$. Como X é convexo, então,

$$a + t(b - a) = (1 - t)a + tb \in X, \forall t \in [0, 1].$$

Usando a desigualdade do Valor Médio, obtemos,

$$\|f(b) - f(a)\| \leqslant \|f'(a + t(b - a))\| \|(b - a)\| \leqslant M\|b - a\|.$$

□

1.6 Teorema da função inversa e da função implícita

Nesta seção, trabalharemos o método das aproximações sucessivas do qual faremos uso para alcançar nosso objetivo principal. Para tal, recorreremos as definições de difeomorfismo, difeomorfismo local e alguns teoremas importantes. Agora, apresentaremos os teoremas da Função Inversa e da Função Implícita, bem como algumas aplicações que os acompanham. Estabeleceremos a equivalência dos teoremas mencionados anteriormente e para tal objetivo faremos o uso de todos os resultados que foram apresentados no trabalho.

1.6.1 Difeomorfismo e difeomorfismo local

Aqui, apresentaremos as definições de difeomorfismo e difeomorfismo local

Definição 1.14. *Um homeomorfismo do conjunto $X \subset \mathbb{R}^n$ sobre o conjunto $Y \subset \mathbb{R}^n$ é uma bijeção contínua, cuja inversa é contínua.*

Definição 1.15 (Difeomorfismo). *Sejam $X,\ Y\ \subset \in \mathbb{R}^n$ conjuntos abertos. Um difeomorfismo $f : X \longrightarrow V$ é uma bijeção diferenciável cuja inversa também é diferenciável.*

Exemplo 1.22. *Considere a aplicação $f : \mathbb{R} \longrightarrow \mathbb{R}$, dada por $f(x) = x^3$. Sabemos que f é diferenciável. Lembre que $f^{-1}(x) = \sqrt[3]{x}$. Temos que f é um homeomorfismo diferenciável , mas não é um difeomorfismo, pois sua inversa não é diferenciável no ponto 0, já que o limite*

$$\lim_{x \longrightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \longrightarrow 0} \frac{\sqrt[3]{x} - 0}{x - 0} = \lim_{x \longrightarrow 0} \frac{\sqrt[3]{x}}{x} = \lim_{x \longrightarrow 0} \frac{1}{\sqrt[3]{x^2}}$$

não existe.

Observação 1.6. *Pela definição 1.15 temos que todo difeomorfismo é um homeomorfismo, já que toda função diferenciável é contínua. A recíproca não é verdadeira, pois o fato de ser contínua em um ponto não garante ser diferenciável neste ponto.*

Definição 1.16. *Seja uma* $f : X \longrightarrow \mathbb{R}$ *uma função definida no aberto* $X \subseteq \mathbb{R}^m$*. Dizemos que* f *é de classe* C^k *e escrevemos* $f \in C^k$ *em que,* $1 \leqslant k \leqslant \infty$*, se em cada ponto de* X*,* f *possui todas as derivadas parciais de ordem* K *e estas são contínuas em* X*. Dizemos que* $f \in C^\infty$ *se* $f \in C^k$*, para todo* $K \in \mathbb{N}$*.*

Exemplo 1.23. *Seja* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *uma aplicação dada por* $f(x,y) = e^{xy}$*, temos que*

$$\frac{\partial f}{\partial x}(x,y) = ye^{xy}, \frac{\partial^2 f}{\partial x \partial x}(x,y) = y^2 e^{xy}, \frac{\partial^2 f}{\partial x \partial x}(x,y) = e^{xy} + xye^{xy} = e^{xy}(1+xy).$$

$$\frac{\partial f}{\partial y}(x,y) = xe^{xy}, \frac{\partial^2 f}{\partial y \partial y}(x,y) = x^2 e^{xy}, \frac{\partial^2 f}{\partial x \partial y}(x,y) = e^{xy} + +xye^{xy} = e^{xy}(1+xy).$$

Logo, $f \in C^2$*. Fazendo essas derivadas de maneira sucessiva, temos a forma clara que* $f \in C^\infty$*.*

Definição 1.17 (Difeomorfismo Local). *Dizemos que uma aplicação diferenciável* $f : X \longrightarrow \mathbb{R}^m$ *, definida no aberto* $X \subset \mathbb{R}^m$ *é um Difeomorfismo Local, se para cada* $x \in X$*, existir um aberto* V_x*, e que* $x \in V_x \subset X$ *e a restrição de* f *a* V_x *é um Difeomorfismo sobre um aberto* W_x *(que contém* $f(x)$*). Quando* f *é de classe* C^k*, dizemos que* f *é um difeomorfismo local de classe* C^k*.*

Observação 1.7. *É importante sabermos que todo difeomorfismo é um difeomorfismo local. Porém, a recíproca não é válida. Este fato será justificado pelo Teorema da Função Inversa.*

Vejamos, agora, como derivar uma composta de funções diferenciáveis.

Teorema 1.6.1 (Regra da Cadeia). *Sejam* $X \subseteq \mathbb{R}^n$*,* $Y \subseteq \mathbb{R}^m$ *conjuntos abertos. Sejam* $f : X \longrightarrow \mathbb{R}^m$ *e* $g : Y \longrightarrow \mathbb{R}^p$ *diferenciáveis em* $a \in X$ *e* $b = f(a) \in Y$*, respectivamente, com* $f(X) \subseteq Y$*. Então*

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$$

Demonstração. Sabemos que

$$f(a+v) = f(a) + f'(a).v + r(v)$$

e,

$$g(b+w) = g(b) + g'(b).w + s(w),$$

onde $\lim\limits_{w\longrightarrow 0}\dfrac{s(w)}{\|w\|}=0$ e $\lim\limits_{v\longrightarrow 0}\dfrac{r(v)}{\|v\|}=0$. Com isso,

$$g(f(a+v))=g(f(a)+f'(a).v+r(v))$$

Consideremos $w=f'(a).v+r(v)$. Logo,

$$\begin{aligned}
g\circ f(a+v) &= g(f(a+v))\\
&= g(f(a)+w)\\
&= g(b+w)\\
&= g(b)+g'(b).w+s(w)\\
&= g(b)+g'(b)[f'(a).v+r(v)]+s(w)\\
&= g(b)+g'(b)\circ f'(a).v+g'(b).r(v)+s(w)
\end{aligned}$$

Seja $p(v)=g'(f(a))r(v)+s(w)$. Devemos provar que $\lim\limits_{v\longrightarrow 0}\dfrac{p(v)}{\|v\|}=0$. Mas

$$\frac{p(v)}{\|v\|}=g'(f(a)).\frac{r(v)}{\|v\|}+\frac{s(w)}{\|w\|}\frac{\|w\|}{\|v\|}$$

Além disso, $\dfrac{\|w\|}{\|v\|}$ é limitado quando v é suficientemente pequeno. De fato, primeiramente observe que

$$v\longrightarrow 0\Rightarrow w=f'(a)v+r(v)=f'(a)v+\frac{r(v)}{\|v\|}\|v\|\longrightarrow 0$$

Pois $f'(a)$ é contínua e $\lim\limits_{v\longrightarrow 0}\dfrac{r(v)}{\|v\|}=0$. Logo,

$$\begin{aligned}
\frac{\|w\|}{\|v\|} &= \frac{\|f'(a)v+r.(v)\|}{\|v\|}\\
&= \left\|f'(a).\frac{v}{\|v\|}+\frac{r(v)}{\|v\|}\right\|\\
&\leqslant \|f'(a)\|+\frac{\|r(v)\|}{\|v\|}.
\end{aligned}$$

Como $\lim\limits_{v\longrightarrow 0}\dfrac{r(v)}{\|v\|}=0$, então $\dfrac{\|w\|}{\|v\|}$ é limitado para v suficientemente pequeno. Por fim,

$$\frac{p(v)}{\|v\|}=g'(b).\frac{r(v)}{\|v\|}+\frac{s(w)}{\|w\|}.\frac{\|w\|}{\|v\|}\longrightarrow 0$$

quando $v \longrightarrow 0$.

Portanto, $g \circ f$ é diferenciável em $a \in X$ e $(g \circ f)'(a) = g'(b) \cdot f'(a) = g'(f(a) \cdot f'(a))$.

$\square$

Exemplo 1.24. *Se $f : \mathbb{R} \longrightarrow \mathbb{R}^2$ é dada por $f(x) = (\cos x, senx)$ e $g : \mathbb{R} \longrightarrow \mathbb{R}$ por $g(x) = x^2$, temos que $f \circ g : \mathbb{R} \longrightarrow \mathbb{R}^2$ é dada por*

$$f(g(x)) = f(x^2) = (\cos(x^2), sen(x^2)).$$

Daí,

$$\begin{aligned}(f \circ g)'(x) &= f'(g(x)).g'(x) \\ &= (-sen(x^2), \cos(x^2))2x \\ &= (-2xsen(x^2), 2x\cos(x^2).)\end{aligned}$$

Corolário 1.2. *Se $f : X \longrightarrow \mathbb{R}^n$, $g : X \longrightarrow \mathbb{R}^p$, onde $X \subseteq \mathbb{R}^n$ e $Y \subseteq \mathbb{R}^n$ são abertos e $f(X) \subseteq Y$, são ambas classe C^k , então $g \circ f : X \longrightarrow \mathbb{R}^n$ também é de classe C^k.*

Demonstração. Sabemos que $g \circ f$ é k vezes diferenciável e que

$$(g \circ f)' \equiv (g' \circ f) \circ f' : X \longrightarrow L(\mathbb{R}^n, \mathbb{R}^p)$$

Se $f, g \in C^1$, então

$$(g \circ f)' = (g' \circ f).f' \in C^0$$

Logo, $f, g \in C^1$. Se $f, g \in C^2$, obtemos

$$(g \circ f)' = (g' \circ f).f' \in C^1$$

Portanto, $g \circ f \in C^2$. Por indução, supondo que $f, g \in C^k$, a igualdade acima garante $(g \circ f)' \in C^{k-1}$. Portanto, $(g \circ f)' \in C^k$.

$\square$

Teorema 1.6.2. *Seja $X \subseteq \mathbb{R}^m$ e $f : X \longrightarrow \mathbb{R}^n$ um difeomorfismo local de classe C^k, $k \geqslant 1$, de X sobre $Y = f(X)$. Então para cada $x \in X$, a derivada $Df(x) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ é um isomorfismo.*

Demonstração. Pelo enunciado do Teorema, temos que $f : X \longrightarrow \mathbb{R}^m$ é um difeomorfismo local de X sobre $V = f(X)$, logo para cada $x \in X$, existe aberto V_x contendo x e um aberto W_x, contendo $f(x)$, tais que, $V_x \subseteq X$ e $g = f|Y_x : Y_x \longrightarrow W_x$ é um difeomorfismo. Com g é um difeomorfismo, temos que

$$g^{-1} \circ g = Id_{V_x}.$$

Derivando , obtemos

$$D(g^{-1} \circ g) = D(Id_{V_x}).$$

Usando o Teorema da Regra da Cadeia, segue que

$$D(g^{-1})(g(x)) \cdot D(g(X)) = Id_{\mathbb{R}^m},$$

pois Id_{V_x} é linear. Portanto, se $v \in \ker D(g(x))$, onde $Dg(x) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ é linear, temos que

$$Dg(x) \cdot v = 0 \text{ e}$$
$$v = Id_{\mathbb{R}^m} = D(g^{-1})(g(x)) \cdot (D((g(x) \cdot v)) = D(g^{-1})(g(x)) \cdot 0 = 0.$$

já que, $D(g^{-1})(g(x))$ é linear. Logo, temos que $v = 0$ e , pelo Teorema do Núcleo e da Imagem, segue que $D((g(x)) = D(f|Y_x)(x) = Df(x) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ é um isomorfismo. □

Faremos agora um exemplo que ilustra o Teorema acima.

Exemplo 1.25. *Dada a aplicação $f : \mathbb{R} \longrightarrow$ (0,∞), definida por $f(x) = e^x$, temos que f é um difeomorfismo de classe C^∞ . Logo f é um difeomorfismo local de classe C^∞, em que$Df(x) : \mathbb{R} \longrightarrow \mathbb{R}$ é dada por $Df(x) \cdot \lambda = \lambda\ f'(x) = \lambda e^x$. Sabemos que $Df(x)$ é linear,mas mostraremos este fato na prática. Veja que*

$$Df(x)(\lambda_1 + \lambda_2) = (\lambda_1 + \lambda_2)e^x = \lambda_1 e^x + \lambda_2 e^x = Df(x) \cdot \lambda_1 + Df(x) \cdot \lambda_2$$

Além disso,

$$Df(x) \cdot \alpha\lambda = (\alpha\lambda)e^x = \alpha Df(x) \cdot \lambda$$

Agora, vamos provar que $Df(x)$ é um isomorfismo. Com efeito, se $\lambda \in \ker Df(x)$, então $Df(x) \cdot \lambda = 0$, isto é, $\lambda e^x = 0$. Como $e^x \neq 0$, segue que λ = 0. Utilizando o Teorema do Núcleo e da Imagem, temos que $Df(x) : \mathbb{R} \longrightarrow \mathbb{R}$ é um isomorfismo.

Teorema 1.6.3. *Todo difeomorfismo local $f : X \longrightarrow \mathbb{R}^m$ é um aplicação aberta, isto é, transforma cada abeto $Y \subseteq X$ num aberto $f(X) \subseteq \mathbb{R}^m$.*

Demonstração. Seja $Y \subseteq X$ um aberto. Fixe $x \in$X. Se $Y \subseteq Y_x$, onde $Y_x \subseteq X$ é um aberto contendo x e $f : Y_x \longrightarrow f(Y_x)$ é um difeomorfismo, então $f(Y_x)$ é aberto, pois f é um homeomorfismo sobre Y_x. Em geral

$$Y = X_{x \in X}(Y_x \cap Y)$$

Portanto, $f(Y) = X_{x \in X}$. Mas $Y_x \cap Y \subseteq Y_x$ é um aberto, logo, pelo que foi feito anteriormente $f(Y_x \cap Y)$ é aberto em $\mathbb{R}^m$. Portanto $f(Y)$ é aberto em $\mathbb{R}^m$ como união arbitrária de abertos.

□

Contração

É importante definirmos contrações, pois esse resultado é primordial para demonstrarmos alguns resultados que estão inteiramente relacionados com o Teorema da Função Inversa.

Definição 1.18. *Seja $X \subseteq \mathbb{R}^m$. Uma aplicação $f : X \subset \mathbb{R}^m \to \mathbb{R}^n$ chama-se uma λ-contração quando existe $0 \leq \lambda < 1$ tal que*

$$\|f(x) - f(y)\| \ \|x\|_{\mathbb{R}^m} \leq \lambda \|x - y\| \ \|f(x)\|_{\mathbb{R}^m}, \text{ para todo } x, y \in X.$$

Exemplo 1.26. *Considere uma aplicação $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = \dfrac{1}{2}x + \sqrt{2}$. É fácil ver que essa função é uma $\frac{1}{2}$-contração. Com efeito,*

$$\begin{aligned} \|f(x) - f(y)\| &= \left\| \frac{1}{2}x + \sqrt{2} - \left(\frac{1}{2}y + \sqrt{2} \right) \right\| \\ &= \left\| \frac{1}{2}(x - y) \right\| \\ &= \frac{1}{2}\|x - y\|, \text{ para todo } x, y \in \mathbb{R}^m. \end{aligned}$$

Observação 1.8. *Por definição, toda λ-contração é uma função função Lipschitziana. Como toda função Lipschitziana é uniformemente contínua, segue que toda λ-contração é uniformemente contínua.*

Ponto fixo

Nesta subseção, daremos nome aos pontos que são mantidos pela lei de transformação de uma função. Estes serão denominados pontos fixos. Mais precisamente,

Definição 1.19. *Um ponto fixo de uma aplicação* $f : X \subseteq \mathbb{R}^m \to \mathbb{R}^m$ *é um ponto* $x \in X$ *tal que* $f(x) = x$.

Exemplo 1.27. *Seja* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *uma aplicação definida por*

$$f(x, y) = \left(sen(x), \frac{\pi}{2} - \cos(y)\right).$$

O ponto $\left(0, \frac{\pi}{2}\right)$ *é um ponto fixo de* f*, pois,*

$$f\left(0, \frac{\pi}{2}\right) = \left(sen(0), \frac{\pi}{2} - \cos\left(\frac{\pi}{2}\right)\right) = \left(0, \frac{\pi}{2} - 0\right) = \left(0, \frac{\pi}{2}\right).$$

Por outro lado, é fácil ver que $(0, \pi)$ *não é ponto fixo de* f*, já que,*

$$f\left(0, \pi\right) = \left(sen(0), \frac{\pi}{2} - \cos\left(\pi\right)\right) = \left(0, \frac{\pi}{2} + 1\right) = \left(0, \frac{\pi + 2}{2}\right).$$

1.6.2 Aproximações sucessivas

A seguir demonstraremos o Teorema do Ponto Fixo para Contrações que tem uma importância significativa para o Teorema da Função Inversa. O método que apresentaremos a seguir é conhecido como Método das Aproximações Sucessivas.

Teorema 1.6.4. *(Teorema do Ponto Fixo para Contrações): Sejam* $F \subseteq \mathbb{R}^m$ *um conjunto fechado e* $f : F \to F$ *uma* λ*-contração. Dado* $x_0 \in F$*, a sequência*

$$x_1 = f(x_0), \cdots, x_{k+1} = f(x_k), \cdots,$$

converge para um ponto $a \in F$*, e este é o único ponto fixo de* f*.*

Demonstração. Como $f : F \to F$ é uma λ-contração, então existe $\lambda \in (0, 1)$ tal que

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|, \text{ para todo } x, y \in F.$$

Portanto,

$$\begin{aligned}
\|x_{k+1} - x_k\| &= \|f(x_k) - f(x_{k-1})\| \\
&\leq \lambda\|x_k - x_{k-1}\| \\
&= \lambda\|f(x_{k-1}) - f(x_{k-2})\| \\
&\leq \lambda^2\|x_{k-1} - x_{k-2}\| \\
&= \lambda^2\|f(x_{k-2}) - f(x_{k-3})\|.
\end{aligned}$$

Prosseguindo este processo, chegamos a

$$\|x_{k+1} - x_k)\| \leq \lambda^k\|x_1 - x_0\|.$$

Assim segue, da desigualdade triangular,que

$$\begin{aligned}
\|x_{k+p} - x_k)\| &= \|x_{k+p} - x_{k+p-1} + x_{k+p-1} - x_{k+p-2} + x_{k+p-2} - x_{k+1} + x_{k+1} - x_k\| \\
&\leq \|x_{k+p} - x_{k+p-1}\| + \|x_{k+p-1} - x_{k+p-2}\| + \cdots + \|x_{k+1} - x_k\| \\
&\leq \lambda^{k+p-1}\|x_1 - x_0\| + \lambda^{k+p-2}\|x_1 - x_0\| + \cdots + \lambda^k\|x_1 - x_0\| \\
&= \sum_{i=0}^{p-1} \lambda^{k+i}\|x_1 - x_0\| \\
&= \|x_1 - x_0\| \sum_{i=0}^{p-1} \lambda^{k+i} \\
&\leq \frac{\lambda^k}{1-\lambda}\|x_1 - x_0\|,
\end{aligned}$$

pois,

$$\sum_{i=0}^{p-1} \lambda^{k+i} = \sum_{i=0}^{p-1} \lambda^k.\lambda^i = \lambda^k \sum_{i=0}^{p-1} \lambda^i \leq \lambda^k \sum_{i=0}^{\infty} \lambda^i = \lambda^k \frac{1}{1-\lambda} = \frac{\lambda^k}{1-\lambda}.$$

É importante destacar que $1 - \lambda > 0$ e $\sum_{i=0}^{\infty} \lambda^i$ é uma série geométrica. Como $\lambda \in (0, 1)$, então $\|x_{k+p} - x_k\| \longrightarrow 0$, quando $p \longrightarrow \infty$. Assim (x_k) é uma sequência de Cauchy em $\mathbb{R}^m$. Como $\mathbb{R}^m$ é completo, então existe $a \in \mathbb{R}^m$ tal que $\lim\limits_{k\longrightarrow\infty} x_k = a$. Mas F é fechado e $(x_k) \subseteq F$, então $a \in F$. Vamos provar agora que a é o ponto fixo de f. Como f é uma λ-contração então f é uniformemente contínua, logo

$$a = \lim_{k\longrightarrow\infty} x_{k+1} = \lim_{k\longrightarrow\infty} f(x_k) = f(a),$$

ou seja, $f(a) = a$, daí a é ponto fixo de f. Vamos provar que a é o único ponto fixo de f. Seja b um ponto fixo de f, então $b = f(b)$, logo

$$\|b - a\| = \|f(b) - f(a)\| \leq \lambda\|b - a\|,$$

pois f é uma λ-contração. Desse modo,

$$\|b-a\| - \lambda\|b-a\| \leq 0 \Longrightarrow \|b-a\|(1-\lambda) \leq 0.$$

Como $(1-\lambda) > 0$, temos que $\|b-a\| = 0 \Longrightarrow a = b$. Ou seja, a é o único ponto fixo de f.

□

Lema 1.1. *Seja $f : X \to \mathbb{R}^n$ uma λ-contração. Se X contém a bola fechada $B_r[a]$ tal que*

$$\|f(a) - a\| \leq (1-\lambda)r,$$

então f admite um ponto fixo em $B_r[a]$.

Demonstração. Vamos provar que, $f(B_r[a]) \subseteq B_r[a]$. De fato, seja $y \in f(B_r[a])$ então $y = f(x)$, onde $x \in B_r[a]$. Daí,

$$\begin{aligned} \|y-a\| &= \|f(x)-a\| \\ &= \|f(x) - f(a) + f(a) - a\| \\ &\leq \|f/(x) - f(a)\| + \|f(a) - a\| \\ &\leq \lambda\|x-a\| + (1-\lambda)r, \end{aligned}$$

pois f é uma λ-contração e $\|f(a) - a\| \leq (1-\lambda)r$. Portanto $\|y-a\| \leq r$, daí $y \in (B_r[a])$, logo $f(B_r[a]) \subseteq B_r[a]$. Segue que $f|_{B_r[a]} : B_r[a] \longrightarrow B_r[a]$. Como $B_r[a]$ é fechado segue pelo Teorema do Ponto Fixo para Contrações que f admite um ponto fixo em $B_r[a]$.

□

1.6.3 Pertubação da identidade

Demonstraremos, nessa subseção, o Teorema da Pertubação da Identidade, e para que tal fato seja finalizado utilizaremos o Lema 1.1 juntamente com as definições de λ-contração, homeomorfismo e conjuntos abertos.

Teorema 1.6.5. *Seja $\varphi : X \subseteq \mathbb{R}^m \to \mathbb{R}^m$ uma λ-contração definida no aberto X. A aplicação $f : X \to \mathbb{R}^m$ dada por $f(x) = x + \varphi(x)$ é um homeomorfismo de X sobre o conjunto aberto $f(X) \subseteq \mathbb{R}^m$. Além disso, se $X = \mathbb{R}^m$, tem-se que $f(X) = \mathbb{R}^m$.*

Demonstração. Primeiramente provaremos que f é bijetiva sobre $f(U)$. De fato,

$$\begin{aligned} \|f(x) - f(y)\| &= \|x + \varphi(x) - (y + \varphi(y))\| \\ &= \|x - y + \varphi(x) - \varphi(y)\| \\ &\geq \|x - y\| - \|\varphi(x) - \varphi(y)\| \end{aligned}$$

.

Como φ é uma λ-contração, segue que

$$\|f(x) - f(y)\| \geq \|x - y\| - \lambda\|x - y\| = (1 - \lambda)\|x - y\|, \ \textit{para todo } x, y \in U.$$

Assim se $x \neq y$ temos que,

$$\|f(x) - f(y)\| \geq (1 - \lambda)\|x - y\| > 0.$$

Logo $f(x) \neq f(y)$. Portanto f é injetiva. Ou seja, f é uma bijeção sobre $f(X)$. Agora vamos verificar que f é um homeomorfismo. Veja que,

$$\|f^{-1}(w) - f^{-1}(z)\| \leq \frac{1}{1 - \lambda}\|w - z\|, \ \textit{para todo } w, z \in f(X).$$

Por conseguinte $f^{-1} : f(X) \longrightarrow X$ é contínua. Como φ é uma λ-contração (logo, uniformemente contínua), então $f : X \longrightarrow f(X)$ é uma aplicação contínua, pois, f é a soma de funções contínuas. Portanto como f é uma bijeção contínua e sua inversa, f^{-1}, também é contínua, segue que f é um homeomorfismo. Vamos provar agora que $f(X)$ é aberto. Tomemos $b \in f(X)$, devemos mostrar que existe $s > 0$ tal que $B_s(b) \subseteq f(X)$. Mas, existe $a \in X$ tal que $b = f(a) = a + \varphi(a)$. Como X é aberto, então existe $r > 0$ tal que $B_r[a] \subseteq X$. Vamos definir $\xi_y : B_r[a] \to \mathbb{R}^m$, onde $y \in \mathbb{R}^m$ é fixo, por $\xi_y(x) = y - \varphi(x)$, *para todo* $x \in B_r[a]$. Então

$$\xi_y(x) = x \Leftrightarrow x = y - \varphi(x) \Leftrightarrow x + \varphi(x) = y \Leftrightarrow f(x) = y, x \in B_r[a].$$

É fácil ver que, ξ_y é λ-contração. De fato,

$$\begin{aligned} \|\xi_y(x) - \xi_y(x')\| &= \|y - \varphi(x) - (y - \varphi(x'))\| \\ &= \|\varphi(x) - \varphi(x')\| \\ &\leq \lambda\|x - x'\|, \ \textit{para todo } x \in B_r[a], \end{aligned}$$

pois φ é uma λ-contração. Seja $y \in B_{(1-\lambda)r}[b]$, então

$$\begin{aligned} \|\xi_y(a) - a\| &= \|y - \varphi(a) - a\| \\ &= \|y - (\varphi(a) + a) \\ &= \|y - b\| \\ &\leq (1 - \lambda)r, \end{aligned}$$

pois, $y \in B_{(1-\lambda)r}[b]$. Logo, pelo Lema 1.1, ξ_y tem um ponto fixo em $B_r[a]$. Ou seja, $\xi_y(x) = x se, e somente se, f(x) = y, x \in B_r[a]$. Daí $B_{(1-\lambda)r}[b] \subseteq f(B_r[a]) \subseteq f(X)$, pois, $B_r[a] \subseteq X$. Seja

$$s = (1-\lambda)\frac{r}{2}, \text{ então } B_s(b) \subseteq f(U).$$

Ou seja, $f(X)$ é aberto.

Agora provaremos que se $X = \mathbb{R}^m$, então $f(X) = \mathbb{R}^m$. Sabemos que $f(X)$ é aberto, então basta mostrar que $f(X)$ é fechado. Seja $(y_k) \subseteq f(X)$ uma sequência tal que $y_k \longrightarrow y \in \mathbb{R}^m$, então existe $(x_k) \subseteq X = \mathbb{R}^m$ tal que $f(x_k) = y_k$ e vimos que

$$\|y_k - y_m\| = \|f(x_k) - f(x_m)\| \geq (1-\lambda)\|x_k - x_m\|,$$

com $(1-\lambda) > 0$. Como (y_k) é convergente, (logo, de Cauchy) então (x_k) é de Cauchy em $X = \mathbb{R}^m$, que é um conjunto completo. Logo $\exists \ \ x \in X = \mathbb{R}^m$ tal que $x_k \longrightarrow x$. Como f é contínua em $X = \mathbb{R}^m$, então $y_k = f(x_k) \longrightarrow f(x)$. Logo $y = f(x) \in f(X)$. Portanto $f(X)$ é fechado. Dessa forma $f(X) = \mathbb{R}^m$, já que $\mathbb{R}^m$ é conexo e $f(X) \neq \emptyset$.

□

1.6.4 Pertubação de um isomorfismo

Utilizaremos o Teorema da Pertubação da Identidade bem como as definições de aplicação linear e λ-contração para demonstrarmos o Teorema da Pertubação de um Isomorfismo.

Corolário 1.3. *Sejam $X \subseteq \mathbb{R}^m$ um aberto e $f : X \to \mathbb{R}^m$ uma aplicação definida por $f(x) = Tx + \varphi(x)$, onde $T : \mathbb{R}^m \to \mathbb{R}^m$ é uma aplicação linear invertível e $\varphi : X \to \mathbb{R}^m$ satisfaz*

$$\|\varphi(x) - \varphi(y)\| \leq \lambda\|x - y\|, \ \textit{para todo } x, y \in X,$$

com $\lambda\|T^{-1}\| < 1$. Então f é um homeomorfismo de X sobre o conjunto aberto $f(X) \subseteq \mathbb{R}^m$. Se $X = \mathbb{R}^m$, tem-se $f(X) = \mathbb{R}^m$.

Demonstração. Defina $g : X \to \mathbb{R}^m$ por $g(x) = T^{-1} \circ \varphi(x), \ para\ todo x \in X$. Afirmamos que g é uma $\lambda\|T^{-1}\|$-contração. De fato,

$$\begin{aligned}
\|g(x) - g(y)\| &= \|T^{-1} \circ \varphi(x) - T^{-1} \circ \varphi(y)\| \\
&= \|T^{-1}(\varphi(x) - \varphi(y))\| \\
&\leq \|T^{-1}\| \|\varphi(x) - \varphi(y)\| \\
&\leq \lambda\|T^{-1}\| \|x - y\|, \ \textit{para todo } x, \ y \in X,
\end{aligned}$$

pois, φ é uma λ-contração e T é linear invertível. Por outro lado, obtemos

$$T^{-1} \circ f(x) = T^{-1}(Tx + \varphi(x)) = T^{-1} \circ Tx + T^{-1} \circ \varphi(x) = x + g(x), \ \textit{para todox} \in X.$$

Assim, pelo Teorema da Pertubação de Identidade, $T^{-1} \circ f$ é um homeomorfismo de X sobre $T^{-1}f(X)$. Logo, $f = T \circ (T^{-1}f) : X \to \mathbb{R}^m$ é um homeomorfismo de X sobre $f(X)$, pois T é um homeomorfismo (já que T é linear). Se $X = \mathbb{R}^m$, então pelo Teorema da Pertubação da Identidade $T^{-1} \circ f(X) = \mathbb{R}^{\geqslant}$. Portanto, $f(X) = T \circ T^{-1} \circ f(X) = T(\mathbb{R}^m)= \mathbb{R}^m$.

□

1.6.5 Diferenciabilidade do homeomorfismo inverso

Estudaremos quando um homeomorfismo diferenciável tem inversa diferenciável. Neste caso, existe uma relação entre derivadas da função e sua inversa.

Lema 1.2. *Seja $f : X \to Y$ um homeomorfismo entre abertos $X, Y \subseteq \mathbb{R}^m$. Se f é diferenciável num ponto $a \in X$ e a derivada $Df(a) : \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo, então o homeomorfismo inverso $f^{-1} : Y \to X$ é diferenciável no ponto $b = f(a)$.*

Demonstração. Denotaremos por g a aplicação f^{-1}. O único candidato para $Dg(b)$ é $Df(a)^{-1}$. Assim, tomemos,

$$g(b + w) - g(b) = Df(a)^{-1}.w + s(w), w \in \mathbb{R}^m.$$

Vamos mostrar que, $\displaystyle\lim_{w \longrightarrow 0} \frac{s(w)}{\|w\|} = 0$. Seja $v = g(b - w) - g(b)$. Então,

$$f(a + v) - f(a) = f(a + g(b + w) - g(b)) - b = f(g(b - w)) - b = b + w - b = w,$$

onde $f(a) = b$ e $g(b) = a$. Como f é um homeomorfismo, então, $w = [f(a+v) - f(a)] \longrightarrow 0$ se $v \longrightarrow 0$. Reciprocamente, $v = [g(b+w) - g(b)] \longrightarrow 0$ se $w \longrightarrow 0$. Assim $v \longrightarrow 0 \Leftrightarrow w \longrightarrow 0$. Como f é diferenciável em a, então

$$w = f(a+v) - f(a) = Df(a).v + r(v), \text{ onde } \lim_{v \longrightarrow 0} \frac{r(v)}{\|v\|} = 0.$$

Logo,

$$\begin{aligned} v &= g(b+w) - g(b) \\ &= Df(a)^{-1}.w + s(w) \\ &= Df(a)^{-1}.(Df(a).v + r(v)) + s(w) \\ &= v + Df(a)^{-1}.r(v) + s(w). \end{aligned}$$

Daí, $s(w) = -Df(a)^{-1}.r(v) \Rightarrow \dfrac{s(w)}{\|w\|} = -Df(a)^{-1}.\dfrac{s(w)}{\|w\|}$. Como $Df(a)$ é isomorfismo, então

$$\|v\| = \|Df(a)^{-1} \circ Df(a).v\| \leq \|Df(a)^{-1}\| \|Df(a).v\| \; \textit{para todo } v \in \mathbb{R}^m.$$

Seja $c = \dfrac{1}{2\|Df(a)^{-1}\|}$, então, $\|Df(a).v\| \geq 2c\|v\|$, *para todo* $v \in \mathbb{R}^m$. Como $\lim\limits_{v \longrightarrow 0} \dfrac{r(v)}{\|v\|} = 0$, então existe $\delta > 0$ tal que, $\|v\| < \delta \Rightarrow \|r(v)\| < c\|v\|$. Logo,

$$\begin{aligned} \|w\| &= \|f(a+v) - f(a)\| \\ &= \|Df(a).v + r(v)\| \\ &\geq |Df(a).v\| - \|r(v)\| \\ &\geq 2c\|v\| - c\|v\| \\ &= c\|v\|, \; \textit{se } \|v\| < \delta. \end{aligned}$$

Ou seja, $\dfrac{\|v\|}{\|w\|} \leq \dfrac{1}{c}$, se $\|v\| < \delta$. Daí, $\lim\limits_{w \longrightarrow 0} \dfrac{r(v)}{\|v\|}.\dfrac{\|v\|}{\|w\|} = 0$, pois, $\dfrac{r(v)}{\|v\|} \longrightarrow 0$ e $\dfrac{\|v\|}{\|w\|}$ é limitado. Portanto,

$$\begin{aligned} \lim_{w \longrightarrow 0} \frac{s(w)}{\|w\|} &= -\lim_{w \longrightarrow 0} Df(a)^{-1} \left(\frac{r(v)}{\|w\|} \right) \\ &= -\lim_{w \longrightarrow 0} Df(a)^{-1} \left(\frac{r(v)}{\|v\|} \right) \left(\frac{\|v\|}{\|w\|} \right) = 0. \end{aligned}$$

Ou seja, $g = f^{-1}$ é diferenciável em $b = f(a)$ e $D(f^{-1}(b)) = Df(a)^{-1}$.

$\square$

1.6.6 Teorema da Função Inversa

Este teorema trata da possibilidade de inverter uma função, mesmo que localmente. Além disso, fala sobre a perturbação do isomorfismo e da diferenciabilidade de um homeomorfismo inverso.

A seguir, enunciaremos e demonstraremos esse Teorema para que através do mesmo, possamos estabelecer uma relação com o Teorema da Função Implícita.

Teorema 1.6.6. *Sejam* $X \cap \mathbb{R}^m$ *um aberto e* $f : X \longrightarrow \mathbb{R}^m$ *de classe* $C^k (k \geq 1)$ *tal que* $Df(x_0)$*:* $\mathbb{R}^m \longrightarrow \mathbb{R}^m$ *é um isomorfismo, onde* $x_0 \in X$*. Então* f *é um difeomorfismo de classe* C^k *de uma vizinhança* V *de* x_0 *sobre uma vizinhança* W *de* $f(x_0)$*.*

Demonstração. Para simplificar a notação, suponhamos $x_0 = f(x_0) = 0$ (o resultado geral se obtém por translação). Como $f \in C^k(k \geq 1)$, temos

$$f(x) = f(x) - f(0) = Df(0).x + r(x),$$

onde $\lim_{x \to 0} \frac{r(x)}{\|x\|} = 0$ Portanto, $r(x) = f(x) - Df(0)x$ é de classe C^k, pois f é de classe C^k e $Df(0)$ é linear. Além disso,

$$Dr(0) = Df(0) - Df(0) = 0.$$

Seja λ tal que $0 < \lambda < \dfrac{1}{\|Df^{-1}(0)\|}$ ($Df(0)$ é um isomorfismo). Como $r \in C^k$, então $Dr : X \longrightarrow L(\mathbb{R}^m)$ é contínua e assim,

$$\lim_{x \to 0} Dr(x) = Dr(0) = 0.$$

Com isso, existe $\delta > 0$ tal que $\|x\| < \delta$ então $\|Dr(x)\| < \lambda$. Isto é, existe $V = B_\delta(0)$ vizinhança de $x_0 = 0$ tal que,

$$\|Dr(x)\| < \lambda, \ para \ todo \ x \in Y.$$

Pela Desigualdade do Valor médio, temos que

$$\|r(x) - r(y)\| \leqslant \lambda \|x - y\|, \ para \ todo \ x, y \in Y (convexo).$$

Pelo Corolário da perturbação do Isomorfismo, $f|_Y$ é um homeomorfismo de Y sobre um aberto W que contém $f(x_0)$. Como $f \in C^k$, então $Df : X \longrightarrow \mathrm{L}(\mathbb{R}^m)$ é contínua. Além disso, $Df(x_0)$ é um isomorfismo, isto é $[Df(x_0)]$ é uma matriz invertível. Portanto, diminuindo V se necessário, $[Df(x)]$ é invertível (basta usar a continuidade do determinante), ou seja, $Df(x)$ é um isomorfismo, *para todo* $x \in Y$. Pelo lema da Diferenciabilidade do Homeomorfismo Inverso, segue que $g = f^{-1}$: $W \longrightarrow Y$ é diferenciável em $y = f(x)$ e $Dg(y) = Df^{-1}$. Em particular, $f|_Y : Y \longrightarrow W$ é um difeomorfismo.

Vamos agora mostrar que, $g \in C^k$. Note que $Dg = i \circ Df \circ g$, onde $i(x) = x^{-1}$ é de classe C^∞. Como $f \in C^1$ e i,Df, g são contínuas, então Dg é contínua, isto é, $g \in C^1$. Se $f \in C^1$ então $i, Df, g \in C^1$, logo $Dg \in C^1$, isto é, $g \in C^2$. Logo o resultado segue.

□

Exemplo 1.28. *Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $f(x, y) = (e^x \cos y, e^x sen\ y)$. Daí, a matriz Jacobiana é dada por*

$$[Df(x, y)] = \begin{pmatrix} e^x \cos y & -e^x sen\ y \\ e^x sen\ y & e^x \cos y \end{pmatrix}$$

. e, portanto, $\det[Df(x, y)] = e^{2x} \neq 0$, *para todo $(x, y) \in \mathbb{R}^2$. Logo, $Df(x, y)$ é um isomorfismo para cada $(x, y) \in \mathbb{R}^2$. Pelo Teorema da função Inversa, f é um difeomorfismo local, em particular, vimos que f não é um difeomorfismo. Isso justifica que nem sempre um difeomorfismo local é um difeomorfismo.*

Exemplo 1.29. *Lembre que uma matriz X é uma raiz quadrada de uma matriz Y quando $X^2 = Y$. Vamos utilizar o Teorema da Função Inversa para provar que, próximo da identidade, toda matriz tem uma raiz quadrada. Seja $f : M_n\ (\mathbb{R}) \longrightarrow M_n(\mathbb{R})$ dada por $f(X) = X^2$, onde $M_n(\mathbb{R})$ é o conjunto das matrizes reais $n \times n$. Assim,*

$$f(I) = I^2 = I,\ f \in C^2 \text{ e } Df(X)H = XH + HX.$$

Em particular,

$$Df(I)H = 2H, \forall H \in M_n(\mathbb{R}).$$

Logo, $Df(I) = 2I$ é um isomorfismo. Portanto, pelo Teorema da Função Inversa, existe uma vizinhança V de I tal que $f : V \longrightarrow f(V)$ é um difeomorfismo de classe C^1. Logo, para todo $I \in f(V)$, existe um único $x \in V$ tal que $f(X) = Y$.

Isto é, $X^2 = Y$. Portanto, X é uma raiz quadrada de $Y \in f(V)$ (aberto contendo $I = f(I)$).

A seguir, provaremos que o Teorema da Função Inversa implica o Teorema da Função Implícita.

1.6.7 Teorema da função implícita

Teorema 1.6.7. *Sejam $X \subseteq \mathbb{R}^{m+n}$ um aberto e $f : X \longrightarrow \mathbb{R}^n$ uma aplicação de classe C^k ($k \geqslant 1$). Suponha que $\mathbb{R}^{m+n} = E \oplus F$ é uma decomposição em soma direta tal que $\partial_2 f(z_0) : X \longrightarrow \mathbb{R}^n$ é um isomorfismo, onde $z_0 = (x_0, y_0) \in X$. Seja $c = f(z_0) \in \mathbb{R}^n$. Então, existem abertos $V \subset E$*

contendo x_0 e $Z \subset X$ contendo y_0 com a seguinte propriedade: Para cada $x \in V$ há um único $\xi(x) \in F$ tal que $(x, \xi(x)) \in Z$ e $(F(x), \xi(x)) = c$. A aplicação $\xi : V \longrightarrow F$ assim definida é de classe C^k e sua derivada é dada por

$$D\xi(x) = -[\partial_2 f(x, \xi(x))]^{-1} \circ \partial_1 f(x, \xi(x)), \ \textit{para todo } x \in V$$

Demonstração. Defina $\phi : U \longrightarrow E\times \mathbb{R}^n$ por $\phi(x, y) = (x, f(x, y))$. Note que $\phi \in C^k$, pois f $\in C^k$. Logo, ϕ é diferenciável e

$$D\emptyset(z_0)(h, k) = (h, \partial_1 f(x_0, y_0)h + \partial_2 f(x_0, y_0)k), \ \textit{para todo } (h, k) \in \mathbb{R}^{m+n}$$

Como $\partial_2 f(z_0) : F \longrightarrow \mathbb{R}^n$ é um isomorfismo, então a $\dim F = n$. Mas $\mathbb{R}^{m+n} = E \oplus F$, logo, $\dim E = m$. Daí, $D\phi(z_0) : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}^{m+n}$ é um isomorfismo se,

$$\ker D\phi(z_0) = \{0\}.$$

Se $D\phi(z_0)(h, k) = (0, 0)$, então,

$$(h, \partial_1 f(z_0)h + \partial_2 f(z_0)k) = (0, 0).$$

Logo, $h = 0$ e $\partial_2 f(z_0)k = 0$. Como $\partial_2 f(z_0)$ é um isomorfismo, então $k = 0$. Portanto, $(h, k) = (0, 0)$, isto é, $\partial_2 f(z_0)$ é um isomorfismo. Pelo Teorema da Função Inversa, ϕ é um difeomorfismo de classe C^k de uma vizinhança z_0 em uma vizinhança de (x_0, c). Esta última pode ser escolhida da forma $V \times W$, onde V é um aberto contendo E, contendo x_0 e W é um aberto de $\mathbb{R}^n$ contendo $f(z_0) = c$. Sejam $Z = \phi^{-1}(V \times W)$ e $h = \phi^{-1}(V \times W \longrightarrow Z$. Observe que $z_0 \in Z$ é um aberto contendo X. Afirmamos que,

$$f \circ h(x, w) = w, \ \textit{para todo } (x, w) \in V \times W.$$

Ou seja, f é localmente uma projeção. De fato,

$$\phi(x, y) = (x, f(x, y))$$

Portanto, $h(x, w) = (x, h_2(x, w))$. Logo,

$$\begin{aligned}(x,w) &= \phi\circ h(x,w)\\ &= \phi(x,h_2(x,w))\\ &= (x,f(x,h_2(x,w)))\\ &= (x,f\circ h(x,w)),\ \textit{para todo}\ \in V\times W.\end{aligned}$$

Daí segue que,

$$f\circ h(x,w)),\ \textit{para todo}\ \in V\times W.$$

Além disso, $\xi = h_2(.,c) : V \longrightarrow F$. Como $h=\phi^{-1}$ é de classe C^k, então $\xi\in C^k$. Além disso,

$$(x,\xi(x)) = f(x,h_2(x,c)) = f\circ h(x,c) = c,\ \textit{para todo}\ x\in V$$

Agora, vamos provar a unicidade de $\xi(x)$. Seja $(x,y)\in Z$ tal que $f(x,y)=c$, então

$$\begin{aligned}(x,y) &= h\circ\xi(x,y)\\ &= (x,h_2(x,f(x,y)))\\ &= (x,h_2(x,c))\\ &= (x,\xi(x)),\ \textit{para todo}\ xi\in Y\end{aligned}$$

Assim, $y=\xi(x)$. Derivando $f(x,\xi(x))=c$, obtemos

$$\partial_1 f(x,\xi(x)) + \partial_2 f(x,\xi(x))\circ D\xi(x) = 0 \Rightarrow D\xi(x) = -[\partial_2 f(x,\xi(x))]^{-1}\circ\partial_1 f(x,\xi(x)).$$

□

1.6.8 Equivalência entre o teorema da função inversa e a função implícita

Suponhamos que o Teorema de Função Implícita seja válido, vamos provar, a partir deste, que o Teorema da Função Inversa é verdadeiro. Assim, como já provamos a recíproca, estes dois Teoremas são equivalentes.

Demonstração. De fato, suponhamos que $f\in X\subset\mathbb{R}^m\longrightarrow\mathbb{R}^m$ seja uma aplicação de classe $C^k(k\geqslant 1)$ onde $X\subseteq\mathbb{R}^m$ é um aberto e seja $Df(x_0):\mathbb{R}^m\longrightarrow\mathbb{R}^m$ um isomorfismo, para algum x_0 em X. Defina $F:\mathbb{R}^m\times X\longrightarrow\mathbb{R}^n$, por

$$F(x,y) = f(y) - x.$$

Logo, F é de classe C^k, pois, $f \in C^k$. Note que, se $F(x_0) = y_0$, então

$$F(x_0, y_0) = f(x_0) - y_0 = y_0 - y_0 = 0.$$

Além disso, $\partial_2 F(y_0, x_0) = Df(x_0) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ é um isomorfismo. Assim, pelo Teorema da Função Implícita, existe uma vizinhança V de y_0 tal que para cada $x \in V$, existe um único $\xi : V \longrightarrow \mathbb{R}^m$ $\in C^k$ tal que

$$\xi(y_0) = x_0 \; e \; F(x, \xi(x)) = F(x_0, y_0) = 0.$$

Além disso,

$$D\xi(x) = -[\partial_2 f(x, \xi(x))]^{-1} \circ \partial_1 f(x, \xi(x), \; para \; todo \; x \in V.$$

Portanto,

$$f(\xi(x)) - x = F(x, \xi(x)) = 0 \; para \; todo \; x \in V.$$

Também temos que,

$$D\xi(x) = -[\partial_2 f(x, \xi(x))]^{-1} \circ Id_v = [Df(x, \xi(x)^)]^{-1} \circ Id_v, \; para \; todo \; x \in V$$

. Em particular,

$$D\xi(y_0) = [Df(\xi(y_0))]^{-1} \circ Id_v = Df(x_0)^{-1} \circ Id_v.$$

Por conseguinte, $D\xi(x)$ é um isomorfismo pois $Df(x_0)$ o é . Analogamente, aplicando o processo para ξ , e temos que existe $\varphi \in C^k$ tal que

$$\xi \circ \varphi(x) = x, \; para \; todo \; x \in W,$$

onde W é uma vizinhança de $\xi(y_0) = x_0$. Mas,

$$\varphi = Id_v \circ \varphi = (f \circ \xi) \circ \varphi = f \circ (\xi \circ \varphi) = f \circ Id_v = f \in W.$$

Logo,

$$\xi \circ f = Id_w e f \circ \xi = Id_v.$$

Portanto, f é um difeomorfismo de uma vizinhança W de x_0 a uma vizinhança V de $y_0 = f(x_0)$. □

Exemplo 1.30. *Seja* $f : M_n(\mathbb{R}) \times M_n(\mathbb{R}) \longrightarrow M_n(\mathbb{R})$ *dada por*

$$f(X,Y) = X^2 - Y, \text{ para todo } (X,Y) \in M_n(\mathbb{R}) \times M_n(\mathbb{R}),$$

Assim, com esta aplicação, podemos provar que próximo a identidade qualquer matriz tem uma raiz quadrada, desta vez utilizaremos o Teorema da Função Implícita . Com efeito

$$f \in C^1, f(I,I) = I^2 - I = 0 \text{ e } \partial_1 f(X,Y)H = XH + HX.$$

Logo,

$$\partial_1 f(I,I)H = 2H, \text{ para todo } H \in M_n(\mathbb{R}).$$

Daí, $\partial_1 f(I,I)$ *é um isomorfismo. Assim, pelo Teorema da Função Implícita , existe uma vizinhança* V *de* I *tal que para cada* $Y \in V$*, existe um único* ϕ*:* $V \longrightarrow M_n(\mathbb{R}) \in C^1$ *satisfazendo*

$$\phi(I) = I \text{ e } f(\phi(Y), Y) = f(I,I) = 0.$$

Com isso,

$$\phi(Y)^2 - Y = 0, \text{ para todo } Y \in V,$$

e, portanto,

$$\phi(Y)^2 = Y, \forall\ Y \in V.$$

Ou seja, Y *possui uma raiz quadrada em* V*.*

1.7 Derivadas parciais de ordem superior

Vimos na seção 1.5.1, como $\dfrac{\partial f}{\partial x}$ e $\dfrac{\partial f}{\partial y}$ são funções, sendo $z = f(x,y)$ uma função de duas váriaveis que é o estudo das derivadas parciais de 1ª ordem e que tal conceito pode ser estendido.

Da mesma forma, podemos agora definir as derivadas parciais das funções $\dfrac{\partial f}{\partial x}$ e $\dfrac{\partial f}{\partial y}$, obtendo quatro novas funções que são chamadas **derivadas parciais de segunda odem de f**, a saber:

a)

$$f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x,y) = \lim_{\triangle_x \to 0} \frac{\dfrac{\partial f}{\partial x}(x + \triangle_x, y) - \dfrac{\partial f}{\partial x}(x,y)}{\triangle_x},$$

b)

$$f_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x,y) = \lim_{\triangle_y \to 0} \frac{\frac{\partial f}{\partial x}(x, y+\triangle_y) - \frac{\partial f}{\partial x}(x,y)}{\triangle_y},$$

c)

$$f_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x,y) = \lim_{\triangle_x \to 0} \frac{\frac{\partial f}{\partial x}(x+\triangle_x, y) - \frac{\partial f}{\partial x}(x,y)}{\triangle_x},$$

d)

$$f_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)(x,y) = \lim_{\triangle_y \to 0} \frac{\frac{\partial f}{\partial y}(x, y+\triangle_y) - \frac{\partial f}{\partial y}(x,y)}{\triangle_y}$$

se tais limites existirem.

As definições das derivadas parciais de segunda ordem de funções de três variáveis são análogas.

Exemplo 1.31. Calcule todas as deriavadas parciais de segunda ordem da função $f(x,y) = xy - e^x \cos y$.

Solução:

As derivadas parciais de primeira ordem são:

$\frac{\partial f}{\partial x}(x,y) = y - e^x \cos y$ *e* $\frac{\partial f}{\partial y}(x,y) = x + e^x sen\ y.$

Derivando estas funções em relação a x e y, obtemos:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y) &= -e^x \cos y, \ e \ \frac{\partial^2 f}{\partial y \partial x}(x,y) = 1 + e^x sen\ y, \\ \frac{\partial^2 f}{\partial y^2}(x,y) &= e^x \cos y, \ e \ \frac{\partial^2 f}{\partial x \partial y}(x,y) = 1 + e^x sen\ y. \end{aligned}$$

Note que, no exemplo A.19, as derivadas parcias mistas $\frac{\partial^2 f}{\partial y \partial x}$ e $\frac{\partial^2 f}{\partial x \partial y}$ são iguais. O teorema a seguir nos dá uma condição que garante a igualdade das derivadas parcias mistas.

Observação 1.9. *Se as derivadas parciais de segunda ordem de f são contínuas, dizemos que f é de classe C^2.*

Vamos provar o teorema a seguir para funções de duas váriaveis, mas a demonstração pode ser adaptada para funções de várias variáveis.

Teorema 1.7.1. ***(Teorema de Schwarz)***

Se $f : X \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ é duas vezes diferenciável (ou classe C^2) no ponto $(x_0, y_0) \in X$, então:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0). \tag{1.10}$$

Demonstração. Tomemos um ponto (x_0, y_0) no domínio de f. Considere $\triangle_x$ e $\triangle_y$ o incremento da função

$$g(x) = f(x, y_0 + \triangle_y) - f(x, y_0).$$

Veja que

$$g(x_0 + \triangle_x) - g(x_0) = f(x_0 + \triangle_x, y_0 + \triangle_y) - f(x_0, y_0 + \triangle_y) - f(x_0, y_0 + \triangle_y) + f(x_0, y_0). \tag{1.11}$$

Aplicando o teorema do valor médio para funções de uma variável **??** à função g, obtemos que

$$\begin{aligned} g(x_0 + \triangle_x) - g(x_0) &= f(x_0 + \triangle_x, y_0 + \triangle_y) - f(x_0 + \triangle_x, y_0) - f(x_0, y_0 + \triangle_y) + f(x_0, y_0) \\ &= [f(x_0 + \triangle_x, y_0 + \triangle_y) - f(x_0, y_0 + \triangle_y)] - [f(x_0 + \triangle_x, y_0) - f(x_0, y_0)] \\ &= \frac{\partial f}{\partial x}(\overline{x}, y_0 + \triangle_y)\triangle_x - \frac{\partial f}{\partial x}(\overline{x}, y_0)\triangle_x. \end{aligned}$$

para algum $\overline{x}$ entre x_0 e $x_0 + \triangle_x$.
Portanto se definirmos $h : \{\overline{x}\} \times \mathbb{R} \to \mathbb{R}$ por $(\overline{x}, y) \mapsto h(\overline{x}, y) = h(\overline{x}, y_0 + \triangle_y) - h(\overline{x}, y)$, segue de 1.11 que

$$\begin{aligned} g(x_0 + \triangle_x) - f(x_0) &= h_y(\overline{x}, \overline{y})\triangle_y \\ &= \left[\frac{\partial f}{\partial x}(\overline{x}, y_0 + \triangle_y) - \frac{\partial f}{\partial x}(\overline{x}, y_0)\right]\triangle_x \\ &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(\overline{x}, \overline{y})\right)\triangle_x\triangle_y. \end{aligned}$$

para algum $\overline{y}$ entre y_0 e $y_0 + \triangle_y$, ou seja, a expressão (1.11) é igual a $\dfrac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y})\triangle_x\triangle_y$.

Como $\dfrac{\partial^2 f}{\partial y \partial x}$ é contínua em (x_0, y_0), segue que

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \lim_{(\triangle_x, \triangle_y) \to (0,0)} \frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y}).$$

Logo,

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \lim_{(\triangle_x, \triangle_y) \to (0,0)} \frac{1}{\triangle_x \triangle_y}[f(x_0 + \triangle_x, y_0 + \triangle_y) - f(x_0 + \triangle_x, y_0) - f(x_0, y_0 + \triangle_y) + f(x_0, y_0)] \quad (1.12)$$

De modo anàlogo, mostramos que o limite que aparece à direita da igualdade 1.12 também é igual a $\dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$, o que prova o teorema.

$\square$

Capítulo 2

Otimização em funções de várias variáveis

Neste capítulo, estudaremos problemas de otimização em funções de várias variáveis. Tal área da matemática é extremamente crucial e tem como objetivo, modelar problemas do cotidiano. Diversos exemplos são dados para ilustrar a aplicação de conceitos e preposições para a resolução de problemas práticos.

2.1 Máximos e mínimos não-condicionados

Nesta seção, estudaremos como se comportam funções não-lineares. Como o assunto é bem amplo, observaremos, por enquanto, o estudo de máximos e mínimos sem restrições.

Trataremos de maximização e minimização em funções de várias variáveis. Com isso, vamos definir formalmente, os valores de máximos e mínimos dessas funções.

2.1.1 Extremos de funções em várias variáveis

Seja $X \subset \mathbb{R}^m$ um aberto. Se $f : X \subset \mathbb{R}^m \longrightarrow \mathbb{R}$, então sabemos que a derivada de f em x_0 é uma transformação linear,

$$Df(x_0) : \mathbb{R}^m \longrightarrow \mathbb{R}.$$

Existindo $Df(x_0)(u)$, então a mesma é dada pela expressão:

$$Df(x_0)(u) = \sum_{i=1}^{m} D_i f(x_0) u_i.$$

Para ver detalhes da demonstração vide **ALDO**.

Definição 2.1. *Seja* $f : X \subset \mathbb{R}^m \longrightarrow \mathbb{R}$*. Um ponto* x_0*;* $Df(x_0) = 0$ *é chamado ponto crítico de* f*.*

Exemplo 2.1. *Determine os pontos críticos da função* $f(x,y) = 3 + 2e^{-\frac{(5x^2+2y^2-10x-8y+13)}{10}}$*.*

Solução:

Temos que resolver o seguinte sistema de equação:

$$\frac{\partial f}{\partial x} = 2e^{-\frac{(5x^2+2y^2-10x+13)}{10}} \cdot (-10x + 10) = 0 \ \text{e},$$

$$\frac{\partial f}{\partial y} = 2e^{-\frac{(5x^2+2y^2-10x+13)}{10}} \cdot (-4y + 8) = 0.$$

Logo, $x = 1$ *e* $y = 2$ *e, portanto, o ponto* $(1,2)$ *é o único ponto crítico da função* f*.*

Nesse ponto, a função assume o valor $f(1,2) = 2e^{\frac{-(5\cdot 1^2+2\cdot 2^2-10\cdot 1+13)}{10}} = 3 + 2e^0 = 5$*. A figura 2.1 apresenta uma visualização do gráfico da função* f*.*

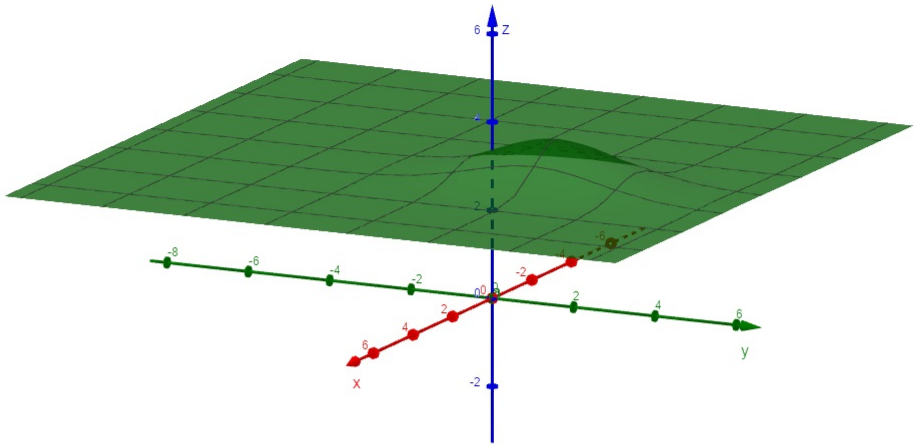

Figura 2.1: Gráfico da função $z = 3 + 2e^{-\frac{(5x^2+2y^2-10x-8y+13)}{10}}$

Definição 2.2. *Seja* f *uma função definida num subconjunto* $X \subset \mathbb{R}^m$ *com valores reais:*

1) *Um ponto* x_0 *de* X *é um ponto de máximo absoluto ou global de* f *se, para todo* x *de* X*, temos,*

$$f(x_0) \geq f(x).$$

Neste caso, o valor $f(x_0)$ chama-se valor máximo absoluto ou global de f.

Exemplo 2.2. *Considere a função $f(x,y) = 4 - (x^2 + 3y^2)$, o ponto $(0,0)$, onde a função assume o valor 4, é um ponto de máximo absoluto de f, de acordo com a sua representação gráfica apresentada na figura 2.2.*

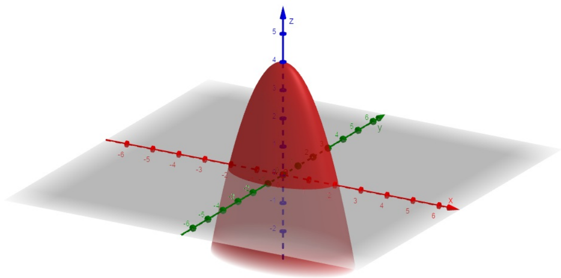

Figura 2.2: Gráfico da função $z = 4 - (x^2 + 3y^2)$

2) *Um ponto $x_)$ de X é um ponto de mínimo absoluto ou global de f se, para todo (x) de X, temos,*

$$f(x_0) \leq f(x).$$

Neste caso, o valor $f(x_0)$ chama-se valor mínimo absoluto ou global de f.

Exemplo 2.3. *Se $f(x,y) = x^2 + y^2$, o ponto $(0,0)$, onde a função assume o valor 0, é um ponto de mínimo absoluto de f, de acordo com a sua representação gráfica apresentada na figura 2.3.*

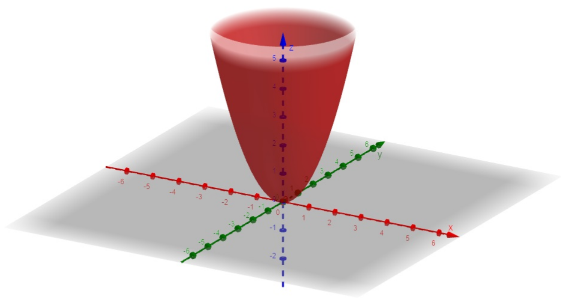

Figura 2.3: Gráfico da função $z = x^2 + y^2$

3) *Um ponto x_0 de X é um ponto de mínimo relativo de f se existe um $\delta > 0$ tal que:*

$$f(x_0) \leq f(x) \text{ para todo } x \in X, \text{com } \|x - x_0\| < \delta.$$

Um ponto x_0 de X é um ponto de mínimo relativo estrito de f se existe um $\delta > 0$ tal que

$$f(x_0) < f(x) \text{ para todo } x \in X, \text{com } 0 < \|x - x_0\| < \delta.$$

4) *De maneira análoga definimos um ponto de máximo relativo de f e ponto máximo relativo estrito de f.*

Se $x_0 \in X$ é um ponto de mínimo ou máximo relativo de f, dizemos que x_0 é um ponto de extremo relativo de f, ou que f tem um extremo relativo em x_0.

Se $x_0 \in X$ é um ponto de mínimo ou máximo relativo estrito de f, dizemos que x_0 é um ponto de extremo relativo de f, ou que f tem um extremo relativo estrito em x_0.

Observação 2.1. *Os valores de máximo e mínimo relativos de funções de duas váriaveis podem ser considerados, respectivamente, como picos e vales de funções, de acordo com a figura 2.4.*

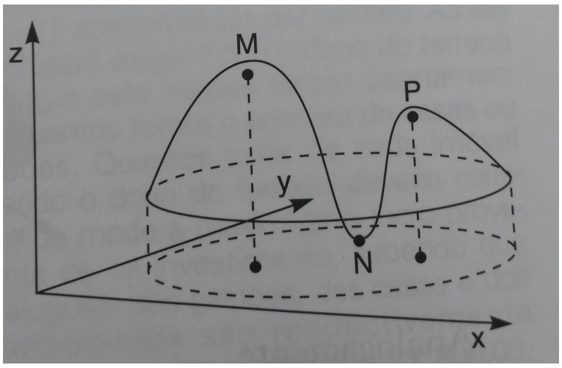

Figura 2.4: Gráfico com pontos que representam máximos e mínimos relativos

Na figura 2.4 o ponto N representa um mínimo local ou um vale, que é global, enquanto P representa um máximo local ou um pico, que não é global. O ponto M é máximo local, que também é global.

Teorema 2.1.1. ***(condição necessária de máximo e de mínimo)***

Seja $f : X \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}$. Se um ponto interior x_0 de X é um ponto de extremo relativo de f, e se a derivada direcional $D_u f(x_0)$ de f em relação ao vetor $u \in \mathbb{R}^m$ existe, então $D_u f(x_0) = 0$.

Demonstração. Defina $F : I \longrightarrow \mathbb{R}$ por

$$F(t) = f(x_0 + tu),$$

onde I é um intervalo $\mathbb{R}$. Sendo $F'(t) = Df(x_0 + tu)(u)$ e por hipótese x_0 é um ponto de extremo relativo de f, segue-se que $t = 0$ é um ponto de extremo relativo F e, portanto, $F'(0) = 0$. Logo, $Df(x_0)(u) = 0$ e, portanto pelo Teorema citado no ALDO, obtemos $D_u f(x_0) = Df(x_0)(u) = 0$.

□

Corolário 2.1. *Seja $f : X \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}$. Se um ponto interior x_0 de X é um ponto de extremo relativo de f e, se a derivada $Df(x_0)$ existe, então $Df(x_0) = 0$.*

Demonstração. Segue que a partir do Teorema citado no ALDO que , se a derivada parcial $D_j f(x_0)$, $j = 1, ..., m$ existe e se $u = (u_1, ..., u_m) \in \mathbb{R}^m$, então:

$$Df(x_0)(u) = \sum_{j=1}^{m} u_j D_j f(x_0).$$

Pelo Teorema anterior temos que $D_j f(x_0) = 0$, para $j = 1, ..., m$ e, portanto, $Df(x_0)(u) = 0$.

□

Comentários:

(i) Uma função f pode ter um extremo relativo em um ponto interior x_0 de X no qual a derivada $Df(x_0)$ não exista.

(ii) Uma função f pode ter um extremo relativo em um ponto $x_0 \in X$, o qual não é um ponto interior de X.

Nos casos (i) e (ii) acima, o ponto x_0 não será ponto crítico de f.

Exemplo 2.4. Determine os pontos críticos da função $f(x, y) = 1 + \dfrac{12 - 12y}{x^2 + y^2 - 4x - 2y + 7}$.

Solução:

Temos que resolver o seguinte sistema de equação:

$$\frac{\partial f}{\partial x} = \frac{-(12-12y)(2x-4)}{(x^2+y^2-4x-2y+7)^2} = 0 \Rightarrow y = 1 \text{ ou } x = 2 \text{ e,}$$

$$\begin{aligned}
\frac{\partial f}{\partial y} = \frac{-12(x^2+y^2-4x-2y+7)-(12-12y)(2y-2)}{(x^2+y^2-4x-2y+7)^2} &= 0 \quad \Rightarrow \\
\frac{-12x^2-12y^2+48x+24y-84-24y+24+24y^2-24y}{(x^2+y^2-4x-2y+7)^2} &= 0 \quad \Rightarrow \\
\frac{-12x^2+12y^2+48x-24y-60}{(x^2+y^2-4x-2y+7)^2} &= 0 \quad \Rightarrow
\end{aligned}$$

Se $y = 1 \Rightarrow -12x^2+12+48x-24-60 = 0 \Rightarrow -12x^2+48x-72 = 0 \Rightarrow x^2-4x+6 = 0 \Rightarrow$ *não há solução.*

Se $x = 2 \Rightarrow -12\cdot 2^2+12y^2+48\cdot 2-24y-60 = 0 \Rightarrow 12y^2-24y-12 = 0 \Rightarrow y^2-2y-1 = 0 \Rightarrow$ $y = \dfrac{2\pm\sqrt{8}}{2} = 1\pm\sqrt{2}.$

Logo, $(2, 1+\sqrt{2})$ *e* $(2, 1-\sqrt{2})$ *são pontos críticos da função* f*. Nesses pontos, a função assume valores respectivamente iguais a:*

$$f(2, 1+\sqrt{2}) = 1 + \frac{12-12(1+\sqrt{2}}{(2^2+(1+\sqrt{2})^2-4\cdot 2-2(1+\sqrt{2})+7} = 1-3\sqrt{2},$$

$$f(2, -+\sqrt{2}) = 1 + \frac{12-12(1-\sqrt{2}}{(2^2+(1-\sqrt{2})^2-4\cdot 2-2(1-\sqrt{2})+7} = 1+3\sqrt{2}.$$

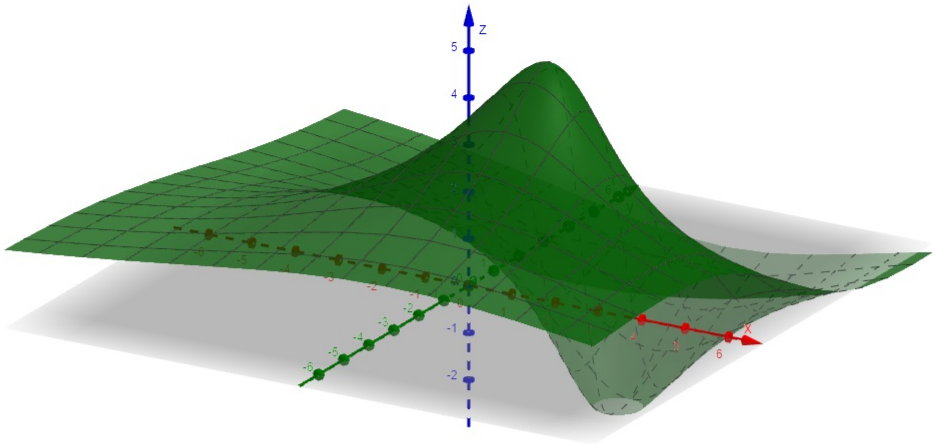

Figura 2.5: Gráfico da função $z = 1 + \dfrac{12 - 12y}{x^2 + y^2 - 4x - 2y + 7}$

Os pontos críticos são $(2, 1 + \sqrt{2})$ *e* $(2, 1 - \sqrt{2})$.

Exemplo 2.5. Determine os pontos críticos da função $f(x, y) = 3 + x^2 - y^2 - 2(x - y)$.

Solução: *Temos que resolver o seguinte sistema de equação:*

$$\frac{\partial f}{\partial x} = 2x - 2 = 0 \Rightarrow x = 1 \ e,$$

$$\frac{\partial f}{\partial y} = -2y + 2 = 0 \Rightarrow y = 1.$$

Logo$(1, 1)$ *é o único ponto crítico de* f. *Nesse ponto a função assume o valor* $f(1, 1) = 3 + 1^2 - 1^2 - 2(1 - 1) = 3$.

O gráfico de f *pode ser representado pela figura 2.5:*

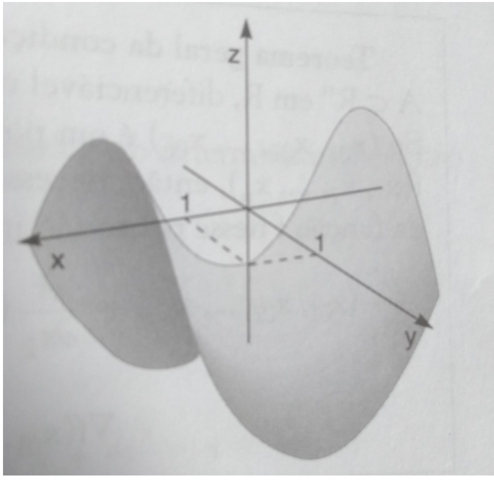

Figura 2.6: Gráfico da função $z = 3 + x^2 - y^2 - 2(x - y)$

Observe que o ponto crítico não representa nem máximo e nem mínmo de f, denominando-se ***ponto de sela*** *(em referência à sela do cavalo).*

Embora as condições colocadas até o momento não caracterizem ainda os máximos e mínmos, elas já restringem significativamente os candidatos. Precisaremos de algo que garanta a natureza dos pontos críticos encontrados.

2.1.2 Condição suficiente do máximo e de mínimo de funções em várias variáveis

Depois de ter calculado os possíveis valores para máximos e mínimos interiores através das condições necessárias apresentadas na seção anterior, devemos agora especificar a natureza dos pontos críticos encontrados. Para encontrar as condições suficientes para que tais pontos sejam máximos ou mínimos locais, vamos desenvolver um teste, que possui como base das derivadas parciais de segunda ordem.

Teorema 2.1.2. ***(Teorema de Weierstrass):***

Seja $f : K \longrightarrow \mathbb{R}$ contínua no compacto K de $\mathbb{R}^n$. Então, f assume valor máximo absoluto e valor mínimo absoluto em K.

Demonstração. Inicialmente, mostraremos que f é limitada.

• **f é limitada.** Suponhamos, por absurdo, f ilimitada. o compacto K está contido em um cubo compacto C_0, de arestas paralelas aos eixos coordenados e de comprimento L. O cubo C_0 é a união de 2^n cubos compactos de arestas de comprimento $\dfrac{L}{2}$. Então f é ilimitada na intersecção de K com alguns destes 2^n cubos. Seja C_1 um tal cubo. Iterando o procedimento obtemos uma sequência $C_0, C_1, ...$ de cubos compactos com arestas de comprimento $\dfrac{L}{2^n}$, com C_{j+1} contido em C_j para todo $j \geq 0$, com f ilimitada em cada intersecção $K \cap C_n$. Pelo Príncipio dos Intervalos Encaixantes temos

$$\bigcap_{n=0}^{\infty} C_n = \{p\}, \; com \; p \; em \; \mathbb{R}^n.$$

Temos

$$|x_n - p| \leq \frac{L\sqrt{2}}{2^n} \; e \; lim \; x_n = p.$$

Ainda, como K é fechado, $p \in K$. Pela continuidade de f segue

$$|f(p)| = \lim_{n \longrightarrow \infty} |f(x_n)| = \infty.$$

• **Os pontos de máximo / mínimo.** Como $f(K)$ é limitado em $\mathbb{R}$, pela propriedade do supremo existe $M = sup \; f(K)$, o supremo de $f(K)$. Suponhamos por absurdo, $f(X) < M$ para todo x em K. Então,

$$\frac{1}{M - f(x)}, \; com \; x \; variando \; em \; K,$$

é contínua e, pela definição de supremo, ilimitado.

Logo, existe a em K com $f(a) = M$. O valor mínimo de f é o oposto do valor máximo de $-f$. $\square$

Consideremos $X \subset \mathbb{R}^n$ um aberto. Dada a função $f : X \longrightarrow \mathbb{R}$ de classe C^2, a *forma quadrática hessiana* $H(x) = (Hf)(x)$ de f no ponto $x \in X$ é aquela, cuja a matriz é:

$$[h_{ij}] = \Big[\frac{\partial^2 f}{\partial \; x_i \partial x_j}\Big].$$

Assim,

$$H(x)(u) = \sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial \; x_i \partial x_j}(x) u_i u_j;$$

onde $u = (u_1, ..., u_n)$.

A forma *Hessiana* é usada para determinar a natureza dos pontos críticos da função f.

Teorema 2.1.3. ***(Teorema da condição suficiente):***

Seja $X \subset \mathbb{R}^n$ um aberto e $x_0 \in X$ um ponto crítico da função $f : X \longrightarrow \mathbb{R}$ de classe C^2.
(i) Se $H(x_0) > 0$ então x_0 é um ponto de mínimo local de f;
(ii) Se $H(x_0 < 0$ então x_0 é um ponto de máximo local de f;
(iii) Se $H(x_0)$ for indefinida então nada se pode afirmar.

Demonstração. **(i)** Por simplicidade, escrevemos H em vez de $H(a)$. Seja

$$S = \{v \in \mathbb{R}^m; \|v\| = 1\}.$$

Tem-se que S é um conjunto compacto em $\mathbb{R}^m$. como H é uma função contínua positiva, pelo Teorema de Weierstrass 2.1.2, H assume um valor mínimo $2c > 0$ no conjunto S. Isto é,

$$2c \leq H(u), \text{ } para \text{ } todo \text{ } u \in S.$$

Desde que a é um ponto de crítico de f, a Fórmula de Taylor $T(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$ se resume a

$$f(a + v) - f(a) = \frac{1}{2}H(v) + \rho(v)\|v\|^2, \text{ } com \lim_{v \longrightarrow 0} \rho(v) = 0. \tag{2.1}$$

Como $\dfrac{v}{\|v\|} \in S$, obtemos

$$\frac{1}{2}H(v) = \frac{\|v\|^2}{2} H\left(\frac{v}{\|v\|}\right) \geq \frac{\|v\|^2}{2} 2c = \|v\|^2 c. \tag{2.2}$$

Substituindo 2.2 em 2.1, obtemos

$$f(a + v) - f(a) \geq \|v\|^2 (c + p(v)).$$

Pela definição de limite, existe $\delta > 0$ tal que $a + v \in A$ e $0 < \|v\| < \delta$ o que implica $|p(v)| < c$ e, consequentemente, $c + p(v) > 0$. Logo,

$$f(a+v) - f(a) > 0,$$

isto é, $f(a+v) > f(a)$ para todo v tal que $a+v \in A$ e $0 < \|v\| < \delta$.

Portanto, a é um ponto de mínimo local para f.

(ii) Segue o reciocínio análogo ao feito na parte (i).

(iii) Dado $v \in \mathbb{R}^m$, tem-se $a+tv \in A$ para todo t suficientemente pequeno. Como $H(tv) = t^2 H(v)$, temos pela Fórmula de Taylor

$f(a+tv) - f(a) = t^2\|v\|^2|H(v) + \rho(tv)|$, com $\lim\limits_{t \longrightarrow 0} \rho(tv) = 0$.

Seguiremos, como acima, que para todo t suficientemente pequeno, $f(a+tv) - f(a)$ tem o mesmo sinal de $H(v)$. Assim, se H é indefinida, tem-se que $H(v) > 0$ e $H(w) < 0$ em qualquer bola de centro em a. Portanto, existem pontos $a+tv$ e $a+tw$ tais que

$$f(a+tv) > f(a) \text{ e } f(a+tw) < f(a).$$

Assim, f não tem máximo nem mínimo local no ponto a.

$\square$

Para $n = 2$ o Teorema 2.1.3 aparece da seguinte maneira:

Teorema 2.1.4. *Seja $X \subset \mathbb{R}^2$ um aberto e $(x_0, y_0) \in X$ um ponto crítico da função $f : X \longrightarrow \mathbb{R}$ de classe C^2.*
(i) Se det $H(x_0, y_0) > 0$ e $\dfrac{\partial^2 f}{\partial x^2} > 0$, o ponto (x_0, y_0) é um ponto de mínimo local de f;
(ii) Se det $H(x_0, y_0) > 0$ e $\dfrac{\partial^2 f}{\partial x^2} < 0$, o ponto (x_0, y_0) é um ponto de máximo local de f;
(iii) Se det $H(x_0, y_0) < 0$, então (x_0, y_0) não é ponto de máximo, nem mínimo de f. Nesse caso, dizemos que é um ponto de sela;
(iv) Se det $H(x_0, y_0) = 0$, nada se pode afirmar.

Observação 2.2. *det $H(x, y) = AC - B^2$, onde*

$$A = \frac{\partial^2 f}{\partial x^2},\ B = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \text{ e } C = \frac{\partial^2 f}{\partial y^2}.$$

Exemplo 2.6. *Determine os pontos de máximo e mínimo relativos da função $f(x, y) = x^3 - y^2 - 12x + 6y + 7$.*

solução:

Vamos calcular as derivadas parciais de primeira ordem:

$$\frac{\partial f}{\partial x} = 3x^2 - 12,$$
$$\frac{\partial f}{\partial y} = -2y + 6.$$

Com isso, determinaremos os pontos críticos:

$$3x^2 - 12 = 0 \Rightarrow x = \pm 2.$$

e,

$$-2y + 6 = 0 \Rightarrow y = 3$$

logo, os pontos crítico são $(2,3)$ *e* $(-2,3)$.

agora, vamos calcular as derivadas parciais de segunda ordem:

$A = \dfrac{\partial^2 f}{\partial x^2} = 6x.$

$C = \dfrac{\partial^2 f}{\partial y^2} = -2.$

$B = \dfrac{\partial^2 f}{\partial x \partial y} = 0.$

$B = \dfrac{\partial^2 f}{\partial y \partial x} = 0.$

Temos que, $AC - B^2 = -12x.$

Agora, vamos classificar os pontos críticos:

i) Ponto $(2,3)$*:*

$AC - B^2 = -12x = -12 \cdot 2 = -24 < 0$, *logo,* $(2,3)$ *é um ponto de sela. A função assume nesse ponto o valor* $f(2,3) = -1$

ii) Ponto $(-2,3)$*:*

$AC - B^2 = -12x = -12 \cdot (-2) = 24 > 0$ *e* $A = 6x = 6 \cdot (-2) = -12 < 0$*, logo ponto* $(-2,3)$ *é um ponto de máximo relativo. A função assume nesse ponto o valor* $f(-2,3) = 32$*.*

Exemplo 2.7. Determine as dimensões do paralelepípedo retangular que tem três faces sobre os planos coordenados e um vértice no primeiro octante sobre o plano $x+y+z=1$, que tenha volume máximo.

Solução:

Como queremos encontrar um paralelepípedo de volume máximo, podemos considerar o paralelepípedo de arestas $x > 0$*,* $y > 0$ *e* $z > 0$*, pois caso contrário o volume seria zero. Seu volume é* $V = xyz$*. Como o vértice* $P = (x,y,z)$ *está no plano* $x+y+z=1$*, tem-se que* $z = 1-x-y$ *e, portanto, o volume* V *se torna uma função de 2 variáveis,*

$$V = V(x,y) = xy(1-x-y) = xy - x^2y - x^2.$$

Assim, para encontrar seus pontos críticos, resolvemos as equações

$$\frac{\partial V}{\partial x} = y - 2xy - y^2 = 0 \ e \ \frac{\partial V}{\partial y} = x - x^2 - 2xy = 0$$

obtendo

$$x = \frac{1}{3} \ e \ y = \frac{1}{3}.$$

Calculando as derivadas de segunda ordem

$$A = \frac{\partial^2 V}{\partial x^2} = -2y, \ C = \frac{\partial^2 V}{\partial xy} = 1 - 2x - 2y \ e \ B = \frac{\partial^2 V}{\partial y^2} = -2x,$$

temos,

$$AC - B^2 = \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right) > 0 \ e \ A = -\frac{2}{3} < 0.$$

Pelo teste da derivada segunda, o ponto $\left(\frac{1}{3}, \frac{1}{3}\right)$ *é um ponto de máximo local e como temos um único ponto crítico, pela própria natureza do problema, este ponto é, também, absoluto. Como* $z = 1 - x - y$, *para* $x = \frac{1}{3}$ *e* $y = \frac{1}{3}$, *tem-se que* $z = \frac{1}{3}$. *Portanto, o paralelepípedo de arestas* $x = \frac{1}{3}$, $y = \frac{1}{3}$ *e* $z = \frac{1}{3}$ *é o paralelepípedo de volume máximo.*

Vamos procurar encontrar agora os **extremos globais**. Supondo que a função em questão assuma máximo ou mínimo finito e, tendo em vista que todos os pontos dos conjuntos de escolhas em problemas de otimização não condicionada são pontos interiores, já que as variáveis independentes modificam-se livremente, os únicos candidatos a máximo ou mínimo global da função tratada são respectivamente seus máximos e mínimos locais. Para encontrá-los, basta verificar dentre os correspondentes locais, aquele em que a função assume respectivamente maior e menor valor, bem como seu comportamento quando suas variáveis independentes tendem a infinito.

Exemplo 2.8. *Encontre, caso haja, os valores máximos e mínimos globais de* $f(x, y) = x^2 - 2x + y^2 - 4y + 5$, *bem como os pontos onde* f *assume tais valores.*

Solução:

Trataremos de um problema de otimização global não condicionada, pois o domínio de f *é* $\mathbb{R}^2$. *Logo, precisamos encontrar os pontos críticos de* f.

Calculando as derivadas parciais, temos:

$\frac{\partial f}{\partial x}(x, y) = 2x - 2$ *e* $\frac{\partial f}{\partial y}(x, y) = 2y - 4$.

Resolvendo o sistema de equação:

$2x - 2 = 0 \Rightarrow x = 1$ *e* $2y - 4 = 0 \Rightarrow y = 2$.

Portanto, o ponto crítico de f *é* $(1, 2)$. *Agora, vamos calcular as derivadas de segunda ordem:*

$A = \frac{\partial^2 f}{\partial x^2} = 2$, $B = \frac{\partial^2 f}{\partial y^2} = 2$ *e* $C = \frac{\partial^2 f}{\partial xy} = 0$.

Dessa forma:

Ponto $(1, 2)$*:*

$AC - B^2 = 4 > 0$. *Como* $A = 2 > 0$ *o ponto* $(1, 2)$ *é um ponto de mínimo relativo. A função assume nesse ponto o valor* $f(1, 2) = 0$.

Desde que a função cresce indefinidamente quando x ou y tendem a infinito e $(1,2)$ é o único ponto crítico em um domínio não restrito, temos que
o valor mínimo global de $f(x,y)$ ocorre em $(1,2)$ e vale 0. A figura 2.7 apresenta um esboço do gráfico de f, com $P = (1,2,0)$.

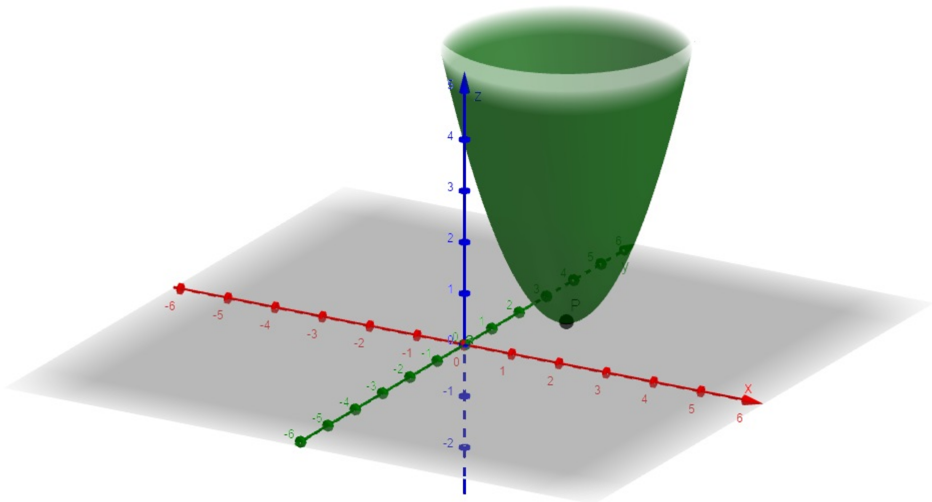

Figura 2.7: Gráfico da função $z = x^2 - 2x + y^2 - 4y + 5$

2.2 Máximos e mínimos condicionados

Neste capítulo, vamos tratar de restrições determinadas por igualdades, aumentando a quantidade de técnicas para que possamos resolver problemas de otimização envolvendo funções quaisquer (lineares ou não) sujeitas a restrições quaisquer.

Na maioria das aplicações, temos de lidar com problemas de otimização condicionada. Com isso, vamos precisar determinar os pontos de máximos ou de mínimos de uma função $f(x,y)$ de modo que os pontos (x,y) satisfaçam, por exemplo, uma condição do tipo $g(x,y) = 0$. A condição $g(x,y) = 0$ é chamada **restrição**.

Em geral, a natureza dos pontos críticos da função não restrita não condiciona a natureza dos pontos críticos da função restrita. Pode existir funções que não assumem máximos, quando não restritas, mas que passam a assumir, quando restritas.

2.2.1 Multiplicadores de Lagrange

Uma das importantes aplicações do *Teorema da Função Implícita* é o método dos multiplicadores de Lagrange para o cálculo de extremos locais de funções sujeitas as retrições.

Seja $B = B[x_0, r] = \overline{B(x_0, r)} \subset \mathbb{R}^n$ e $f : B \longrightarrow \mathbb{R}$ contínua. Desejamos determinar o mínimo global de f sobre a bola B.

Como B é compacto e f é contínua, então existe $\overline{x} \in B$ tal que

$$f(x) \geq f(\overline{x}), \forall x \in B.$$

Se f é diferenciável e $\overline{x} \in int(B)$, então a solução pode ser determinada dentro dos pontos críticos de f. Mas como determinar a solução se $\overline{x} \in B$?

Teorema 2.2.1. ***(Multiplicador de Lagrange)*** *Sejam $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ e $g : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ funções definidas e de classe C^1 e,*

$$S = \{u \in \mathbb{R}^n / g(u) = 0\}.$$

Suponha $u_0 \in S$ tal que:

$$g'(u_0) \neq 0 \text{ e } f(u_0) = min\{f(u)/u \in S\}.$$

Então, existe $\lambda \in \mathbb{R}$ tal que

$$\nabla\ f(u_0) = \lambda \nabla\ g(u_0).$$

Demonstração. Se $g'(u_0) \neq 0$, podemos supor sem perda de generalidade que $\frac{\partial g(u_0)}{\partial x_n} \neq 0$. Seja $\lambda \in \mathbb{R}$ tal que

$$\frac{\partial f(u_0)}{\partial x_n} = \lambda \frac{\partial g(u_0)}{\partial x_n}.$$

Para continuarmos a prova, mostremos que:

$$\frac{\partial f(u_0)}{\partial x_n} = \lambda \frac{\partial g(u_0)}{\partial x_n}, \text{ para } i = 1, 2, ..., n.$$

Denotemos por $(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R}$, $u_0 = (x_0, y_0)$, então g é de classe $g \in C^1$, $g(x_0, y_0) = 0$ e

$$\frac{\partial f(u_0)}{\partial x_n} = \lambda \frac{\partial g(u_0)}{\partial x_n} \neq 0.$$

Segue-se do Teorema da Função Implícita que existe uma vizinhança aberta $A \subset \mathbb{R}^{n-1}$ de x_0 é uma função $\varphi : A \longrightarrow \mathbb{R}$ de classe C^1 tal que

$$\varphi(x_0) = y_0.$$

e

$$g(x, \varphi(x)) = 0, \ para \ todo \ x \in A.$$

Além disso, como

$$f(x_0, \varphi(x_0)) = f(x_0, y_0) \leq f(x, \varphi(x)), \ para \ todo \ x \in A,$$

pois $f(u_0) = min\{f(u)/u \in A\}$, segue que

$x_0 \in A$ é um ponto de mínimo para a função diferenciável dada por: $\psi : A \longrightarrow \mathbb{R}$ tal que $\psi(x) = f(x, \varphi(x))$.

Portanto, $\psi'(x_0) = 0$ e pela regra da cadeia, temos:

$$\psi'(x_0) = \frac{\partial f(x_0)}{\partial x} + \frac{\partial g(x_0)}{\partial y}\varphi'(x_0) = 0.$$

Multiplicando a equação acima por λ e subtraindo na equação anterior, temos:

$$\frac{\partial f(x_0)}{\partial x} = \lambda \frac{\partial g(x_0)}{\partial y}.$$

como queríamos demonstrar.

□

Observação 2.3. *O escalar $\lambda \in \mathbb{R}$ é chamado multiplicador de Lagrange.*

Exemplo 2.9. *Determine as dimensões de um retângulo com área máxima cujo perímetro é igual a 10.*

Solução:

Sejam $x > 0$ e $y > 0$ o conprimento e a largura do retângulo, respectivamente. A área do retângulo é então $f(x, y) = xy$.

Trata-se de um problema de máximo condicionado:

maximizar $f(x, y) = xy$ sujeito à condição $2x + 2y = 10 \Longleftrightarrow g(x, y) = x + y - 5 = 0$.

Note que f e g são de classe C^1, o vetor $\nabla g(x,y) = (1,1)$ e $\nabla f(x,y) = (y,x)$.

Pelo teorema do multiplicador de Lagrange, para (x,y) pertencente a x+y-5=0, devemos resolver o sistema de equações

$$\begin{cases} (y,x) = \lambda(1,1) \\ x+y=5 \end{cases}$$

Assim $x = y = \lambda$, substituindo na segunda equação, temo que $\lambda = \dfrac{5}{2}$.

Logo, $\left(\dfrac{5}{2}, \dfrac{5}{2}\right)$ é a única solução do sistema e o retângulo de área máxima para um dado perímetro é um quadrado.

Exemplo 2.10. Determinar o paralelepípedo retangular de volume máximo inscrito em uma esfera de raio a.

Solução:

Se (x,y,z) é o vértice do paralelepípedo no primeiro octante, as dimensões da caixa são $2x$, $2y$ e $2z$, e seu volume é dado por

$$V(x,y,z) = 8xyz.$$

Assim, o problema consiste em maximizar a função $V(x,y,z)$, sujeita à restrição de que o ponto $P = (x,y,z)$ pertença à esfera de raio a, ou seja, satisfaz a condição $x^2+y^2+z^2 = a^2$, com $x > 0$, $y > 0$ e $z > 0$. Aplicando o métodos dos multiplicadores de Lagrange temos que resolver o sistema

$$\begin{cases} 8yz = \lambda 2x \\ 8xz = \lambda 2y \\ 8xy = \lambda 2z \\ x^2+y^2+z^2 = a \end{cases} \Leftrightarrow \begin{cases} 4xyz = \lambda x^2 \\ 4xyz = \lambda y^2 \\ 4xyz = \lambda z^2 \\ x^2+y^2+z^2 = a \end{cases}$$

Vemos que $\lambda x^2 = \lambda y^2 = \lambda z^2$ e como $\lambda > 0$, pois caso contrário teríamos ou x ou y ou z igual a zero, temos que $x = y = z$. Substituindo na quarta equação, obtemos $x = \dfrac{a}{\sqrt{3}}$. Portanto, o paralelepípedo com volume máximo inscrito em um a esfera de raio a tem arestas de comprimento $\dfrac{2a}{\sqrt{3}}$ e seu volume é $\dfrac{8\sqrt{3}a^3}{9}$.

Multiplicadores de Lagrange para funções de 3 variáveis com 2 restrições

Suponha agora, que queiramos determinar os pontos extremos de uma função de 3 variáveis $w = f(x, y, z)$ sujeita às restrições $g(x, y, z) = 0$ e $h(x, y, z) = 0$. Mostra-se, de modo análogo ao caso de duas variáveis, que se f tem um ponto extremo em $P = (x_0, y_0, z_0)$ e o produto vetorial $\nabla g(P) \times \nabla h(P)$ é não nulo. Assim, existem constantes λ e μ, chamados de Multiplicadores de Lagrange, tais que,

$$\nabla f(P) + \lambda \nabla g(P) + \mu \nabla h(P) = 0$$

Neste caso o método do multiplicadores de Lagrange se resume a resolver um sistema de 5 equações e 5 incógnitas x, y, z, λ e μ.

Exemplo 2.11. Considere a curva γ intersecção do plano $x+2y+z = 1$ com o cilindro $\frac{x^2}{4}+z^2 = 1$. Determine as distâncias máxima e mínima dos pontos de γ ao plano $y = 0$.

Solução:

Determinar os pontos da curva que dão a maior e a menor distância ao plano $y = 0$ é determinar os pontos que tem o maior e o menor valor de $|y|$. Isto é equivalente a determinar os pontos extremos da função $g(x, y, z) = \sqrt{y^2}$, o que é equivalente a determinar os pontos extremos da função $f(x, y, z) = y^2$ com duas restrições: $x + 2y + z = 1$ e $\frac{x^2}{4} + z^2 = 1$. Aplicando o método dos multiplicadores de Lagrange para duas restrições, temos que resolver on sistema de 5 equações e 5 incógnitas x, y, z, λ e μ :

$$\begin{cases} 0 = \lambda + \frac{1}{2}x\mu \\ 2y = 2\lambda \\ 0 = \lambda + \mu 2z \\ x + 2y + z = 1 \\ \frac{x^2}{4} + z^2 = 1 \end{cases} \Leftrightarrow \begin{cases} \lambda + \frac{1}{2}\mu x = 0 \\ \lambda = y \\ x + 2y + z = 1 \\ \frac{x^2}{4} + z^2 = 1 \end{cases} \Leftrightarrow \begin{cases} y + \frac{1}{2}\mu x = 0 \\ x + 2y + z = 1 \\ \frac{x^2}{4} + z^2 = 1 \end{cases}$$

Pela primeira e segunda equações, tem-se $\mu(x - 4z) = 0$. Então, $\mu = 0$ ou $x = 4z$. Se $\mu = 0$ a primeira equação também implica que $y = 0$. Com isto, resolvendo o sistema

$$\begin{cases} x + z = 1 \\ \frac{x^2}{4} + z^2 = 1 \end{cases}$$

obtém-se as soluções:

$$\left\{x = 0, z = 1\} \; e \; \{x = \frac{8}{5}, z = \frac{-3}{5}\right\}$$

ou seja, encontramos os pontos

$$P_1 = (0, 0, 1) \; e \; P_2 = \left(\frac{8}{5}, 0, -\frac{3}{5}\right)$$

Agora, se $x = 4z$. Substituindo-se na última equação obtém-se

$$z = \pm\frac{1}{\sqrt{5}} \; e \; x = \pm\frac{4}{\sqrt{5}}.$$

Substituindo estes valores na equação $x + 2y + z = 1$ obtém-se os pontos

$$P_3 = \left(\frac{4\sqrt{5}}{5}, \frac{1-\sqrt{5}}{2}, \frac{\sqrt{5}}{5}\right) \; e \; P_4 = \left(\frac{-4\sqrt{5}}{5}, \frac{1+\sqrt{5}}{2}, \frac{-\sqrt{5}}{5}\right).$$

Portanto, como os pontos P_1 e P_2 estão no plano $y = 0$, eles são os pontos de mínimo e o ponto P_4 é o ponto de máximo, pois tem o maior valor de y.

A seguir, aplicando o Método de Multiplicadores de Lagrange, também provaremos a importância da desigualdade de *Young*, provaremos as importantes desigualdades de *Hölder* e a desigualdade de*Minkowski*.

2.2.2 Aplicações do método de multiplicadores de lagrange em análise funcional.

Como aplicação do Método dos multiplicadores de Lagrange, provaremos que a média geométrica de uma coleção de números não negativos $\{x_1, x_2, ..., x_n\}$ não excede a média aritmética destes números, isto é, mostraremos

$$\sqrt[n]{x_1 \cdot \; \cdots \; \cdot x_n} \leq \frac{1}{n}(x_1 + x_2 + \cdots + x_n).$$

A seguir, aplicando o Método de Multiplicadores de Lagrange, também provaremos a importantes desigualdades de Young, de Hölder e de Minkowski.

Desigualde de Young

Sejam p e q tais que $1 < p, q < \infty$ e $\frac{1}{p} + \frac{1}{q} = 1$. Então, para todo número real $x, y \geq 0$ tem-se:

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Demonstração. A desigualde fica evidente, quando $x = 0$ ou $y = 0$.
Se a desigualdade é válida para os números x e y, também será válida para os números $xt^{\frac{1}{p}}$ e $yt^{\frac{1}{q}}$, onde t é um número positivo arbitrário.

De fato, seja $t > 0$ e $1 < p, q < \infty$ tal que $\frac{1}{p} + \frac{1}{q} = 1$. Então:

$$\left(xt^{\frac{1}{p}}\right)\left(yt^{\frac{1}{q}}\right) \leq \frac{1}{p}\left(xt^{\frac{1}{p}}\right)^p + \frac{1}{q}\left(yt^{\frac{1}{q}}\right)^q \Leftrightarrow$$

$$\left(xy\right)t^{\frac{1}{p}+\frac{1}{q}} \leq \frac{1}{p}x^p t + \frac{1}{q}y^q t \Leftrightarrow \left(xy\right)t \leq t\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \Leftrightarrow xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q.$$

Portanto, é suficiente considerarmos somente os valores de x e y para os quais $xy = 1$. De fato, se $xy \neq 1$, como $xy > 0$, existe $t > 0$ tal que

$$(xy)t = 1 \Leftrightarrow (xy)^{\frac{1}{p}+\frac{1}{q}} = 1 \Leftrightarrow \left(xt^{\frac{1}{p}}\right)\left(yt^{\frac{1}{q}}\right) = 1.$$

Logo, devemos mostrar que a desigualdade

$$1 \leq \frac{1}{p}x^p + \frac{1}{q}y^q,$$

se verifica para todos os números positivos x e y tais que $xy = 1$. Para tanto, devemos determinar o mínimo da função

$$f(x, y) = \frac{1}{p}x^p + \frac{1}{q}y^q,$$

sujeito a restrição $xy = 1$. Este mínimo existe e ocorre em um ponto (x, y), onde $x \neq 0$ e $y \neq 0$.

Seja

$$F(x, y) = \frac{1}{p}x^p + \frac{1}{q}y^q - \lambda xy$$

Por derivação parciais, obtemos,

$$\begin{cases} x^{p-1} - \lambda y = 0 \\ y^{q-1} - \lambda x = 0. \end{cases} \tag{2.3}$$

Multiplicando estas duas últimas equações por x e y respectivamente, obtemos:

$$x^p = \lambda \text{ e } y^q = \lambda.$$

Considerando estes resultados, como fato que $xy = 1$ resulta que $x = y = 1$. Portanto, o valor mínimo da função $f(x, y)$ é igual a $\frac{1}{p} + \frac{1}{q} = 1$. Assim,

$$1 \leq f(x, y) = \frac{1}{p}x^p + \frac{1}{q}y^q,$$

ou seja,

$$1 \leq \frac{1}{p}x^p + \frac{1}{q}y^q,$$

quando $xy = 1$.

Portanto,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

para todo $x, y \geq 0$ e $p, q > 0$ tais que $\frac{1}{p} + \frac{1}{q} = 1$.

$\square$

Desigualdade de Hölder

Sejam p e q tais que $1 < p, q < \infty$ e $\frac{1}{p} + \frac{1}{q} = 1$. Sejam $(x_i), (y_i),\ i = 1, ..., n$ números reais positivos. Provemos a desigualdade de Holder

$$\sum_{i=1}^{n} x_i y_i \leq \Big(\sum_{i=1}^{n} x_i^p\Big)^{\frac{1}{p}} \cdot \Big(\sum_{i=1}^{n} y_i^q\Big)^{\frac{1}{q}}.$$

Demonstração. Para o caso em que pelo menos um x_i ou y_i for igual a zero a desigualdade é clara. Suponhamos que pelo menos um dos x_i e y_i seja diferente de zero.

Considere

$$x = \frac{x_i}{\Big(\sum_{i=1}^{n} x_i^p\Big)^{\frac{1}{p}}} \quad \text{e} \quad y = \frac{y_i}{\Big(\sum_{i=1}^{n} y_i^q\Big)^{\frac{1}{q}}}.$$

Pela desigualdade de Young, obtemos:

$$\frac{x_i}{\Big(\sum_{i=1}^{n} x_i^p\Big)^{\frac{1}{p}}} \cdot \frac{y_i}{\Big(\sum_{i=1}^{n} y_i^q\Big)^{\frac{1}{q}}} \leq \frac{1}{p}\Big[\frac{x_i}{\Big(\sum_{i=1}^{n} x_i^p\Big)^{\frac{1}{p}}}\Big]^p + \frac{1}{q}\Big[\frac{y_i}{\Big(\sum_{i=1}^{n} y_i^p\Big)^{\frac{1}{q}}}\Big]^q.$$

Somando as desigualdades acima para $i = 1, ..., n$, obtemos:

$$\frac{1}{\left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}} \cdot \frac{1}{\left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}}} \Big(\sum_{i=1}^n x_i y_i\Big) \leq \frac{1}{p} + \frac{1}{q}.$$

Desta desigualdade e do fato de ser $\frac{1}{p} + \frac{1}{q} = 1$, obtemos:

$$\sum_{i=1}^n x_i y_i \leq \Big(\sum_{i=1}^n x_i^p\Big)^{\frac{1}{p}} \cdot \Big(\sum_{i=1}^n y_i^q\Big)^{\frac{1}{q}}.$$

$\square$

Desigualdade de Minkowski

Sejam $p > 1$ e $(x_i), (y_i), i = 1, ..., n$, números reais. Prove a *desigualdade de Minkowski*

$$\Big(\sum_{i=1}^n |x_i + y_i|^p\Big)^{\frac{1}{p}} \leq \Big(\sum_{i=1}^n |x_i|^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^n |y_i|^p\Big)^{\frac{1}{p}}.$$

Demonstração. Sejam $p, q > 1$, tais que $\frac{1}{p} + \frac{1}{q} = 1$. Daí temos $p = 1 + \frac{p}{q}$.

Note inicialmente que para quaisquer $x_i, y_i \in \mathbb{R}$, temos:

$$|x_i + y_i|^p = |x_i + y_i||x_i + y_i|^{\frac{p}{q}} \leq [|x_i| + |y_i|]\,|x_i + y_i|^{\frac{p}{q}} = |x_i||x_i + y_i|^{\frac{p}{q}} + |y_i||x_i + y_i|^{\frac{p}{q}}.$$

Assim, temos

$$\sum_{i=1}^n |x_i + y_i|^p \leq \sum_{i=1}^n |x_i|\;|x_i + y_i|^{\frac{p}{q}} + \sum_{i=1}^n |y_i|\;|x_i + y_i|^{\frac{p}{q}}. \tag{2.4}$$

Agora, usando a desigualdade Hölder, vamos estimar cada parcela do segundo membro de 2.4.

De fato,

$$\sum_{i=1}^n |x_i|\;|x_i + y_i|^{\frac{p}{q}} \leq \Big(\sum_{i=1}^n |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^n \Big(|x_i + y_i|^{\frac{p}{q}}\Big)^q\Big)^{\frac{1}{q}} = \Big(\sum_{i=1}^n |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^n |x_i + y_i|^p\Big)^{\frac{1}{q}},$$

isto é,

$$\sum_{i=1}^{n} |x_i|\ |x_i + y_i|^{\frac{p}{q}} \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{\frac{1}{q}}. \tag{2.5}$$

Analogamente, obtemos:

$$\sum_{i=1}^{n} |y_i|\ |x_i + y_i|^{\frac{p}{q}} \leq \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{\frac{1}{q}}. \tag{2.6}$$

Substituindo 2.5, 2.6 em 2.4 obtemos

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left[\Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{\frac{1}{p}}\right] \cdot \Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{\frac{1}{q}}.$$

ou

$$\Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{1-\frac{1}{q}} \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{\frac{1}{p}}.$$

Agora, usando o fato que $1 - \dfrac{1}{q} = \dfrac{1}{p}$, obtemos desta última desigualdade

$$\Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{\frac{1}{p}} \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{\frac{1}{p}}.$$

□

Abaixo, consta uma aplicação do Método Multiplicadores de Lagrange, utilizada no Dimensionamento Econômico de Reservatórios e Vasos de Pressão Segundo uma determinada norma de medida. Tal problema, foi retirado de um artigo.

2.2.3 Aplicação principal: emprego da ferreamenta "Multiplicadores de Lagrange"no dimensionamento econômico de reservatórios e vasos de pressão segundo normas API (American Petroleum Institute)

Um dos grandes desafios da Engenharia não está somente em dimensionar máquinas e equipamentos seguros e eficientes, ma também que sejam idealizados dentro de parâmetros econômicos. Por outro lado, a busca incessante de projetos de baixo custo não deve afetar o desempenho e demais itens de segurança. Desta forma, esta análise exige uma visão global dos estudos técnico e econômico procurando a otimização do conjunto.

Classifica-se neste estudo as variáveis envolvidas para determinação das dimensões ideais de reservatórios atmosféricos ou vasos de pressão construídos conforme normas API, gerando expressões matemáticas que resultam nas dimensões ideais para obtenção da mínima área externa em função de uma dado volume desejado.

Faz se finalmente a análise desse estudo atual, mostrando que mesmo sendo esta ferramenta amplamente conhecida no meio acadêmico, pode gerar uma contribuição técnica referente a fabricação.

Aplicação da ferramenta "Multiplicadores de Lagrange"teoria dos máximos e mínimos condicionados.

Utilizando-se da teoria dos máximos e mínimos condicionados, com emprego do Método dos Multiplicadores de Lagrange (λ), para a resolução de problemas frequentes nos quais se deseja o máximo ou o mínimo de uma função que serão apresentados dois exemplos.

Descrição resumida do método:

Toma-se a função dada $z = f(x, y, z)$ e a condição desejada qualquer $g(x, y, z) = 0$ e procura-se formar uma nova função $F(x, y, z, \lambda)$ onde (λ) é uma nova variável chamada parâmetro, denominada Multiplicador de Lagrange de modo que

$$F(x, y, z, \lambda) = f(x, y, z) + \lambda \cdot g(x, y, z).$$

O número de Multiplicadores de Lagrange (λ) serão em mesmo número quanto o são os das condição. Para descrever melhor o método faz-se a realização de um problema passo a passo.

Problema 2.1. *Reservatório cilindrico vertical com tampo inferior (fundo) e tampo superior (teto)* *do tipo cônico com ângulos distintos, construídos conforme norma API. Observe as figuras 2.1 e 2.1.*

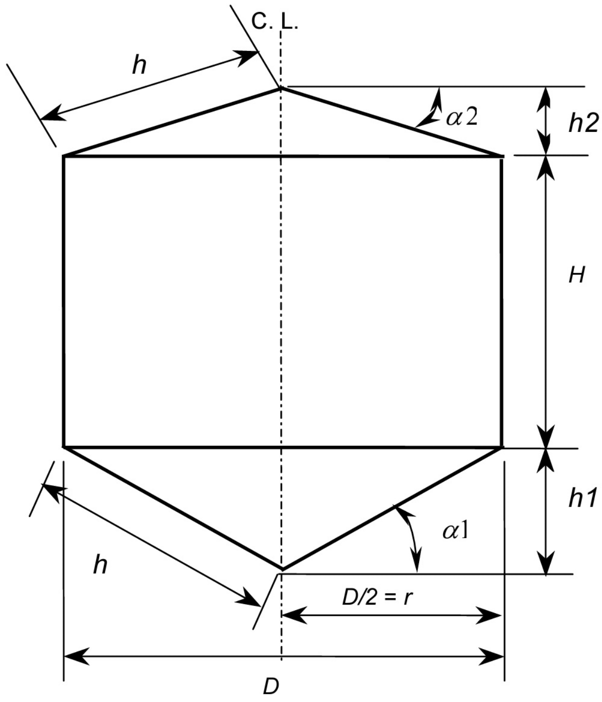

Figura 2.8: Reservatório atmosférico construído conforme norma API

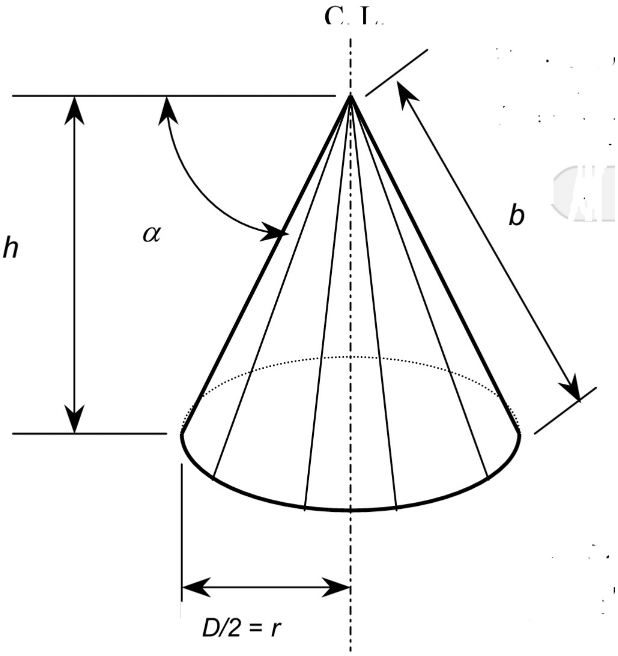

Figura 2.9: Nomeclatura identificadora das dimensões dos tampos do reservatório segundo API

Neste problema vamos focar no dimensionamento econômico do reservartório construídos segundo normas API, com interesse na mínima área em função de uma condição desejada, neste caso o volume útil.

Solução:

Observe que a figura 2 representa os tampos da figura 1, e tem o formato de um cone, assim, vamos determinar o volume do cone (V_c) *e a área do cone* (S_c).

Volume do cone:

$$V_c = \frac{\pi r^2 h}{3} \tag{2.7}$$

$$= \frac{\pi D^2}{4} \cdot \frac{D}{6} \cdot \tan(\alpha). \tag{2.8}$$

Área do cone:

$$
\begin{aligned}
S_c &= \pi r(r+g) && (2.9)\\
&= \pi r^2 + \pi r g && (2.10)\\
&= \pi \cdot \frac{D^2}{4} + \pi \cdot \frac{D}{2} \cdot g. && (2.11)
\end{aligned}
$$

Como não vamos utilizar a área da base, então, só precisaremos da área lateral do cone (Sl_c), ou seja:

$$
Sl_c = \pi \cdot \frac{D}{2} \cdot g. \tag{2.12}
$$

Note que:

$$
\begin{aligned}
\cos(\alpha) &= \frac{\frac{D}{2}}{g} && (2.13)\\
g &= \frac{D}{2\cos(\alpha)}. && (2.14)
\end{aligned}
$$

Daí, temos que:

$$
Sl_c = \frac{\pi D^2}{4\cos(\alpha)}. \tag{2.15}
$$

1º Passo:

Determinar a função a que se deseja minimização.

Em nosso caso especifico deseja-se a mínima área do tanque para que contenha um determinado volume.

Área Total (S_T)= Área do costado + Área dos tampos, ou seja:

$$
S_T = \pi D H + \left(\frac{\pi D^2}{4}\right) \cdot \left(\frac{1}{\cos(\alpha_1)} + \frac{1}{\cos(\alpha_2)}\right), \tag{2.16}
$$

2º ***Passo:***

Determinar agora a condição para que a função S_T deve ser atendida.

Em nosso caso específico deseja-se que o tanque possua determinado volume útil.

Volume útil (V)= Volume da parte cilíndrica + Volume do fundo, ou seja:

$$V = \frac{\pi D^2 H}{4} + \frac{\pi}{D}^{2} 4 \cdot \frac{D}{6} \cdot \tan(\alpha_1) \quad (2.17)$$

$$V = \frac{\pi D^2}{4} \cdot \left(H + \frac{D}{6} \tan(\alpha_1) \right). \quad (2.18)$$

3º ***Passo:***

Determinar as dimensões do tanque no espaço se referenciando aos eixos ortogonais x, y e z. Determinação das dimensões básicas em coordenadas retangulares.

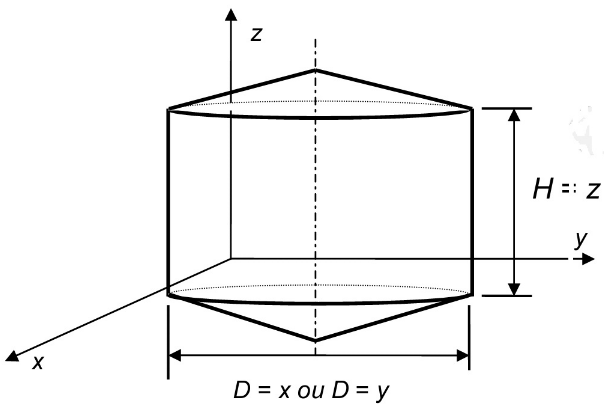

Figura 2.10: Dimensões do tanque no espaço se referenciando aos eixos x, y e z

4º ***Passo:***

Determinar a função e a condição fornecidas no 1º passo e 2º passo obedecendo as dimensões do tanque se referenciando aos eixos x, y e z.

Função (Área Total):

$$
\begin{aligned}
f(x,y,z) &= \pi xz + \left(\frac{\pi x^2}{4}\right)\cdot\left(\frac{1}{\cos(\alpha_1)}+\frac{1}{\cos\alpha_2)}\right), && (2.19)\\
f(x,y,z) &= \pi xz + \frac{\pi x^2}{4\cos(\alpha_1)} + 4\cos(\alpha_2). && (2.20)
\end{aligned}
$$

Condição (Volume):

$$
\begin{aligned}
V &= \frac{\pi x^2}{4}\cdot\left(z+\frac{x}{6}\tan(\alpha_1)\right), && (2.21)\\
V &= \frac{\pi x^2 z}{4} + \frac{\pi x^3\tan(\alpha_1)}{24}. && (2.22)
\end{aligned}
$$

5º Passo:

Trata-se de um problema de mínimo condicionado. Vamos minimizar $f(x,y,z) = \pi xz + \frac{\pi x^2}{4\cos(\alpha_1)} + 4\cos(\alpha_2)$ *sujeito a condição:*

$$
V = \frac{\pi x^2 z}{4} + \frac{\pi x^3\tan(\alpha_1)}{24} \Leftrightarrow g(x,y,z) = \frac{\pi x^2 z}{4} + \frac{\pi x^3\tan(\alpha_1)}{24} - V.
$$

6º Passo:

Note que f *e* g *são de* C^1, *assim, aplicando o método do multiplicador de Lagrange, temos que:*

$$
\nabla f(x,y,z) + \lambda\nabla g(x,y,z) = 0. \tag{2.23}
$$

O resultado físico da expressão 2.23 é que resultará na obtenção do ponto de mínimo que é o objeto de estudo.

Calculando, temos que:

$$
\begin{cases}
\nabla f(x,y,z) = \left(\pi z + \dfrac{\pi x^2}{4\cos(\alpha_1)} + 4\cos(\alpha_2), 0, \pi x\right)\\
\lambda\nabla g(x,y,z) = \lambda\left(\dfrac{2\pi xz}{4} + \dfrac{3\pi x\tan(\alpha_1)}{24}, 0, \dfrac{\pi x^2}{4}\right)
\end{cases}
$$

7º Passo:

Resolver o sistema:

$$\begin{cases} \left(\pi z + \dfrac{\pi x^2}{4\cos(\alpha_1)} + 4\cos(\alpha_2), 0, \pi x\right) = \lambda\left(\dfrac{2\pi x z}{4} + \dfrac{3\pi x\tan(\alpha_1)}{24}, 0, \dfrac{\pi x^2}{4}\right) \\ V = \dfrac{\pi x^2 z}{4} + \dfrac{\pi x^3\tan(\alpha_1)}{24} \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} \pi z + \dfrac{2\pi x}{4\cos(\alpha_1)} + 4\cos(\alpha_2) = \dfrac{2\pi x z\lambda}{4} + \dfrac{3\pi x\tan(\alpha_1)\lambda}{24} \\ \pi x = \dfrac{\pi x^2\lambda}{4} \\ V = \dfrac{\pi x^2 z}{4} + \dfrac{\pi x^3\tan(\alpha_1)}{24} \end{cases}$$

Daí, temos que:

$$\pi x = \frac{\pi x^2\lambda}{4} \Leftrightarrow \tag{2.24}$$

$$1 = \frac{x\pi}{4} \Leftrightarrow \tag{2.25}$$

$$x = \frac{4}{\lambda}. \tag{2.26}$$

Por outro lado, temos que:

$$\pi z + \frac{2\pi x}{4\cos(\alpha_1)} + \frac{2\pi x}{4\cos(\alpha_2)} = \frac{2\pi x z\lambda}{4} + \frac{3\pi x^2\tan(\alpha_1)\lambda}{24} \Leftrightarrow \tag{2.27}$$

$$\pi z + \frac{2\pi x\frac{4}{\lambda}}{4\cos(\alpha_1)} + \frac{2\pi\frac{4}{\lambda}}{4\cos(\alpha_2)} = \frac{2\pi\frac{4}{\lambda}z\lambda}{4} + \frac{3\pi\frac{16}{\lambda^2}\tan(\alpha_1)\lambda}{24} \Leftrightarrow \tag{2.28}$$

$$\pi z + \frac{2\pi}{\lambda\cos(\alpha_1} + \frac{2\pi}{\lambda\cos(\alpha_2} = 2\pi z + \frac{2\pi\tan(\alpha_1)}{\lambda} \Leftrightarrow \tag{2.29}$$

$$\pi z = \frac{2\pi}{\lambda}\cdot\left(\frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2) - \tan(\alpha_1)}\right) \Leftrightarrow \tag{2.30}$$

$$\tag{2.31}$$

$$z = \frac{2}{\lambda} \cdot \left(\frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2)} - \tan(\alpha_1) \right). \quad (2.32)$$

Substituindo 2.26 e 2.32 em V, temos que:

$$\begin{aligned} V &= \frac{\pi \cdot \frac{16}{\lambda^2} \cdot \left[\frac{2}{\lambda} \cdot \left(\frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2)} \right) - \tan(\alpha_1) \right]}{4} + \frac{\pi \cdot \frac{64}{\lambda^3} \cdot \tan(\alpha_1)}{24} && (2.33) \\ &= \frac{8\pi}{\lambda^3} \left(\frac{1}{\cos(\alpha_1)} + \frac{1}{\cos(\alpha_2)} - \tan(\alpha_1) \right) + \frac{8\pi \cdot \tan(\alpha_1)}{3\lambda^3} && (2.34) \\ &= \frac{8\pi}{\lambda^3} \left(\frac{1}{\cos(\alpha_1)} + \frac{1}{\cos(\alpha_2)} - \tan(\alpha_1) + \frac{\tan(\alpha_1)}{3} \right). && (2.35) \end{aligned}$$

Assim, temos que:

$$V = \frac{8\pi}{\lambda^3} \left(\frac{1}{\cos(\alpha_1)} + \frac{1}{\cos(\alpha_2)} - \frac{2}{3} \cdot \tan(\alpha_1) \right). \quad (2.36)$$

Resolvendo a equação 2.36 para uma determinada situação específica, sendo informado somente Volume Útil (V) desejado e os ângulos de inclinação do Fundo (α_1) e do Teto (α_2), e encontraremos o valor específico de (λ), e consequentemente teremos conhecidos os valores de:

Diâmetro,

$$D = x = \frac{4}{\lambda}. \quad (2.37)$$

Altura,

$$H = z = \frac{2}{\lambda} \cdot \left(\frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2)} - \tan(\alpha_1) \right). \quad (2.38)$$

Como segue exemplificado:

Neste estudo também podemos determinar a realação ideal entre Altura (H) e Diâmetro (D), da seguinte forma:

Substituindo a equação 2.37 na equação 2.38, temos que:

$$z = \frac{2}{\frac{4}{x}} \cdot \left(\frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2)} - \tan(\alpha_1) \right) \tag{2.39}$$

$$\frac{z}{x} = \frac{H}{D} = \frac{1}{\cos(\alpha_1} + \frac{1}{\cos(\alpha_2)} - \frac{1}{2} \cdot \tan(\alpha_1). \tag{2.40}$$

Capítulo 3

Problemas de otimização envolvendo a matemática do ensino médio

Este capítulo, tem como objetivo principal apresentar alguns métodos algébricos acessíveis ao estudante do Ensino Médio, para resolução de problemas simples de otimização. Dentre estes, destacam-se a otimização de funções quadráticas, além de aplicações da desigualdade das médias. A aplicação dos métodos apresentados é exemplificada por meio de alguns problemas, escolhidos de maneira a mostrar uma ampla e significativa diversidade que permite a utilização dos métodos aqui desenvolvidos. Consequentemente, estes métodos podem apresentar alguns conteúdos do Ensino Médio de uma forma interessante, despertando o interesse dos alunos, pois, uma vez bem assimilados podem tornar-se poderosas ferramentas na solução de vários problemas, frequentemente encontrados no próprio cotidiano dos alunos e, inclusive, em olimpíadas de matemática, vestibulares e concursos.

Diante desta realidade, cabe ao professor expor seus alunos a situações-problema que estimulem o desenvolvimento da competência matemática. No desenvolvimento desta competência, os problemas são fundamentais, pois permitem ao aluno colocar-se diante de situações que possibilitem o exercício do raciocínio lógico, pensando por si próprio, sem a utilização de regras e fórmulas padronizadas. O Ensino da Matemática deve, dentre os principais objetivos, desafiar os alunos e incitar a curiosidade através da apresentação de problemas compatíveis com os conhecimentos destes. Assim, o professor deve auxiliá-los por meio de indagações estimulantes que objetivem o desenvolvimento de um pensamento crítico e autônomo. Mais do que o "simples"dever de ensinar, o professor encontra-se diante de um novo contexto sócio-cultural em que o aluno possui fácil acesso à informação, tornando-se desafiador mostrar e ressaltar a importância da Matemática. Utilizar problemas desafiadores que estimulem a curiosidade e o raciocício-lógico pode ser uma maneira de aumentar o interesse pela Matemática.

A seguir serão apresentados alguns métodos algébricos para solução de problemas de otimização, dentre eles, destacam-se a desigualdade das médias, que praticamente não são explorados no Ensino Médio.

3.1 Máximos e mínimos de funções quadráticas

Nas atuais orientações curriculares para o Ensino Médio, os problemas de máximos e mínimos usualmente explorados quase sempre estão ligados às funções quadráticas. Nestes, a tarefa mais difícil é achar a função que modela o problema, feito isso, resolver o problema resume-se a encontrar as coordenadas do vértice do gráfico da função. A seguir será apresentada uma breve análise da forma canônica destas funções com o objetivo de encontrar as coordenadas do vértice, consequentemente, o seu valor máximo ou o seu valor mínimo.

Dada a função quadrática $f(x) = ax^2 + bx + c$ com a, b e c números reais e $a \neq 0$, note que valem as seguintes igualdades:

$$
\begin{aligned}
f(x) &= ax^2 + bx + c \\
&= a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) \\
&= a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} - \frac{c}{a}\right) \\
&= a\left(x^2 + \frac{b}{a}x + \frac{b2}{4a^2}\right) - a\left(\frac{b^2}{4a^2} - \frac{c}{a}\right) \\
&= a\left(x + \frac{b}{2a}\right)^2 - \left(\frac{b^2 - 4ac}{4a}\right)
\end{aligned}
$$

Como $x \in \mathbb{R}$, o primeiro termo na expressão anula-se apenas para $x = \dfrac{-b}{2a}$. Logo, conclui-se que:

(i) Se $a > 0$, o menor valor de $f(x)$ ocorre quando $x = \dfrac{-b}{2a}$.

(ii Se $a < 0$, o maior valor de $f(x)$ ocorre quando $x = \dfrac{-b}{2a}$.

Observe um problema onde este resultado pode ser usado:

Problema 3.1. *Os alunos de uma escola alugaram, para uma festa de formatura, um salão de eventos com capacidade para* 150 *pessoas. Cada aluno comprometeu-se, de início, a pagar R\$* $10,00$. *Caso a lotação do estabelecimento não fosse atingida, o gerente propôs que cada aluno que comparecesse pagasse um adicional de R\$* $0,50$ *por lugar vazio. Qual deve ser a quantidade de alunos presentes a festa de formatura para que a receita seja máxima?*

Solução: Seja x o número alunos na festa, tem-se que a receita(R) é dada, em reais, pela função:

$$R(x) = x[10 + 0,5(150 - x)] = -0,5x^2 + 85x.$$

Logo, a solução do problema se resume a determinar o valor de x para que a função atinja seu maior valor, isto ocorre quando $x = \dfrac{-85}{2 \cdot (-0,5)}$, ou seja, quando $x = 85$.

Portanto, o número de alunos que devem estar presentes na festa para que a receita seja máxima é 85*, neste caso a receita será igual a R\$* $3612,50$.

Nos cursos superiores de Matemática ou áreas afins os problemas de otimização costumam ser resolvidos com o uso de derivadas, já no Ensino Médio a maioria destes problemas conduzem a uma função quadrática, cuja solução foi analisada nesta seção. Porém, existe uma ampla quantidade de problemas que podem ser resolvidos usando outros recursos algébricos, tipicamente expressos por meio de desigualdades. Algumas destas serão discutidas a seguir.

3.2 A Desigualdade das Médias

A desigualdade das médias, por exemplo, mostra-se muito útil na solução de alguns problemas de otimização. Geralmente, no Ensino Médio, estas médias são tratadas no conteúdo de Noções de Estatística, onde as aplicações restringem-se ao seu simples cálculo com base em informações dadas em gráficos ou tabelas. A definição destas médias será apresentada a seguir:

Definição 3.1. *Sejam $a_1, a_2, ..., a_n$ números reais positivos. Define-se:*

(i) *A média aritmética (m_a) de $a_1, a_2, ..., a_n$ como o número* $\dfrac{a_1 + a_2 + ... + a_n}{n}$.

(ii) *A média geométrica (m_g) de $a_1, a_2, ..., a_n$ como o número* $\sqrt[n]{a_1 a_2 ... a_n}$.

(iii) *A média harmônica* (m_h) *de* $a_1, a_2, ..., a_n$ *como o número* $\dfrac{n}{\dfrac{1}{a_1}+\dfrac{1}{a_2}+...+\dfrac{1}{a_n}}$.

(iv) *A média quadrática* (m_q) *de* $a_1, a_2, ..., a_n$ *como o número* $\sqrt{\dfrac{a_1^2+a_2^2+...+a_n^2}{n}}$.

Uma vez definidas, existe uma importante relação entre estas médias, apresentada no seguinte teorema:

Teorema 3.2.1. *Desigualdade das Médias Para toda coleção de números reais positivos* $a_1, a_2, ..., a_{n-1}$ *e* a_n *verificam-se as seguintes desigualdades:*

$$m_h(a_1, a_2, ..., a_n) \leq m_g(a_1, a_2, ..., a_n) \leq m_a(a_1, a_2, ..., a_n) \leq m_q(a_1, a_2, ..., a_n).$$

Além disso, em cada caso a igualdade ocorre se, e somente se, $a_1 = a_2 = ... = a_n$.

Várias e interessantes demonstrações destas desigualdades para n números podem ser encontradas em Oliveira[11]. Porém a demonstração aqui apresentada será restrita ao caso que envolve apenas dois números reais positivos, sendo perfeitamente acessível aos alunos do Ensino Médio.

Demonstração. Sejam a_1 e a_2 dois números reais positivos quaisquer. Para mostrar que $m_g \leq m_a$, basta observar que

$$m_a - m_g = \frac{a_1+a_2}{2} - \sqrt{a_1a_2} = \frac{a_1+a_2-2\sqrt{a_1a_2}}{2} = \frac{(\sqrt{a_1}+\sqrt{a_2})^2}{2} \geq 0.$$

Como $m_a - m_g$ é não negativo, segue que $m_g \leq m_a$ para quaisquer $a_1 \neq a_2$ e a igualdade ocorre quando $a_1 = a_2$. Para mostrar que $m_a \leq m_q$, observa-se que

$$\begin{aligned}
(a_1-a_2)^2 \geq 0 &\iff a_1^2 - 2a_1a_2 + a_2^2 \geq 0\\
&\iff 2a_1^2 + 2a_2^2 \geq a_1^2 + 2a_1a_2 + a_2^2\\
&\iff \frac{a_1^2+a_2^2}{2} \geq \left(\frac{a_1+a_2}{2}\right)^2\\
&\iff \sqrt{\frac{a_1^2+a_2^2}{2}} \geq \frac{a_1+a_2}{2}.
\end{aligned}$$

Observe que esta última implicação é válida porque a_1 e a_2 são números positivos, nota-se também que a igualdade ocorre quando $a_1 = a_2$.

E, finalmente, para mostrar que $m_h \leq m_g$, aplica-se a desigualdade $m_g \leq m_a$ aos números positivos $\dfrac{1}{a_1}$ e $\dfrac{1}{a_2}$, de onde tem-se que

$$\sqrt{\frac{1}{a_1} + \frac{1}{a_2}} \leq \frac{\frac{1}{a_1} + \frac{1}{a_2}}{2} \Longleftrightarrow \frac{1}{\sqrt{a_1 a_2}} \leq \frac{a_1 + a_2}{2a_1 a_2} \Longleftrightarrow \frac{2a_1 a_2}{a_1 + a_2} \leq \sqrt{a_1 a_2}.$$

que completa a demonstração, pois

$$\frac{2a_1 a_2}{a_1 + a_2} = \frac{2}{\frac{1}{a_1} + \frac{1}{a_2}}.$$

□

Problema 3.2. *Uma lata cilíndrica é feita para receber 1 litro de óleo. Encontre as dimensões que minimizarão o custo do metal para produzir a lata.*

Solução:

A ideia é tentar exprimir uma função de acordo com a situação proposta pelo problema, e, se possível, colocá-la em função de uma única variável dependente.

Finalmente, escolhe-se um método algébrico que permita minimizar a função encontrada. O volume da lata cilíndrica é fixo e igual a 1 litro, ou seja, $1000cm^3$. *A expressão que determina o volume do cilindro é* $V = \pi r^2 h$, *onde* r *representa o raio da base e* h *a altura do cilindro. Consequentemente,*

$$\pi r^2 h = 1000. \tag{3.1}$$

Por outro lado, a área da superfície cilíndrica é dada por:

$$A = 2\pi r^2 + 2\pi r h, \tag{3.2}$$

assim isolando h *na equação 3.1 e substituindo em 3.2, tem-se*

$$
\begin{aligned}
A(r) &= 2\pi r^2 + 2\pi r \cdot \frac{1000}{\pi r^2} \\
&= 2\pi r^2 + \frac{2000}{r} \\
&= 2(\pi r^2 + \frac{1000}{r}), com\ r > 0.
\end{aligned}
$$

Logo, o objetivo é determinar o valor de r que minimiza $\pi r^2 + \dfrac{1000}{r}$*, que pode ser adaptada para uma aplicação da desigualdade das médias geométrica e aritmética, observando que*

$$\pi r^2 + \frac{1000}{r} = \pi r^2 + \frac{500}{r} + \pi r^2 + \frac{500}{r}.$$

Assim

$$\frac{\pi r^2 + \frac{500}{r} + \frac{500}{r}}{3} \geq \sqrt[3]{\pi r^2 \cdot \frac{500}{r} \cdot \frac{500}{r}} \Longleftrightarrow \pi r^2 + \frac{1000}{r} \geq 150\sqrt[3]{2\pi},$$

e a igualdade, que minimiza a expressão, vale exatamente quando os três termos são iguais, ou seja, $\pi r^2 = \dfrac{500}{r}$*. Daí,*

$$\pi r^3 = 500 \Longleftrightarrow r^3 = \frac{500}{\pi} \Longleftrightarrow \sqrt[3]{\frac{500}{\pi}}.$$

Substituindo o valor de r encontrado em 3.1 e simplificando, encontra-se $h = 2\sqrt[3]{\dfrac{500}{\pi}}$*, ou seja, a altura do cilindro deve ser o dobro da medida do raio da base para que a área da superfície cilíndrica seja mínima e igual a* $300\sqrt[3]{2\pi}$.

Diante do exemplo anterior, é importante observar que a desigualdade das médias aritmética e geométrica é muito útil para estudar funções que envolvem somas de potências positivas e negativas de x com coeficientes positivos. Neste caso, pode ser conveniente decompor um mesmo termo em duas ou mais parcelas de modo que o produto de todas as parcelas resulte em uma constante.

3.3 Aplicações:

Problema 3.3. *(UE-PI) Um agricultor tem* 140 *metros de cerca para construir dois currais: um deles, quadrado, e o outro, retangular, com comprimento igual ao triplo da largura. Se a soma das* *áreas dos currais deve ser a menor possível, qual é a área do curral quadrado?*

Solução:

Seja x o lado do quadrado, logo para construir o curral retangular restaram $(140-4x)$ metros de cerca, ou seja, o comprimento será $\dfrac{3(140-4x)}{8}$ e a largura $\dfrac{3(140-4x)}{8}$. Assim a função $A(x)$ que determina a soma das áreas dos dois currais é dada por:

$$A(x) = x^2 + \frac{3}{8}(140-4x)\cdot\frac{1}{8}(140-4x).$$

Logo,

$$\begin{aligned} A(x) &= x^2 + \frac{3}{64}(19600 - 1120x + 16x^2) \\ &= x^2 + \frac{3}{4}x^2 - \frac{105}{2}x + \frac{3675}{4} \\ &= \frac{7x^2 - 210x + 3675}{4} \\ &= \frac{7}{4}(x^2 - 30x + 525) \\ &= \frac{7}{4}(x-15)^2 + 525. \end{aligned}$$

Assim analisando a forma canônica de $A(x)$ observamos que a função é mínima quando $x = 15$. Portanto, a área do curral quadrado é $225m^2$.

Neste problema algumas indagações são relevantes: Por que minimizar e não maximizar a área como seria o objetivo de uma situação real? Quando é possível utilizar este procedimento em algum outro problema?

Problema 3.4. *Determinar as dimensões do paralelepípedo de menor diagonal possível, sabendo que a soma dos comprimentos de todas as suas arestas é 12.*

Solução: A desigualdade entre as médias quadrática e aritmética pode ajudar a resolver o problema. Pois, o comprimento da diagonal de um paralelepípedo de dimensões a, b e c é dado por: $d = \sqrt{a^2 2 + b^2 + c^2}$, o que sugere o uso da média quadrática. Por outro lado, a soma de todas as arestas pode ser associada à média aritmética. Sejam a, b e c as dimensões do paralelepípedo, logo pretendemos minimizar

$$d = \sqrt{a^2 + b^2 + c^2}$$

onde d representa o comprimento da diagonal do paralelepípedo. Por outro lado, temos

$$4a + 4b + 4c = 12 \Leftrightarrow a + b + c = 3.$$

Assim, usando a desigualdade das médias quadrática e aritmética temos que

$$\sqrt{\frac{a^2 + b^2 + c^2}{3}} \geq \frac{a + b + c}{3} = 1 \Longrightarrow a^2 + b^2 + c^2 \geq 3.$$

Portanto, para minimizar d usa-se a igualdade que ocorre quando $a = b = c = 1$, ou seja, quando o paralelepipedo for um cubo de lado 1, a diagonal será mínima e igual a$\sqrt{3}$.

Algumas questões são pertinentes a respeito deste problema: Quando usar a desigualdade entre as médias quadrática e aritmética? É possível usar este mesmo resultado para outros poliedros regulares?

Problema 3.5. *(AV2 da disciplina MA12 - PROFMAT - Turma 2011): Uma caixa retangular sem tampa tem dimensões x, y e z representando respectivamente o comprimento, a largura e a altura.*

a) *Exprima a área e o volume da caixa em função de x, y e z.*

b) *Use a desigualdade das médias para mostrar que, se o volume da caixa é igual a 32, então sua área é maior ou igual a 48.*

c) *Determine as medidas das arestas da caixa de área mínima com volume igual a 32.*

Solução:

a) *O volume da caixa é $V = xyz$ e a área total é $A = xy + 2xz + 2yz$.*

b) *Pela desigualdade entre as médias aritmética e geométrica temos,*

$$\frac{xy + 2xz + 2yz}{3} \geq \sqrt[3]{xy \cdot 2xz \cdot 2yz} = \sqrt[3]{4x^2y^2z^2},$$

como $xyz = 32$ *(volume da caixa) conclui-se que:*

$$\frac{xy + 2xz + 2yz}{3} \geq \sqrt[3]{4 \cdot 32^2} = \sqrt[3]{4096} = 16.$$

Consequentemente, $xy + 2xz + 2yz \geq 48$ *como pretendia-se mostrar.*

c) *Como* $xy + 2xz + 2yz \geq 48$ *tem-se pela desigualdade das médias aritmética e geométrica que a igualdade só ocorre quandoxy* $= 2xz = 2yz$*, ou seja,* $x = y = 2z$*. Fazendo* $x = a$ *obtem-se que* $a \cdot a \cdot \frac{a}{2} = 32$*, consequentemente* $a^3 = 64 \Longrightarrow a = 4$*. Portanto, as dimensões do comprimento, da largura e da altura que minimizam a área total da caixa são respectivamente iguais a* 4*,* 4 *e* 2*.*

É importante ressaltar que as indústrias (principalmente do setor alimentício) fazem este tipo de cálculo para minimizar o custo das embalagens de determinados produtos. Pode-se utilizar esta ideia para minimizar o material gasto em alguma embalagem, por exemplo, uma caixa de leite, com volume fixo igual a $1\ l$ *e fazer alguma indagações com base na resposta encontrada.*

Referências Bibliográficas

[1] ENCINAS, ALDO S., **Desigualdades e inecuaciones**, 1a edição, Lumbreras: Lima, Perú, (2012).

[2] ENGINNER WLTER, LUIZ APOLONIO, **Emprego da Ferramenta "Multiplicadores de Lagrange"no Dimensionamento Econômico de Reservatórios e Vasos de Pressão segundo Normas API e ASME**, Faculdade de Engenharia de Bauru-UNESP, Maio de 2003.

[3] FIGUEIREDO, D; NEVES, A. **Equações diferenciais aplicadas**. IMPA, Rio de Janeiro 2002.

[4] HANSER, E. de T. **Equações diferenciais autônomas e aplicações.** Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas. São Paulo, 2016.

[5] LIMA, ELON LAGES, **Curso de análise 2**, IMPA, (2010).

[6] LIMA, E. L.; **Curso de Análise**, vol. 1. Coleção Projeto Euclides, IMPA, 1976.

[7] LIMA, E.L. **Espaços Métricos**, Projeto Euclides, IMPA, Rio de Janeiro, 1997.

[8] MANFRINO, RADMILA BULAJICH; ORTEGA, JOSE ANTONIO GÓMEZ; DELGADO, ROGELIO VALDEZ. **Inequalities: A Mathematical Olympiad Approach**, Birkhauser, (2009).

[9] PINTO, D. e MORGADO, M.C.F. : Cálculo Diferencial e Integral de Funções de Várias Variáveis. Editora UFRJ, 1999.

Apêndice A

Introdução a topologia do $\mathbb{R}^n$

Inicialmente faremos breves definições sobre a topologia do $\mathbb{R}^n$, que contribuirão para o nosso objetivo final, isto é, obter os principais resultados da teoria de máximos e mínimos não condicionados e maximos e mínimos de funções condicionados, através de Multiplicadores de Lagrange. E, enfim, aplicar tal teoria em algumas aplicações para problemas cotidianos.

A.1 O Espaço vetorial $\mathbb{R}^n$

Seja $n \in \mathbb{N}$. O espaço *euclidiano n-dimensional* é o produto cartesiano de n fatores iguais a $\mathbb{R}$:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ vezes}}.$$

Os pontos de $\mathbb{R}^n$ são as n-listas $x = (x_1, \cdots, x_n)$ cujas coordenadas $x_1, \cdots, x_n$ são números reais.

Dados $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$ em $\mathbb{R}^n$ e um número real λ, definimos a *adição* $x + y$ e a *multiplicação por escalar* $\lambda \cdot x$ por

$$\begin{aligned} x + y &= (x_1 + y_1, \cdots, x_n + y_n), \\ \lambda \cdot x &= (\lambda x_1, \cdots, \lambda x_n). \end{aligned}$$

Com estas operações, o $\mathbb{R}^n$ é um espaço vetorial de dimensão n sobre $\mathbb{R}$, no qual o elemento neutro para a adição é $0 = (0, \cdots, 0)$ e o simétrico de $x = (x_1, \cdots, x_n)$ é $-x = (-x_1, \cdots, -x_n)$.

Os elementos de $\mathbb{R}^n$ serão às vezes chamados de pontos e às vezes de vetores. Destaca-se a *base canônica* $\{e_1, \cdots, e_n\}$, formada pelos vetores

$$e_1 = (1, 0, \cdots, 0),\ e_2 = (0, 1, \cdots, 0), \cdots,\ e_n = (0, 0, \cdots, 1),$$

que tem uma coordenada igual a 1 e as outras nulas. Para todo $x = (x_1, ..., x_n)$ em $\mathbb{R}^n$, temos:

$$x = x_1 e_1 + ... + x_n e_n.$$

Teorema A.1.1. *Sejam $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ o conjunto das aplicações lineares $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ e o conjunto $M(n \times m)$ das matrizes reais (a_{ij}) com n linhas e m colunas. Existe uma bijeção natural entre os conjuntos $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ e $M(n \times m)$.*

Demonstração. De fato, dada $T \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, seja $A_T = (a_{ij})$ a matriz cuja j-ésima coluna é o vetor coluna $(Te_j)^t$, onde $\{e_1, ..., e_m\}$ é a base canônica de $\mathbb{R}^m$.

$$T(e_j) = \sum_{i=1}^{n} a_{ij} e_i^* \quad (j = 1, ..., m),$$

onde $\{e_1^*, ..., e_n^*\}$ é base canônica de $\mathbb{R}^n$. Assim, defina a função

$$\begin{aligned} \phi : \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) &\longrightarrow M(n \times m) \\ T &\longmapsto A_T, \end{aligned}$$

dada $A \in M(n \times m)$, seja $T_A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ definida por

$$T_A(x) = \left(\sum_{j=1}^{n} a_{ij} x_j, \cdots, \sum_{j=1}^{n} a_{nj} x_j \right).$$

Como $T_A(e_j) = (a_{ij}, \cdots, a_{nj})$, temos que a aplicação ϕ é sobrejetora.
Além disso, ϕ é injetora, pois se $\phi(T) = \phi(L)$, então $T(e_j) = L(e_j)$, $j = 1, \cdots, m$ e, portanto,

$$T(x) = x_1 \cdot T(e_1) + \cdots + x_m \cdot T(e_m) = L(x),$$

para todo $x = (x_1, \cdots, x_m) \in \mathbb{R}^m$.

Agora, escrevendo as entradas de uma matriz $A \in M(n \times m)$ na forma de lista, podemos identificar A como um ponto do espaço euclidiano $\mathbb{R}^{mn}$. Assim, $M(n \times m)$ pode ser visto como um espaço vetorial real de dimensão nm, no qual as matrizes

$$A_{kl} = (a_{ij}^{kl});\ 1 \leqslant k \leqslant n \ e\ 1 \leqslant m,$$

onde

$$(a_{ij}^{kl}) = \begin{cases} 1, (i,j) = (k,l) \\ 0, (i,j) \neq (k,l) \end{cases}$$

formam uma base natural.

Além disso, como ϕ é uma bijeção, podemos induzir em $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ uma estrutura de espaço vetorial, para a qual T^{lk}, $1 \leqslant k \leqslant n$ e $1 \leqslant m$, onde $T^{lk}(e_l) = e_k^*$ e $T^{lk}(e_j) = 0$ se $j \neq l$ é uma base natural.

Podemos, assim, sempre que for conveniente, substituir $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ ora por $M(n \times m)$, ora por $\mathbb{R}^{mn}$. □

• No caso particular em que $n = 1$, $\mathcal{L}(\mathbb{R}^m;\mathbb{R})$ é o espaço vetorial de dimensão n formado pelos *funcionais lineares* R^m em $\mathbb{R}$, para o qual $\{\pi_1, \cdots, \pi_m\}$ é uma base, onde

$$\pi_i(l_j) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

ou seja,

$$\pi_i \times (x_1, ..., x_m) = \sum_{i=1}^{n} x_i \pi_i(e_j) = x_i$$

é a projeção de $\mathbb{R}^{mn}$ sobre o seu i-ésimo fator.

O espaço $\mathcal{L}(\mathbb{R}^m;\mathbb{R}) = (\mathbb{R}^m)^*$ é chamado espaço dual do $\mathbb{R}^m$ e, a base $\{\pi_1, ..., \pi_m\}$ constituem a *base dual* da base canônica de $\mathbb{R}^m$.

Definição A.1. *Sejam E, F, G espaços vetoriais reais. Uma aplicação $\varphi : E \times F \longrightarrow G$ chama-se bilinear quando é linear em relação a cada uma de suas componentes. Ou seja:*

$$\begin{aligned} \varphi(\lambda x + x', y) &= \lambda\varphi(x, y) + \varphi(x', y) \\ e\ \varphi(x, \alpha y + y') &= \alpha\varphi(x, y) + \varphi(x', y). \end{aligned}$$

Definição A.2. *Uma aplicação bilinear $\varphi : E \times E \longrightarrow G$ é simétrica quando*

$$\varphi(x, y) = \varphi(y, x)$$

Produto interno e norma

Definição A.3. *Seja E um espaço vetorial real. O produto interno em E é uma aplicação $\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{R}$ que satisfaz as seguintes propriedades:*

1. $\langle x, y \rangle = \langle y, x \rangle$,

2. $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$,

3. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = \langle x, \alpha y\rangle$,

4. $\langle x, x\rangle > 0$, *sempre que* $x \neq 0$,

para quaisquer $x,\ x',\ y \in E$ *e* $\lambda \in \mathbb{R}$. *Ou seja, o produto interno sobre* E *é uma função bilinear, simétrica e positiva definida.*

Exemplo A.1. *O produto interno canônico do espaço euclidiano* $\mathbb{R}^n$ *é dado por*

$$\langle x, y\rangle = x_1y_1 + \cdots x_ny_n,$$

onde $x = (x_1, ..., x_n)$ *e* $y = (y_1, ..., y_n)$.

Observação A.1. *Dois vetores* $x, y \in \mathbb{R}^n$ *dizem-se ortogonais quando* $\langle x, y\rangle = 0$. *Deste modo, o vetor nulo é ortogonal a todos os vetores do espaço.*

Definição A.4. *Uma norma num espaço vetorial real* E *é uma função real* $\|\cdot\| : E \longrightarrow \mathbb{R}$ *que satisfaz as seguintes condições:*

1. $\|\lambda x\| = |\lambda|\|x\|$;

2. $\|x + y\| \leqslant \|x\| + \|y\|$

3. $\|x\| > 0$, *sempre que* $x \neq 0$,

para quaisquer $x, y \in E$ *e* $\lambda \in \mathbb{R}$

Proposição A.1 (**Desigualdade de Cauchy -Schwarz**). *Seja* E *um espaço vetorial com produto interno* $\langle , \rangle$. *Então*

$$\langle x, y\rangle \leq \|x\|\|y\|, \forall x, y \in E,$$

e a igualdade é valida se, e somente se, x *e* y *são linearmente dependentes (LD), onde* $\|x\| = \sqrt{\langle x, x\rangle}$ *e,* $\|y\| = \sqrt{\langle y, y\rangle}$.

Demonstração. Suponhamos y $\neq 0$ e seja $\lambda \in \mathbb{R}$. Como

$$\langle x + \lambda y, x + \lambda y\rangle = \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2 \geqslant 0,$$

para todo $\lambda \in \mathbb{R}$, temos que o discriminante

$$\triangle = 4\langle x, y\rangle^2 - 4\|x\|^2\|y\|^2 \leqslant 0$$

Logo,

$$\langle x, y\rangle \leq \|x\|\|y\|.$$

Além disso, $\langle x, y\rangle = \|x\|\|y\|$, se e somente se, $\triangle = 0$, ou seja, se e somente se, existe $\lambda_0 \in \mathbb{R}$ tal que x + $\lambda_0 y = 0$.

Logo, $\langle x, y\rangle = \|x\|\|y\|$ se , e somente se, x e y são LD.

$\square$

Teorema A.1.2. $|\|x\| - \|y\|| \leqslant \|x - y\|$

Demonstração. De fato, como

$$\|x\| = \|(x - y) + y\| \leqslant \|x - y\| + \|y\|$$

e,

$$\|y\| = \|(y - x) + x\| \leqslant \|x - y\| + \|x\|$$

temos que

$$-\|(x - y) + y\| \leqslant \|x\| - \|y\| \leqslant \|(x - y) + y\|$$

donde,

$$|\|x\| - \|y\|| \leqslant \|(x - y) + y\|.$$

$\square$

Observação A.2. *Se$\langle\cdot,\cdot\rangle : E \times E \longrightarrow \mathbb{R}$ é um produto interno em E, então$\|\cdot\| : E \longrightarrow \mathbb{R}$, dado por $\|x\| = \sqrt{\langle x, x\rangle}$ é uma norma em E.*

Exemplo A.2. *Se $\langle\cdot,\cdot\rangle$ é produto interno em $\mathbb{R}^n$,*

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$$

é chamada de norma euclidiana do vetor $x \in \mathbb{R}^n$.

Teorema A.1.3. *Há uma infinidade de normas que podem ser definidas no espaço euclidiano $\mathbb{R}^n$. Dentre elas, temos:*

1. *A norma do máximo :$\|x\|_m = max\{|x_1|, \cdots, |x_n|\}$;*

2. A norma da soma : $\|x\|_s = |x_1| + \cdots + |x_n|$.

Além disso, para todo x *em* $\mathbb{R}^n$,

$$\|x\|_m \leqslant \|x\| \leqslant \|x\|_s \leqslant n\|x\|_m$$

onde $\|x\|$ *é a norma euclidiana.*

Demonstração. De fato, como $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} \geqslant |x_i|$, para todo $i = 1, \cdots, n$, temos que $\|x\|_m \leqslant \|x\|$.

Agora,

$$\|x\|_s = |x_1| + \cdots + |x_n| \leqslant n \cdot max|x_i| = n\|x\|_m$$

Finalmente,

$$\|x\|_s^2 = (|x_1| + \cdots + |x_n|)^2 = |x_1|^2 + \cdots + |x_n|^2 + 2\sum_{i,j=1}^{n} |x_i||y_i| \geqslant |x_1|^2 + \cdots + |x_n|^2 = \|x\|^2,$$

explica que, $\|x\| \leqslant \|x\|_s$.

$\square$

Observação A.3. *As desigualdades no teorema A.1.3 servirão para mostrar que as três normas são equivalentes.*

Observação A.4. *Uma norma pode não provir de um produto interno, ou seja, nem sempre existe um produto interno* $\langle \cdot, \cdot \rangle$ *em* E *tal que*

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Com efeito, se a norma $\|x\|$ *provém de um produto interno* $\langle , \rangle$, *então vale a identidade do paralelogramo:*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|(x\|^2 + \|y\|^2)$$

que diz que a soma dos quadrados das diagonais de um paralelogramo é igual à soma dos quadrados de seus quatro lados.

De fato,

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \\ \|x - y\|^2 &= \langle x - y, x - y \rangle = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle \\ \Longrightarrow \|x + y\|^2 + \|x - y\|^2 &= 2(\|x\|^2 + \|y\|^2) \end{aligned}$$

Com isso, podemos mostrar que as normas $\|\cdot\|_m$ e $\|\cdot\|_s$ em $\mathbb{R}^n$ não provém de um produto interno: se e_1 e e_2 são vetores da base canônica do $\mathbb{R}^n$, temos que

$$\|e_1 + e_2\|_m^2 + \|e_1 - e_2\|_m^2 = 1 + 1 = 2 \neq 4 = 2(\|e_1\|_m^2 + \|e_2\|_m^2)$$

e

$$\|e_1 + e_2\|_s^2 + \|e_1 - e_2\|_s^2 = 4 + 4 = 8 \neq 4 = 2(\|e_1\|_s^2 + \|e_2\|_s^2).$$

Bolas e conjuntos limitados

Definição A.5. *Num espaço normado M, definimos os seguintes conjuntos:*

1. *Bola aberta de centro a em M e raio $r > 0$;$B(a, r) = \{x \in M / \|x - a\| < r\}$;*
2. *Bola fechada de centro a em M e raio $r > 0$;$B[a, r] = \{x \in M / \|x - a\| \leqslant r\}$;*
3. *Esfera de centro a em M e raio $r > 0$;$S[a, r] = \{x \in M / \|x - a\| = r\}$.*

Segue-se que $B[a, r] = S[a, r] \cup B(a, r)$.

Exemplo A.3. *No espaço euclidiano $\mathbb{R}$ de dimensão 1, temos que $B[a, r] = [a - r, a + r], S[a, r] = \{a - r, a + r\}, B(a, r) = (a - r, a + r)$.*

Observação A.5. *A forma geométrica das bolas e esferas, em geral, depende da norma que se usa.*

Por exemplo, se considerarmos o plano $\mathbb{R}^2$, com a métrica euclidiana, teremos: $B((a, b), r) = \{(x, y) \in \mathbb{R}^2 / (x - a)^2 + (y - b)^2 < r\}$ (disco aberto de centro (a,b) e raio $r > 0$.)

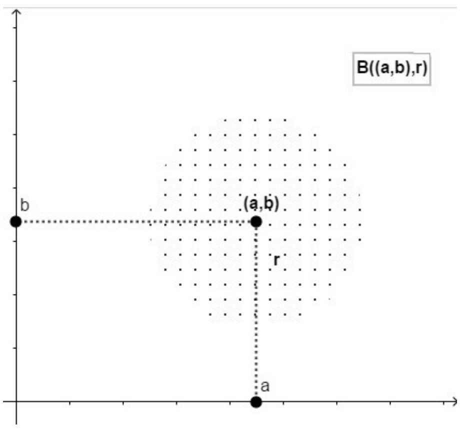

Figura A.1: Bola com a métrica Euclidiana.

E se considerarmos, $\mathbb{R}^2$ *com a métrica do máximo, teremos:* $B_m((a,b),r) = \{(x,y) \in \mathbb{R}^2 / |x - a| < r$ *e* $|y - b| < r\} = (a-r, a+r) \times (b-r, b+r)$

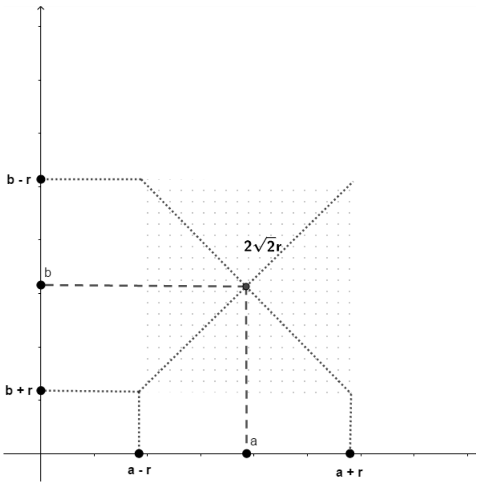

Figura A.2: Bola aberta, em relação à norma do máximo.

Finalmente, se tomarmos $\mathbb{R}^2$ *com a métrica da soma, teremos:* $B((a,b),r) = \{(x,y) \in \mathbb{R}^2 / (x - a)^2 + (y-b)^2 < r\}$ *(disco aberto de centro (a,b) e raio* $r > 0$*.)*

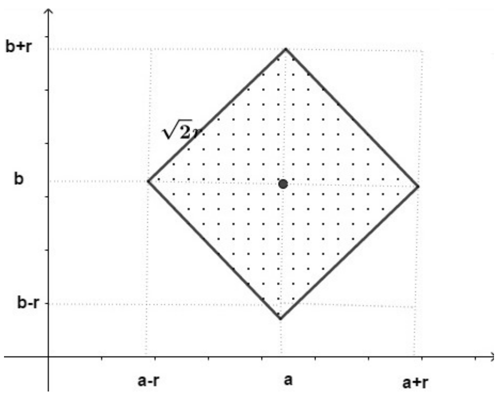

Figura A.3: Bola fechada com relação à métrica da soma.

Teorema A.1.4. *De um modo geral, a bola aberta* $B_m(a,r) \subset \mathbb{R}^n$ *definida pela norma* $\|x\|_m = max\{|x_1|, \cdots, |x_n|\}$ *é o produto cartesiano* $(a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r)$*, onde* $a = (a_1, \cdots, a_n)$*.*

Demonstração. De fato,

$$\begin{aligned} x = (x_1, \cdots, x_n) \in B_m(a, r) &\iff |x_1 - a_1| < r, \cdots, |x_n - a_n| < r \\ &\iff x_1 \in (a_1 - r, a_1 + r), \cdots, x_n \in (a_n - r, a_n + r) \\ &\iff (x_1, \cdots, x_n) \in (a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r). \end{aligned}$$

□

Definição A.6. *Um conjunto $X \subset \mathbb{R}^n$ é convexo quando contém qualquer segmento da reta cujos extremos pertencem a X, ou seja,*

$$x, y \in X \Longrightarrow [x, y] \subset X.$$

Exemplo A.4. *Todo espaço vetorial $E \subset \mathbb{R}^n$ é convexo.*

Teorema A.1.5. *Toda bola aberta ou fechada de $\mathbb{R}^n$, com respeito a qualquer norma, é um conjunto convexo.*

Demonstração. Sejam $x, y \in B(a, r)$. Então, $\|x - a\| < r$ e $\|y - a\| < r$. Logo,

$$\|(1-t)x + ty - a\| = \|(1-t)x + y - (1-t)y - (1-t)a - ta\| \leqslant \|(1-t)(x-a)\| + \|t(y-a)\| = r$$

para todo $t \in [0, 1]$, pois $0 \leq t \leq 1$.

De modo análogo, podemos provar que a bola fechada é convexa.

□

Definição A.7. *Um subconjunto $X \subset \mathbb{R}^n$ é limitado com respeito a uma norma em $\mathbb{R}^n$ quando existe $c > 0$ tal que $\|x\| \leqslant c$ para todo x em X, ou seja, quando existe $c > 0$ tal que $X \subset B[o, c]$.*

A.1.1 Sequências convergentes e de Cauchy

Nesta seção, estudaremos quando uma sequência de pontos em $\mathbb{R}^n$ é convergente ou de Cauchy. Tais sequências serão úteis para obtermos informações que serão estudadas na teoria das aproximações sucessivas.

Definição A.8. *Uma aplicação $x : \mathbb{N} \to \mathbb{R}^n$ é chamada sequência em $\mathbb{R}^n$ e será denotada por*

$$(x_m) \text{ ou } (x_m)_{m \in \mathbb{N}} \text{ ou } (x_1, x_2, \cdots, x_m, \cdots).$$

Observação A.6. *Observe que* $x_m \in \mathbb{R}^n$*, então* $x_m = (x_1^m, x_2^m, \cdots, x_n^m)$*, para todo* $m \in \mathbb{N}$*.*

Exemplo A.5. *Consideremos* $x_m = \left(\frac{1}{m}, sen(m)\right)$*, para todo* $m \in \mathbb{N}$*. Assim,* (x_m) *é uma sequência em* $\mathbb{R}^2$

Definição A.9. *Dizemos que* (x_m)*é limitada se existir* $r > 0$ *tal que* $\|x_m\| < r$*, para todo* $m \in \mathbb{N}$*.*

Mostraremos uma outra maneira de definirmos uma sequência limitada através das coordenadas de seus valores.

Teorema A.1.6. $(x_m)_{m\in\mathbb{N}}$ *é limitada em* $\mathbb{R}^n$ *se, e somente se,* $(x_i^m)_{m\in\mathbb{N}}$ *é limitada em* $\mathbb{R}$ *para todo* $i = 1, 2, \cdots, n$*, onde* $x_m = (x_1^m, x_2^m, \cdots, x_n^m)$*, para todo* $m \in \mathbb{N}$*.*

Demonstração. Se $(x_m)_{m\in\mathbb{N}}$ é limitada $\mathbb{R}^n$, então existe $r > 0$ tal que $\|x_m\| < r$, para todo $m \in \mathbb{N}$. Logo, pela definição de norma euclidiana,

$$(x_1^m)^2 + \cdots + (x_n^m)^2 \leq r^2.$$

Assim,

$$(x_i^m)^2 \leq r^2 \text{ para todo } i = 1, 2, \cdots, n \text{ e para todo } m \in \mathbb{N}.$$

Portanto,

$$|x_i^m|^2 \leq r \text{ para todo } m \in \mathbb{N} \text{ e } i = 1, 2, \cdots, n.$$

Ou seja, $(x_i^m)_{m\in\mathbb{N}}$ é limitada em $\mathbb{R}$, para todo $i = 1, 2, \cdots, n$.

Reciprocamente, seja $(x_i^m)_{m\in\mathbb{N}}$. Daí, existe $r_i > 0$ tal que $|x_i^m| \leq r_i$, para todo $i = 1, 2, \cdots, n$. e para todo $m \in \mathbb{N}$. logo,

$$(x_1^m)^2 + \cdots + (x_n^m)^2 \leq r_1^2 + \cdots + r_n^2 \textit{ para todo } m \in \mathbb{N}.$$

Portanto, $\|x_m\| \leq \sqrt{r_1^2 + \cdots + r_n^2} = r$, para todom $\in \mathbb{N}$.

□

Exemplo A.6. *A sequência* $(x_m) \subseteq \mathbb{R}$*, definida por* $x_m = m \ \forall \ m \in \mathbb{N}$*, não é limitada.*

Exemplo A.7. *Um exemplo de sequência limitada é* (x_m)*, onde* $x_m = \frac{m}{m+1}$ *para todo* $m \in \mathbb{N}$*, pois,* $|x_m| < 1$ *para todo* $m \in \mathbb{N}$*.*

Exemplo A.8. *A sequência* $x_m = \left(\frac{1}{m}, sen(m)\right)$ *para todo* $m \in \mathbb{N}$ *é uma sequência limitada em* $\mathbb{R}^2$*. De fato,*

$\dfrac{1}{m} \leq 1$ *e* $|sen(m)| \leq 1$,*para todo* $m \in \mathbb{N}$. *Logo, pelo Teorema A.1.6,* (x_m) $m \in \mathbb{N}$ *é limitada.*

Exemplo A.9. *Seja* $x_m = (1, m) \in \mathbb{R}^2$, *para todo* $m \in \mathbb{N}$ *uma sequência. Usando o Teorema A.1.6, temos que* (x_m) *não é limitada, já que a segunda coordenada do n-ésimo termo desta sequência não forma uma sequência limitada em* $\mathbb{R}$.

Definição A.10. *Dizemos que uma sequência, em* $\mathbb{R}^n, (x_m)$ *é convergente e converge para* $x \in \mathbb{R}^n$ *se dado qualquer número* $\epsilon > 0$, *é sempre possível encontrar um número* n_0 *tal que*

$\|x_m - x\| < \epsilon$, *sempre que* $m \geq n_0$. *Neste caso, escrevemos* $\lim x_m = x$. *Caso contrário,* (x_m) *é dita divergente.*

Exemplo A.10. *A sequência* (x_m), *definida por* $x_m = \left(\dfrac{1}{m}, 2\right) \in \mathbb{R}^2$, *é convergente e converge para* $(0, 2)$. *De fato, dado* $\epsilon > 0$, *seja* $n_0 \in \mathbb{N}$ *tal que* $n_0 > \dfrac{1}{\epsilon}$ *(Propriedade Arquimediana). Logo, para todo* $m > n_0$, *tem-se* $\dfrac{1}{n_0} \geq \dfrac{1}{m}$. *portanto,*

$$\|x_m - (0,2)\| = \left\|\left(\frac{1}{m}, 2\right) - (0,2)\right\| = \left\|\left(\frac{1}{m}, 0\right)\right\| = \frac{1}{m} \geq \frac{1}{n_0} < \epsilon$$

Daí, $\lim x_m = (0, 2)$.

Podemos caracterizar a convergência de uma sequência em $\mathbb{R}^n$ *através das coordenadas dos valores desta sequência.*

Teorema A.1.7. *Sejam* $(x_m) \subseteq \mathbb{R}^n$ *uma sequência e* $x \in \mathbb{R}^n$. *Então,*

$$\lim x_n = x \Leftrightarrow \lim x_i^m = x_i, \text{ para todo } i = 1, \cdots, n,$$

onde $x = (x_1, x_2, \cdots, x_n)$ *e* $x_m = (x_1^m, x_2^m \cdots, x_n^m)$.

Demonstração. Dado $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para todo $m \geq n_0$, tem-se

$$|x_i^m - x_i| \leq \|x_m - x\| < \epsilon, \forall \; i = 1, \cdots, n,$$

Logo,

$$\lim x_i^m = x_i, \text{ para todo } i = 1, \cdots, n.$$

Reciprocamente, suponha que $\lim x_i^m = x_i$, para todo $i = 1, \cdots, n$. Assim, dado $\epsilon > 0$,existe $k_0^i \in \mathbb{N}$ tal que para todo $m \geq k_0^i$, tem-se

$$|x_i^m - x_i| < \frac{\epsilon}{\sqrt{n}}, \text{ para todo } i = 1, \cdots, n.$$

Logo, para $k_0 = max\{k_0^1, \cdots, k_0^n\}$ e $m \geq k_0$, encontramos

$$\|x_m - x\|^2 = \sum_{i=1}^{n} |x_i^m - x_i|^2 < \frac{n\epsilon^2}{n}.$$

Portanto,

$$\|x_m - x\| < \epsilon, \ para\ todo\ m \geq k_0.$$

□

Definição A.11 (**Sequência de Cauchy**). *Dizemos que uma sequência* $(x_m) \subseteq \mathbb{R}^n$ *é uma sequência de Cauchy se para cada* $\varepsilon > 0$*, existe um* $k_0 \in \mathbb{N}$ *tal que para todo* $k, l \geq k_0$*, tem-se*

$$\| x_k - x_l \| < \varepsilon.$$

Exemplo A.11. *A sequência* $(x_m) = \left(\left(1, \frac{1}{m}\right)\right)_{m \,\in\, \mathbb{N}}$ *é de Cauchy.*

De fato, dado $\epsilon > 0$*, existe* $n_0 \in \mathbb{N}$ *tal que* $n_0 > \frac{2}{\epsilon}$*. Com isso, para todo* $m, k \geqslant n_0$*, obtemos*

$\|x_m - x_k\| < \epsilon = \left\|\left(1, \frac{1}{m}\right) - \left(1, \frac{1}{k}\right)\right\| = \left\|\left(0, \frac{1}{m} - \frac{1}{k}\right)\right\| = \left|\left(\frac{1}{m} - \frac{1}{k}\right)\right| \leq \frac{1}{m} + \frac{1}{k} \leq \frac{1}{n_0} + \frac{1}{n_0} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$

Portanto, $\|x_m - x_k\| < \epsilon$ *sempre que* $m, k \geqslant n_0$*. Por conseguinte,* (x_m) *é uma sequência de Cauchy.*

Exemplo A.12. *A sequência* $((0, m)) \subset \mathbb{R}^{\ 2}$ *não é de Cauchy. De fato, existe* $\epsilon = \frac{1}{2} > 0$ *tal que para todo* $n_0 \in \mathbb{N}$*, encontra-se* $n_0, n_0 + 1 \geqslant n_0$ *que satisfazem*

$$\|(0, n_0) - (0, n_0 + 1)\| = \|(0, -1)\| = 1 > \frac{1}{2}$$

Portanto, (x_m) *não é de Cauchy.*

Mostraremos que podemos verificar se uma sequência é de Cauchy avaliando os valores de suas coordenadas.

Teorema A.1.8. *Seja* $(x_m) \subseteq \mathbb{R}^n$ *uma sequência. Então,* (x_m) *é de Cauchy se, e somente se,* (x_i^m) *é de Cauchy, para todo* $i = 1, \cdots, n$*.*

Demonstração. Suponhamos que (x_m) é de Cauchy. Então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, k \geqslant n_o$ tem-se

$$\|x_m - x_k\| < \epsilon.$$

Consequentemente, para todo $m, k \geqslant k_0$, obtemos

$$|x_i^m - x_i^k| \leqslant \|x_m - x_k\| < \epsilon, \ para \ todo \ i = 1, \cdots, n.$$

Logo, $(x_i^m)_{m \in \mathbb{N}}$ é de Cauchy em $\mathbb{R}$, para tod $i = 1, \cdots, n$.

Por outro lado, suponhamos que $(x_i^m)_{m \in \mathbb{N}}$ é de Cauchy, para todo $i = 1, \cdots, n$. Assim, dado $\epsilon > 0$, existe $n_0^i \in \mathbb{N}$ tal que para todo $m, k \geqslant n_o^i$, tem -se

$$|x_i^m - x_i^k| < \frac{\epsilon}{\sqrt{n}} \forall i = 1, \cdots, n$$

Logo, para $n_0 = \max n_0^i : i = 1, \cdots, n$ e $m, k \geqslant n_o$, obtemos

$$\|x_m - x_k\|^2 = \sum_{i=1}^{n} |x_i^m - x_i^k|^2 < n\frac{\epsilon^2}{n} = \epsilon^2$$

Ou seja, $\|x_m - x_k\| < \epsilon$, para todo $m, n > n_0$. Portanto, $(x_m) \subseteq \mathbb{R}^n$ é de Cauchy. □

Exemplo A.13. *A sequência* $((1 - \cos m, m))_{m \in \mathbb{N}} \subseteq \mathbb{R}^2$ *não é de Cauchy, pois a sequência real* $(m)_{m \in \mathbb{N}}$ *não é de Cauchy.*

Vejamos, agora, que o conceito de sequência de Cauchy é equivalente a ser convergente em $\mathbb{R}^n$.

Teorema A.1.9. *Uma sequência* $(x_k)_{k \in \mathbb{N}}$ *é de Cauchy em* $\mathbb{R}^n$ *se, e somente se,* (x_k) *for convergente.*

Demonstração: Se (x_k) é um sequência de Cauchy em, então pelo teorema anterior, $(x_{ki})_{k \in \mathbb{N}}$ para todo $(i = 1, ..., n)$ é uma sequência de Cauchy em $\mathbb{R}$. Como $\mathbb{R}$ é um conjunto completo, toda sequência de Cauchy é convergente, ou seja, $\lim_{k \to \infty} x_i^k = a_i$; $i = 1, ..., n$. Então segue pelo teorema A.1.8, $\lim_{k \to \infty} x^k = a$, com $a = (a_1, \cdots, a_n)$. A recíproca é imediata.

A.1.2 Conjuntos abertos

Estabeleceremos a definição de conjunto aberto no $\mathbb{R}^n$, bem como alguns exemplos e algumas propriedades que os caracterizam. Pretendemos ilustrar tal trabalho para que o leitor possa ter uma melhor compreensão de alguns resultados que estão por vir. Porém, para definirmos o que significa um conjunto ser aberto, precisamos saber que tipo de conjunto podemos chamar de bola aberta.

Definição A.12. *Seja $X \subseteq \mathbb{R}^n$. Um ponto $x \in X$ é chamado ponto interior a X, e escrevemos $x \in intX$, se existe $\epsilon > 0$ tal que*

$$B_\epsilon(x) \subseteq X.$$

O conjunto $intX = \{x \in \mathbb{R}^n : x \text{ é interior a } X\}$ é chamado conjunto interior de X.

Observação A.7. *É importante destacarmos que $intX \subseteq X$, já que se $x \in intX$, então $\exists\ \epsilon > 0$ tal que $x \in B_\epsilon(x) \subseteq X$. Consequentemente, $x \in X$.*

Definição A.13. *Dizemos que $X \subseteq \mathbb{R}^n$ é aberto se $X = intX$.*

Exemplo A.14. *O intervalo aberto $(a, b) \subseteq \mathbb{R}$ é aberto. Pelo que foi destacado anteriormente temos que $int(a, b) \subseteq (a, b)$. Seja $c \in (a, b)$. Desejamos encontrar $\epsilon > 0$ tal que*

$$c + \epsilon < b \Rightarrow \epsilon < b - c \text{ e}$$

$$c - \epsilon > a \Rightarrow \epsilon < c - a.$$

Basta tomar $\epsilon = min\{b - c, c - a\}$, daí temos que $B_\epsilon(c) = (c - \epsilon, c + \epsilon) \subseteq (a, b)$. Logo, $c \in int(a, b)$. Assim sendo, $(a, b) \subseteq int(a, b)$. Portanto $(a, b) = int(a, b)$, isto é, (a, b) é aberto.

Exemplo A.15. *Considerando o conjunto $[a, b)$, vemos que ele não é aberto, pois $a \in [a, b)$, mas $a \notin\ int[a, b)$, já que $(a - \epsilon, a) \subseteq (a - \epsilon, a + \epsilon)$ contém pontos que não estão em $[a, b)$, para todo $\epsilon > 0$.*

Exemplo A.16. *O conjunto $A = \{(x, 0) \in \mathbb{R}^2 : x \in (0, 1)\}$ não é aberto em $\mathbb{R}^2$. De fato, seja $(x, 0) \in A$, para todo $\epsilon > 0$ temos que $B_\epsilon((x, 0)) \nsubseteq A$, pois $\left(x, \dfrac{-\epsilon}{2}\right) \in B_\epsilon((x, 0))$ e $\left(x, \dfrac{-\epsilon}{2}\right) \notin A$.*

Exemplo A.17. *O conjunto $X = \{(x, y) \in \mathbb{R}^2 : x > 0\}$ é um conjunto aberto em $\mathbb{R}^2$. Com efeito, para $(x, y) \in X$, consideremos $B_r((x, y)) \subseteq X$, com $r = \dfrac{x}{2} > 0$. Para provar esta inclusão escolha $(a, b) \in B_r((x, y))$ e note que*

$$|a - x| \leqslant \|(a, b) - (x, y)\| < r.$$

Desta forma

$$a - x > -r = -\frac{x}{2},$$

donde,

$$a > x - \frac{x}{2} = \frac{x}{2} > 0.$$

Isto nos diz que $(a, b) \in X$.

O próximo resultado nos diz que a união (interseção) qualquer (finita) de conjuntos abertos é um conjunto aberto em $\mathbb{R}^n$.

Teorema A.1.10. *São verdadeiras as seguintes afirmações:*

i) Se $A_1, A_2, \cdots, A_m$ são abertos, então $\bigcap_{i=1}^{m} A_i$ é aberto;

ii)Se $(A_\lambda)_{\lambda \in L}$ é uma família de conjuntos abertos, então $\bigcup_{\lambda \in L} A_\lambda$ é aberto, onde L é o conjunto de índices.

Demonstração. **(i)** Se $x \in \bigcap_{i=1}^{m} A_i$, então $x \in A_i$, para todo $i = 1, 2, \cdots, m$. Com isso, existem $r_1, r_2, \cdots, r_m$ positivos tais que

$$B_{r_i}(x) \subseteq A_i, \text{ para todo } i = 1, 2, \cdots, m,$$

pois A_i é aberto. Seja $r = min\{r_1, r_2, \cdots, r_m\} > 0$. Então

$$B_r(x) \subseteq B_{r_i}(x) \subseteq A_i \text{ para todo } i = 1, 2, \cdots, m,$$

de forma que,$B_r(x) \subseteq \bigcap_{i=1}^{m} A_i$. Ou seja, $x \in int\left(\bigcap_{i=1}^{m} A_i\right)$. Com isso, $\bigcap_{i=1}^{m} A_i$ é aberto.

(ii) Seja $x \in \bigcup_{\lambda \in L} A_\lambda$, então existe $\lambda_0 \in L$ tal que $x \in A_{\lambda_0}$ que é aberto. Assim existe $r > 0$ tal que,

$$B_r(x) \subseteq A_{\lambda_0} \subseteq \bigcup_{\lambda \in L} A_\lambda.$$

Portanto, $x \in int\left(\bigcup_{\lambda \in L} A_\lambda\right)$. Por fim, $\bigcup_{\lambda \in L} A_\lambda$ é aberto.

□

Vejamos um contraexemplo para a seguinte afirmação: Uma interseção qualquer de conjuntos abertos é aberto.

Exemplo A.18. *Seja $A_n = \left(\dfrac{-1}{n}, \dfrac{1}{n}\right)$, para todo $n \in \mathbb{N}$. Os $A_n's$ são abertos (ver exemplo A.14). Note que $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$, mas $\{0\}$ não é aberto, pois, $0 \notin int\{0\}$, já que o intervalo $(-\epsilon, \epsilon)$ contém números diferentes de 0, para todo $\epsilon > 0$. Concluímos que uma interseção qualquer de abertos não necessariamente é aberto.*

A.1.3 Conjuntos fechados

Retomamos a importância sobre alguns conteúdos topológicos de $\mathbb{R}^n$, pois, os mesmos tem uma relação fundamental com o Teorema do ponto fixo para contrações ou ainda com o Método das aproximações sucessivas, um desses resultados é a ideia de conjunto fechado. Assim sendo, nesta seção, mostraremos tal definição como também outros aspectos importantes que estão ligados a esse assunto e ainda algumas propriedades importantes que regem os conjuntos que satisfazem esta denominação.

Definição A.14. *Seja $X \subseteq \mathbb{R}^n$. Dizemos que $x \in \mathbb{R}^n$ é ponto aderente a X se existe $(x_m) \subseteq X$ tal que $\lim x_m = x$.*

Exemplo A.19. *Seja $X = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$. É fácil ver que $\lim \frac{1}{n} = 0$. Daí, 0 é ponto aderente a X. Observe que $0 \notin X$.*

Definição A.15. *O conjunto de todos os pontos aderentes a X será chamado fecho de X. Denotaremos este por $\overline{X} = \{x \in \mathbb{R}^n : x \text{ é aderente a } X\}$.*

Observação A.8. *Veja que $X \subseteq \overline{X}$, para todo $X \subseteq \mathbb{R}^n$. De fato, seja $x \in X$ e defina $x_m = x$, para todo $m \in \mathbb{N}$. Então (x_m) é uma sequência constante e $\lim x_m = x$. Ou seja, $x \in \overline{X}$. Dessa forma, $X \subseteq \overline{X}$. Todavia pode ocorrer $\overline{X} \nsubseteq X$, conforme Exemplo A.19.*

Exemplo A.20. *Seja $\|x\| = r$ então x não pertence à bola aberta $B = B_r(0)$, porém é aderente a ela. Com efeito, pondo $x_k = \left(1 - \frac{1}{k}\right) x$, para todo $k \in \mathbb{N}$. Assim sendo*

$$x_k \in B_r(0), \text{ para todo } k \in \mathbb{N}, \text{ e } \lim x_k = x.$$

Logo $x \in \overline{B_r(0)}$. Reciprocamente, $x \in \overline{B_r(0)}$, então $x = \lim x_k$ com $\|x_k\| < r$, para todo $k \in \mathbb{N}$. Portanto $\|x\| = lim\|x_k\| \leq r$. Concluí-se então que $x \in \overline{B_r(0)}$ se, e somente se $\|x\| \leqslant r$, ou seja, $\overline{B_r(0)} = B_r[0]$. O mesmo argumento mostra que fecho de toda bola aberta $B_r(a)$ é bola fechada $B_r[a]$.

Definição A.16. *Dizemos que $X \subseteq \mathbb{R}^n$ é fechado se $\overline{X} = X$.*

Exemplo A.21. *O conjunto $X = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$, não é fechado, pois, $0 \in \overline{X}$, mas, $0 \notin X$, isto é, $X \neq \overline{X}$.*

Exemplo A.22. *Vamos provar que $\overline{B_r(x)} = B_r[x_0]$. Seja $y \in S_r[x]$, então $\|y - x\| = r$. Vamos provar que $y \in \overline{B_r(x)}$. De fato, seja $(x_m) \subseteq \mathbb{R}^n$ tal que definida por*

$$x_m = x + \left(1 - \frac{1}{m}\right)(y - x), \text{ para todo } m \in \mathbb{N}.$$

Assim,

$$\|x_m - x\| = \left\|\left(1 - \frac{1}{m}\right)(y - x)\right\| = \left(1 - \frac{1}{m}\right)\|(y - x)\| = \left(1 - \frac{1}{m}\right) r < r, \forall m \in \mathbb{N}$$

Ou seja, $(x_m) \subseteq B_r(x)$*. Além disso,*

$$\lim_{m \longrightarrow \infty} x_m = x + (y - x) = y$$

Ou seja, $y \in \overline{B_r(x)}$*. Com isso,* $B_r(x) \subseteq \overline{B_r(x)}$*.Agora, seja* $z \in \overline{B_r(x)}$*, então* $\in (x_m) \subset B_r(x)$ *tal que* $\lim_{m \longrightarrow \infty} \|x_m - x\| = \|z - x\|$ *.Assim,*

$$r \geqslant \lim_{m \longrightarrow \infty} \|x_m - x\|| = \|z - x\|,$$

pois $\|x_m - x\| \geq r$*. Por fim,* $\|z - x\| \leqslant r$*. Isto é,* $z \in B_r(x) = \overline{B_r(x)}$ *.*

Analogamente , prova-se que $\overline{S_r[x]} = s_r[X]$*.*

Teorema A.1.11. *Seja* $X \subseteq \mathbb{R}^n$*. Então* $x \in \overline{X} \Leftrightarrow \forall\ \epsilon > 0$*, tem-se que* $B_\epsilon(x) \cap X \neq \emptyset$*.*

Demonstração:

Se $x \in \overline{X}$, então existe $(x_n) \subseteq X$ tal que $\lim x_n = x$. Assim sendo,dado $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para todo $n > n_0$, temos $x_n \in B_\epsilon(x)$. Consequentemente,

$$B_\epsilon(x) \cap X \neq \emptyset.$$

Reciprocamente, uponha que

$$B_\epsilon(x) \cap X \neq \emptyset, \text{ para todo } \epsilon > 0$$

Assim sendo, temos que

$$B_{\frac{1}{n}}(x) \cap X \neq \emptyset, \text{ para todo } n \in \mathbb{N}.$$

Logo, existe (x_n) em X tal que $\|x_n - x\| < \dfrac{1}{n}$. Passando o limite, quando $n \longrightarrow \infty$, obtemos $0 \leq \lim \|x_n - x\| \leq 0$. Dessa forma, $\lim x_n = x$. Isto nos diz que $x \in \overline{X}$.

Teorema A.1.12. $X \subseteq \mathbb{R}^n$ *é fechado se, e somente se,* $X^c = \mathbb{R}^n - X$ *é aberto.*

Demonstração. Se X é fechado, então $X = \overline{X}$. Seja x $\in X^c$. Então, x não pertence $X = \overline{X}$, ou seja, existe$\epsilon > 0$ tal que $B_\epsilon(x) \cap X = \emptyset$. Logo, $B_\epsilon(x) \subseteq X^c$ e isto mostra que X^c é aberto.

Reciprocamente, Suponha que X^c é aberto e tome $x \in \overline{X}$, como X^c é aberto, então $X^c = intX^c$. Ora, se x não pertence X, então $X \in X^c$, logo existe $epsilon > 0$ tal que $B_\epsilon(x) \subseteq X^c$. Portanto,

$B_\epsilon(x) \cap X = \emptyset$. Isso contradiz o teorema anterior, dessa forma, $x \in X$. Consequentemente, $X = \overline{X}$, isto é, X é fechado.

□

Vamos agora provar que união finita (respectivamente interseção qualquer) de conjuntos fechados é fechado em $\mathbb{R}^n$.

Teorema A.1.13. *As afirmações a seguir são verdadeiras:*

1. *Se $X_1, \cdots, X_m$ são fechados, então $\cup_{i=1}^m X_i$ é fechado*

2. *Se (X_λ) é uma família de conjuntos fechados, então $\cap_{\lambda \in L} X_\lambda$ é fechado*

Demonstração. 1. Sejam $X_1, X_2, \cdots, X_m$ fechados. Então, $X_1^c, \cdots, X_n^c$ são abertos, logo $(\cup_{i=1}^m X_i)^c / X_i^c = \cup_{i=1}^m C X_i$ é aberto. Assim, $\cup_{i=1}^m X_i$ é fechado.

2. (X_λ) uma família qualquer de fechados, então $(\cap_\lambda X_\lambda)^c = \cap_\lambda / X_i^c$ é aberto. Logo, $\cap_\lambda$ é fechado.

□

Não é verdade que união qualquer de fechados é fechada. Vejamos um exemplo:

Exemplo A.23. *Sabemos que $B_1(0) = \cup_x \in B_1[x]$. Assim, não é verdade que uma união qualquer de fechados é fechada, pois $\{x\}$ é fechado,$\forall x \in \mathbb{R}^n$, e $\overline{B_1(0)} = B_1[x]$*