UNIVERSIDADE FEDERAL DE SANTA CATARINA

DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO E SISTEMAS

OTIMIZAÇÃO GLOBAL DETERMINÍSTICA - UM ALGORITMO

DISSERTAÇÃO SUBMETIDA A UFSC PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA

NELSON HEIN

FLORIANÓPOLIS
SANTA CATARINA - BRASIL
ABRIL DE 1994

OTIMIZAÇÃO GLOBAL DETERMINÍSTICA - UM ALGORITMO

DISSERTAÇÃO DE MESTRADO

NELSON HEIN

ESTA DISSERTAÇÃO FOI JULGADA PARA A OBTENÇÃO DO TÍTULO DE

"MESTRE EM ENGENHARIA"

ESPECIALIDADE ENGENHARIA DE PRODUÇÃO E APROVADA EM SUA FORMA FINAL PELO PROGRAMA DE PÓS-GRADUAÇÃO

Prof. Osmar Possamai, Dr.
Coordenador

BANCA EXAMINADORA:

Prof. Plínio Stange, Dr.
Presidente

Prof. Edgar A. Lanzer, Dr.

Prof. Claudio Loesch, Msc.

Ao meu irmão
Wilmar
(in memorium)

RESUMO

O problema da otimização consiste em determinar os valores ótimos (máximos ou mínimos) de uma função, a qual é denominada de função objetivo. Esta função objetivo depende de um conjunto de variáveis as quais são denominadas de variáveis de decisão ou de controle.

A teoria da otimização desenvolve métodos determinísticos (numéricos ou analíticos) e estocásticos (simulação convencional, algoritmos genéticos, etc.), para a obtenção dos valores ótimos das variáveis de decisão dos problemas de otimização. Valores estes que vão determinar a imagem ótima (máxima ou mínima) da função objetivo.

O objetivo da otimização global é determinar o menor dos mínimos (ou o maior dos máximos), que pode ser único ou não, porém sua imagem será única; a este ponto denominamos de mínimo global.

Este trabalha procura localizar deterministicamente o mínimo global de problemas de programação matemática (preferencialmete não-linear), sugerindo um algoritmo que se utiliza de modificações do problema original, usando para isso uma nova função objetivo, e removendo a antiga função objetivo para junto das restrições iniciais.

Para a verificação dos resultados foram usados algoritmos genéticos que são métodos estocásticos na forma de uma simulação inteligente, usando para isso mecanismos que a genética animal utiliza para o melhoramento das espécies, tais como operadores de mutação e crossing-over. A otimização se faz usando a lei de Darwin, ou seja, o algoritmo da prioridade para aqueles que fenotipicamente são mais aptos, deixando em segundo plano os elementos mais fracos, porém não os desprezando completamente, pois devido a sua formação genotípica, eles podem em alguma geração posterior dar contribuições de melhoria.

ABSTRACT

The issue of optimization consists of determinig the optimal values (maximum and minimum) for a function called objective function. This objective function depends on a set of variables called decision or control variables.

This optimization theory develops deterministic (numerical or analytical) and stocastic methods convencional simulation, genetic algorithms, etc), so that optimal values dealing with decision variables regarding the issues of optimization can be obtained. These values will determine the optimal image (maximum or minimum) of the objective function.

The objective of the global optimization is to estimate the lowest minimum (or the highest maximum), wich can be unique or not, but it's image will be unique. Such a point we call global minimum.

The study intends to estimate deterministically the global minimum of problems related to mathematical programing (preferably non-linear), by indicasting an algorithm wich requires changes in the original problem through another objetive function and removal of the previous objective function to it's initial restrictions.

The estimation of the results was done trough genetic algorithms wich are stocastic methods in the shape of an intelligent simulation and mechanisms used by the animal genetics to imporve the species as, for example, mutation operators and crossing-over. The optimization happens with the use of Darwin law, that is, the algorithm gives priority to those phenotypically more qualified ones and lets weaker elements behind withoud neglecting them completely because, due to their phenotypical formation, they can, in any later generation, contribute to improvement.

SUMÁRIO

pagina

LISTA DE FIGURAS....viii

CAPÍTULO 1

1. Introdução....01
1.1. Origem do Trabalho....01
1.2. Objetivo do Trabalho....02
1.3. Importância do Trabalho....02
1.4. Estrutura do Trabalho....03
1.5. Limitações do Trabalho....03

CAPÍTULO 2

2. Aspéctos Topológicos....05
2.1. Espaços Topológicos....05
2.2. Comparação de Topologias....11
2.3. Subespaços....12
2.4. Espaço Produto....13
2.5. Espaços Métricos....16
2.6.Espaços Compactos....21
2.7. Espaços Conexos....27
2.8. Sucessões....32
2.9. Continuidade....38
2.10. Comentário Final....45

CAPÍTULO 3

3. Teoria de Otimização....46
3.1. Introdução....46
3.2. Cones Convexos....46
3.3. Cones Complementares....49
3.4. Hiperplanos e Teoremas de Separação....52
3.5. Funções Côncavas e Semi-côncavas....54
3.6. Funções Semi-côncavas Diferenciáveis....57
3.7. Maximização de Funções Semi-côncavas....61
3.8. O Teorema de Kuhn-Tucker - Caso Geral....64
3.9. Comentário Final....67

CAPÍTULO 4

4. Condições de Otimalidade Global....68
4.1. Introdução....68
4.2. Otimização sem Restrições....68
4.3. Otimização de Funções com restrições de Igualdades....70
4.4. Otimização de Funções com Restrições de Desigualdades....72
4.5. Dualidade Lagrangeana....74
4.6. Condições de Otimalidade Global....74
4.7. Comentário Final....75

CAPÍTULO 5

5. Algoritmo de Otimização Global....76
5.1. Introdução....76
5.2. O Algoritmo....76

CAPÍTULO 6

6. Aplicações....81
6.1. Apresentação dos Modelos....81
6.2. Problemas....81
6.3. Resultados....84
6.4. Limitações....85

REFERÊNCIAS BIBLIOGRÁFICAS....86

ANEXO 1 Pesquisa Multidimensional Discreta....88
Pesquisa Mutidimensional Contínua

ANEXO 2 Otimização Restrita....89

ANEXO 3 Problemas....90

ANEXO 4 Algoritmos genéticos....99

LISTA DE FIGURAS

pagina
Figura 1..48
Figura 2..57
Figura 3..58
Figura 4..60
Figura 5..62

1. INTRODUÇÃO

1.1. Origem do Trabalho

A Pesquisa Operacional, segundo Ehrlich, é "uma metodologia de estruturar processos aparentemente não estruturados por meio da construção de modelos. Utiliza um conjunto de técnicas quantitativas com o intuito de resolver aspectos matemáticos dos modelos"[10].

As técnicas quantitativas dividem-se em grandes grupos, como: métodos estatísticos, simulação, PERT/CPM, teoria do inventário, teoria das filas, programação dinâmica, programação heurística, programação linear e programação não-linear. Estas técnicas nós classificamos normalmente como clássicas.

A programação matemática (linear, não-linear, dinâmica, heurística), tem por objetivo o estudo teórico dos problemas de otimização finita, assim como concepção e implementação de métodos numéricos para resolvê-los.

A programação linear geralmente é apresentada em separado do caso não-linear, por motivos históricos e práticos. Os motivos históricos não merecem comentários aqui, porém os práticos advém do problema da convergência, que na programação linear pode ser obtida num número finito de passos, o que não ocorre tão facilmente nos problemas de programação não-linear, além de que o ferramental matemático necessário no caso linear é muito menor do que no caso não-linear.

Acreditamos que tão importante quanto o assunto a ser pesquisado é a sua escolha, pois ele deve estar ligado ao pesquisados e atender às suas inclinações e limitações. Ele não deve ser escolhido como objetivo e como meio de atender uma exigência regulamentar, deve ser escolhido sim, como a iniciação dentro de uma linha de pesquisa e em busca de uma realização profissional. Assim, devido a nossa formação acadêmica admitimos que, um forte laço nos une aos aspectos matemáticos envolvidos em aplicações práticas. Aspectos tão numerosos numa área tão fértil que é a Engenharia de Produção. Com isso fica legitimado o nosso compromisso quase que sentimental com programação não-linear, por ela permitir e ao mesmo tempo ímpor grande habilidade com esta Ciência que é a matemática.

O motivo que nos levou a escolher a Otimização Global, dentro da programação não-linear, foi a vontade que temos em dar continuidade aos trabalhos já iniciados por parte de nosso professor orientador, que em outra ocasião chegou a trabalhar com um dos maiores pesquisadores na área, que é o respeitável Dr. Professor R. Horst

(Universidade de Trier - Trier - Alemanha). Trazendo-nos desta convivência subsídios bibliográficos, além do seu entusiasmo por este tema.

1.2. Objetivo do Trabalho

O objetivo fundamental do presente trabalho divide-se em duas partes:

(I) Construir um algoritmo que permita a localização do ponto de ótimo global dentro de problemas de programação matemática (não-linear) multimodais;

(II) Apresentar subsídios matemáticos já existentes em literatura específica, que fundamente cientificamente o algoritmo proposto.

1.3. Importância do trabalho

A otimização global é uma área recente de pesquisa dentro da programação matemática. Não haveria a princípio nenhuma necessidade que justificasse o motivo de iniciar pesquisas neste sentido, pois frente aos modelos didáticos que costumeiramente são utilizados dentro da programação não-linear, os conceitos do cálculo diferencial são suficientes para a localização dos pontos de ótimo. Porém, quando partimos para modelos profissionais práticos de porte mais considerável, a localização do ponto de ótimo global pode significar uma economia considerável, independente de sua especificação (monetária, recursos naturais, recursos materiais, ou outro).

Heuristicamente parece razoável admitir que, para localizarmos o ótimo global de um problema de pragramação não-linear, seria partir de um conjunto não-vazio e não-unitário de pontos de partida (startpoints), e utilizar como ótimo global aquele que verificar os melhores valores de retorno. Este tipo de técnica pode eventualmente até vir a localizar o ótimo global, porém não há garantia nenhuma de que isto se verifique.

Atualmente, além dos problemas práticos de otimização de utilização e alocação de recursos naturais, temos a inteligência aplicada que tem como objetivo simular em computadores (hardware) o comportamento do raciocínio humano, usando para isso software (programas) específicos. Para isso lança-se mão de uma ferramenta, denominada pelos profissionais da área de inteligência artificial, como redes neurais. Estas redes além de simularem a função dos neurônios humanos, são também bastante numerosos. Estas redes precisam ser otimizadas a ponto de cometer o menor erro quando estiverem diante de uma situação, e frente a ela tiver a responsabilidade de tomar decisões. Evidente que a decisão é na verdade a solução ótima de um problema de programação matemática. Acreditamos que com o algoritmo a ser apresentado, estaremos de alguma forma contribuindo na resolução deste problema, fazendo com que a possibilidade de erro por parte de uma rede neural seja minimizado globalmente.

1.4. Estrutura do Trabalho

O presente trabalho foi dividido em seis seções.

Esta seção visa definir os objetivos do trabalho apresentado, bem como sua importância e limitações.

A seção seguinte, denominada "Aspéctos Topológicos" tem por objetivo apresentar um sumário conceitual sobre topologia, que vem de encontro as necessidades de entendimento do algoritmo a ser exposto. Serão vistos conceitos como espaços topológicos, conjuntos fechados, conjuntos abertos, vizinhança, ponto interior, fecho, aderência, ponto de fronteira, ponto de acumulação, espaços separáveis, métrica, conjunto limitado, bola aberta, bola fechada, espaços compactos, coberturas, espaços conexos, sucessões, convergência e divergência de sucessões, função contínua, espaços homeomorfos, bem como boa parte dos teoremas obtidos através do uso destes conceitos.

O terceiro capítulo visa apresentar conceitos e teoremas utilizados dentro da teoria de otimização matemática. Estes teoremas e conceitos permitirão um tratamento unificado e rigoroso de várias questões, que aparecem como:, cones convexos, hiperplanos, funções côncavas (diferenciáveis), maximização de funções, bem como teoremas clássicos como o de Farkas e o enunciado das condições necessárias e suficientes de Kuhn-Tucker.

O quarto capítulo estabelece a caracterização do ótimo global, isto é feito mediante a teoria da programação não-linear preescrita em manuais específicos, bem como a apresentação de definições e teoremas recentemente criados por estudiosos da otimização global.

O quinto capítulo é o ponto alto do trabalho. Nele é apresentado, um algoritmo não-numérico, para a obtenção do ponto de ótimo global de problemas de programação não-linear. Caracterizado pela modificação estrutural do problema inicial, o algoritmo cria uma nova função objetivo, incorporando a antiga as restrições já existentes.

O sexto capítulo apresenta os resultados obtidos com o uso do algoritmo, em problemas de programação não-linear, obtidos em bibliografias, que tratam da obtenção de resultados ótimos globais. Estes problemas serão denominados por nós de problemas patogênicos, devido a sua construção que dificulta enormemente a obtenção do ótimo global.

Além destes seis capítulos, acompanham em anexo a documentação dos problemas resolvidos, bem como apresenta um resumo das técnicas mais freqüentes utilizadas pela programação não-linear, para a otimização dos problemas apresentados. Traz também listagens dos programas computacionais que utilizam conceitos de estudos recentes denominados de algoritmos genéticos, que são técnicas estocásticas na forma de uma simulação inteligente, utilizando para tanto mecanismos naturais de melhoramento genético, idêntico ao utilizado pelos seres vivos.

1.5. Limitações do Trabalho

O algoritmo possui algumas variações que não foram devidamente testadas, são elas:

(1) O comportamento do algoritmo frente a problemas de grande porte.

(2) O algoritmo somente foi testado em problemas que inicialmente possuiam restrições convexas, com o uso do nosso algoritmo a única restrição não-convexa, fica sendo a antiga função objetivo, que é incorporada as restrições dadas inicialmente.

2. ASPECTOS TOPOLÓGICOS

2.1. Espaços Topológicos

Ao se tratar com otimização, há a necessidade de se fazer algumas considerações topológicas as quais mesmo que não se autoexpliquem no decorrer do texto, fundamentam intuitivamente o assunto.

Definição 2.1

Denomina-se por topologia (ou estrutura topológica) o conjunto não vazio X, toda a família τ de subconjuntos de X satisfazendo:

(i) X e Ø são elementos de τ;

(ii) A intersecção de um número finito de elementos de τ é ainda um elemento de τ;

(iii) Toda a união de elementos de τ é um elemento de τ.

Os elementos de τ serão designados por abertos de X (para a topologia τ).

O par (X,τ) é chamado de espaço topológico.

Observação: Por abuso de linguagem designaremos X por espaço topológico sempre que não haja possibilidade de confusão sobre a topologia a considerar.

Definição 2.2

Seja X um espaço topológico. Um subconjunto F de X é um fechado se o seu complemento, que notaremos por X\F, for um aberto.

Teorema 2.1

A família dos fechados de um espaço topológico X tem as seguintes propriedades:

(p_1) X e Ø são fechados;

(p_2) A união finita de fechados é um fechado;

(p_3) A intersecção de fechados é um fechado.

A demonstração é imediata a partir das definições de fechado e aberto. É necessário notar que: (Leis de Morgan)

$$(1) X \setminus \left[\bigcap_{i=1}^{n} G_i \right] = \bigcap_{i=1}^{n} [X \setminus G_i];$$

$$(2) X \setminus \left[\bigcap_{i=1}^{n} G_i \right] = \bigcup_{i=1}^{n} [X \setminus G_i],$$

com G_i subconjunto de X, utilizados para as demonstrações.

Definição 2.3

Um subconjunto v de um espaço topológico X, é uma vizinhança de um elemento x de X, se existir um aberto A tal que $x \in A$ e $A \subset v$.

Designaremos a família de vizinhanças de x por $\vartheta(x)$.

Observações:

- $\vartheta(x)$ não é vazia;
- x é elemento de qualquer das suas vizinhanças;
- Se $v \in \vartheta(x)$ e $v \subset w$, então $w \in \vartheta(x)$;
- Toda intersecção finita de vizinhanças de x é ainda elemento de $\vartheta(x)$;

É evidente que um aberto é vizinhança de qualquer dos seus elementos. Se supusermos que um conjunto A é vizinhança de todos os seus elementos, então para todo o elemento x de A, A é vizinhança de x, logo existe um aberto A tal que:

$x \in A_x$ e $A_x \subset A$.

Como $A = \bigcup_{x \in A} \{x\}$, então $A \subset \left[\bigcup_{x \in A} A_x\right]$. Por outro lado, $\left[\bigcup_{x \in A} A_x\right] \subset A$ (pois $A_x \subset A$), donde $A = \bigcup_{x \in A} A_x$. Por (iii), podemos agora afirmar que: um subconjunto A de um espaço topológico X é aberto se e só se $A \in \vartheta(x)$, para todo x em A.

Definição 2.4

Seja A um subconjunto do espaço topológicoX. Um elemento x de X é ponto interior de A, se A for vizinhança de x. Designa-se por interior de A o conjunto de pontos interiores de A. Notaremos o interior de A por $\overset{o}{A}$.

Observação: É evidente que $\overset{o}{A} \subset A$.

Teorema 2.2

Seja X um espaço topológico. Se A e B forem subconjuntos de X, então:

(a) $\overset{o}{A}$ é o "maior" aberto contido em A, isto é, $\overset{o}{A}$ é aberto e todo aberto A, contido em A, satisfaz $A \subset \overset{o}{A} \subset A$.

(b) A é aberto se e só se $A = \overset{o}{A}$.

(c) Se A é subconjunto de B então $\overset{o}{A}$ é subconjunto de $\overset{o}{B}$.

(d) $\overset{o}{\overbrace{A \cap B}} = \overset{o}{A} \cap \overset{o}{B}$

(e) $\overset{o}{\overbrace{A \cup B}} = \overset{o}{A} \cup \overset{o}{B}$

Demonstração

(a) Seja x um elemento qualquer de $\overset{o}{A}$. Como $A \in \vartheta(x)$, existe um aberto A_1 tal que $x \in A_1 \subset A$. Vamos provar que todo aberto A, contido em A, está contido em $\overset{\circ}{A}$ (em particular, $A_1 \subset \overset{\circ}{A}$). Seja y um elemento arbitrário em A. Como A é aberto, então A é vizinhança de y. Mas $A \subset A$, logo $A \in \vartheta(y)$. Por definição de ponto interior, $y \in \overset{\circ}{A}$, donde $A_1 \subset \overset{\circ}{A}$. Podemos então afirmar que $x \in A_1 \subset \overset{\circ}{A}$, isto é, $\overset{\circ}{A} \in \vartheta(x)$. x era arbitrário em $\overset{\circ}{A}$, logo $\overset{\circ}{A}$ é aberto por ser vizinhança de qualquer dos seus pontos.

A demonstração deste item termina se observarmos que $\overset{\circ}{A} \subset A$.

(b) Se A é aberto, então é o "maior" aberto contido em A. Logo pelo item anterior, A em $\overset{\circ}{A}$ são iguais. Suponhamos agora $A = \overset{\circ}{A}$. Ainda pelo item anterior, $\overset{\circ}{A}$ é aberto. Logo A é aberto.

As demais demonstrações não julgamos necessárias observando apenas que em (e) a inclusão pode ser estrita.

<u>Definição 2.5</u>

Designa-se por exterior de um subcojunto A de um espaço topológico X, e nota-se por ext(A), o interior de X\A.

<u>Definição 2.6</u>

Seja A um subconjunto de um espaço topológico X.

(a) Um elemento x de X é <u>aderente</u> a A, se $\forall\, v \in \vartheta(x)$, $v \cap A \neq \emptyset$.

(b) Designa-se por <u>fecho</u> ou <u>aderência</u> de A, o conjunto dos pontos aderentes a A.

Notaremos o fecho de A por $\overline{A}$.

<u>Observações:</u>

- Um ponto aderente a A pode não pertencer a A;
- É evidente que $A \subset \overline{A}$.

<u>Teorema 2.3</u>

Sejam A e B subconjuntos de um espaço topológico X.

(a) $\overline{A} = X\backslash(\text{ext}(A))$.

(b) $\overline{A}$ é o "menor" fechado contendo A, isto é, $\overline{A}$ é fechado e qualquer fechado C, contendo A, é tal que $C \supset \overline{A} \supset A$.

(c) A é fechado se e só se $A = \overline{A}$.

(c)$\overline{A \cup B} = \overline{A} \cup \overline{B}$

(e)$\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

(f) Se A é um subconjunto de B, então $\overline{A}$ é subconjunto de $\overline{B}$.

Demonstração

(a) Demonstrar que $\overline{A} = X\backslash ext(A)$ é equivalente a provar que

$X\backslash\overline{A} = ext(A)$.

Vamos demonstrar esta última igualdade. Seja x um elemento qualquer de $X\backslash\overline{A}$. Então $x \notin \overline{A}$, isto é, existe $v \in \vartheta(x)$ tal que $v \cap A = \emptyset$. Logo $v \subset X\backslash A$, pertencendo x ao interior de $X\backslash A$. Então $x \in ext(A)$, o que nos permite afirmar que $X\backslash\overline{A} \subset ext(A)$. Fazendo o raciocínio no sentido inverso, provaríamos que $ext(A) \subset X\backslash\overline{A}$. Provamos assim que $X\backslash\overline{A}$ e ext(A) são iguais.

(b) Pelo item anterior sabemos que $\overline{A}$ é fechado, pois pelo Teorema 2.2, $\overset{o}{\overbrace{X\backslash A}}$ é aberto.

Notamos que $A \subset \overline{A}$. Falta provar que se trata do "menor" fechado que contém A. Seja C um fechado que contém A. Seja C um fechado contendo A. Então X\C é aberto e está contido em X\A. Logo pelo Teorema 2.2, $(X\backslash C) \subset \overset{o}{\overbrace{X\backslash A}}$. Mas $\overset{o}{\overbrace{X\backslash A}} = ext(A)$. Então pelo item anterior, $(X\backslash C) \subset (X\backslash\overline{A})$, donde $C \supset \overline{A}$.

Novamente deixamos de lado as demais alíneas.

Definição 2.7

Seja A um subconjunto de um espaço topológico X. A fronteira de A, que notaremos por fr(A), é a intersecção dos fechos de A e X\A.

Observação: A fronteira de A é o conjunto dos elementos x de X, tais que toda a vizinhança v de x, tem intersecção não vazia com A e X\A. Estes pontos são também denominados de pontos de fronteira.

Teorema 2.4

Seja A um subconjunto de um espaço topológico X.

(a) $\overline{A} = A \cup fr(A) = \mathring{A} \cup fr(A)$.

(b) $fr(A) = \overline{A}\backslash A$.

(c) $fr(A) = X\backslash(\mathring{A} \cup ext(A))$.

(d) A é fechado se e só se $fr(A) \subset A$.

(e) A é simultaneamente fechado e aberto se e só se $fr(A) = \emptyset$.

(f) fr(A) é fechado.

<u>Demonstração</u>

(a) $A \cup fr(A) = A \cup (\overline{A} \cap \overline{X\backslash A}) = A \cup (\overline{A} \cap (X\backslash ext.(X\backslash A))) =$

$= A \cup (\overline{A} \cap (X\backslash\mathring{A})) = (A \cup \overline{A}) \cap (A \cup (X\backslash\mathring{A})) = \overline{A} \cap X = \overline{A}$.

A penúltima igualdade resulta de $A \subset \overline{A}$ e $(X\backslash\mathring{A}) \supset (X\backslash A)$.

$\mathring{A} \cup fr(A) = = \overline{A}$ demonstra-se de modo análogo.

(b) $fr(A) = \overline{A} \cap \overline{X\backslash A} = \overline{A} \cap (X\backslash A)) = \overline{A} \cap (X\backslash\mathring{A}) = \overline{A}\backslash\mathring{A}$.

(c) Demonstração análogo à das alíneas superiores.

(d) Se A é fechado então $fr(A) = \overline{A} \cap (\overline{X\backslash A}) = A \cap (\overline{X\backslash A}) \subset A$.

Se $fr(A) \subset A$, então $\overline{A} = A \cup fr(A) = A$. Logo é fechado.

(e) Se A é fechado, então $A = \overline{A}$. Se A é aberto, então $A = \mathring{A}$. Lgo, por (b), $fr(A) = \overline{A}\backslash\mathring{A} = A\backslash A = \emptyset$. Se $fr(A) = \emptyset$, então por (a), $\overline{A} = A \cap fr(A) = A$. Assim demonstramos que A é fechado. Ainda por (a), $\overline{A} = \mathring{A}$. Esta última igualdade prova que A é aberto.

(f) $\overline{fr(A)} = \overline{\overline{A}} \cap \overline{\overline{(X\backslash A)}} = \overline{\overline{(A)}} \cap \overline{\overline{(X\backslash A)}}$, mas $\overline{A}$ e $\overline{X\backslash A}$ são fechados. Logo $\overline{fr(A)} = \overline{A} \cap \overline{X\backslash A} = fr(A)$, isto é, fr(A) é fechado.

<u>Definição 2.8</u>

Sejam A um subconjunto do espaço topológico X e x um ponto de X.

(a) x é <u>ponto de acumulação</u> de A se, para toda a vizinhança v de x, se verifica $(v\backslash\{x\}) \cap A = \emptyset$.

(b) x é ponto isolado de A se existir uma vizinhança v de x tal que $v \cap A = \{x\}$.

(c) Chamaremos derivado de A ao conjunto A' dos pontos de acumulação de A.

Observações:

- É evidente que se x é ponto de acumulação de A, então $x \in \overline{A}$.
- Nem todo ponto de acumulação de A tem que pertencer a A.
- Todo ponto isolado do conjunto A pertence a A.
- A é a união do derivado de A e do conjunto dos pontos isolados de A.
- Todo elemento de A é ponto isolado ou de acumulação de A.

Logo $\overline{A} = A \cup A'$.

Teorema 2.5

Um subconjunto A de um espaço topológico X é fechado se e só se $A' \subset A$.

Demonstração

Se A é fechado, então $\overline{A} = A = A' \cup A$. Logo $A' \subset A$.

Se $A' \subset A$, então $\overline{A} = A \cup A' = A$, o que, pelo Teorema III, nos permite afirmar que A é fechado.

Definição 2.9

Um subconjunto A de um espaço topológico X é denso em X se $\overline{A} = X$.

2.2. Comparação de Topologias

Definição 2.10

Sejam (X,τ_1) e (X,τ_2) espaços topológicos. τ é uma topologia menos fina que τ_1 $(\tau_2 \leq \tau_1)$ se todo aberto de τ_2 é aberto de τ_1.

2.3. Subespaços.

Teorema 2.6

Seja A um espaço topológico (X,τ) e $A \subset X$, $A \neq \emptyset$. Então

$\tau_A = \{A \cap \vartheta, \vartheta \in \tau\}$ é uma topologia em A.

Demonstração

(a) Como $\emptyset \in \tau$, então $\emptyset = (A \cap \emptyset) \in \tau_A$.

Sabemos também que $X \in \tau$, logo $A = (A \cap X) \in \tau_A$.

(b) Sejam A_1 e A_2 elementos de T_A.

Então $A_1 = A \cap A_1$ e $A_2 = A \cap A_2$, com A_1 e A_2 elementos de T.

Logo $A_1 \cap A_2 = A \cap (A_1 \cap A_2) \in \tau_A$, pois $(A_1 \cap A_2) \in T$.

(c) Seja $\{A_i, i \in N^*\}$ uma família de elementos de τ_A. Para cada índice i existe $A_1 \in \tau$ tal que $A_i = A \cap A_i$. Como $\bigcup_{i \in N^*} A_i$ é elemento de τ e $\bigcup_{i \in N^*} = A \cap (\bigcup_{i \in N^*} A_i)$, então $\bigcup_{i \in N^*} A_i$, ainda é um elemento de τ_A.

Deste modo provamos o que pretendíamos,que τ_A é uma topologia em A.

Definção 2.11

Nas condições do teorema anterior, τ_A será designada por topologia induzida em A, por τ e (A,τ_A) é subespaço topológico de (X,τ).

Por comodidade, escreveremos A em vez de (A,τ), sempre que daí não advenha qualquer confusão relativamente à topologia em questão.

Teorema 2.7

Se A for um subespaço de (X,τ), então:

(a) Todo o fechado de A' é da forma $A \cap F$, com F fechado em τ.

(b) Se x for um elemento arbitrário de A, se $\vartheta(x)$ e $\vartheta(x)$ forem, respectivamente, as famílias de vizinhança de x em τ_A e τ, podemos afirmar que todo o elemento v de $\vartheta_A(x)$ pode ser escrito na forma $v = A \cap v_1$, com $v_1 \in \vartheta(x)$.

Demonstração

(a) Seja F um fechado em (A,τ_A). Então $A\backslash F = A\cap\vartheta$, com $\vartheta\in\tau$. Logo $F = A\backslash(A\cap\vartheta) = A\cap(X\backslash(A\cap\vartheta)) = A\cap(X\backslash A)\cup(X\backslash\vartheta)) = A\cap(X\backslash\vartheta) = A\cap F_x$, com $F_x = X\backslash A$ em fechado em $(X\backslash\tau)$.

(b) Se $v \in \vartheta_A(x)$, então existe $A = (A \cap \vartheta) \in \tau_A$ tal que $x \in \vartheta_A \subset v$ e A aberto em $(X\backslash\tau)$.

Logo $x \in A \cap \vartheta \subset v$.

É evidente que $v_1 = v \cup \vartheta$ é um elemento de $\vartheta(x)$. Além disso, $v_1 \cap A=$ $= (v \cap A) \cup (\vartheta \cap A) = v \cup \vartheta_A = v$.

2.4. Espaço Produto

Definição 2.12

Sejam $\{(X_k,\tau_k)\}_{1\leq k\leq n}$ uma família finita de espaços topológicos e

$$X=\prod_{k=1}^{n} X_k.$$

Para k entre 1 e n, consideremos abertos w_k em X_k.

$W=\prod_{k=1}^{n} W_k$ é designado por paralelepípedo aberto de bases $(w_k)_{1\leq k\leq n}$.

Teorema 2.8

Nas condições da definição anterior, a família τ das uniões de paralelepípedos abertos de X, é uma topologia em X.

Demonstração

(i)$X = \prod_{k=1}^{n} X_k \in \tau$

$\emptyset = \emptyset \times X_2 \times \ldots \times X_n \in \tau$

(ii) Sejam A e B elementos de τ, então:

$$A = \bigcup_{i\in I}\left[\prod_{k=1}^{n} W_{ik}\right] \text{ e } B = \bigcup_{j\in J}\left[\prod_{k=1}^{n} W_{jk}\right]$$

com W_{ik} e W_{jk} elementos de τ_k, para $i \in I$ e $j \in J$ e k entre 1 e n.

$$A \cap B = \left[\bigcup_{i\in I}\left(\prod_{k=1}^{n} W_{ij}\right)\right] \cap \left[\bigcup_{j\in J}\left(\prod_{k=1}^{n} W_{jk}\right)\right] =$$

$$= \bigcup_{(i,j)\in I x J}\left[\left(\prod_{k=1}^{n} W_{ik}\right) \cap \left(\prod_{j=1}^{n} W_{jk}\right)\right] =$$

$$= \bigcup_{(i,j)\in I x J}\left[\prod_{k=1}^{n}\left(W_{ik} \cap W_{jk}\right)\right]$$

Logo $A \cap B$ pertence a τ, pois W_{ik} e W_{ij} são elementos de T_k.

(iii) Seja A_j elementos de τ, para j e J. Então:

$$A_j = \bigcup_{l\in L}\left[\prod_{k=1}^{n} W_{lk}\right]$$

Logo:

$$\bigcup_{j\in J} A_j = \bigcup_{j\in J}\left[\bigcup_{l\in L}\left(\prod_{k=1}^{n} W_{lk}\right)\right] \in \tau$$

<u>Definição 2.13</u>

A topologia definida no teorema anteiror é designada por <u>topologia produto</u> da família $\{\tau_k\}_{1\leq k\leq n}$.

(X,τ) será designado por espaço topológico produto ou espaço produto da família de espaços $(X_k, \tau_k)_{1\leq k\leq n}$.

<u>Definição 2.14</u>

Seja X um espaço topológico

(a) Dois pontos x e y de X são <u>separáveis</u> se existem $v \in \vartheta(x)$ e

$w \in \vartheta(y)$ tais que $v \cap w = \emptyset$.

(b) Se todos os pares de elementos distintos de X são constituídos por pontos separáveis, então X é separado.

Teorema 2.9

Todo subespaço A de um espaço separado X, é separado.

Teorema 2.10

Se $(X_i)_{1 \leq k \leq n}$ for uma família de espaços separados, o espaço produto $X = \prod_{j=1}^{n} X_i$ é um espaço separado.

Demonstração

Sejam $x = (x_i)_{1 \leq i \leq n}$ e $y = (y_i)_{1 \leq i \leq n}$ dois elementos distintos pertencentes a X.

Como $x \neq y$, existe k entre 1 e n tal que $x_k \neq y_k$.

Mas X_k é separado. Logo existem $v_k \in \vartheta(x_k)$ e $w_k \in \vartheta(y_k)$ tais que $v_k \cap w_k = \emptyset$.

Consideremos $v = X_1 \times X_2 \times \ldots \times X_{k-1} \times v_k \times \ldots \times X_n \in \vartheta(x)$

e

$w = X_1 \times X_2 \times \ldots \times X_{k-1} \times w_k \times \ldots \times X_n \in \vartheta(y)$.

Então $v \cap w = \emptyset$, pois $v_k \cap w_k = \emptyset$.

Como x e y eram arbitrários, X é separado.

Corolário 2.1

$\Re^n$ é separado.

2.6. Espaços Métricos

Definição 2.15

Seja X um conjunto não vazio.

(a) Uma aplicação d: XxX $\rightarrow \Re^+$ é uma distância (ou métrica) em X', se satisfaz os seguintes três axiomas:

(i) $\forall$ (x,y) $\in X^2$, (d(x,y) = 0) $\leftrightarrow$ (x = y).

(ii)$\forall$ (x,y) $\in X^2$, d(x,y) = d(y,x).

(iii) $\forall$ (x,y,z) $\in X^3$, d(x,y) $\leq$ d(x,y) + d(y,z).

(b) Se d for uma distância em X, então (X,d) é um espaço métrico.

É chegado o momento de observarmos uma série de itens, por exemplo:

(1) d: $\underset{(x,y)\rightarrow|x-y|}{\Re x \Re \rightarrow \Re^+}$, é uma métrica em $\Re$.

(2) Consideremos as seguintes aplicações:

$$d_1: \underset{(x,y)\rightarrow\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2},\text{com}\begin{cases}x=(x_i)_{1\leq i\leq n}\\ y=(y_i)_{1\leq i\leq n}\end{cases}}{\Re^n x \Re^n \rightarrow \Re_+}$$

$$d_2: \underset{(x,y)\rightarrow\sum_{i=1}^{n}|x_i-y_i|,\text{com}\begin{cases}x=(x_i)_{1\leq i\leq n}\\ y=(y_i)_{1\leq i\leq n}\end{cases}}{\Re^n x \Re^n \rightarrow \Re_+}$$

$$d_3: \underset{(x,y)\rightarrow\max_{1\leq i\leq n}|x_i-y_i|,\text{com}\begin{cases}x=(x_i)_{1\leq i\leq n}\\ y=(y_i)_{1\leq i\leq n}\end{cases}}{\Re^n x \Re^n \rightarrow \Re_+}$$

d_1, d_2 e d_3 são distâncias em $\Re^n$.

Sempre que nada seja dito em contrário, consideraremos definida em $\Re^n$ a distância d.

A única dificuldade que poderá existir para demonstrar que d_1, d_2 e d_3 são distâncias em $\Re^n$, deverá residir em provar que d satisfaz (iii).

Vamos fazer esta demonstração.

Sejam $x = (x_i)_{1 \leq i \leq n}$, $y = (y_i)_{1 \leq i \leq n}$ e $z = (z_i)_{1 \leq i \leq n}$ elementos de $\Re^n$.

Comecemos por provar que se a_k e b_k são elementos de $\Re$ $(1 \leq k \leq n)$, então:

$$\left[\sum_{k=1}^{n}(a_k b_k)\right]^2 = \left[\sum_{k=1}^{n} a_k^2\right].\left[\sum_{k=1}^{n} b_k^2\right] - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2.$$

$$.\left[\sum_{k=1}^{n} a_k^2\right].\left[\sum_{k=1}^{n} b_k^2\right] - \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n}(a_i b_j - a_j b_i)^2 =$$

$$= \left[\sum_{k=1}^{n} a_k^2\right].\left[\sum_{k=1}^{n} b_k^2\right] - \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i^2 b_j^2 - 2a_i a_j b_i b_j + a_j^2 b_i^2)\right] =$$

$$= \left[\sum_{k=1}^{n} a_k^2\right].\left[\sum_{k=1}^{n} b_k^2\right] - \frac{1}{2}\left[\left(\sum_{i=1}^{n} a_i^2\right).\left(\sum_{j=1}^{n} b_j^2\right) - 2\left(\sum_{i=1}^{n} a_i b_i\right).\right.$$

$$\left. = \left(\sum_{j=1}^{n} a_j b_j\right) + \left(\sum_{j=1}^{n} a_j^2\right).\left(\sum_{i=1}^{n} b_i^2\right)\right] = \left[\sum_{i=1}^{n} a_i b_i\right].\left[\sum_{j=1}^{n} a_j b_j\right] = \left[\sum_{k=1}^{n}(a_k b_k)\right]^2$$

Obtivemos, assim uma igualdade que nos permite concluir a desigualdade de Cauchy-Buniakovski, que é da forma:

$$\left[\sum_{k=1}^{n} a_k b_k\right]^2 \leq \left[\sum_{k=1}^{n} a_k^2\right].\left[\sum_{k=1}^{n} b_k^2\right]$$

Consideremos $a_k = y_k - x_k$ e $b_k = z_k - y_k$.

Então:

$$[d_1(x,z)]^2 = \sum_{k=1}^{n}(z_k - x_k)^2 = \sum_{k=1}^{n}(a_k + b_k)^2 = \sum_{k=1}^{n} a_k^2 + 2\sum_{k=1}^{n}(a_k b_k) +$$

$$+\sum_{k=1}^{n} b_k^2 \leq \sum_{k=1}^{n} a_k^2 + 2\sqrt{\sum_{k=1}^{n} a_k^2}\sqrt{\sum_{k=1}^{n} b_k^2} + \sum_{k=1}^{n} b_k^2 =$$

$$= \left[\sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} a_k^2}\right]^2$$

Logo:

$$d(x,z) \leq \sqrt{\sum_{k=1}^{n}(y_k - x_k)^2} + \sqrt{\sum_{k=1}^{n}(z_k - y_k)^2} = d(x,y) + d(y,z)$$

(3) Seja X um espaço vetorial sobre K ($\Re$ ou C).

Uma aplicação $\| \,.\, \|: X \to \Re^+$ é uma <u>norma</u> em X, se satisfaz os seguintes axiomas:

(i) $\forall\, x \in X, \forall\, \lambda \in K, \| \lambda x \| = | \lambda | \, \| x \|$.

(ii) $\forall\, (x,y) \in X^2, \| x+y \| \leq \| x \| + \| y \|$.

(iii) $\| x \| = 0$ se e somente se $x = 0$.

É fácil verificar que a aplicação d: XxX $\to \Re^+$ é uma distância em X.

Designa-la-emos por <u>distância associada</u> à norma ou <u>distância induzida pela norma</u>.

<u>Definição 2.16</u>

Seja (X,d) um espaço métrico e A um subconjunto não vazio de X.

(a) o <u>diâmetro</u> $\delta(A)$ de A, é definido $\delta(A) = \sup.\{d(x,y)/(x,y) \in A^2\}$.

(b) Se $\delta(A) < +\infty$, então A é dito limitado.

(c) Se $\delta(A) = +\infty$, então A é dito não-limitado.

<u>Observação:</u> Se $A = \emptyset$, convenciona-se $\delta(A) = -\infty$.

<u>Definição 2.17</u>

Sejam (X,d) um espaço métrico, A e B subconjuntos não vazios de X.

Designemos por <u>distância</u> de A e B (que notaremos d(A,B)), o seguinte valor $d(A,B)= \inf_{(x,y)AxB} d(x,y)$.

<u>Definição 2.18</u>

Seja (X,d) um espaço métrico:

(a) <u>Bola aberta</u> de centro x e raio r, com $x \in X$ e $r \geq 0$, é o conjunto $D(X,r) = \{y \in X : d(x,y) < r\}$.

(b) <u>Bola fechada</u> de centro x e raio r, com x E X e $r \geq 0$, é o conjunto $D(X,r) = \{y \in X : d(x,y) \leq r\}$.

(c) <u>Esfera</u> de centro x e raio r, $x \in X$ e $r \geq 0$, é o conjunto

$S(x,r) = \{y \in X : d(x,y) = r\}$.

<u>Teorema 2.11</u>

Sejam (X,d) um espaço métrico e A um subconjunto não vazio de X. Então A é limitado e e somente se estiver contido em alguma bola aberta de X.

<u>Demonstração</u>

Consideremos A limitado.

Então $A \subset D'(x,\delta(A))$, com x um qualquer elemento de A. Seja agora A um subconjunto de $D'(y,r/2)$.

Logo para x_1 e x_2 arbitrários em $A(x_1,x_2) \leq d(x_1,y)+d(y,x_2) \leq r$.

Podemos afirmar que A é imitado, pois $\delta(A) \leq r$.

<u>Teorema 2.12</u>

Seja (X,d) um espaço métrico. Consideremos τ_d a família constituída por $\emptyset$ e pelos subconjuntos A de X tais que, para todo elemento x de A, existe r > 0 satisfazendo $D(x,r) \subset A$.

Então τ_d é topologia em X.

<u>Demonstração</u>

É evidente que $\emptyset$ e X são elementos de τ_d. Sejam A e B elementos de τ_d. Vamos provar que $A \cap B$ ainda é elemento de τ_d. A demonstração é imediata de $A \cap B = \emptyset$. Seja x um qualquer elemento de $A \cap B$. Sabemos que existem r_1 e r_2 reais positivos, tais que $B(x_1,r_1) \subset A$ e $B(x_1,r_2) \subset B$.

Portanto $(A \cap B) \in \tau_d$, pois, para $r=\min(r_1,r_2)$, $B(x,r) \subset (A \cap B)$. Consideremos um conjunto I de índices de $(A_i)_{i \in I}$ uma família de elementos de τ_d.

Se x for um elemento de $D=\bigcup_{i \in I} A_i$, existem reais positivos $r_i (i \in I)$ tais que $B(x,r_i) \subset A_i \subset D$. Logo $D \in \tau_d$.

<u>Definição 2.19</u>

Consideremos (X,d) e τ_d nas condições do teorema anterior. τ_d designa-se por <u>topologia associada</u> à estrutura métrica de X e (X,τ_d) é o <u>espaço topológico associado</u> ao espaço métrico (X,d).

<u>Teorema 2.13</u>

Num espaço métrico (X,d), toda bola aberta é um aberto e toda a bola fechada é um fechado.

<u>Demonstração</u>

Seja B(x,r) uma bola aberta e y um seu elemento. Seja

$\varepsilon_1 = (r-d(x,y)) > 0$. Então é fácil provar que $B(y,\varepsilon_1) \subset B(x,r)$.

Teorema 2.14

Todo espaço métrico é separado.

Não demonstraremos o teorema.

Definição 2.20

Duas distâncias d' e d'' são equivalentes em X, se existirem $m > 0$ e $M > 0$ tais que para todo o x e y em X

$$md'(x,y) \leq d'(x,y) \leq Md''(x,y).$$

Teorema 2.15

Duas métricas equivalentes em X, tem a mesma topologia associada.

Não iremos fazer a demonstração deste teorema. Ela se baseia em provar que toda bola aberta D(x,y) para a topologia associada a cada uma das métricas contém uma bola aberta D(x,r') para a topologia associada à outra métrica.

2.7. Espaços Compactos

Definição 2.21

Sejam X um conjunto não vazio e $\{X_i\}_{i\in I}$ uma família de subconjuntos de X.

(a) Se $\bigcup_{i\in I} X_i = X$, então $\{X_i\}_{i\in I}$ é uma cobertura de X. Se I é finito a cobertura diz-se finita.

Caso todos os subconjunto X_i sejam abertos, a cobertura designa-se por aberta.

(b) Se $\{X_i\}_{i\in I}$ é cobertura e se existe $J \subset I$ tal que J finito e $\bigcup_{i\in J} X_i = X$, então $\{X_i\}_{i\in J}$ é uma subcobertura.

Definição 2.22

Seja X um espaço topológico. Se toda a cobertura aberta de X, é possível extrair uma cobertura finita, então X é um espaço compacto.

Teorema 2.16

Seja X um espço topológico. X é compacto se e somente se toda a família $\{F_i\}_{i\in I}$ de fechados de X, tal que a intersecção de qualquer sua subfamília é não vazia satisfaz $\bigcup_{i\in I} F_i \neq \emptyset$.

Demonstração

Suponhamos que x é compacto e toda a subfamília finita de $\{F_i\}_{i\in I}$ tem intersecção não vazia. Se $\bigcup_{i\in I} F_i \neq \emptyset$ então $\bigcup_{i\in I} (X\backslash F_i) = X\backslash(\bigcup_{i\in I} F_i) = X$.

Logo $\{X\backslash F_i\}_{i\in J}$, isto é, $\bigcup_{i\in J} (X\backslash F_i) = X$, com J finito.

Isto levar-nos-ia a afirmar que $\bigcup_{i\in J} F_i = \emptyset$, com J finito, o que contradiz a hipótese formulada. Então $\bigcup_{i\in I} F_i \neq \emptyset$.

Suponhamos agora que, se toda subfamília finita de uma família $\{F_i\}_{i\in I}$ de fechados tem intersecção vazia, então $\bigcup_{i\in I} F_i \neq \emptyset$.

Seja $\{A_s\}_{s\in S}$ uma cobertura aberta de X, escolhida arbitrariamente.

Suponhamos que $\{A_s\}_{s\in S}$ não admite qualquer

subcobertura finita de X.

Logo, para todo o subconjunto finito de S, $\{A_s\}_{t\in T}$ é tal que

$\bigcup_{t\in T} A_t \neq X$.

Então $\bigcap_{t\in T} (X\backslash A_t) \neq \emptyset$.

Corolário 2.2

Num espaço compacto X, toda a parte infinita A admite um ponto de acumulação.

Demonstração

Suponhamos que A não possui qualquer ponto de acumulação. Como A é infinita, A contém um subconjunto numerável $F = \{f_1, f_2, ..., f_n, ...\}$, com $f_i \neq f_j$, se $i \neq j$. F não tem pontos de acumulação, logo $F' = \emptyset$.

Sabemos que $\overline{F} = F \cup F'$. Então $\overline{F} = F$, isto é, F é fechado. Qualquer subconjunto G de F é fechado, pois $G' \subset F' = \emptyset$. Para todo número natural n, consideremos $F_n = \{f_n, f_{n+1}, ...\}$. Obtemos assim uma família $\{F_n\}_{n \in N}$ de fechados tal que a intersecção de umaqualquer subfamília finita é não vazia.

Pelo teorema anterior, a intersecção dos elementos da família não vazia, o que constItui um absurdo. Logo A possui um ponto de acumulação.

Teorema 2.17

Seja X um espaço compacto. Em X, todo fechado é compacto.

Demonstração

Em X, consideremos um fechado A. seja $\{A_i\}_{i \in I}$ uma cobertura aberta de A, $\{A_i\}_{i \in I}$ é uma cobertura aberta de X. Logo admite uma subcobertura da forma

$\{A_i\}_{i \in J} \cup \{X \backslash A\}$ com subconjunto finito de T.

Então $\{A_i\}_{i \in I}$ admite a subcobertura finita $\{A_i\}_{i \in J}$ de A. Como a cobertura $\{A_i\}_{i \in J}$ foi arbitrariamente escolhida, A é compacto.

Teorema 2.18

Seja X um espaço topológico separado. Em X, todo o compacto é fechado.

Demonstração

Seja A um compacto em X. Seja y um elemento arbitrário de X\A. Como X e separado, podemos afirmar que para todo o elemento x de A, existem vizinhanças, v_x e w_x respectivamente de x e y, cuja intersecção é vazia.

Vamos sem perda de generalidade, supor v_x e w_x abertos. $\{v_x\}_{x\in A}$ é uma cobertura aberta de A admitindo, visto A ser compacto, uma subcobertura finita $\{v_{xj}\}_{1\leq i\leq n}$.

$W=\bigcap_{j=1}^{n} W_{x_j}$ é uma vizinhança de y, cuja intersecção com $v_{xj}(1\leq j\leq n)$ é vazia.

Logo, $w\cap A\subset w\cap\left(\bigcup_{j=1}^{n} W_{x_j}\right)=\bigcup_{j=1}^{n}\left(W\cap V_{x_j}\right)=\emptyset$.

Então X\A é vizinhança de y, pois $w\subset X\backslash A$.

Como y era arbitrário em X\A, podemos afirmar que X\A é aberto, o que equivale a dizer que A é fechado.

<u>Corolário 2.3</u>

Num espaço compacto e separado, uma parte A é compacta se e somente se é fechada.

<u>Demonstração</u>

Não será apresentada a demonstração.

<u>Teorema 2.19 (Teorema de Tychonoff)</u>

Sejam X e Y espaços topológicos. Então XxY é compacto se e somente se X e Y são compactos.

<u>Demonstração</u>

A demonstração é bastante extensa, razão pela qual a não apresentamos. O leitor mais interessado poderá consultar o livro de Elon Lages Lima, que está indicado na bibliografia.

Corolário 2.4

Se $X_1, X_2, \ldots, X_n$ forem espaços topológicos, então $X_1 x X_2 x \ldots x X_n$ é compacto se e somente se $X_1, X_2, \ldots, X_n$ forem compactos.

Demonstração

Por indução matemática do corolário III.

Teorema 2.20 (Teorema de Heine-Borel-Lebesgue)

Em $\Re$, todo intervalo fechado e limitado [a,b] é compacto.

Demonstração

Se a = b, então [a,b] = {a} é evidentemente compacto.

Se A<b, seja $\{\vartheta_i\}_{i \in I}$ uma cobertura aberta de [a,b]. Designemos por E o conjunto dos elementos x de [a,b], tais que existe uma subcobertura finita $\{\vartheta_i\}_{i \in J}$ de [a,x]. É evidente que E não vazio pois a é elemento de E. b é majorante de E, logo E é um subconjunto não vazio e majorado de $\Re$. Por um dos axiomas da construção dos reais, existe, em [a,b] um supremo c de E.

Vamos provar que c é elemento de E e c = b.

(i) Como $c \in [a,b]$, existe k em I tal que c é elemento de A. A_k é aberto logo $A_k \in \vartheta(c)$, existindo então $\varepsilon > 0$ tal que $(c-2\varepsilon, c+2\varepsilon) \subset A_k$.

Portanto $[c-\varepsilon, c+\varepsilon] \subset A_k$. Por definição de supremo, existe pelo menos um elemento y na intersecção de $[c-\xi, c]$ e E.

Como $y \in E$, existe J_y c I, J_y finito, tal que $\bigcup_{i \in J_y} A_i \supset [a,b]$. Logo $[a,c] \subset$ $(\bigcup_{i \in J_y} A_i \cup A_k)$, isto é, c é elemento de E.

(ii) Se $c \neq b$, então $(\bigcup_{i \in J_y} A_i \cup A_k) \supset [a,d]$, com $d = \inf\{b, c+\varepsilon\} > c$, o que permite dizer que d pertence a E. Tal é um absurdo, atendendo ao fato de c ser supremo de E. Logo $c = b$.

Provamos, assim, que $\{A_i\}_{i \in J}$ admite uma subcobertura finita de [a,b], logo [a,b] é compacto.

Teorema 2.21 (Teorema de Bolzano-Weistrass)

Todo conjunto infinito M, contido num intervalo fechado e limitado [a,b] de $\mathfrak{R}$, tem um ponto de acumulação.

Demonstração

Basta aplicar o Corolário 2.2 e reparar que [a,b] é compacto.

Observação: Um subconjunto A de $\mathfrak{R}$ é limitado se e somente se estiver contido em alguma bola de $\mathfrak{R}$. Isto equivale a dizer que A é limitado se e somente se estiver contido em algum intervalo fechado e limitado [a,b].

Do mesmo modo $A \subset \mathfrak{R}^n$ é limitado se e somente se estiver contido em alguma bola do $\mathfrak{R}^n$. Se a distância considerada em $\mathfrak{R}^n$ for d_3, as bolas serão do tipo:

$$B((x_1,x_2,\ldots,x_n),r)=\{(y_1,y_2,\ldots,y_n)\in\mathfrak{R}^n: \max_{1\le i\le n}|y_i-x_i|<r\} =$$

$$= (x_1-r, x_1+r)x(x_2-r,x_2+r)x\ldots x(x_n-r,x_n+r).$$

Então podemos dizer que $A \subset \mathfrak{R}^n$ é limitado se e somente se A é fechado e limitado.

Teorema 2.22

Um subconjunto A de $\mathfrak{R}^n$ é compacto se e somente se A é fechado e limitado.

Demonstração

Seja A compacto. $\mathfrak{R}^n$ é separado. Então atendendo ao Teorema 1.17, A é fechado. Falta provar que A é limitado.

Consideremos $w_{pk} = (-p-1,p+1)$, $k = 1, 2, \ldots, n$. É evidente que $\left\{\prod_{k=1}^{n} W_{p_j k}\right\}$ constitui uma cobertura aberta de A.

Como A é compacto, existe uma subcobertura finita $\left\{\prod_{k=1}^{n} W_{p_j k}\right\}_{1\leq j\leq m}$.

Seja M = $\max_{1\leq j\leq m}(p_j+1)$. Então $\prod_{k=1}^{n}[-M_k,M_k]$, com $M_1=M_2=...=M_n=M$, contém A, o que nos permite pela observação anterior afirmar que A é limitado.

Assim, A está contido em algum conjunto P do tipo $\prod_{i=1}^{n}[a_i,b_i]$. Pelo Teorema 2.20, cada intervalo $[a_i,b_i]$ é compacto. Então pelo Corolário 2.4, P é compacto, o que prova atendendo ao Teorema 2.17, a compacidade de A.

Teorema 2.23

Um subconjunto A de $\Re$ é compacto se e somente se A é fechado e limitado.

Corolário 2.5 (Teorema de Baire)

Toda a sucessão de intervalos fechados e limitados $([a_n,b_n])_{n\in N}$ satisfazendo, para todo $n \in N$, $[a_n,b_n] \supset [a_{n+1}, b_{n+1}]$, tem intersecção não vazia.

Demonstração

Basta aplicar o Teorema 2.23 a cada um dos intervalos e, em seguida usar o Teorema 2.16.

2.8. Espaços Conexos

Teorema 2.24

Seja X um espaço topológico. São equivalentes as três afirmações:

(i) Não existem dois abertos de X, não vazios e disjuntos, tais que sua reunião seja igual a X.

(ii) Não existem dois fechados de X, não vazios e disjuntos, cuja reunião seja X.

(iii) Os únicos conjuntos simultaneamente fechados e abertos em X são $\emptyset$ e X.

Demonstração

(a) Vamos começar a provar (i) $\Rightarrow$ (ii)

Consideremos verdadeiro (i).

Se existirem dois fechados F_1 e F_2 em X, não vazios disjuntos e tais que $F_1 \cup F_2 = X$, então $X = X \setminus (F_1 \cap F_2) = (X \setminus F_1) \cup (X \setminus F_2)$, com $X \setminus F_1$ e $X \setminus F_2$ abertos, não vazios e disjuntos, o que contraria (i)

(b) (ii) $\Rightarrow$ (i)

Demonstração análoga à anterior.

(c) (ii) $\Rightarrow$ (iii)

Consideremos verdadeiro (ii).

Se existisse um subconjunto A de X, que verificasse simultaneamente:

(c_1) $A \neq \emptyset$

(c_2) $A \neq X$

(c_3) A fechado

(c_4) A aberto

então A e $X \setminus A$ seriam dois fechados, não vazios, disjuntos e de reunião igual a X, o que contraria a veracidade de (ii).

Logo, (iii) é verdadeiro.

(d) (iii) $\Rightarrow$ (ii)

Consideremos (iii) verdadeiro.

Se existissem fechados F_1 e F_2, disjuntos, não vazios e tais que $F_1 \cup F_2 = X$, então F_1 seria um aberto, pois $F_2 = X \setminus F_1$ é fechado. Como F_1 não é vazio nem é todo o espaço X, chegaríamos a uma contradição com o que é afirmado em (iii).

Logo, (ii) é verdadeiro.

Mostramos assim, a equivalência de (i), (ii) e (iii).

Definição 2.23

Um espaço topológico X, no qual seja verdadeiro um dos axiomas do teorema anterior, é designado por conexo.

Definição 2.24

Seja A uma parte não vazia de um espaço topológico X. A é uma parte conexa ou, simplesmente, um conexo de X, se não existirem dois abertos U e V de X verificando simultaneamente:

(1) $U \cap A \neq \emptyset$ e $V \cap A \neq \emptyset$

(2) $(U \cap V) \cap A = \emptyset$

(3) $A \subset U \cup V$.

Teorema 2.25

Seja X um espaço topológico e A um subconjunto não vazio de X. A é conexo se e somente se, para quaisquer abertos U e V de X satisfazendo em simultâneo:

(a) $(U \cap V) \cap A = \emptyset$

(b) $U \cup V \supset A$

(c) $A \cap U \neq \emptyset$

se pode concluir que A é subconjunto de U.

Demonstração

Suponhamos que A é conexo e que existem abertos U e V de X, satisfazendo (a), (b) e (c).

Se A e V tivessem intersecção não vazia, então A não seria conexo.

Logo, $V \cap A = \emptyset$.

Por (b), podemos observar que A é um subconjunto de U. Para provarmos a recíproca condição, vamos supor que quaisquer abertos U e V de X, satisfazendo (a), (b), e (c), são tais que A é subconjunto de U.

Se A não for conexo, existem abertos U_1 e V_1 de X, tais que:

(1') $U_1 \cap A \neq \emptyset$ e $V_1 \cap A \neq \emptyset$

(2') $(U_1 \cap V_1) \cap A = \emptyset$

(3') $A \subset U_1 \cup V_1$.

Mas se (1'), (2'), e (3') são verdadeiras, então, por um lado $A \subset U_1$, por outro $A \subset V_1$.

Como A é não vazio e $(U_1 \cap V_1) \cap A = A$, obtemos uma contradição em (2').

Logo A é conexo.

Teorema 2.26

Seja X um espaço topológico e $(A_i)_{i \in I}$ uma família de conexos de X. Se $\bigcap_{i \in I} A_i \neq \emptyset$, então $A = \bigcup_{i \in I} A_i$ é conexo.

Demonstração

Seja $(A_i)_{i \in I}$, uma família de conexos de X, verificando $\bigcap_{i \in I} A_i \neq \emptyset$. Sejam U e V abertos de X, satisfazendo as condições (a), (b), e (c) do Teorema 1.25. Por (c), existe j $\in$ I tal que A_j e U tem intersecção não vazia. Como além disso:

(a') $(U \cup V) \cap A_j = \emptyset$

(b') $A_j \subset U \cup V$

(c') A_j é conexo

então pelo Teorema 2.25, A_j é subconjunto de U.

Assim o conjunto $\bigcap_{i \in I} A_i$ que, por hipótese, é não vazio, está contido em U.

Logo, para todo o elemento i de I:

(a'') $(U \cup V) \cap A_i = \emptyset$

(b'') $A_i \subset U \cup V$

(c'') $A_i \cap U \neq \emptyset$

(d'') A_i é conexo

Então, ainda pelo Teorema 2.25, podemos afirmar que todo conexo A_i é subconjunto de U.

Portanto A está contido em U. De novo pelo Teorema 2.25 concluímos que A é conexo.

Teorema 2.27

Seja A um subconjunto não vazio de $\Re$. Se A é conexo, então é um intervalo de $\Re$.

Demonstração

Seja A um conexo de $\Re$. Suponhamos que A não é um intervalo. Então há dois elementos a e b, em A, distintos e tais que existe um elemento c de (a,b), que não pertence a A.

Nesse caso, os abertos $(-\infty,c)$ e $(c, +\infty)$ de $\Re$, satisfazem simultaneamente:

(1) $(-\infty,c) \cap A \neq \emptyset$ e $(c,+\infty) \cap A \neq \emptyset$

(2) $((-\infty,c) \cap (c,+\infty)) \cap A \neq \emptyset$

(3) $A \subset (-\infty,c) \cup (c,+\infty)$.

Logo , A é não conexo, o que contraria a hipótese. Então A terá que ser um intervalo em $\Re$.

2.9. Sucessões

Definição 2.25

(a) Uma sucessão num conjunto x é uma aplicação

$$x: \underset{n \to x(n) = x_n}{N \to X}$$

Notaremos a sucessão por $x = (x_n)_{n \in N}$ e designaremos x_n por termo geral ou termo de ordem n de $(x_n)_{n \in N}$.

(b) Uma subsucessão de $(x_n)_{n \in N}$ é uma aplicação $x \circ \varphi$: $N \to x$, com φ: $N \to N$ uma aplicação estritamente crescente.

Notaremos a subsucessão por $(x_{\varphi(n)})_{n \in N}$.

Definição 2.26

Seja $x(n)_{n \in N}$ uma sucessão num espaço topológico X. Diz-se que $(x_n)_{n \in N}$ converge para um elemento a de X ou que $(x_n)_{n \in N}$ admite o ponto a por limite quando n tende para $+\infty$, se: $\forall\, v \in \vartheta(a)\ \exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow x_n \in v$.

Notaremos $a = \lim\limits_{n \to \infty} x_n$.

Se existir a tal que $\lim\limits_{n \to \infty} x_n = a$, então $(x_n)_{n \in N}$ é convergente, caso contrário diz-se divergente.

Observação: O bom entendimento desta colocação, resultará na boa compreensão do algoritmo proposto pelo nosso trabalho.

Se (X,d) é um espaço métrico, então $a = \lim\limits_{n \to \infty} x_n$ se e somente se $\forall\, \varepsilon > 0$, $\exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \geq x_n \in B(a,\varepsilon)$. (1)

É evidente que a condição (1) equivale a $\forall\, \varepsilon > 0$, $\exists\, n_0 \in N$: $\forall\, n \in N$, $n \geq n_0 \geq d(x_n,a) < \varepsilon$.

Em particular se $X = \Re$, então $a = \lim\limits_{x \to \infty} x_n$ se e somente se $\forall\, E > 0$, $\exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow | x_n - a | < \varepsilon$ (2).

A condição (2) ainda é equivalente a $\forall\ \varepsilon > 0,\ \exists\ n_0 \in N : \forall\ n \in N,$ $n \geq n_0 \Rightarrow x_n \in (a-\varepsilon,\ a+\varepsilon)$.

Em (1) poderíamos ter usado bolas fechadas.

Teorema 2.28

Se X for um espaço separado e $(x_n)_{n\in N}$ uma sucessão em X, então, caso exista, o limite da sucessão é único.

Demonstração

Seja $(x_n)_{n\in N}$ uma sucessão convergente em X. Suponhamos que existem af e bt distintos, tais que $\lim\limits_{n\to\infty} x_n = \alpha$ e $\lim\limits_{n\to\infty} x_n = \beta$.

Como X é separado existem vizinhanças v e w, respectivamente de α e β, com intersecção vazia.

Por definição de limite,

(a) $\exists\ n_0 \in N : \forall\ n \in N,\ n \geq n_0 \Rightarrow x_n \in v$

e

(b) $\exists\ n_1 \in N : \forall\ n \in N,\ n \geq n_1 \Rightarrow x_n \in w$.

Mas (a) e (b) contradizem o fato de v e w terem intersecção não vazia, pois x_n é elemento de $v \cap w$ se n é igual ou maior que o máximo entre n_0 e n_1.

Esta contradição adveio de supormos que α e β eram diferentes, logo o limite da sucessão é único.

Teorema 2.29

Seja $(x_n)_{n\in N}$ uma sucessão de espaço métrico euclidiano. Então $\lim\limits_{n\to\infty} x_n =$ a se e somente se $\lim\limits_{n\to\infty} d(x_n, a) = 0$.

Demonstração

Basta reparar que são equivalentes as quatro condições:

(1) $\lim\limits_{n\to\infty} x_n = a$

(2) $\forall\, \varepsilon > 0, \exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow d(x_n,a) < \varepsilon$

(3) $\forall\, \varepsilon > 0, \exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow |d(x_n,a) - 0| < \varepsilon$

(4) $\lim\limits_{n\to\infty} d(x_n,a) = 0.$

Teorema 2.30

Sejam (X,d) um espaço métrico e A um subconjunto de X. Então a é elemento de $\overline{A}$ se e somente se existe uma sucessão $(x_n)_{n\in N}$, com elemento em A e $\lim\limits_{x\to\infty} x_n = a$.

Demonstração

Se $\lim\limits_{n\to\infty} x_n = a$ e x_n é elemento de A, então $\forall\, v \in \vartheta(a)$,

$\exists n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow x_n \in v.$

Como toda vizinhança de a tem intersecção não vazia com a, a é elemento de $\overline{A}$.

Vamos provar a outra condição. Suponhamos que a é elemento de $\overline{A}$.

Para qualquer valor inteiro n, B(a,1/n) é vizinhança de A. Logo, $B(a,1/n) \cap A \neq \emptyset$.

Em cada bola B(a,1/n), podemos escolher um elemento x_n de A, com $\lim\limits_{n\to\infty}(1/n) = 0$, então:

$\forall\, \varepsilon > 0, \exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow d(x_n,a) < 1/n < \varepsilon.$

Corolário 2.6

A é fechado se e somente se contém todos os limites das suas sucessões convergentes.

Demonstração

Pelo Teorema 2.3, A é fechado se e somente se $\overline{A} = A$. Concluímos a demonstração aplicando o Teorema 2.30.

Definição 2.27

Num espaço métrico (X,d), uma sucessão $(x_n)_{n\in N}$ é limitada se

$\{x_n, n \in N\}$ é limitado.

Teorema 2.31

Num espaço métrico, qualquer sucessão convergente é limitada.

Demonstração

Seja (X,d') um espaço métrico. Em X consideremos uma sucessão $(x_n)_{n\in N}$ satisfazendo $\lim_n x_n = a$.

Então:

$\exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow d(x_n,a) < 1.$

Seja $r = \max.\{d(x_1,a), d(x_2,a), ..., d(x_{n-1},a), 1\}$. Logo, $\{x_n, n \in N\} \subset B'(a,r)$. Aplicando o Teorema 1.11, concluímos que $\{x_n, n \in N\}$ é limitado, isto é, a sucessão $(x_n)_{n\in N}$ é limitada.

Teorema 2.32

Sejam $\{x_i\}_{i\in N}$ uma família finita de espaços topológicos e $X=\prod_{i=1}^{n} X_i$ o espaço produto. Uma sucessão $(x_n)_{n\in N}$ tende, em X para $a=(a_i)_{1\leq i\leq n}$ se e somente $(X(i)_n)_{n\in N}$ tende, em X_i, para a_i, com $1\leq i\leq n$ e $x(i)_n$ a i-ésima coordenada de x_n.

Demonstração

Seja $(x_n)_{n\in N}$ uma sucessão convergente, em X, para $a=(a_i)_{1\leq i\leq n}$. Para k entre 1 e n, seja $\vartheta_{Xk}(a_k)$ a família das vizinhanças em X_k, de a_k.

Consideremos v_k um aberto pertencente a $\vartheta_{Xk}(a_k)$.

Seja $P = X_1 \times X_2 \times \ldots \times X_{k-1} \times v_k \times X_{k+1} \times \ldots \times X_n$.

Se $\vartheta_X(a)$ for a família das vizinhanças, em X, de a, é evidente que

$P \in \vartheta_X(a)$.

Logo, $\exists\, n_0 \in N : \forall\, n \in N, n \geq n_0 \Rightarrow x_n \in P$.

Podemos, portanto, afirmar que $x(k)_n$ é elemento de v_k, o que nos permite concluir (porquê?) que $\lim_{n\to\infty} x(i)_n=a_i$.

Para i entre 1 e n, vamos supor que $\lim_{n\to\infty} x(i)_n=a_i$.

Consideremos um elemento w em $\vartheta_X(a)$, com $a = (a_i)_{1\leq i\leq n}$. Por definição de vizinhança existem abertos U_i de X_i, tais que $U=\prod_{i=1}^{n} U_i$ satisfaz $a \in P \subset w$.

Como $\lim_n x(i)_n=a_i$, então:

$\exists\, n(i)_0 \in N : \forall\, n \in N, n \geq n(i)_0 \Rightarrow x(i)_n \in U_i$.

Para $n_0=\max_{1\leq i\leq n} \{n(i)_0\}$, temos:

$\forall\, n \in N, n \geq n_0 \Rightarrow x_n \in \prod_{i=1}^{n} U_i = P \subset w$.

Por conseguinte, $\lim_n x_n = a$.

Definição 2.28

Sejam (X,d) um espaço métrico e $(x_n)_{nEN}$ uma sucessão em X. $(x_n)_{nEN}$ é uma sucessão de Cauchy se:

$\forall\ \varepsilon>0, \exists\ m\in N: \forall\ (p,q)\in N^2, (p\geq m \wedge q\geq m)\Rightarrow d(x_p,x_q)<\varepsilon$.

Teorema 2.33

Toda sucessão convergente é de uma sucessão de Cauchy

Mesmo sendo esta demonstração importantísima, ela não será feita.

Teorema 2.34

Se é convergente uma subsucessão de uma sucessão de Cauchy é convergente, então a própria sucessão é convergente e tem o mesmo limite que a subsucessão.

Demonstração

Seja $(x_n)_{n\in N}$ uma sucessão de Cauchy. Suponhamos que existe uma subsucessão $(x_{\varphi(n)})_{n\in N}$ tal que $\lim\limits_{n\to\infty} \varphi(n) = a$.

Como $(x_n)_{n\in N}$ é de Cauchy, então:

$\forall\varepsilon_1>0, \exists n_0\in N: \forall(p,q)\in N^2, (p\geq n_0 \wedge q\geq n_0)\Rightarrow d(x_p,x_q)<\varepsilon_1$. Por outro lado, $\lim\limits_{n}\varphi(n)=a$. Logo, $\forall\varepsilon_2>0,\exists n_1\in N: \forall n\in N, n\geq n_1\Rightarrow d(x_{\varphi(n)},a)<\varepsilon_2$.

Fixemos, arbitrariamente, um valor real e positivo E.

Se $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$ e $n_2 = \max(n_0,n_1)$, então, notando que $\varphi(n_2) \geq n_2$, podemos concluir que:

$\forall\ n \in N, n \geq n_2 \Rightarrow d(x_n,x_{\varphi(n2)}) + d(x_{\varphi(n2)}, a) < \varepsilon$.

Por conseguinte, $\lim\limits_{n\to\infty} x_n = a$.

Definição 2.29

Um espaço métrico (X,d) é completo, se toda a sucessão de Cauchy em X, é convergente em X.

Teorema 2.34 (Teorema do Ponto Fixo de Banach)

Sejam (X,d) um espaço métrico completo de $f:X \rightarrow X$ uma contração, isto é, uma aplicação satisfazendo:

$$\exists\, \alpha \in \Re,\ 0 \leq \alpha \leq 1:\ \forall (x,y) \in X^2 \Rightarrow (d(f(x),f(y)) \leq \alpha d(x,y)).$$

Então f admite um único ponto fixo, isto é, um único elemento p verificando a equação $f(x) = x$.

A demonstração deste teorema foi omitida e poderá ser consultada em obras citadas na bibliografia.

2.10. Continuidade

Definição 2.30

Sejam X e Y espaços topológicos e f uma aplicação de X em Y.

Dizemos que f é contínua num ponto x_0 e X_1, se $\forall\, v \in \vartheta(f(x_0))$,

$\exists\, u \in \vartheta(x_0) : f(u) \subset v$ (1).

Teorema 2.36

É equivalente afirmar que f é contínua em x_0 ou que toda vizinhança v de $f(x_0)$ é tal que $f^{-1}(v)$ é vizinhança de x_0.

Demonstração

Se f é contínua, então (1) verifica-se. Logo,

$f^{-1}(v) \supset f^{-1}(f(u)) \supset u$.

Como u é vizinhança de x_0, então $f^{-1}(v)$ também é. Noutro sentido basta considerar $u = f^{-1}(v)$ e notar que $v \supset f(f^{-1}(v)) = f(u)$.

Observações.

(1) Num espaço métrico, se $v \in \vartheta(x)$, então existe $B(x,\delta)$, com $\delta>0$, tal que $B(x,is) \subset v$.

Sabemos ainda que $B(x,\delta)$ é vizinhança de X, logo, $f:(X,d) \to (X',d')$ é contínua em x_0 de X se e somente uma das condições é satisfeita:

(1.1) $\forall\, \varepsilon > 0, \exists\, \delta > 0 : f[B(x_0,\delta)] \subset B(f(x_0,\varepsilon))$.

(1.2) $\forall\, \varepsilon > 0, \exists\, \delta > 0 : x \in B(x_0,\delta) \Rightarrow f(x) \in B(f(x_0),\varepsilon)$.

(1.3) $\forall\, \varepsilon > 0, \exists\, \delta > 0 : d\,(x,x_0) < \delta \Rightarrow d'(f(x),f(x_0)) < \varepsilon$.

(2) Por conseguinte $f:\mathfrak{R} \to \mathfrak{R}$ é contínua em $x_0 \in \mathfrak{R}$ se e somente se:

$\forall\, \varepsilon > 0, \exists\, \delta > 0 : |x-x_0| < is \Rightarrow |f(x)-f(x_0)| < \varepsilon$. (2)

A condição anterior ainda equivale a:

$\forall\, \varepsilon > 0, \exists\, \delta > 0 : x \in (x_0-\delta\, , x+\delta) \Rightarrow f(x) \in (f(x)-\varepsilon\, , f(x_0)+\varepsilon))$.

(3) Como se sabe, dizemos que o limite, quando x tende para x_0,de f(x) é $f(x_0)$ se e somente se (2) se verificar.

Se notarmos esse limite por $\lim\limits_{x \to x_0} f(x_0)$, podemos afirmar que f é contínua em x_0 se e somente se $\lim\limits_{x \to x_0} = f(x_0)$, o que corresponde à habitual noção de continuidade num ponto x_0.

Teorema 2.36

Sejam X, Y e Z espaços topológicos, $f:X \to x$ e $g:Y \to Z$. Se f é contínua em x_0 e g em $f(x_0)$, então $g^\circ f$ é contínua em x_0.

Demonstração

Seja x uma vizinhança de $(g^\circ f)(x_0)$. Como g é contínua em $f(x_0)$, $g^{-1}(w) \in V(f(x_0))$. Mas f é contínua em x_0, logo, $f^{-1}(g^{-1}(w)) \in \vartheta(x_0)$.

Então $(g^\circ f)^{-1}(w) = f^{-1}(g^{-1}(w))$ é vizinhança de x_0, o que nos permite afirmar que $g^\circ f$ é contínua em x_0.

Teorema 2.38

Se X e Y forem dois espaços topológicos e se f for uma aplicação de X em Y, então são equivalentes as seguintes propriedades:

(1) Se A é um aberto de Y, então $f^{-1}(A)$ é aberto em X.

(2) Se F é um fechado, então $f^{-1}(F)$ é fechado em X.

(3) $\forall$ x $\in$ X, $\forall$ v $\in$ V(f(x)) => $f^{-1}(v) \in \vartheta(x)$.

(4) f é contínua em todos os pontos de X.

Demonstração

Pelo Teorema 2.37, temos que (3) e (4) são equivalentes. Se (1) é verdadeira, provemos que (2) também é.

Seja F um fechado de Y. Logo Y\F é aberto em Y. Aplicando (1),

$f^{-1}(Y\backslash F)$ é aberto em X. Logo Y\F é aberto em Y. Mas $f^{-1}(Y\backslash F)=X\backslash f^{-1}(F)$, o que nos permite concluir que $f^{-1}(F)$ é fechado.

De modo semelhante provaríamos que da veracidade de (2), decorre a de (1), verificando assim que (1) e (2) são equivalentes.

Vejamos que (3) se pode provar supondo (1) verdadeira.

Sejam x um elemento arbitrário em X e v uma vizinhança de f(x). Por definição de vizinhança, existe um aberto A de Y, tal que $f(x) \in A \subset \varpi$.

Então $x \in f^{-1}(A) \subset f^{-1}(v)$. Como $f^{-1}(A)$ é aberto, $f^{-1}(v)$ é vizinhança de x. Provemos , finalmente, que (1) é verdadeira se (3) o for.

Seja A um aberto de Y. Se $f^{-1}(A) = \emptyset$, a demonstração termina, pois $\emptyset$ é um aberto de X. Se $f^{-1}(A) \neq \emptyset$, consideremos um qualquer elemento x de $f^{-1}(A)$. Como A é aberto e $f(x) \in A$, então A é vizinhança de f(x). Logo por (3), $f^{-1}(A) \in \vartheta(x)$.

$f^{-1}(A)$ é aberto, pois é vizinhança de qualquer um dos seus elementos. Por conseguinte (3) é equivalente a (1).

Assim provamos a equivalência das quatro proposições.

Definição 2.31

Se f for uma aplicação de um espaço topológico X num outro Y, então f é contínua se satisfazer algumas das quatro propriedades do teorema anterior.

Teorema 2.39

Sejam X, Y e Z três espaços topológicos; sejam $f:X\rightarrow Y$ e $g:Y\rightarrow Z$ aplicações contínuas.

Então $g^\circ f : X\rightarrow Z$ é contínua.

Definição 2.32

Uma aplicação $f:X \to Y$, com X e Y espaços topológicos, diz-se um homeomorfismo se:

(1) biunívoca e

(2) f e f^{-1} são contínuas.

X e Y são homeomorfos se existir um homeomorfismo $f:X \to Y$.

Teorema 2.40

Sejam X e Y espaços topológicos e $f:X \to Y$ contínua. Seja ainda $(x_n)_{n \in N}$ uma sucessão em X tal que l $\lim_{n \to \infty} x_n = a$.

Então $\lim_{n \to \infty} f(x_n) = f(\lim_{n \to \infty} x_n) = a$.

Demonstração

Seja v uma vizinhança de f(a). Como f é contínua em a, existe $u \in \vartheta(a)$ tal que $f(u) \subset v$. Por outro lado, $\lim_{n \to \infty} x_n = a$. Logo existe $n_0 \in N$ tal que, para $n \geq n_0$, x_n é elemento de u.

Observação:

Nem sempre o recíproco do teorema anterior é verdadeiro.

Teorema 2.41

Se (X,d) for um espaço métrico e Y um espaço topológico, então $f:X \to Y$ é contínua se e somente se para toda sucessão $(x_n)_{n \in N}$ convergente, em X, para a, sucessão $(f(x_n))_{n \in N}$ converge para f(a).

Demonstração

No teorema anterior já mostramos uma das condições. Demonstraremos outra.

Suponhamos que, para toda sucessão (xn)n∈N convergente, em X, a sucessão $(f(x_n))_{n \in N}$ tem f(a) por limite.

Pretendemos provar que f é contínua.

Suponhamos que não é, então existe $w \in \vartheta(f(a))$ tal que, para todo o n em N, $f(B(a,1/n)) \not\subset w$.

Seja x_n um elemento de $B(a,1/n)$ tal que $f(x_n)$ não pertence a w.

Então $\lim\limits_{n\to\infty} x_n = a$, pois:

$\forall\, U \in \vartheta(a), \exists\, V_{n0} = B(a,1/n) \subset U : \forall\, n \geq n_0 \Rightarrow x_n \in v_n \subset U.$

Logo $\lim\limits_{n\to\infty}(f(x)) = f(a)$, o que é um absurdo, atendendo ao fato de nenhum $f(x_n)$ pertencer a w.

Portanto f tem que ser contínua.

Teorema 2.42

Sejam X e Y dois espaços topológicos, A um compacto de X e $f{:}X \to Y$ uma aplicação contínua.

Então f(A) é um compacto de Y.

Demonstração

Seja $\{A_i\}_{i \in N}$ uma cobertura aberta de f(A), escolhida arbitrariamente.

Então $f^{-1}(A_i)_{i \in I}$ é uma cobertura aberta de A.Como A é compacto, admite uma subcobertura finita $f^{-1}(A_i)_{i \in J}$.

Por conseguinte,

$\bigcup\limits_{i \in J} A_i \supset f(f^{-1}(\cup_{i \in J} A_i)) = f(\cup_{i \in J}(f^{-1}(A_i))) \supset f(a).$

Logo, $\{A_i\}_{i \in J}$ é uma sucobertura finita de f(a).

Podemos então afirmar que f(a) é compacto.

Definição 2.33

Seja f uma aplicação do conjunto não vazio X no espaço métrico (Y,d). Dizemos que f é limitada se o conjunto:

$\{z \in Y : z = f(x), x \in X\}$

é limitado.

O conjunto atrás referido, pode ser notado por f(x).

Corolário 2.7

Seja X um espaço compacto e $f:X \to \Re$ uma aplicação contínua.

Então f é limitada em X, e atinge, em X, máximo e mínimo.

Pelo Teorema 2.30, f(x) é compacto em $\Re$. Atendendo ao Teorema 2.23, f(x) é limitado.

Então f é limitada.

Seja $\overline{f(X)}$ e $M = \sup_{x \in X} \{f(x)\}$.

Por definição de ínfimo e supremo,

$\forall \varepsilon > 0, \exists x \in X : f(x) \in [m, m+\varepsilon]$

e

$\forall \varepsilon_1 > 0, \exists x_1 \in X : f(x_1) \in [M-\varepsilon, M]$

Logo m e M são elementos do fecho de f(x).

Como f(x) é um compacto no espaço separado $\Re$, então pelo Teorema 2.19, f(x) é fechado, isto é, $f(x) = \overline{f(X)}$.

Por conseguinte, existem a e b, em X, tais que $m = f(a)$ e $M = f(b)$, o que permite afirmar que f atinge, em X, mínimo e máximo (global).

Teorema 2.43

Sejam X um espaço conexo, Y um espaço topológico e $f:X \to Y$ uma aplicação contínua. Então f(x) é conexo em Y.

Demonstração

Suponhamos que f(x) não é conexo. Então existem dois abertos U e V de Y, tais que:

(i) $(U \cap V) \cap f(x) = \emptyset$

(ii) $(U \cap f(x)) \cup (V \cap f(x))) = f(x)$ (isto é, $f(x) \subset U \cup V$)

(iii) $U \cap f(x) \neq \emptyset$ e $V \cap f(x) \neq \emptyset$

Logo os abertos $U_1 = f^{-1}(U)$ e $V_1 = f^{-1}(V)$ verificam:

(i) $(U_1 \cap V_1) \cap X = \emptyset$

(ii) $X = U_1 \cup V_1$

(iii) $U_1 \neq \emptyset$ e $V_1 \neq \emptyset$.

Deste modo concluímos que X não é conexo, o que contraria a hipótese.

Podemos então afirmar que f(x) é um conexo.

Corolário 2.8

Um subconjunto A de $\Re$ é conexo se e somente se A é um intervalo.

Definição 2.34

Um subconjunto A de $\Re^n$, é convexo se, para todo par (a,b) de elementos de A e todo o real α em [0,1], $\alpha a+(1-\alpha)b$ é elemento de A.

Corolário 2.9

Todo subconjunto convexo do $\Re^n$ é conexo.

Não demonstraremos este corolário pois envolveria a necessidade de definir novos termos.

Corolário 2.10

$\Re^n$ é conexo.

Corolário 2.11

Seja f uma aplicação contínua num intervalo I de $\Re$, em $\Re$. Então f(I) é um intervalo de $\Re$.

Corolário 2.12 (Teorema do valor Médio)

Seja f uma função contínua num intervalo I de $\Re$, em $\Re$. Se a e b são dois pontos de I e y pertence ao intervalo [f(a),f(b)], então existe um elemento c de I, verificando f(c) = y.

Corolário 2.13

Toda aplicação contínua no intervalo [a,b] de $\Re$, em $\Re$, é limitada, atinge máximo global e mínimo global, em [a,b] e ainda todos os valores entre o seu máximo global e seu mínimo global.

2.11 Comentário Final

Com o que acabamos de apresentar, finalizamos o Capítulo 2, o qual nos dará subsídios suficientes para trabalhar os próximos capítulos.

3. TEORIA DE OTIMIZAÇÃO

3.1. Introdução

O objetivo deste capítulo é de demonstrar os principais teoremas sobre conjuntos convexos, sobre programação linear e não linear envolvidos na teoria de otimização. Esses teoremas permitirão um tratamento unificado e rigiriso de várias questões, com o critério de otimização na programação linear e o teorema do Kuhn-Tucker que será amplamente trabalhado no capítulo 4.

A notação será dada como na maioria dos compêndios de Álgebra Linear, além de não ser necessário colocar que serão importantes os itens estudados no Capítulo 2, que tratava de topologia.

3.2. Cones Convexos

Definição 3.1

Um conjunto C de vetores do $\Re^n$ é dito um cone, quando $x \in C$ e $\lambda \geq 0$ implicar $\lambda x \in C$. Em outras palavras, C é um cone quando for formado de semi-retas partindo da origem.

De acordo com a noção geral de convexidade, um cone C é dito convexo quando dados dois quaisquer de seus pontos, o segmento que os une está contido no cone. É fácil verificar que um conjunto C é um cone convexo se e somente se satisfizer a seguinte condição: "quaisquer que sejam x e $y \in C$, se $\alpha \geq 0$ e $\beta \geq 0$ forem números reais não negativos, $\alpha x + \beta y \in C$." Em suma, C é um cone convexo se e somente se contiver todas as combinações lineares com coeficientes não negativos de seus elementos.

Teorema 3.1

Seja C um cone convexo. Então seu fecho $\overline{C}$ é também um cone convexo.

Demonstração

Seja x e y dois elementos quaisquer de $\overline{C}$, α e β dois números reais não negativos e $z = \alpha x + \beta y$. Basta provarmos que $z \in C$. Com efeito, seja V uma vizinhança qualquer de z com raio igual a ε. Sejam também x' e y' dois vetores tais que $d(x,x') < \frac{\partial}{2\alpha}$ e $d(y,y') < \frac{\partial}{2\beta}$, e seja $z' = \alpha x' + \beta y'$. É fácil verificar que $z' \in V$, o que significa que $z \in C$.

Teorema 3.2

Seja C um cone convexo do $\Re^n$; $x_1, x_2, ..., x_m, ...$ uma sucessão de valores em C cuja soma seja convergente, isto é, tal que $x_1+x_2+...+x_n+... \rightarrow x$. Então x pertence a C.

Demonstração

Seja p o número máximo de vetores linearmente independentes na sucessão em questão (p é necessariamente menor ou igual a n). Podemos supor sem perda de generalidade que esses são os p primeiros veetores da sucessão. Seja agora: $y_m = x_m + x_{m+1} + ...$. Podemos escrever que $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x! + ... + \lambda_p x_p$, os λ_i tendendo a zero quando m tender para o infinito. Escolhamos m de modo que $|\lambda_i| < 1$, para todo índice i. Então: $x = (1-\lambda_1)x_1 + (1 - \lambda_2)x_2 + ... + (1 - \lambda_p)x_p + x_{p+1} + ... + x_{m-1}$.

O segundo membro representa uma combinação linear, com coeficientes positivos de vetores do cone. Logo $x \in C$.

Teorema 3.3

Seja C um cone convexo cujo fecho coincida com o espaço $\Re^n$. Então $C = \Re^n$.

Demonstração

Seja x um vetor qualquer do $\Re^n$. Por hipótese, será possível tomar duas sucessões de vetores, $x_1, x_2, ..., x_m, ...$ e $y_1, y_2, ..., y_m, ...$, tais que:

$$x = x_1 + y_1$$

$$y_1 = x_2 + y_2$$

$y_2 = x_3 + y_3$

...

sendo $x_1, x_2, x_3, ...$ pertencentes ao cone C e $\|y_m\| < \frac{1}{2^m}$. Segue-se que x $= x_1 + x_2 + x_3 + ...$. Pelo Teorema 3.2, x pertencerá a C, o que prova que o cone em questão é todo o espaço.

Cencentremo-nos agora nos cones convexos fechados. De um mode geral, os cones que forem definidos por desigualdades do tipo $\geq$ ou $\leq$ satisfarão a essa condição. Assim, por exemplo o conjunto de vetores do $\Re^n$ de coordenadas maiores ou iguais a zero, forma um cone convexo fechado. Já o conjunto dos vetores de coordenadas estritamente positivas, constituiria um cone convexo aberto.

Nosso interesse especial em relação aos cones convexos fechados resulta da validade da seguinte proposição:

"Seja C um cone convexo fechado do $\Re^n$, e z um vetor qualquer do $\Re^n$. Então existe um vetor d $\in$ C tal que, para qualquer outro vetor y do cone tenha:

d(z,d) < d(z,y)."

No caso, d é o vetor do cone à mínima distância de z; sua existência resulta simplesmente do fato de o cone em questão ser um conjunto fechado no sentido topológico, como já se viu anteriormente. A existência desse vetor à mínima não poderia ser garantida no caso de um cone não fechado.

A respeito desses vetores à mínima distância vale o seguinte.

Teorema 3.4

Seja C um cone convexo fechado do $\Re^n$; z um vetor qualquer do mesmo $\Re^n$; d o vetor de C à mínima distância de z, u = d-z. Então:

(i) [u,d] = 0

(ii) [u,y] $\geq$ 0, qualquer que seja y $\in$ C.

Demonstração

Em primeiro lugar é interessante visualizar geometricamente o teorema em questão. Isso se apresenta na figura a seguir. É intuitivamente óbvio pelo gráfico 1, que o vetor u = d-z, forma com d um ângulo de 90°, isto é, que [u,d] = 0; e que u forma um ângulo agudo ou reto com qualquer outro vetor do cone, isto é, que [u,v] $\geq$ 0 para qualquer y $\in$ C.

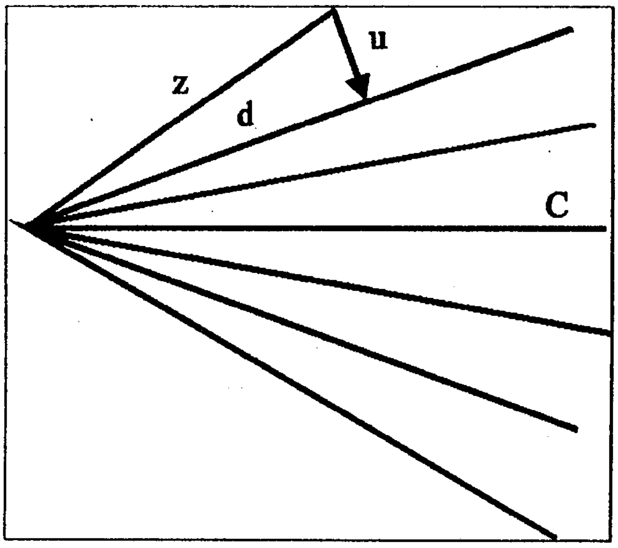

Figura 1

Para demonstrar formalmente o teorema, suponhamos que y seja um vetor qualquer do cone C, e construamos a função de duas variáveis reais α e β definida por:

$$F(\alpha,\beta) = \{d(\alpha d + \beta y, z)\}$$

Como para $\alpha \geq 0$ e $\beta \geq 0$ e $\alpha d + \beta y \in C$, conclui-se que o valor mínimo de $F(\alpha,\beta)$ com a condição $\alpha \geq 0$ e $\beta \geq 0$ deve ocorrer quando $\alpha = 1$ e $\beta = 0$ (já que d é o vetor de C a mínima distância de z). Daí se conclui que para $\alpha = 1$ e $\beta = 0$ deve-se ter:

$$\frac{\partial F}{\partial \alpha} = 0 \quad \text{e} \quad \frac{\partial F}{\partial \beta} \geq 0$$

Note-se a assimetria as condições de mínimo, resultantes de estarmos operando sob a condição de $\alpha \geq 0$ e $\beta \geq 0$. A condição $\frac{\partial F}{\partial \alpha} = 0$ pode ser escrita pelo fato de o mínimo dever ocorrer para $\alpha = 1$, e de portanto se poder variar α num e noutro sentido dentro do campo de definição da função; já para a derivada parcial $\frac{\partial F}{\partial \beta}$ só se pode estabelecer a condição de ser ela ≥ 0 no ponto em questão ($\alpha = 1$; $\beta = 0$), já que, no campo de definição β só pode variar no sentido crescente.

O resto da demonstração segue-se por simples desenvolvimento de cálculo.

$F(\alpha,\beta) = [\alpha d + \beta y\text{-}z\ ,\ \alpha d + \beta y\ \text{-}z]$, ou seja,

$$F(\alpha,\beta) = \|d\|^2\alpha^2 + \|y\|^2\beta^2 + 2\alpha\beta[d,y] - 2\alpha[d,z] - 2\beta[y,z] + \|z\|^2.$$

Logo:

$$\frac{\partial F}{\partial \alpha} = 2\alpha\|d\|^2 + 2\beta[d,y] - 2[d,z]$$

$$\frac{\partial F}{\partial \alpha} = 2\beta\|y\|^2 + 2\,\alpha[d,y] - 2[y,z]$$

No ponto $\alpha = 1$; $\beta = 0$ tem-se pois:

$$\frac{\partial F}{\partial \beta} = 2\|d2\| - 2[d,z] = 2[d,d\text{-}z] = 2[d,u]$$

$$\frac{\partial F}{\partial \beta} = 2[d,y] - 2[y,z] = 2[d\text{-}z,y] = 2[u,y]$$

Entretanto com as condições de que nesse ponto $\frac{\partial F}{\partial \alpha} = 0$ e $\frac{\partial F}{\partial \beta} \geq 0$, resulta:

(i) $[d,u] = 0$

(ii) $[u,y] \geq 0$

Aqui termina a demonstração do teorema, porém a recíproca do teorema verifica-se facilmente. Se C é um cone convexo fechado (de fato à demonstração agora não exige nem que se trate de cone, nem de convexo, nem de fechado), z um vetos qualquer do espaço $\Re^n$, d um vetor de C tal que se $u = d\text{-}z$ se tem:

(i) $[d,u] = 0$

(ii) $[u,y] \geq 0$ para qualquer que seja $y \in C$.

Então d é o vetor de C à mínima distância de z, isto é, qualquer que seja $y \in C$ se tem:

$d(z,d) \leq d(z,y)$.

Para demonstração, observemos preliminarmente que:

$[d(x,d)]^2 = [z\text{-}d,\, z\text{-}d] = \|u\|^2$.

Por outro lado, se $y \in C$, tem-se:

$$[d(z,y)]^2 = [z\text{-}y,z\text{-}y] = [d\text{-}u\text{-}y\,,\, d\text{-}u\text{-}y] =$$

$$= \|u\|^2 + \|d\text{-}y\|^2 - 2[u,d\text{-}y] =$$

$$= \|u\|^2 + \|d\text{-}y\|^2 - 2[u,d] + 2[u,y].$$

Como $[u,d] = 0$ e $[u,y] \geq 0$ segue-se que: $[d(z,y)]^2 =$

$=\|u\|^2+\|d\text{-}y\|^2\geq\|u\|^2=[d(z,d)]^2$.

3.3. Cones Complementares

Definição 3.2

Seja C um conjunto qualquer de vetores do $\Re^n$ (Não necessariamente um cone). Denominamos cone complementar desse conjunto a coleção C' dos vetores $y \in \Re^n$, tais que:

$$[x,y] \geq 0$$

para todo $x \in C$.

Teorema 3.5

C' é um cone, convexo e fechado.

Demonstração

Se $y \in C'$, então para todo $x \in C$ se tem $[x,y] \geq 0$; logo se $\lambda \geq 0$ se tem $[x,\lambda y] \geq 0$, ou seja $\lambda y \in C'$. Isso prova que C' é um cone. Para provar que C' é convexo basta observar que se y_1 e y_2 pertencem a C' então, qualquer que seja $x \in C$, se tem $[x,y_1] \geq 0$ e $[x,y_2] \geq 0$. Logo se $0 \leq \alpha \leq 1$, $[x,(1-\alpha)y_1 + \alpha\, y_2] = (1-\alpha)[x,y_1] + \alpha\, [x,y_2] \geq 0$, o que prova que:

$$(1-\alpha)y_1 + \alpha y_2 \in C'.$$

Para provar que C' é fechado, basta mostrar que seu complemento (isto é, o conjunto dos vetores do $\Re n$ não pertencentes a C') é aberto. Suponhamos que $z \notin C$. Isso significa que existe algum $x \in C$ tal que:

$$[z,x] < 0$$

$$\text{Seja } \delta = -\frac{[z,x]}{2\|x\|}$$

Então se $\|v-z\| \leq \delta$ tem-se:

$$[v,x] = [z,x] + [v\text{-}z,x].[z,x] + \|v\text{-}z\|.\|x\| \leq^{[1]}$$

$$\leq [z,x] + \delta\|x\|.1/2[z,x] < 0.$$

Isso significa que $v \notin C'$, ou seja, que existe uma vizinhança de z contida no complemento de C'. Logo, o complemento de C' é aberto e, portanto, C' é fechado.

Teorema 3.6

Seja C um cone convexo fechado. Então $(C')' = C$.

[1] pela desigualdade de Schwarz

Demonstração

Se $z \in (C')'$ então $z \in C$ já que a inclusão oposta vale genericamente. Seja $z \in (c')'$ e d o vetor de C mais próximo de z. Seja ainda $u = d - z$. Provaremos que $u = 0$, donde resulta que $z = d \in C$.

Com efeito, pelo Teorema 2.4 sabemos que $[u,y] \geq 0$, qualquer que seja $y \in C$. Isso significa que $u \in C'$. Como $z \in (C')'$, segue-se que:

$$[z,u] \geq 0.$$

Porém

$$[z,u] = [d-u,u] = [d,u] - \|u\|^2$$

Pelo Teorema 3.4, $[d,u] = 0$, logo:

$$[z,u] = -\|z\|^2 \leq 0.$$

Como $[z,u] \geq 0$ ($z \in (C')'$ e $u \in C'$) segue se que $u = 0$.

Teorema 3.7 (Teorema de Farkas)

Seja A uma matriz de m linhas por n colunas; C_1 um conjunto de vetores $x \in \Re^n$ tais que $Ax \geq 0$; C_2 o conjunto dos vetores do mesmo espaço da forma A^tp, sendo $p \in \Re^m$ e $p \geq 0$. Então C_1 e C_2 são cones fechados e mutuamente complementares.

Demonstração

A verificação de que C_1 e C_2 são cones convexos e fechados é inteiramente trivial. É fácil também verificar que se $x \in C_1$ e $y \in C_2$ então $[x,y] \geq 0$. Com efeito, $y = A^tp$ sendo $p \geq 0$; logo $[x,y] = [x,A^tp] = [Ax,p]$. Como também $Ax \geq 0$, resulta $[Ax,p] \geq 0$.

Falta apenas demonstrar que todo vetor que tenha produto escalar não negativo com todos os elementos de C_1 pertence a C_2, e vice-versa. Na realidade, em virtude do Teorema 3.4 anterior basta provar a proposição num dos sentidos. Mostraremos que se z é um vetor tal que $[z,x] \geq 0$ para todo $x \in C_1$ então $z \in C_2$.

Com efeito, seja d o vetor de C_2 à mínima distância de z e seja $u = d-z$. Pelo Teorema 3.4 resulta que $u \in C'_2$; isso significa que se $p \geq 0$, sendo $p \in \Re^m$, então:

$$[u,A^tp] \geq 0.$$

A expressão acima equivale a dizer que $[Au,p] \geq 0$, para todo $p \geq 0$ de $\Re^m$. Isso implica em que se tenha $Au \geq o$, ou seja, $u \in C_1$.

É fácil verificar agora que [z,u]=[d-u,u]=[d,u]-[u,u]= = -[u,u], já que [d,u] = 0, pelo Teorema 3.4 Como u $\in$ C, por hipótese [z,u] $\geq$ 0. Tendo em vista a igualdade anteiror, isso só será possível se u = 0 e portanto z = d $\in C_2$.

3.4. Hiperplanos e Teoremas de Separação

Definição 3.3

Sejam x_0 e u dois vetores do espaço $\Re^n$. O hiperplano que passa por x_0 e normal a u $\neq$ 0 é, por definição, o conjunto dos vetores x tais que:

$[x-x_0,u] = 0$

Teorema 3.8

Seja C um cone convexo do $\Re^n$, que não cubra todo o espaço. Então C está inteiramente contido em algum semi-espaço passando pela origem, isto é, existe algum vetor u $\neq$ 0 tal que [u,y] $\geq$ 0 para todo y $\in$ C.

Demonstração

Pelo Teorema 3.3 deve existir algum vetor z do $\Re^n$, que não pertença ao fecho $\overline{C}$ do cone dado. Seja d o vetor de $\overline{C}$ à mínima distância de z. Então, se u = d-z, pelo Teorema 3.4, [u,y] $\geq$ 0 para todo y $\geq$ 0.

Teorema 3.9

Seja S um conjunto convexo de vetores do $\Re^n$, e x_0 um vetor não pertencente a S. Então existe algum hiperplano que passa por x_0 e que deixe S num só dos seus semi-planos.

Demonstração

Seja C o conjunto dos vetores z = $\lambda(x-x_0)$, sendo x $\in$ S e $\lambda \geq 0$. É fácil verificar que se x $\in$ S, então x_0-x não pertence a C, sem o que não valeria a hipótese de que $x_0 \notin$ S. Logo, pelo Teorema 3.8 existirá um vetor u diferente de zero tal que [z,u] $\geq$ 0 para todo z pertencente a C. Isso implica em se ter $[x-x_0,u] \geq 0$ para todo x pertencente a S, o que prova o teorema.

Teorema 3.10

Sejam S_1 e S_2 dois conjuntos do $\Re^n$ tendo em comum um pnonto x_0. Suponhamos ainda que S_2-x_0 (isto é, o conjunto dos pontos exceto x_0) também seja convexo e não vazio. Então existe um hiperplano passando por x_0 e que se separa S_1 e S_2. Mais precisamente, existe um vetor $p \neq 0$ tal que:

$[x-x_0,p] \leq 0$ para todo $x \in S_1$

$[y-x_0,p] \geq 0$ para todo $y \in S_2$

Demonstração

Seja C o conjunto dos vetores da forma $\alpha(x-x_0) - \beta(y-x_0)$, sendo $\alpha \geq 0$, $\beta \geq 0$, $x \in S_1$ e $y \in (S_1-x_0)$. É fácil verificar que C é um cone convexo. Por outro lado, se y_0 é um vetor de S_2-x_0 então y_0-x não pertence a C. Com efeito, suponhamos que se tivesse, nas condições acima:

$$y_0-x_0 = \alpha(x-x_0) - \beta(y-y_0).$$

Transformando concluímos que:

$$\frac{1}{1+\alpha+\beta}[y_0+\alpha x_0+\beta] = \frac{1}{1+\alpha+\beta}[x_0+\alpha x+\beta x_0]$$

é fácil verificar que o primeiro membo representaria um elemento de S_2-x_0 e o segundo membro, um elemento de S_1-x_0 o que contradiz a hipótese de S_1 e S_2-x_0 serem disjuntos. Daí se conclui que o cone C não cobre todo o espaço. Logo, pelo Teorema 3.8, existe um vetor $u \neq 0$ tal que $[\alpha(x-x_0) - \beta(y-x_0)u] \geq 0$ para todo $x \in S_1$, $y \in S_2$-x_0, $\alpha \geq 0$ e $\beta \geq 0$. Tomando $p = -u$, e particularizando para $\alpha = 1$, $\beta = 0$ e depois para $\alpha = 0$, $\beta = 1$, obtem-se imediatamente o teorema.

Teorema 3.11

Sejam S_1 e S_2 doi conjuntos convexos de vetores do $\Re^n$, tendo em comum um único ponto x_0. Suponhamos que S_2-x_0 também seja convexo e não vazio. Seja K o conjunto dos vetores p tais que:

$[x-x_0,p] \leq 0$ para todo $x \in S_1$

$[y-x_0,p] \geq 0$ para todo $y \in S_2$.

Então K é um cone convexo e fechado e que contém algum ponto distinto da origem.

Demonstração

Para a demonstração observemos em primeiro lugar que se p_1 e p_2 pertencem a K então $\alpha p_1 + \beta p_2$ pertencerá a K desde que $\alpha \geq 0$ e $\beta \geq 0$. Isso prova que K é um cone convexo. O Teorema 3.11, por outro lado, nos assegura que K contém algum ponto distinto da origem. resta demonstrar que K é fechado, o que é o mesmo que dizer que seu complemento é aberto. Com efeito, suponhamos que u pertença ao complemento de K. Isso é o mesmo que dizer que:

$$[z,u] > 0$$

(x_0+z) representando um ponto de S_2 ou x_0+z representando um ponto de S_1. É fácil verificar que uma vizinhança de u de raio igual a $\left(\frac{[z,u]}{\|z\|}\right)$ está inteiramente contida no complemento de K. Com efeito, se $\|v,u\| \leq \left(\frac{[z,u]}{\|z\|}\right)$ então, aplicando a desigualdade de Schwarz:

$$[z,v] = [z,u] + [z,v\text{-}u] \geq [z,u] - \|z\| - \|v\text{-}u\| > 0.$$

Isso significa que o complemento de K é aberto, ou seja, que K é fechado.

3.5. Funções Côncavas e Semi-côncavas

Definição 3.4

Seja f(x) uma função real definida num subconjunto convexo C, do $\Re^n$ (isto é, uma função real de n variáveis reais). Diz-se que:

(a) f(x) é côncava no sentido lato quando, quaisquer que sejam x e y pertencentes a C e $a \leq \alpha \leq 1$ tem:

$$f[\alpha x+(1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y);$$

(b) f(x) é côncava no sentido estrito quando, quaisquer que sejam $x \neq y$ pertencentes a C e $0 < \alpha < 1$ se tem:

$$f[\alpha x + (1-\alpha)y] > \alpha f(x) + (1-\alpha)f(y);$$

(c) f(x) é convexa no sentido lato, quando -f(x) for côncava no sentido lato;

(d) f(x) é convexa no sentido estrito quando -f(x) for côncava no sentido lato;

(e) f(x) é semi-côncava quando para todo número real α valer a propriedade: "Se $f(x) \geq \alpha$ e $f(y) > \alpha$ então, para todo $0 < \alpha < 1$:

$$f[\alpha x + (1-\alpha)y] > \alpha";$$

(f) f(x) é semi-cônvexa quando -f(x) for semi-côncava.

As seguintes propriedades se verificam imediatamente,

(i) Toda função côncava no sentido estrito é também côncava no sentido lato;

(ii) Toda função linear $f(x) = [c,x] = c_1x_1 + c_2x_2 + ... + c_nx_n$ é côncava no sentido lato (embora não no sentido estrito);

(iii) Toda função côncava no sentido lato é semi-côncava;

(iv) Se f(x) é semi-côncava, o conjunto dos vetores y tais que f(y) > α, sendo α um número real qualquer é vazio ou convexo;

(v) Se f(x) é semi-côncava, o conjunto dos vetores y tais que f(y) $\geq \alpha$, sendo α um número real qualquer, é vazio ou convexo.

As primeiras quatro propriedades se demonstram pela aplicação direta das definições. A última decorre da quarta, observando-se que:

$\{y/f(y) \geq \alpha\} = \bigcap_{n=1}^{\infty}\left\{y \setminus f(y) > \alpha - \frac{1}{n}\right\}$ e que a intersecção de uma família de conjuntos convexos é também um conjunto convexo.

Do mesmo modo, valem as seguintes propriedades sobre funções convexas e semi-convexas:

(vi) Toda função convexa no sentido estrito é também convexa no sentido lato;

(vii) Toda função linear é convexa no sentido lato (embora não no sentido estrito);

(viii) Toda função convexa no sentido lato é semi-convexa;

(ix) Se f(x) é semi-convexa, o conjunto dos vetores y tais que f(y) < α, sendo α um número real qualquer, é vazio ou convexo.

Teorema 3.12

Seja U(x) uma função contínua e seccioalmente côncava, isto é, tal que U(x) = U(y) implique em $U[\alpha x + (1-\alpha)y] > U(x)$, para todo $0 < \alpha < 1$. Então U(x) é semi-côncava.

Demonstração

Provaremos que para todo $0 < \alpha < 1$ se tem $U[\alpha x + (1-\alpha)y] > \min \{U(x), U(y)\}$. Isso é imediato no caso em que U(x) = U(y), pela definição de concavidade seccional. Suponhamos então que $U(x) \neq U(y)$. Podemos, sem perda de generalidade admitir que U(x) < U(y)

É fácil provar que nesse caso, em que U(x) < U(y), tem-se $U(z) \geq U(x)$ para todo z interno ao segmento de extremos x e y. Com efeito, suponhamos que se tivesse. $U(z) < U(x)$, para algum $z = \alpha x + (1-\alpha)y$, sendo $0 < \alpha < 1$. Como U é contínua, existiria um ponto w entre z e y tal que $U(w) = U(x)$; mas z seria um ponto do segmento de extremos x e w, e $U(z) < U((x) = U(w)$, o que contradiz a hipótese de concavidade seccional.

Podemos ir pouco mais adiante e demonstrar que, para todo z interno ao segmento de extremos x e y se tem $U(z) > U(x)$. Já demonstramos a impossibilidade de ser $U(z) < U(x)$. Com efeito, suponhamos que isso ocorre em algum ponto z do segmento de extremos x e y. Então, pela concavidade seccional temos $U(v) > U(z)$, para qualquer v entre x e w. Temos também $U(z) < U(y)$, pela hipótese de que $U(x) < U(y)$. Mas z seria ponto interno do segmento de extremos v e y onde a função U assume valor inferior ao de ambos os extremos. Vemos que isso contradiz a hipótese de concavidade seccional.

Provamos assim, que U é contínua e seccionalmente côncava para todo $0 < \alpha < 1$ se tem $U[\alpha x + (1-\alpha)y] > \min\{U(x),U(y)\}$. segue-se daí que se $U(x) \geq \alpha$ e $U(y) > \alpha$, então para todo $0 < \alpha < 1$ teremos $U[\alpha x + (1-\alpha)y) > \alpha$, o que prova que U é semi-côncava.

Teorema 3.13

Seja F(x) uma função contínua e semi-côncava, definida em todo o espaço do $\Re^n$. Seja C um subconjunto convexo do $\Re n$. Então; para que o máximo de f no conjunto ocorra no ponto $x_0 \in C$ é necessário e suficiente que exista um vetor $p \in \Re^n$ tal que:

$[x-x_0,p] \geq 0$ para todo $x \in C$.

$[y-x_0,p] > 0$ para todo y tal que $f(x) > f(x0)$.

Demonstração

Que a condição é suficiente, é imediato. Mostramos que a condição é necessária. Há dois casos a distinguir:

(i) f(x) é o máximo absoluto da função f em todo o espaço $\Re^n$. Neste caso basta tomar $p = 0$;

(ii) Exite algum ponto $y \in \Re^n$-C tal que $f(y) > f(x_0)$. Nesse caso designemos por S_0 o conjunto dos vetores y tais que $f(y) > f(x_0)$, e por S_1 ao conjunto formado pelos pontos de S_0 e pelo ponto x_0 de máximo condicionado em C. É fácil verificar que:

(1) S_0 é não vazio e convexo (pois f é semi-côncava);

(2) S_1 é convexo (pois f é semi-côncava)

(3) S_1 e C só tem em comum o ponto x_0.

Nessas condições pelo Teorema 3.12, exixtirá um vetor $p \neq 0$ tal que:

$[x-x_0,p] \leq 0$ para todo $x \in C$

$[y-x_0,p] \geq 0$ para todo $y \in S_0$

Precisamos agora provar que esta última desigualdade funciona no sentido estrito, isto é, que $[y-x_0,p] > 0$ para todo y tal que $f(y)>f(x_0)$. Para isso basta observar que, f sendo contínua, S_0 será aberto. Segue-se que se $y \in S$, $y-\lambda p$ também pertencerá a S_0 para valores positivos e suficientemente pequenos de λ. Isso exige que se tenha $[y-\lambda p-x_0,p] \geq 0$, o que implica em $[y-x_0,p]>0$.

O conteúdo geométrico do teorema demonstrado é o que sugere na figura. Se o máximo da função f(x) no conjunto C ocorre no ponto x_0 e se S_0 é o conjunto dos pontos nos quais $f(x) > f(x_0)$ então deve existir um hiperplano passando por x_0, que separe S_0 e C, todos os pontos de S_0 se situando à direita desse hiperplano.

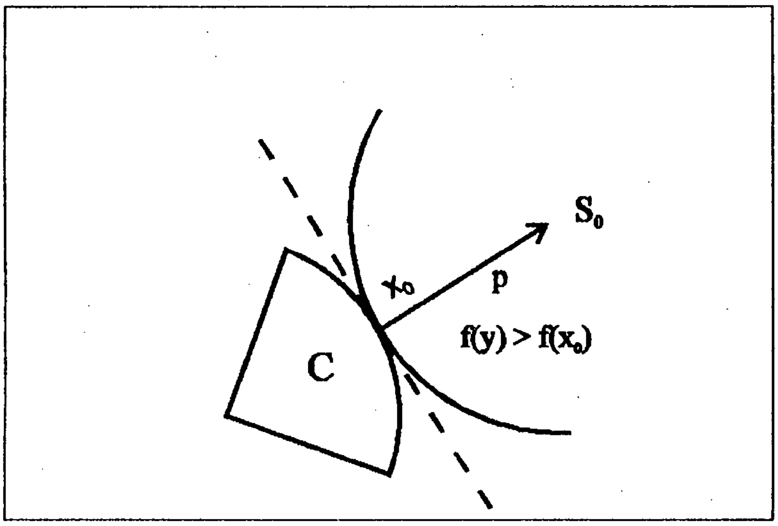

Figura 2

3.6. Funções Semi-côncavas Diferenciáveis

Cuidaremos agora de particularizar os resultados da seção anterior pra funções semi-côncavas diferenciáveis. Como sempre uma função real $f(x) = f(x_1,x_2,x_3,\ldots,x_n)$, se dirá difrenciável no ponto x_0 quando existir um valor grad $f(x_0)$ tal que:

$$f(x_0 + \Delta x) - f(x_0) = [\text{grad}\, f(x_0),\Delta x] + E\|\Delta x\|$$

sendo $\lim\limits_{\Delta x \to 0} E=0$.

O vetor grad f(x), se existir, terá por componentes as derivadas parciais, da função f calculadas no ponto x_0, isto é:

$$\text{grad}\, f(x_0) = \left\{\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right\}$$

Teorema 3.14

Seja f(x) uma função semi-côncava, contínua em todo o espaço $\Re^n$, e diferenciável no ponto $f(x_0) \neq 0$. Então, se $f(y) > f(x_0)$ ter-se-á:

$$[y-x_0, \text{grad } f(x_0)] > 0$$

Demonstração

Suponhamos que $f(y) > f(x_0)$. Como f é semi-côncava, para todo $0 < \lambda < 1$ teremos:

$$f[x_0 + \lambda(x-x_0)] > f(x_0)$$

Como f é diferenciável no ponto x_0, isso equivale a se ter:

$$f[x_0 + \lambda(y-x_0)] - f(x) = \lambda[y-x_0, \text{grad } f(x_0) + \lambda\varepsilon\|y-x_0\|>0,$$

sendo $\lim\limits_{\Delta x \to 0} \varepsilon=0.$

Isso exige que se tenha $[y-x_0, \text{grad } f(x_0)] \geq 0$ para todo y tal que $f(y) > f(x_0)$. Para mostrarmos que a desigualdade funciona no sentido estrito observamos que, f sendo sontínua, o conjunto dos valores de y tal que $f(y) > f(x_0)$ deve ser aberto. Isso implica em que para μ positivo e suficientemente pequeno se tenha:

$f[y-\mu\text{grad } f(x_0)] > f(x_0)$, ou seja:

$$[y-\mu\text{grad } f(x_0) - x_0, \text{grad } f(x_0)] \geq 0$$

o que conduz a

$$[y-x_0, \text{grad } f(x_0)] > 0$$

tendo em vista a hipótese de que grad $f(x_0) \neq 0$.

O conteúdo geométrico do teorema é simples. Se f é contínua e semi-côncava em todo $\Re^n$, diferenciável no ponto x_0, e grad $f(x_0) \neq 0$, então, se $f(y) > f(x_0)$, $y-x_0$ formará ângulo agudo com o vetor grad $f(x_0)$, como na figura:

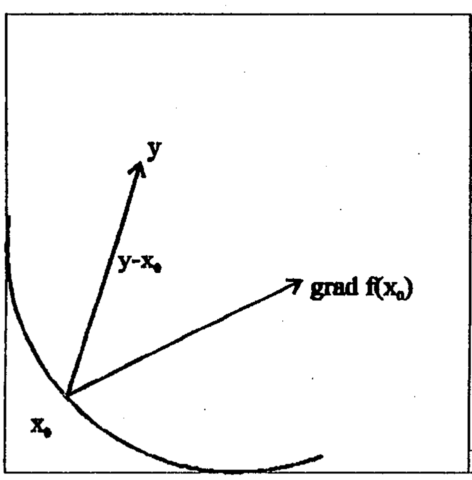

Figura 3

Teorema 3.15

Seja f(x) uma função semi-côncava, contínua em todo o $\Re^n$ e diferenciável no ponto x_0, sendo grad $f(x_0) \neq 0$. Então se $f(y) = f(x_0)$ ter-se-á:

$[y\text{-}x_0, \text{grad } f(x_0)] \geq 0.$

Demonstração

Seja z um vetor tal que $f(z) > f(x_0)$; um tal vetor deve existir visto que grad $f(x_0) \neq 0$. Como f é semi-côncava devemos ter $f[\lambda z + (1\text{-}\lambda)y] > f(x_0)$ para todo $0 < \lambda < 1$.

Pelo Teorema 3.14 isso exige que para todo $0 < \lambda < 1$ se tenha:

$[\lambda z + (1\text{-}\lambda)y - x_0, \text{ grad } f(x_0)] > 0$

o que exige:

$[y\text{-}x_0, \text{grad } f(x_0)] \geq 0.$

Teorema 3.16

Seja f(x) uma função semi-côncava, contínua em todo o $\Re^n$ e diferenciável no ponto x_0, sendo grad $f(x_0) \neq 0$. Seja y um vetor tal que $[y\text{-}x_0, \text{grad } f(x_0)] > 0$. Então existe um número real $\Delta > 0$, tal que para todo $\lambda < \Delta$ se tenha:

$f[x_0 + \lambda(y\text{-}x_0)] > f(x_0)$

Demonstração

A demonstração envolve apenas a diferenciabilidade da função f no ponto x_0. Basta observar que:

$f[x_0 + \lambda(y\text{-}x_0)] - f(x_0) =$

$= \lambda[\lambda - x_0, \text{grad } f(x_0)] + \lambda\xi \, \|y\text{-}x_0\|$, sendo $\lim\limits_{\Delta x \to 0} \varepsilon = 0.$

Teorema 3.16

Seja f(x) uma função semi-côncava, contínua em todo o $\Re^n$, e diferenciável no ponto x_0, sendo grad $f(x_0) \neq 0$. Seja $p \neq 0$ um vetor tal que $[y\text{-}x_0, p] > 0$ para todo y tal que $f(y) > f(x_0)$. Então $p = \lambda \text{grad } f(x_0)$, sendo λ um número real positivo. Reciprocamente, todo vetor dessa forma satisfaz as condições do teorema.

Demonstração

Suponhamos que $p = \lambda \text{grad } f(x_0)$, sendo λ um número real positivo. Então pelo Teorema 3.14, segue-se que $[y-x_0,p] > 0$, para todo y tal que $f(y) > f(x_0)$.

Reciprocamente admitmos que $p \neq 0$ seja tal que $[y-x_0,p] > 0$ para todo y tal que $f(y) > f(x_0)$. Projetando grad $f(x_0)$ sobre p pode-se escrever grad $f(x_0) = \alpha p + u$, sendo $\alpha \geq 0$ e $[p,u] = 0$. Essa última condição nos garante que $f(x_0+\lambda u) \leq f(x_0)$, pelas hipóteses feitas quando ao vetor p. Por outro lado, se $u \neq 0$, teremos $[\text{grad } f(x_0),u] = [u,u] > 0$ o que, pelo Teorema 3.14, implica em se ter $f(x_0 + \lambda u) > f(x_0)$ para algum $\lambda > 0$. Segue-se que para $u = 0$ e que $p = \dfrac{1}{\alpha \text{grad} f(x_0)}$.

O conteúdo geométrico do teorema geométrico demonstrado também é simples. Ele afirma que, se f é contínua e semi-côncava em todo espaço $\Re^n$ e diferenciável no ponto x_0 onde grad $f(x_0) \neq 0$, então existe um e único hiperplano que passa por x_0 e que separa esse ponto do conjunto dos vetores y tais que $f(y) > f(x_0)$. A normal a esse hiperplano é o gradiente da função f no ponto x_0, conforme vemos na figura:

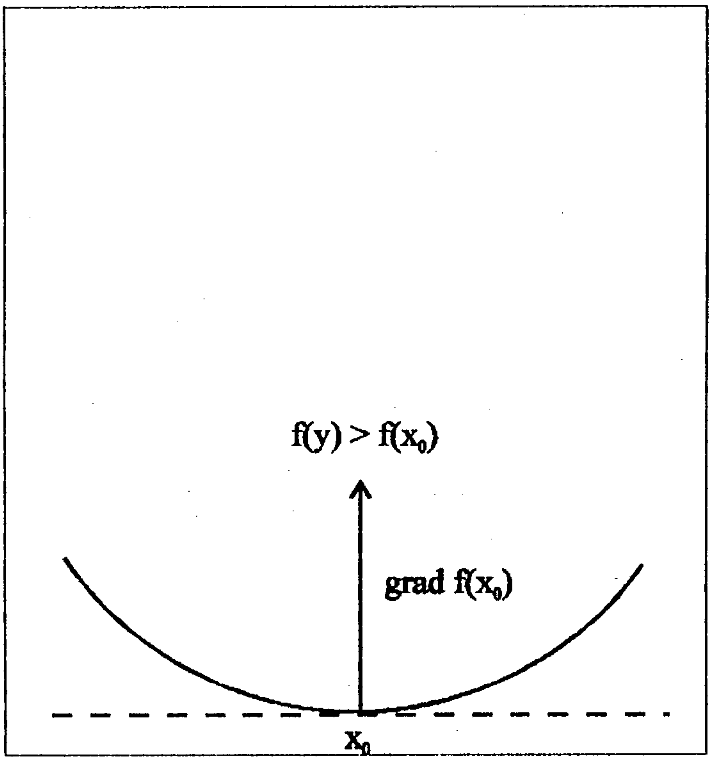

Figura 4

Teorema 3.17

Seja f(x) uma função côncava (no sentido lato), definida no espaço $\Re^n$ e diferenciável no ponto x_0. Então, para que a função por um único máximo absoluto nesse ponto x_0 é necessário e suficiente que grad $f(x_0) = 0$.

Demonstração

Que a a condição é necessária é algo que resulta da própria definição de diferenciabilidade. Como:

$$f(x_0 + \Delta x) - f(x_0) = [\text{grad } f(x_0), \Delta x] + \varepsilon\|\Delta x\|$$

sendo $\lim\limits_{\Delta x \to 0} \varepsilon = 0$, e como $f(x_0 + \Delta x) - f(x_0) \geq 0$ se f passa por um máximo absoluto no ponto x_0, conclui-se que necessariamente grad $f(x_0) = 0$.

Reciprocamente, suponhamos que grad $f(x_0) = 0$. Seja y um outro vetor qualquer do $\Re^n$, e consideremos a função de uma variável real:

$$F(t) = f[(1-t)x_0 + ty]$$

é imediato que F(t) é côncava e diferenciável no intervalo [0,1]. Segue-se daí que:

$$F(1) - F(0) \leq F'(0).$$

Como:

$$F(1) = F(y),\ F(0) = F(x_0) \text{ e } F'(0) = [y-x_0, \text{grad } f(x_0)] = 0.$$

Segue-se que:

$f(y) \leq f(x_0)$ para qualquer y E $\Re^n$, o que prova que f passa por um máximo o ponto x_0.

3.7. Maximização de funções Semi-côncavas

Vejamos agora alguns teoremas sobre maximização de funções semi-côncavas com restrições lineares. Esses teoremas nos conduzirão aos principais resultados sobre o primeiro tipo de programação não linear e sobre programação linear.

Teorema 3.18

Sejam $c_1, c_2, ..., c_m$ vetores do $\Re^n$; $a_1, a_2, ..., a_m$ números reais. Se o conjunto (convexo) dos vetores x tais que:

$$[c_1, x] \leq a_1$$

$$[c_2, x] \leq a_2$$

$$...$$

$$[c_m, x] \leq a_m.$$

Seja x_0 um ponto de S, g(x) uma função contínua e semi-côncava em todo o $\Re^n$ e diferenciável no ponto x_0, sendo grad $g(x_0) \neq 0$. Então, para que o máximo de g em S ocorra no ponto x_0, é necessário e suficiente que:

$[u, grad(x_0)] \leq 0$

para todo vetor u tal que $[c_i,u] \leq 0$ para todo índice i tal que:

$[c_i,x_0] = a_1$.

Demonstração

Suponhamos que x_0 seja o ponto máximo de g no conjunto (convexo) S. Combinando os Teoremas 3.13 e 3.17 conclui-se que para todo $x \in S$, se terá $[x-x_0, \text{grad } g(x_0)] \leq 0$. Seja agora u um vetor $[a_i,x_0] = a_i$. Então é fácil verificar que, para algum λ real positivo, $x_0 + \lambda u \in S$. Segue-se que $[u, \text{grad } g(x_0)] \leq 0$.

Reciprocamente, suponhamos que $[u, \text{grad } g(x_0) \leq 0$ para todo u tal que $[u,c_i] \leq 0$ quando $[c_i,x_0] = a_i$. Seja x um outro ponto qualquer de S. Segue-se então que $[x-x_0, \text{grad } g(x_0)] \leq 0$. Pelo Teorema 3.16, deveremos ter $g(x) \leq g(x_0)$, o que prova que o máximo de g em S ocorre no ponto x_0.

O conteúdo geométrico desse teorema é interessante, e se ilustra na figura a seguir. Seja x_0 um ponto de um conjunto S definido por uma intersecção de hiperplanos (H_1, H_2, H_3 e H_4), Seja S o cone formado pelos hiperplanos que passam por x_0 (no caso H_1 e H_2). Então para que a função contínua, semi-côncava e diferenciável g(x) passe em x_0, por um máximo em S, é necessário e suficiente que nesse ponto ele passe por um máximo em $\bar{S}$. E para isso, é necessário e suficiente que grad $g(x_0)$ forme ângulo reto ou obtuso com todo ponto de $\bar{S}$.

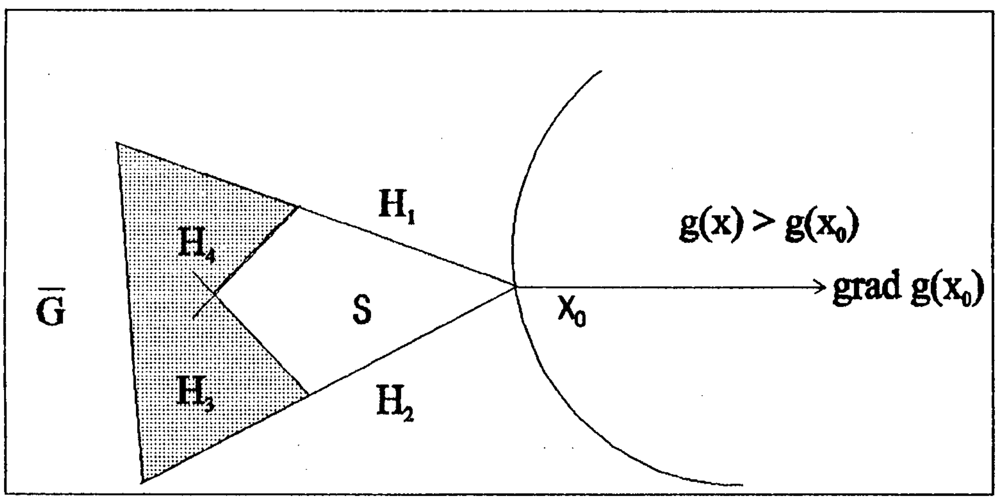

Figura 5

Conjuguemos agora esse teorema com o de Farkas no seu enunciado final. Designaremos por A a matriz cujas linhas são os vetores $-c_i$ tais que $[c_i,x_0] =$

a_i; por c o vetor -grad $g(x_0)$. Então para todo vetor u tal que $u \geq 0$ se terá $[c,u] \geq 0$. Pelo teorema de Farkas segue-se que existirá um vetor $p \geq 0$ tal que $A^t p = c$.

Teorema 3.19

Sejam $c_1, c_2, ..., cm$ vetores do $\Re^n$, $a_1, a_2, ..., a_m$ números reais, g(x) uma função contínua e semi-côncava em todo o espaço. Seja S o conjunto dos vetores x tais que:

$[c_1,x] \leq a_1$

$[c_2,x] \leq a_2$

...

$[c_m,x] \leq a_m$.

Seja x_0 um ponto de S no qual g(x) é diferenciável com gradiente não nulo. Então para que g passe por um máximo em S no ponto x_0 é necessário e suficiente que existam números reais não negativos $p_1, p_2, ..., p_m$, tais que:

(i) grad $g(x_0) = p_1c_1 + p_2c_2 + ... + p_mc_m$

(ii) $p_1 = 0$ se $[c_i,x_0\,] < a_i$

Demonstração

A demonstração não será efetuada, informando apenas que o teorema acima fornece uma extensão do método dos multiplicadores de Lagrange, indicando que os multiplicadores devem ser não negativos em geral, e nulo para as desigualdades estritamente observadas. Passando da notação vetorial para a analític e introduzindo condições de não negatividade das variáveis ($x_j \geq 0$ ou $-x_j \leq 0$) chega-se ao seguinte:

Teorema 3.20 (Kuhn e Tucker)

Consideremos o problema: Maximizar a função contínua e semi-côncava: $g(x_1, x_2, ..., x_n)$ com as condições:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \leq b_2 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \leq b_n \\ x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0 \end{cases}$$

Seja $(z_1,z_2, ..., z_n)$ um ponto do conjunto de soluções no qual a função f é diferenciável, uma pelo menos das derivadas parciais sendo diferente de zero nesse ponto. Então para que esse seja o ponto máximo de f no conjunto de soluções dadas é necessário e suficiente que existam números reais $p_1, p_2, ..., p_m$ tais que:

(i) $p_1 \geq 0; p_2 \geq 0; ...; p_m \geq 0$

(ii) $p_1 = 0$ se $a_{i1}z_1 + a_{i2}z_2 + ... a_{in}z_n < b_i$

(iii) $a_{ij}p_1 + a_{sj}p_2 + ... + a_{mj}p_m \geq \dfrac{\partial f}{\partial x_j}$, se $z_j = 0$

(iv) $a_{1j}p_1 + a_{2j}p_2 + ... + a_{mj}p_m = \dfrac{\partial f}{\partial x_j}$, se $z_j > 0$

as derivadas parciais sendo calculadas no ponto $(z_1, z_2, ..., z_n)$.

Demonstração

A demonstração não será apresentada, porém como aplicação do teorema em questão, para o caso particular em que a função maximizar é linear, chegamos ao próximo teorema.

Teorema 3.21

Consideremos o problema: Maximizar a função linear: $c_1x1 + c_2x_2 + ... + c_nx_n$, com as condições:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \leq b_2 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \leq b_n \\ x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0 \end{cases}$$

Então, para que um ponto $(z_1, z_2, ..., z_n)$ do conjunto de soluções seja solução ótima do problema, é necessário e suficiente que existam números reais p_1, p_2, ..., p_m, tais que:

(i) $p_1 \geq 0; p_2 \geq 0; ...; p_m \geq 0$

(ii) $p_1 = 0$ se $a_{i1}z_1 + a_{i2}z_2 + ... + a_{in}z_n < b_1$

(iii) $a_{ij}p_1 + a_{2j}p_2 + ... + a_{mj}p_m \geq c_j$, se $z_j \geq 0$

(iv) $a_{1j}p_1 + a_{2j}p_s + ... + a_{mj}p_m = cj$, se $zj > 0$

Demonstração

A demonstração do teorema não será efetuada, pois o interesse geral no trabalho vai além de um tratamento puramente linear.

3.8. O Teorema de Kuhn-Tucker - Caso Geral

Trataremos agora o Teorema de Kuhn-Tucker para o caso em que as condições de restrição, ao invés de lineares, são apenas semi-convexas.

Teorema 3.22

Seja g(x) uma função contínua e semi-côncava, em todo $\Re^n$; $g_1(x)$, $g_2(x)$, ..., $g_m(x)$ funções contínuas e semi-convexas em todo o espaço; a_1, a_2, ..., a_m, números reais, C o conjunto dos vetores x tais que:

$$\begin{cases} g_1(x) \leq a_1 \\ g_2(x) \leq a_2 \\ \dots \\ g_m(x) \leq a_m \\ x \geq 0 \end{cases}$$

Suponhamos que em C exista algum vetor y para o qual as desigualdades acima se observem estritamente, isto é, $g_1(y) < a_1$; ...; $g_m(y) < a_m$. Seja x_0 um ponto de C no qual as funções f(x), $g_1(x)$, ..., $g_m(x)$ sejam diferenciáveis e todas com gradientes diferentes de zero. Então, para que o máximo de f(x) em C ocorra no ponto x_0, é necessário e suficiente que nesse ponto ocorra também o máximo de f(x) de f(x) no poliedro C_{fecho} definido pelas desigualdades:

$[x\text{-}x_0, \text{grad}_i(x_0)] \leq 0$ para todo índice i tal que $g_i(x_0) = a_i x \geq 0$.

Demonstração

Suponhamos que o máximo de f(x) em C ocorra em x_0. Isso significa que, para todo $z \in C$, se tem $f(z) \leq f(x_0)$. Seja x um vetor não negativo tal que $[x\text{-}x_0, \text{grad } g_i(x_0)] \leq 0$ para todo índice i tal que $g_i(x_0) = a_i$ (isto é, um vetor de $\overline{C}$). Seja agora y um vetor de C tal que $g_i(y) < a_i$ para todo índice i. Como $-g_i$ é semi-côncava e como $-g_i(y) > -g_i(x_0)$ para todo índice i tal que $g_i(x_0)$ = ai, segue-se pelo Teorema 3.14 que:

$[y\text{-}x_0, \text{grad}_i(x_0)]<0$ para todo índice i tal que $gi(x_0)=a_i$.

Seja ainda α um número real qualquer tal que $0<\alpha<1$ e $z = \alpha y + (1-\alpha)x$. Então, para todo ínidice i tal que $g_i(x_0) = a_i$ teremos:

$[z\text{-}x_0, \text{grad } g_1(x_0)] < 0$.

Seja Δ_i um número real positivo tal que para todo $\lambda \leq \Delta_i$ se tenha $g_i[x_0 + \lambda(z\text{-}x_0)] \leq a_i$. Pela desigualdade acima, Δ_i existirá se $g_i(x_0) = a_i$, em virtude do Teorema 3.16; se $g_i(x_0) < a_i$, Δ_i existirá pela continuidade da função g_i. Segue-se, tomando λ igual ao menor dos Δ_i, que $x_0 + \lambda(z\text{-}x_0) \in C$ para algum λ positivo. Isso significa que $f[x_0 + \lambda(z\text{-}x_0)] \leq f(x_0)$, para algum λ positivo. Isso significa que $f[x_0 + \lambda(z\text{-}x_0)] \leq f(x_0)$, para algum λ positivo já que, por hipótese, x_0 é o máximo de f em C. Como f é semi-côncava, isso exige que se tenha também $f(z) \leq f(x_0)$. Mas isso exige que se tenha $f(z) \leq f(x_0)$. Mas isso imnplica em se ter $f[\alpha y + (a-\alpha)x] \leq f(x_0)$ para todo $0 < \alpha < 1$. Como f é contínua, devemos ter $f(x) \leq f(x_0)$. Isso prova que x_0 é o ponto que maximiza f no conjunto $\overline{C}$.

A recíproca do teorema é bastante simples. Suponhamos que x_0 seja o máximo de f(x) em $\overline{C}$, isto é, que $f(x) \leq f(x_0)$ para todo x tal que $[x-x_0, \text{grad } g_i(x_0)] \leq 0$ para todo índice i tal que $g_i(x_0) = a_i$. Pelos Teoremas 3.14 e 3.15 é imediato que C é subconjunto de $\overline{C}$. Como por hipótese, x_0 pertence a C, o máximo de f em C também ocorrerá em x_0.

Observação: Em termos analíticos, o teorema demonstrado pode enunciar-se da seguinte forma:

Consideremos o problema: Maximizar a função contínua e semi-côncava $f(x_1, x_2, ..., x_n)$, com as restrições:

$$\begin{cases} g_1(x_1, x_2, ..., x_n) \leq a_1 \\ g_2(x_1, x_2, ..., x_n) \leq a_2 \\ ... \\ g_m(x_1, x_2, ..., x_m) \leq a_m \\ x_1 \geq 0; x_2 \geq 0; ...; x_m \geq 0 \end{cases}$$

As diversas funções g_1, g_2, ..., g_m sendo contínuas e semi-convexas. Admitamos que exista um ponto para o qual essas desigualdades se observam todas estritamente. Seja agora $(z_1, z_2, ..., z_n)$ um ponto do conjunto de soluções do sistema de desigualdades no qual as funções f, g_1, g_2, ..., g_m sejam todas diferenciáveis e como gradientes diferentes de zero. Designemos por $f_j = \dfrac{\partial f}{\partial x_j}$ e $g_{ij} = \dfrac{\partial g_i}{\partial x_j}$ as derivadas parciais das funções nesse ponto $(z_1, z_2, ..., z_n)$. Então esse ponto será a solução ótima do problema se e somente se for a solução ótima do problema seguinte: Maximizar $f(x_1, x_2, ..., x_n)$, com as condições:

$(x_1-z_1)g_{i1} + (x_2-z_2)g_{i2} + ... + (x_n-z_n)g_{in} \leq 0$, para todo índice i tal que $g_i(z_1, z_2, ..., z_n) = a_i$;

$x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0.$

Podemos agora observar que este último problema tem restrições lineares, aplicando-se por ela a forma simplificada do Teorema de Kuhn e Tucker. Aplicando então o Teorema 3.16 e atribuindo valores $p_i = 0$ para as desigualdades tais que $g_i(z_1, z_2, ..., z_n) < a_i$ chegamos ao seguinte teorema:

Teorema 3.23 (Forma Geral do Teorema de Kuhn e Tucker)

Consideremos o problema: Maximizar a função contínua e semi-côncava $f(x_1, x_2, ..., x_n)$, com as condições:

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n) \leq a_1 \\ g_2(x_1, x_2, \ldots, x_n) \leq a_2 \\ \ldots \\ g_m(x_1, x_2, \ldots, x_m) \leq a_m \\ x_1 \geq 0; x_2 \geq 0; \ldots; x_m \geq 0 \end{cases}$$

onde as funções g_1, g_2, ..., g_m são contínuas e semi-côncavas em todo o $\Re^n$, supondo-se que exista algum ponto $(y_1, y_2, \ldots, y_n)$ tal que $g_i(y_1, y_2, \ldots, y_n) < a_i$ (i = 1, 2, ..., m). Então para que um ponto $(z_1, z_2, \ldots, z_n)$, que atenda ao sistema de desigualdades, e no qual f, g_1, g_2, ..., g_m sejam diferenciáveis com gradientes não nulos, seja solução ótima do problema, é necessário e suficiente que existam números reais p_1, p_2, ..., p_m tais que:

(i) $p_1 \geq 0; p_2 \geq 0; \ldots; p_m \geq 0$

(ii) $p_1 = 0$ se $g_i(x_1, x_2, \ldots, x_n) < a_i$

(iii) $\dfrac{\partial f}{\partial x_j} = p_1 \dfrac{\partial g_1}{\partial x_j} + p_2 \dfrac{\partial g_2}{\partial x_j} + \cdots + p_m \dfrac{\partial g_m}{\partial x_j}$ se $z_j > 0$

(iv) $\dfrac{\partial f}{\partial x_j} \leq p_1 \dfrac{\partial g_1}{\partial x_j} + p_2 \dfrac{\partial g_2}{\partial x_j} + \cdots + p_m \dfrac{\partial g_m}{\partial x_j}$ se $z_j = 0$

as diversas derivadas parciais sendo calculadas nos pontos $(z_1, ,z_2, \ldots, z_n)$.

3.9. Comentário Final

Com estamos tratando de otimização, fica explicada a apresentação dos rudimentos, definições e teoremas, dentro do capítulo 3; pelo algoritmo que será apresentado no capítulo 5 é o suficiente, porém ainda se torna imperativa a colocação de mais um capítulo, pois o termo global ainda não está justificado, logo incluímos como próximo item de trabalho as condições de otimidade global.

OTIMALIDADE GLOBAL

4.1. Introdução

Antes de apresentar qualquer método numérico, é sempre inteligente apresentar uma série de condições, as quais deem respaldo ao algoritmo que se pretende desenvolver, algoritmos esses que geram subseqüencias convergentes para pontos que satisfazem a essas condições.

Iremos considerar o problema de minimização: minimizar f(x) com as condições:

$$\begin{cases} g_i(x) \leq 0, i = 1, \ldots, m \\ h_j(x) = 0, j = 1, \ldots, p \\ x \in X \subset \Re^n \end{cases}$$

Daonde veremos que a diferenciabilidade do problema tem influência sobre a natureza dessas condições.

4.2. Otimização sem Restrições

<u>Definição 4.1</u>

Definimos como <u>valor ótimo</u> de um problema de minimização, como sendo z* real definido por:

$z^* = \text{Inf}\{f(x) \mid x \in \Re^n\}$ quando o ínfimo existe.

<u>Definição 4.2</u>

Seja uma função $f:\Re^n \rightarrow \Re$. Diremos que x^*é um <u>mínimo global</u> de f se $f(x^*) \leq f(x)$, para qualquer $x \in \Re^n$. Diremos que x^* é um <u>mínimo local</u> de f se existe $\varepsilon > 0$, tal que $f(x^*) \leq f(x)$, para qualquer $x \in \Re^n$ tal que $\|x-x^*\| < \varepsilon$.

<u>Definição 4.3</u>

Um vetor $d \in \Re^n$ é uma <u>direção de descida</u> da função f em $x^* \in \Re^n$ se existe $\delta > 0$ tal que:

$f(x^* + \lambda d) < f(x^*)$, para qualquer $\lambda \in (0,\delta)$

<u>Teorema 4.1</u>

Suponha f diferenciável em x^*. Qualquer vetor de $\Re^n$, tal que $[\text{grad } f(x^*)]^t d < 0$ é uma direção de descida de f em x^*.

<u>Demonstração</u>

Não será feita.

<u>Teorema 4.2</u>

Uma condição necessária para que x^* seja um mínimo local da função f, diferenciável em x^* é:

$\text{grad } f(x^*) = 0$.

<u>Demonstração</u>

A partir da definição de mínimo local, vemos que se $\bar{x}$ é um mínimo local, não existem direções d descida de f em x^*, ou seja, $[\text{grad } f(x^*)]^t d \leq 0$; para qualquer $d \in \Re^n$. Logo $\text{grad } f(x^*) = 0$.

O Teorema 4.2 é uma condição de 1ª ordem, em caso de funções duas vezes diferenciáveis podemos enunciar uma condição de 2ª ordem, na forma do Teorema 4.3.

<u>Teorema 4.3</u>

Uma condição de 1ª ordem necessária para que x^* seja um mínimo local da função f, duas vezes diferenciável em x^*, é:

(i) $\text{grad } f(x^*) = 0$

(ii) $H(x^*) = (\text{grad } f(x^*))^2$ é positiva e semi-definida.

Demonstração

Como f é duas vezes diferenciável em x^*, temos a relação (1)

$f(x) = f(x^*) + [\text{grad } f(x^*)]^t(x-x^*) + 1/2(x-x^*)^t H(x^*)(x-x^*) + |x-x^*\|^2 \in (x^*; x)$, onde $\varepsilon(x^*,x) \rightarrow 0$ quando $x \rightarrow x^*$.

Se x^* é um mínimo local de f, temos grad $f(x^*) = 0$ e $f(x) \geq f(x^*)$, para qualquer x tal que $\|x-x^*\| < \varepsilon_1$.

Como $\varepsilon(x^*-x) \rightarrow 0$ quando $x \rightarrow x^*$ vemos que $(x-x^*)^t H(x^*)(x-x^*) \geq 0$, para qualquer x tal que $\|x-x^*\| < \varepsilon_1$, ou seja, $H(x^*)$ é positiva e semi-definida.

Essas condições são apenas necessárias porque os termos de primeira e de segunda ordem podem estar nulos e teremos ainda uma dúvida sobre a natureza de x^*. Por outro lado, teremos uma condição suficiente se fortalecermos a hipótese feita sobre Hessiano. A condição suficiente de segunda ordem pode ser colocada na forma do Teorema 4.4.

Teorema 4.4

Seja f duas vezes diferenciável em x^* tal que:

(i) grad $(x^*) = 0$

(ii) $H(x^*)$ é positiva definida.

Então x^* é um mínimo local estrito de f.

Demonstração

Com as condições (i) e (ii), a relação (1) torna-se

$f(x) - f(x^*) > 0$, para qualquer x, tal que

$\|x-x^*\|^2 \mid E(x^*,x)| < 1/2(x-x^*)^t(H(x^*)(x-x^*)$.

4.3. Otimização de Funções com Restrições de Igualdades

Estaremos agora tratando do problema: Minimizar $f(x)$ sob as condições:

$$\begin{cases} h_j(x) = 0, j = 1,\dots,p \\ x \in X \subset \Re^n \end{cases}$$

Faremos duas hipóteses sobre a superfície S no ponto $x^* \in S$, onde estão estabelecidas as condições de otimalidade.

Hipóteses

(1) $h_j \in C_1$, $j = 1, ..., p$. Diremos que a superfície S é "suave" em x^*.

(2) Os gradientes (grad $h_j(x^*)$, $j = 1, ..., p$) são linearmente independentes .

Conseqüências

(1) Podemos definir analiticamente o plano tangente d superfície S em x^*, ou seja, o conjunto de todos os vetores derivados as curvas diferenciáveis contidas em S e passando por x^*. Será o subespaço

$$M = \{d \in \Re n \mid \text{grad}\, h(x^*)d = 0\}$$

onde grad $h(x^*)$ é o jacobiano da função vetorial h em x^*.

Teorema 4.5 (Condição Necessária de Primeira Ordem)

Uma condição necessária para que o ponto $x^* \in S$, satisfazendo as hipóteses (1) e (2); seja um mínimo local de um função f continuamente diferenciável em x^*, e que existam multiplicadores de Lagrange $\lambda_1, \lambda_2, ..., \lambda_m$ tais que:

$$\text{gradf}(x^*) + \sum_{j=1}^{p} l_j \text{gradh}_j(x^*) = 0$$

A demonstração deste teorema é sem dúvida alguma muito elegante, porém devido a sua extensão passaremos adiante, deixando ao leitor interessado a referência bibliográfica, onde o mesmo se encontra. [10]

Definição 4.4

A função lagrangiana associada ao problema de minimização com restrições de igualdade, é a função de $\Re^n x \Re^m \rightarrow \Re$ definida por:

$$L(x, \lambda) = f(x) + \sum_{k=1}^{m} \lambda_k h_k(x)$$

Teorema 4.6 (Condição Necessária de 2ª Ordem)

Suponha que $x^* \in S$, e que as hipóteses (1) e (2) são satisfeitas em x^*. Suponha ainda que $f, h_1, ..., h_p \in C^2$. Então se x^* é um mínimo local de f sobre S, existem reais $\lambda_1, ..., \lambda_p$ tais que:

(i) grad $L(x,\lambda) = 0$

(ii) $\Delta x^t \mathsf{L}(x^*,\lambda)\, \Delta x \geq 0$, para qualquer Δx e $M = \{y \in \Re^n \mid \Delta h(x^*)y = 0\}$ onde L é o hessiano de $L(x,\lambda)$ em relação a x:

$$\mathsf{L}(x,\lambda) = [\text{gradf}(x)]^2 + \sum_{k=1}^{m} \lambda_k [\text{gradh}_k(x)]^2$$

Novamente optamos em não apresentar a demonstração.

4.4. Otimização de Funções com Restrições de Desigualdades

Consideremos agora o problema do tipo: Minimizar f(x) sob as condições:

$g_i(x) \leq 0,\ i = 1, ..., m.$

Hipóteses

(1) $g_1 \in C_1,\ i = 1, ..., m$

(2) Os gradientes {grad $g_i(x^*)$, $i = 1, ..., m \mid g_i(x^*) = 0$} são linearmente independentes.

Definição 4.5

As restrições associadas ao conjunto $i = 1, ..., m$ serão ditas ativas, as outras serão com folga.

Teorema 4.7 (Condição Necessária de Primeira Ordem - Condições de Kuhn-Tucker)

Seja x^* um mínimo local de $f \in C_1$ no domínio $x \subset \Re^n$ definido na introdução. Se as hipóteses (1) e (2) forem satisfeitas em $x^* \in X$, então existem multiplicadores $\mu_1, ..., \mu_m$ tais que:

$$\text{(i)}\, \text{gradf}(x^*) + \sum_{i=1}^{m} \mu_i \text{gradg}_i(x^*) = 0$$

(ii) $\mu_i >= 0, i = 1, ..., m$

(iii) $\mu_i g_i(x^*) = 0, i = 1, ..., m$

Demonstração

Vide [10]

Teorema 4.8 (Condição Necessária de Segunda Oredem)

Suponha que $x^* \in X$ e que as hipóteses (1) e (2) são satisfeitas em x^*. Suponha ainda que f, $g_1, ..., g_m \in C^2$. Então, se x^* é um mínimo local de f na região X, existem $\mu_1, ..., \mu_m$ tais que:

$$(i) f(x^*) + \sum_{i=1}^{m} \mu_i \text{grad} g_i(x^*) = 0$$

(ii) $\mu_i \geq 0$ e $\mu_i g_i(x^*) = 0$

(iii) $\Delta x^t L(x^*,r)\Delta x >= 0$ para qualquer $\Delta x \in M$, onde $L(x,r) = [\text{grad } f(x)]^2 + \sum_{i=1}^{m} \mu_i g_i(x^*)$, e $M = \{y \in \Re^n \mid [\text{grad } g_i(x^*)]^t y=0, i=1, ..., m\}$.

Demonstração

Conforme o Teorema 4.6, observemos que o plano tangente em x^* é definido apenas em função das restrições ativas em x^*.

Observação

Para obter uma condição suficiente, devemos incluir as direções Δx tais que $[\text{grad } g_i(x^*)]^t \Delta x \leq 0$ para $i = 1, ..., m$ e $\mu_i = 0$. Por outro lado, se f é convexa e cada g_i é convexa, então qualquer solução satisfazendo as condições do Teorema 4.7 (Kuhn-Tucker) é um mínimo global do problema restrito.

O caso geral já estudado no capítulo 2 de um problema de minimização de igualdades e desigualdades reúne os Teoremas 4.6 e 4.7:

Min f(x)

Sujeito a: $g_i(x) \leq 0$

$hj(x) = 0$

A função lagrangiana é do tipo:

$L(x,\mu,\lambda) = f(x) + \mu^t g(x) + \lambda^t h(x)$

onde: $x \in \Re^n$; $\mu \in \Re^m$, $\lambda \in \Re^p$; $\mu \geq 0$

4.5. Dualidade Lagrangiana

Vamos considerar o seguinte problema de minimização no $\Re^n$ onde a região factível é definida como intersecção de uma região suposta "complicada" defiida explicitamente por m restrições de desigualdades.

Problema Primal

Minimizar $f(x)$

Sujeito a $g_i(x) \leq 0$, $i = 1, ..., m$

$x \in X \subset \Re^n$

Definimo a função lagrangiana associada ao problema como:

$L(x,u) = f(x) + u^t g(x)$

onde $u \in \Re^m$, $u \geq 0$.

$L(u)$ é a função dual associada ao problema, definida para $u \geq 0$ tais que $\inf_{x \in X} L(x,u) > -\infty$.

4.6. Condições de Otimalidade Global

Definição 4.5

O par (x^*, u^*) com $x^* \in X$ e $u^* >= 0$ satisfaz as condições de otimalidade global do problema primal se:

$$(i) L(x^*, u^*) = L(u^*) = \min_{x \in X}\{f(x) + [u^*]^t g(x)\}$$

(ii) $[u^*]tg(x^*) = 0$

(iii) $g(x^*) \leq 0$

Teorema 4.9

Se (x^*,u^*) satisfazem as condições de otimalidade global, então x^* é solução do problema primal.

Demonstração

É fácil verificar que x^* é factível e $x^* \in X$. Temos $f(x^*) = f(x^*) + [u^*]^t g(x^*) \leq f(x) + [u^*]^t g(x) \leq f(x)$ sendo que para qualquer x com $g(x) \leq 0$.

4.7. Comentário Final

Visto que até o momento o conteúdo vem evoluindo gradativamente, estando agora a frente da apresentação de um algoritmo que nos permita conclir nosso objetivo, que é o de otimizar funções não-lineares globalmente, sujeitas a restrições de igualdade e desigualdade.

5. Algoritmo de Otimização Global

5.1. Introdução

O algoritmo não é nenhuma técnica numérica, destas que aparecem com grande freqüência em compêndios de programação não-linear, mas sim a apresentação de modificações na estrutura do problema.

5.2. O Algoritmo

Seja dado um problema de programação matemática:

Problema (1)

Minimizar: f(x)

Sujeito a:

$$g_i(x) \leq 0, \; i=1, 2, ..., m$$

$$h_j(x)= 0, \; j=1, 2, ..., p$$

$$x \in X \subset \Re^n$$

Sua solução pode ser obtida por algum dos numerosos métodos numéricos existentes, muitos explorando características inerentes ao problema. Por exemplo, se todas as funções mesmo que não-lineares, forem convexas e diferenciáveis, existem métodos muitos eficientes para resolver o problema. Caso contrário, surge a necessidade de se escolherem métodos que comportem a não diferenciabilidade e não convexidade do problema., No caso da otimização global, a técnica mais razoável ainda parece ser a de resolver o problema várias vezes, partindo cada vez de um ponto inicial diferente, aceitando como ótimo global a solução que apresentar o menor valor para a função objetivo do **problema (1)**. Esta técnica, apesar de fornecer soluções boas, não garante que o mínimo global tenha sido atingido.

Numa tentativa de resolver o problema da otimização global, propomos aqui:

Primeiro Passo: Construir um hiperplano **H**, paralelo ao domínio da função objetivo, a uma altura **k**, isto é, **H(x)=k, k** $\in \Re$, ou seja, fazer: **f(x)≤k, k** $\in \Re$;

Segundo Passo: Verificar se dentro do domínio viável o hiperplano **H** corta a função objetivo, isto é, se a inequação **f(x)≤k** tem solução viável;

Terceiro Passo: Se sim, diminuir o valor de **k** e repetir a verificação, isto é, repetir o segundo passo com o novo valor de **k**;

Quarto Passo: Os sucessivos ajustes no valor de **k** devem ser conduzidos na forma que o intervalo de variação de **k**, seja cada vez menor, até que para cada **k=k***, o problema tenha solução e, para **k=k+ε, ε>0**, esse problema não tenha solução.

Desta forma, o valor mínimo global viável da função objetivo estará no intervalo:

$$L=(k^*,k^*+\varepsilon]\subset\Re$$

E o ponto de mínimo global viável do problema será qualquer ponto **x** $\in \Re^n$, tal que:

$$k^*<f(x)\le k^*+\varepsilon$$

A partir daí devemos verificar se a inequação **f(x)≤k** tem solução dentro do domínio viável. Evidentemente que isto traz suas dificuldades, devido a verificação de uma inequação à multiplas variáveis seguidas vezes.

Esta dificuldade porém, pode ser contornada construindo-se um novo problema do tipo:

Problema (2)

Minimizar: $\sum_{i=1}^{n} x_i$

Sujeito a :

$g_i(x) \leq 0, i=1,2, ..., m$

$h_j(x) = 0, j=1,2, ..., p$

$f(x) \leq k$

$x \in X \subset \Re^n$

Assim, está-se procurando verificar se existe algum $\mathbf{x} \in \Re^n$ que, simultaneamente satisfaça:

(i) As restrições do problema original (**problema (1)**), isto é **x** é viável;

(ii) **f(x)≤k**, isto é, se esta inequação tem solução.

Deve-se notar que a função objetivo do **problema (2)** foi escolhida arbitrariamente, e poderia muito bem ser de outra forma, como por exemplo:

$$f(x) = \sum_{i=1}^{n} x_i^2$$

Teoricamente, este método conduz ao ótimo global do problema, mas tem a desvantagem de exigir muitas iterações isto é, muitas resoluções sucessivas do mesmo problema, uma para cada valor de **k**, além disso acrescido de uma restrição.

Na evolução do presente raciocínio, surge então a idéia de, ao invés de , para cada iteração, fixar um novo valor de **k** e resolver novamente o **problema (2)**, transformar **k** em uma nova variável do problema, isto é:

$k = x_{n+1}$

e transformar a função objetivo do **problema (2)** em:

$H(x) = x_{n+1}$

Assim, passa-se a outro problema de programação matemática, desta vez:

Problema (3)

Minimizar: x_{n+1}

Sujeito a:

$$g_i(x) \leq 0, \; i = 1, 2, ..., m$$

$$h_j(x) = 0, \; j = 1, 2, ..., p$$

$$f(x) \leq x_{n+1}$$

$$x \in X \subset \Re^{n+1}$$

Isto significa que ao problema de programação matemática original **problema (1)**, foi acrescida uma variável x_{n+1} e uma restrição $f(x) \leq x_{n+1}$, além de a função objetivo também ter sido modificada, passando a ser linear.

Evidentemente, nenhuma das duas idéias acima garante que se obtenha facilmente o mínimo global para o **problema (1)**. Porém, a conjugação de ambas as idéias, como apresentado abaixo, pode trazer bons resultados.

Dado um problema de programação matemática na forma do **problema (1)**, transformá-lo em outro problema, agora na forma de **problema (3)**, e resolvê-lo.

Seja $\mathbf{x}^* \in \Re^{n+1}$ a solução ótima obtida, que pode ser global ou não. A fim de prosseguir na busca de uma solução ótima global para o **problema (1)**, pode-se tentar uma nova redução no valor da função objetivo do **problema (1)**, acrescentando ao **problema (3)** uma nova restrição, esta da forma:

$$f(x) \leq x^*_{n+1} - \varepsilon, \; \varepsilon > 0$$

que pode ser incorporada à restrição $\mathbf{f(x)} \leq \mathbf{x}_{n+1}$, o que resulta em:

$$f(x) \leq (x^*_{n+1} - \varepsilon + x_{n+1})/2, \; \varepsilon > 0$$

Em outras palavras,passa-se do **problema (3)** para:

Problema (4)

Minimizar: x_{n+1}

Sujeito a:

$$g_i(x) \leq 0, \; i = 1, 2, ..., m$$

$h_j(x) = 0, j = 1, 2, ..., p$

$f(x) \leq (x_{n+1} + x^*_{n+1} - \varepsilon)/2$, $\varepsilon > 0$

$x \in X \subset \Re^{n+1}$

Obs.: Notar que $\mathbf{x}^*_{n+1}$ é uma constante.

Conclusão:

(i) Se o **problema (4)** revelar-se inviável é porque não é mais possível reduzir o valor de **f(x)** abaixo de $\mathbf{x}^*_{n+1}$, é o último valor $\mathbf{x}^* \in \Re^{n+1}$ obtido é uma solução ótima global.

(ii) Caso seja ainda possível resolver o **problema (4)** repetir o procedimento, para isto construindo-se um novo problema de programação matemática, agora com o novo valor obtido para $\mathbf{x}^*_{n+1}$.

6. APLICAÇÕES

6.1. Apresentação dos Modelos

O algoritmo foi utilizado na resolução de problemas de *construção patogênica*, encontrados nas referências [5], [12] e [14]. Estes problemas além da suas dificuldades naturais são multimodais, o que reforça os motivos da utilização desta nossa técnica.

Citaremos alguns dos problemas pesquisados, apresentando também os mesmos em sua nova forma, após terem passado pelo pelas transformações que o algoritmo apregoa.

6.2. Problemas

<u>Teste 1.- Forma Original:</u> *(Levy-Gomez(1985, p.243, Problem 2)* [5]

Minimizar: $f(x) = 0.1(x_1^2 + x_2^2)$

Sujeito a: $2\text{sen}(2\pi x_2) - \text{sen}(4\pi x_1) \leq 0$

Observação: o problema enunciado trabalha sobre o domínio $([-1,1],[-1,1])^T$, possuindo 24 mínimos locais e tem apenas um mínimo global em $(0,0)^T$, com f(x*)=0.

<u>*Teste 1.- Pelo Algoritmo:*</u>

Minimizar: $f(x) = x_3$

Sujeito a: $$\begin{aligned} &0.1(x_1^2 + x_2^2) \le x_3 \\ &2\text{sen}(2\pi x_2) - \text{sen}(4\pi x_1) \le 0 \end{aligned}$$

<u>*Teste 2.- Forma Original :*</u> *(Zowe (1985), p.350)* [5]

Minimizar: $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$

Sujeito a: $$\begin{aligned} &x_1^2 - x_2 \le 0 \\ &x_1 + x_2 - 2 \le 0 \end{aligned}$$

*Observação:*com $x \in \Re^2$, onde o mínimo está localizado no ponto $\mathbf{x^*}=(1,1)^T$, cuja imagem ótima global é $\mathbf{f(x^*)}=1$.

<u>*Teste 2.- Pelo Algoritmo:*</u>

Minimizar: x_3

Sujeito a: $$\begin{aligned} &(x_1 - 2)^2 + (x_2 - 1)^2 \le x_3 \\ &x_1^2 - x_2 \le 0 \\ &x_1 + x_2 - 2 \le 0 \end{aligned}$$

<u>*Teste 3.- Forma Original:*</u> *(Branin (1972)).* [12]

Minimizar: $f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$

Observação: o domínio é dado por $([-5,5],[-5.5])^T$, os mínimos globais estão estão $\mathbf{x^*}$=(0.08983,-0.7126) e $\mathbf{x^*}$=(-0.08983, 0.7126).

<u>*Teste 3.- Pelo Algoritmo:*</u>

Minimizar: $f(x) = x_3$

Sujeito a: $4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 \le x_3$

<u>*Teste 4.- Forma Original:*</u> *(Rastrigin (1974)).* [12]

Minimizar: $f(x) = x_1^2 + x_2^2 - \cos 18x_1 - \cos 18x_2$

Observação: o domínio é dado entre $([-1,1],[-1,1])^T$. A função possui um total de 50 mínimos locais , tendo como mínimo global único as coordenadas **x**=(0,0), sendo **f(x*)**=-2.

Teste 4. Pelo Algoritmo:

Minimizar: $f(x) = x_3$

Sujeito a: $x_1^2 + x_2^2 - \cos 18x_1 - \cos 18x_2 \leq x_3$

Teste 5.- Forma Original: (Horst (1990)). [14]

Minimizar: $f(x) = 4x_1^4 + 2x_2^2 - 4x_1^2$

Sujeito a: $x_1^2 - 2x_1 - 2x_2 - 1 \leq 0$

Observação: o domínio do problema é dado $([-1,1],[-1,1])^T$, sendo o ótimo dado em **x***=(0.7197,0.0000), com **f(x*)**=-0.9987.

Teste 5.- Pelo Algoritmo:

Minimizar: $f(x) = x_3$

$$\text{Sujeito a: } \begin{array}{l} 4x_1^2 + 2x_2^2 - 4x_1^2 \leq x_3 \\ x_1^2 - 2x_1 - 2x_2 - 1 \leq 0 \end{array}$$

Teste 6.- Forma Original: (Horst (1990)). [14]

Minimizar: $f(x) = (x_1^2 + x_2 + x_3) - (x_1 + x_2^2 - x_3)$

$$\text{Sujeito a: } \begin{array}{l} (x_1 - x_2 - 1.2)^2 + x_2 \leq 4.4 \\ x_1 + x_2 + x_3 \leq 6.5 \\ x_1 \leq 1.4 \\ x_2 \leq 1.6 \\ x_3 \leq 1.8 \end{array}$$

Observação: o mínimo global é encontrado em **x***=(1.400, 1.8128, 1.800).

Teste 6.- Pelo Algoritmo:

Minimizar: $f(x) = x_4$

$$\text{Sujeito a: } \begin{array}{l} (x_1^4 + x_2 + x_3) - (x_1 + x_2^2 - x_3) \le x_4 \\ (x_1 - x_2 - 1.2)^2 + x_2 \le 4.4 \\ x_1 + x_2 + x_3 \le 6.5 \\ x_1 \le 1.4 \\ x_2 \le 1.6 \\ x_3 \le 1.8 \end{array}$$

Teste 7.- Forma Original: (Imagem 3D-Surface).

Minimizar: $f(x) = e^{-0.5r}.\cos r$, onde $r = \sqrt{x_1^2 + x_2^2}$

*Observação:*o mínimo é procurado em $([-2,2],[-2,2])^T$ e o mínimo global está em **x***=(0,0), com **f***=1.

Teste 7.- Pelo Algoritmo:

Minimizar: $f(x) = x_3$

Sujeito a: $e^{-0.5r}.\cos r \le x_3$, onde $r = \sqrt{x_1^2 + x_2^2}$

Observação: Todos os problemas acima apresentados possuem duas versões, uma na sua forma original a outra na sua forma modificada pelo algoritmo em questão. Os problemas modificados poderiam ainda sofrer modificações para para a forma : Minimizar : $f(x) = x_i^2, x_i^3, \ldots.$

6.3. Resultados

Em todos os testes efetuados o algoritmo mostrou-se eficaz, resultando sempre em valores também obtidos nas referências consultadas (vide referenciação bibliográfica).

A caracterização do mínimo global inicialmente segue o critério geral dos mínimos locais, que é o "critério do gradiente zero". A garantia de que os mínimos globais foram atingidos se baseia primeiramente nas condições de otimalidade de Kunh-Tucker, que foram atendidas, mesmo quando iniciamos o problema sem que todas as restrições fossem convexas. Com o início dos cálculos essa hipótese é naturalmente forçada com a minimização feita pelo hiperplano. Forçando os pontos de mínimo a se localizarem dentro de regiões convexas, dando-nos assim garantia de otimalidade. Como já colocamos anteriormente as condições de Kunh-Tucker não são suficientes para classificar um mínimo como sendo global assim, apelamos aos conjuntos de nível de Zang e Avriel, que consideram um conjunto de níveis $L_f(\alpha)=\{x / x \in A,\ f(x) \le \alpha\}$, onde o domínio efetivo está mapeado por $G_f=\{\alpha / \alpha \in \Re, L_f(\alpha) \neq 0\}$. O ponto mapeado para o conjunto $L_f(\alpha)$ é dito *semi-continuo-superior,* para $\overline{\alpha} \in G_f$ se $x \in L_f(\overline{\alpha}), \{\alpha^i\} \subset G_f, \{\alpha^i\} \to \overline{\alpha}$ implica na existência de um número natural K e da seqüência $\{x^i\}$ tal que $\{x^i\} \to x$ e $x^i \in L_f(\alpha^i)$ para i≥K. Isto caracteriza condição suficiente para a classificação do mínimo como sendo mínimo global. O leitor verificará que nosso algoritmo sugere este teste ao final da apresentação do algoritmo.

Para a otimização não-linear utilizamos o pacote matemático ***Mercury®***, que é da criação de Roger Schlafly (Real Software, PO BOX 1680, Soquel, CA 95073 USA), que usa um método iterativo não especificado para a resolução. Um fato interessante aqui constatado, foi o de quando as novas funções objetivo sofriam modificações para a nova variável, esta com um expoente maior a 2 (dois), o tempo computacional reduzia-se sensivelmente. Como não temos nenhuma explicação suficientemente bem elaborada, acreditamos que se tratem de questões internas de pacote em uso . Para o caso da otimização linear foi utilizado o pacote ***PLM®***, que é da criação de Cláudio Loesch (FURB-Blumenau-SC), que utiliza o método Simplex sobre uma planilha de cálculos, e para a visualização (quando possível) dos problemas foi usado o pacote ***3Dsurf®***.

Para a verificação dos resultados usamos algoritmos genéticos sobre os problemas em teste, porém deixando claro que esta verificação não serve de comprovação de otimalidade global.

6.4. Limitações

Mesmo o algoritmo tendo dado bons resultados, não podemos considerá-lo como totalmente confiável, pois o mesmo necessitaria ser utilizado em problemas de porte maior, os quais não encontramos nas literaturas consultadas, porém confiamos que não deva ser difícil serem construídos a partir de situações práticas.

Outro problema que deve ser levado em consideração, é o detalhe da utilização do algoritmo em problemas de otimização, cujas restrições não sejam convexas, sendo bem possível nesses casos o algoritmo apresentar fracassos, porém nada queremos adiantar neste sentido .

Por último tivemos alguns problemas de inicialização, ou seja, com o uso de um *startpoint* conveniente. Teoricamente isto não deveria ser nenhum impecílio, pois conforme nosso hiperplano é otimizado o problema acaba sendo relaxado das restrições iniciais do problemas original. Acreditamos pois, ser isto um problema de algoritmo numérico usado pelo computador.

Referências Bibliográficas

[1]-BAJPAI, A.C. e outros. Matemática Avançada para Engenharia. Tradução: Luiz Roberto de Godoi Vidal, São Paulo, Editora Hemus, 1980.

[2]-BAZARAA. M.S., SHETY, C.M. Nonlinear Programming-Theory and Algorithms USA, John Wiley and Sons,1979.

[3]-EHRLICH, Jacques P. Introdução a Pesquisa Operacional. São Paulo, Editora Atlas, 1991.

[4]-GOLDBERG, David E. Genetic Algorithms in Search, Optimization and Machine Learnig.Massachusets-USA, Addison Wesley Publishing Company Inc.,

[5]-HORST, R. e outros. On Solving a D.C. Programming Problem by a Sequence of Linear Programs.Journal of Global Optimization, vol.1, No.2, 1991.

[6]-KREYZIG, Erwin. Advanced Engeneering Mathematics. USA. VI Edition, John Wiley and Sons, 1988.

[7]-LANZER, Edgar A. Programação Linear: Conceitos e Aplicações. Rio de Janeiro, IPEA/INPES, 1988.

[8]-LIMA. Elon Lages. Topologia Geral. Rio de Janeiro, Livros Técnicos e Científicos Editora S.A., 1970.

[9]-LIPSCHUTZ, Seymour. Topologia Geral. Tradução: Alfredo Alves Farias. São Paulo, McGraw-Hill do Brasil, 1979.

[10]-MAHEY, Philippe. Programação Não-Linear. Rio de Janeiro, Editora Campus; CNPq, 1987.

[11]-McNITT, Lawrence L. Simulação em Basic. Tradução: Maria Cristina de Lima Villares, Rio de Janeiro, Livros Técnicos e Científicos Editora, 1985.

[12]-RATSCHECK, H. and ROJNE, J. New Computer Methods for Global Optimization. Great Bretain, John Wiley and Sons, 1988.

[13]-STANGE, Plínio. Algumas Considerações Sobre Técnicas Matemáticas de Otimização. (Monografia).UFSC-Florianópolis-SC, 1993.

[14]-TÖRN, A.and ZILINSKAS, A. Global Optimization. Heidelberg, Springer-Verlag, 1989.

ANEXOS

ANEXO 1

Pesquisa Multidimensional Discreta

1. O método de Hoocke e Jeeves

Algoritmo:

Inicialização: Sejam $d_1,...,d_n$ as coordenadas das direções. Escolha um escalar ε>0 que será usado como tolerância para terminar o algoritmo. Também escolha o tamanho do passo inicial, Δ≥ε, e um fator de aceleração α>0. Escolha um ponto se sáida x_1, faça $y_1 = x_1$, faá k=j=1, e vá para o passo principal.

Passo Principal:
1. Se $f(y_j + \Delta d_j) < f(y_j)$, denominamos este caminho de *sucesso*; fazemos $y_{j+1} = y_j + \Delta d_j$, e seguimos para o passo 2. Se porém, $f(y_j + \Delta d_j) \geq f(y_j)$, este caminho denominamos de *fracasso*. Neste caso vá para o passo 2; se $f(y_j - \Delta d_j) \geq f(y_j)$, fazemos $y_{j+1} = y_j - \Delta d_j$, e vá para o passo 3.
2. Se j<n, passe j para j+1, e repita o passo 1. Caso contrário vá para o passo 3 se $f(y_{n+1}) < f(x_k)$, e vá para o passo 4 se $f(y_{n+1}) \geq f(x_k)$.
3. Faça $x_{k+1} = y_{n+1}$ e $y_1 = x_{k+1} + \alpha(x_{k+1} - x_k)$. Passe k para k+1,faça j=1 e vá para o passo 1.
4. Se Δ≤ε pare; x_k é a solução. Caso contrário, passe Δ para Δ/2. Faça $y_1 = x_k, x_{k+1} = x_k$, passe k para k+1, faça j=1 e repita o passo 1.

Pesquisa Multidimensional Contínua

2. O Método Steepest Descent

Algoritmo:

Inicialização: Seja ε>0 uma tolerância dada. Escolha um ponto de partida x_1 , faça k =1, e vá para o passo principal.

Passo Principal: Se $\|\nabla f(x_k)\| \leq \varepsilon$, pare; caso contrário, faça $d_k = -\nabla f(x_k)$, e determine λ_k ótimo do problema de de minimização $f(x_k + \lambda d_k)$ sujeito a λ≥0. Faça $x_{k+1} = x_k + \lambda d_k$, passe k para k+1, e repita o passo principal.

ANEXO 2

Otimização Restrita

Várias técnicas de otimização são especificamente planejadas para trabalhar o problema restrito; entre elas estão os métodos da projeção do gradiente e do gradiente reduzido.

Outras técnicas envolvem modificações dos métodos não restritos. A primeira categoria destas leva em consideração restrições explicitamente em cada passo. Nesta categoria podemos distinguir duas subcategorias: os métodos de pesquisa direta e os métodos dos gradientes. Os métodos de pesquisa direta fazem sentido quando o ponto presente está próximo de um contorno restritivo e modificam conseqüentemente as direções de pesquisa. Os métodos dos gradientes são variações do método de maior descida e modificam, se necessário, a direção do gradiente de forma que o novo ponto esteja dentro da região provável. Entretanto, os gradientes dentro da região não provável podem ser confiáveis. se as restrições forem complicadas então ambos os tipos de métodos nesta primeira categoria podem tornar-se ineficientes.

A segunda categoria compreende aqueles métodos que utilizam funções de penalidades. Aqui a idéia essencial é adicionar a seqüência de funções penalidades à função objetiva básica de modo que as soluções para a seqüência de novas funções objetivas não restritas tendam à solução dos problemas restritos. A dificuldade aqui é que as funções de penalidades podem introduzir características inusuais no problema. Quando uma função que penalize violações de restrição for adicionada, será chamada de função de penalidade exterior. Para restrições do tipo $h_i(x) = 0$ podemos tomar funções de penalidades da forma $r^{-1}[h_i(x)]^2$. Uma seqüência de valores de r é escolhida, desde que tenda a zero, entretanto a pesquisa prossegue. Para restrições da forma $g_i(x) \leq 0$, funções de barreiras são escolhidas, sendo diferentes de zero no interior da região viável e infinitas em seus contornos.

Seria impossível fazer justiça ao tema dos métodos restritos sem entrar no tema extensivamente. Aconselhamos ao leitor interessado a leitura de [2].

ANEXO 3

PROBLEMAS

```
      C:GLOBMIN.EKA  Line 10   Col 1   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
;Problema 1
Minimize f
f =z
0.1(x^2+y^2)<=z
2*sin(2*pi*y)-sin(4*pi*x)<=0
-1<=x<=1
-1<=y<=1
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f        = z
           = +4.8114612091430000E-14
  z        = +4.81146120914300E-14
  x        = +1.019849010798355E-10
  y        = -1.019332269492119E-10

Residuals and derived equations:
  { 0 }  f=z
  { INF }  0.1*(x^2+y^2)<=z
  { INF }  2*SIN(6.28318530717959*y)-SIN(12.56637061435917*x)<=0
  { 0 }  -1<=x
  { 0 }  x<=1
  { 0 }  -1<=y
  { 0 }  y<=1
  { 4.8E-14 }  0.1*(x^2+y^2)+T2-z=0
  { 1.3E-14 }  2*SIN(6.28318530717959*y)-SIN(12.56637061435917*x)+T3-0=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
      C:GLOBMIN.EKA  Line 8    Col 9   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
;Problema 2
Minimize f
f =z
(x-2)^2+(y-1)^2<=z
x^2-y<=0
x+y-2<=0
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f          = z
             = +0.99999999994983024
  z          = +0.999999999949830
  x          = +1.00000000002506
  y          = +1.00000000005800

Residuals and derived equations:
  { 0 }  f=z
  { INF }  (x-2)^2+(y-1)^2<=z
  { INF }  x^2-y<=0
  { INF }  x+y-2<=0
  { 4.1E-14 }  (x-2)^2+(y-1)^2+T2-z=0
  { 7.9E-12 }  x^2-y+T3-0=0
  { 8.3E-11 }  x+y-2+T4-0=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
      C:GLOBMIN.EKA  Line 7    Col 1   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
;Problema 3
Minimize f
f =z
4*x^2-2.1*x^4+(1/3)*x^6+x*y-4*y^2+4*y^4<=z
-5<=x<=5
-5<=y<=5
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f          = z
             = -1.0316284534898774
  z          = -1.03162845348988
  x          = -0.0898420131003181
  y          = +0.712656403020740

Residuals and derived equations:
  { 0 }  f=z
  { INF }  4*x^2-2.1*x^4+0.333333333333333*x^6+x*y-4*y^2+4*y^4<=z
  { 0 }  -5<=x
  { 0 }  x<=5
  { 0 }  -5<=y
  { 0 }  y<=5
  { 6.7E-17 }  4*x^2-2.1*x^4+0.333333333333333*x^6+x*y-4*y^2+4*y^4+T2-z=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
      C:GLOBMIN.EKA  Line 9    Col 1   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
; Problema 4
Minimize f
f =z
x^2+y^2-cos(18*x)-cos(18*y)<=z
-1<=x<=1
-1<=y<=1
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f        = z
           = -2.0000000000000000
  z        = -2.000000000000000
  x        =  0.0000000000000000
  y        =  0.0000000000000000

Residuals and derived equations:
  { 0 }  f=z
  { INF }  x^2+y^2-COS(18*x)-COS(18*y)<=z
  { 0 }  -1<=x
  { 0 }  x<=1
  { 0 }  -1<=y
  { 0 }  y<=1
  { 0 }  x^2+y^2-COS(18*x)-COS(18*y)+T2-z=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
      C:GLOBMIN.EKA  Line 10   Col 1   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
; Problema 5
Minimize f
f =z
4*x^4+2*y^2-4*x^2<=z
x^2-2*x-2*y-1<=0
-1<=x<=1
-1<=y<=1
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f          = z
             = -0.99999999999999944
  z          = -1.0000000000000
  x          = +0.707106781186406
  y          = +3.46382110700264E-14

Residuals and derived equations:
  { 0 }  f=z
  { INF }  4*x^4+2*y^2-4*x^2<=z
  { INF }  x^2-2*x-2*y-1<=0
  { 0 }  -1<=x
  { 0 }  x<=1
  { 0 }  -1<=y
  { 0 }  y<=1
  { 5.5E-16 }  4*x^4+2*y^2-4*x^2+T2-z=0
  { 6.4E-16 }  x^2-2*x-2*y-1+T3-0=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
     C:GLOBMIN.EKA  Line 5    Col 7   Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
; Problema 6
Minimize f
f =z^2
(x^4+y+w)-(x+y^2-w)<=z
(x-y-1.2)^2+y<=4.4
x+y+y<=6.5
x>=1.4
y>=1.6
w>=1.8
```

```
     C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f          = z^2
             = +20.946405419164421
  z          = +4.57672431102906
  x          = +1.40000000000000    { = +7 / 5 }
  y          = +1.809542459198577
  w          = +1.80000000000000    { = +9 / 5 }

Residuals and derived equations:
  { 0 }  f=z^2
  { INF }  x^4+y+w-(x+y^2-w)<=z
  { INF }  (x-y-1.2)^2+y<=4.4
  { INF }  x+y+y<=6.5
  { 0 }  1.4<=x
  { 0 }  1.6<=y
  { 0 }  1.8<=w
  { 2.6E-05 }  x^4+y+w-(x+y^2-w)+T2-z=0
  { 0.00017 }  (x-y-1.2)^2+y+T3-4.4=0
  { 0.00096 }  x+y+y+T4-6.5=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

```
      C:GLOBMIN.EKA  Line 6    Col 10  Insert Indent Tab
; Otimizacao Global Deterministica - Um Algoritmo.
; Autor: Nelson Hein.
; Problema 7
Maximize f
f =z
cos( r )*exp( -0.5*r ))>=z
r=sqrt( x^2+y^2 )
-2<=x<=2
-2<=y<=2
```

```
      C:SOLUTION.EKA Line 1    Col 1   Insert Indent Tab
Variables:
  f         = z
            =  0.0000000000000000
  z         =  0.0000000000000000
  r         = SQRT( x^2+y^2 )
            =  0.0000000000000000
  x         =  0.0000000000000000
  y         =  0.0000000000000000

Residuals and derived equations:
  { 0 }  f=z
  { INF }  z<=COS( SQRT( x^2+y^2 ))*EXP( -0.5*SQRT( x^2+y^2 ))
  { 0 }  r=SQRT( x^2+y^2 )
  { 0 }  -2<=x
  { 0 }  x<=2
  { 0 }  -2<=y
  { 0 }  y<=2
  { 0 }  z+T2-COS( SQRT( x^2+y^2 ))*EXP( -0.5*SQRT( x^2+y^2 ))=0

Largest residual is INF
Method: Iterative
Numeric processor: Emulator
```

ANEXO 4

PROGRAMAS

ALGORITMOS GENÉTICOS

```
program SGA1;
{
  Implementacao do Algoritmo Genetico SGA1, conforme capitulo 3 do livro
  'Genetic Algorithms in Search, Optimization, and Machine Learning', de
  David E. Goldberg. Adaptado por: Nelson Hein.
}
uses wincrt;

{ global.pas }

const
  MAXPOP    = 100;        { Tamanho maximo da populacao }
  MAXSTRING = 40;         { Tamanho maximo do cromossomo (maior locus)}

type
  allele     = boolean;                    { allele = posicao do bit }
  chromosome = array[1..MAXSTRING] of allele; { cadeia de alelos (bits) }
  individual = record
      chrom  :chromosome;           { Genotipo = string de bits }
      x1, x2  :real;              { Fenotipo = inteiro sem sinal }
      fitness :real;             { Valor da funcao objetivo }
      parent1, parent2, xsite :integer; { pais e ponto de cruzamento }
    end;
  population = array[1..MAXPOP] of individual; { uma populacao }

var
  oldpop, newpop  : population; { Duas populacoes nao-sobrepostas}
  popsize,  { tamanho da populacao }
  lchrom,   { comprimento do cromossomo }
  gen,      { variavel contadora do n§ de geracoes }
  maxgen    { n§ total de geracoes }               : integer;
  pcross,   { probabilidade de um overcross }
  pmutation, { probabilidade de uma mutacao }
  sumfitness { soma das adequacoes de uma populacao }    : real;
  nmutation, ncross  :integer;      { Numero de mutacoes e de overcross }
  avg, max, min     :real;        { Estat¡sticas: media, m ximo, m¡nimo}
  filename          :string[8];   { Nome do arquivo de saida }
  buffer     : text;             { Ponteiro para arquivo texto de sa¡da }

{ interfac.pas }

function objfunc(x1, x2 : real):real;
{
  Usa a funcao de adequacao que retorna x ao quadrado.
}
  var
    z, r: real;
  begin
    r := sqrt(x1*x1 + x2*x2);
    z := exp(-0.5*r)*cos(r);
    if z < 0 then z := 0;
    objfunc := z;
```

```
  end;

function decode(indvar : integer; chrom:chromosome; lbits:integer):real;
{
   Calcula o fen¢tipo de um individuo, pegando a string de bits do
   cromossomo e retornando o seu valor em decimal.
}
  var
    j :integer;
    accum, powerof2  :real;
  begin
    accum := 0;
    powerof2 := 1;
    if indvar = 1 then
      begin
        for j := 2 to 20 do
          begin
            if chrom[j] then
              accum := accum + powerof2;
            powerof2 := powerof2 / 2;
          end;
        if chrom[1] then accum := -accum;
      end;
    if indvar = 2 then
      begin
        for j := 22 to 40 do
          begin
            if chrom[j] then
              accum := accum + powerof2;
            powerof2 := powerof2 / 2;
          end;
        if chrom[21] then accum := -accum;
      end;

    decode := accum;
  end;

{ randomic.pas }

function flip(probability:real):boolean;
{
  Retorna verdadeiro se valor aleatorio sorteado ocorreu com sucesso
  numa distribuicao de Bernoulli de probab. probability.
}
  begin
    if ( probability = 1.0) then
      flip := true
    else
      flip := (random <= probability);
  end;

function rnd(low,high:integer):integer;
```

```
{
  Retorna um numero inteiro aleatorio, com distribuicao equiprovavel,
  tal que low <= rnd <= high.
}
  begin
    if (low < high) then
      rnd := random(high-low+1) + low
    else
      rnd := low;
  end;

{ triops.pas }

function select(popsize:integer; sumfitness:real; pop:population):integer;
{
  Retorna o numero de um individuo selecionado na populacao, sorteado em
  proporcao ao seu valor de adequacao na populacao. Finalidade:
  implementar o operador de reprodução, favorecendo os mais adequados.
  Paramtros:  popsize = tamanho da população;
              sumfitness = soma dos valores de adequacao da populacao;
              pop = populacao de onde se far  a selecao.
}
  var
    j                : integer;
    rand, partsum    : real;
  begin
    partsum := 0;
    rand := random * sumfitness;
    j := 0;
    repeat
      j := j + 1;
      partsum := partsum + pop[j].fitness; { fornece a distr. acumulada }
    until (partsum >= rand) or (j >= popsize);
    select := j;
  end;

function mutation(alleleval:allele; pmutation:real;
                  var nmutation:integer):allele;
{
  Sorteia um numero aleatorio para ver se h  mutacao, de acordo com a
  probabilidade pmutation. Se houver, o valor logico do alelo a trocado
  e o contador nmutation e incrementado. Retorna o novo valor do alelo.
}
  var
    mutate : boolean;
  begin
    mutate := flip(pmutation);
    if (mutate) then
      begin
        nmutation := nmutation + 1;
        mutation := not(alleleval);
      end
```

```
    else
      mutation := alleleval;
  end;

procedure crossover(var parent1, parent2, child1, child2:chromosome;
            var lchrom, nchrom, nmutation, jcross:integer;
            var pcross, pmutation:real);
{
  gera dois cromossomos descendentes a partir de dois cromossomos pais,
  atraves do crossover. O local do crossover e previamente sorteado.
  Durante a copia dos segmentos, em cada alelo sao tambem sorteadas ( e
  eventualmente efetuadas ) possiveis mutacoes.
}
  var
    j : integer;
  begin
    if (flip(pcross)) then
      begin
        jcross := rnd(1,lchrom-1);
        ncross := ncross+1;
      end
    else
      jcross := lchrom;
    for j:=1 to jcross do begin
      child1[j] := mutation(parent1[j],pmutation,nmutation);
      child2[j] := mutation(parent2[j],pmutation,nmutation);
    end;
    if (jcross <> lchrom) then
      for j := jcross + 1 to lchrom do
      begin
        child1[j] := mutation(parent2[j],pmutation,nmutation);
        child2[j] := mutation(parent1[j],pmutation,nmutation);
      end;
  end;

{ generate.pas }

procedure generation;
{
  Gera pares de descendentes para a nova populacao aos pares, a partir de
  pares de pais selecionados da populacao velha. Para cada descendente,
  completa as informacoes de adequacao de acordo com o seu genotipo, quais
  sao os pais, e em que ponto ocorreu o crossover.
}
  var
    j, mate1, mate2, jcross : integer;
  begin
    j:=1;
    repeat
      mate1 := select(popsize,sumfitness,oldpop);
      mate2 := select(popsize,sumfitness,oldpop);
      crossover(oldpop[mate1].chrom, oldpop[mate2].chrom,
```

```
              newpop[j].chrom, newpop[j+1].chrom,
              lchrom, ncross, nmutation, jcross, pcross, pmutation);
        with newpop[j] do begin
          x1 := decode(1,chrom,lchrom);
          x2 := decode(2,chrom,lchrom);
          fitness := objfunc(x1,x2);
          parent1 := mate1;
          parent2 := mate2;
          xsite := jcross;
        end;
        with newpop[j+1] do begin
          x1 := decode(1,chrom,lchrom);
          x2 := decode(2,chrom,lchrom);
          fitness := objfunc(x1,x2);
          parent1 := mate1;
          parent2 := mate2;
          xsite := jcross;
        end;
        j := j + 2;
      until (j > popsize);
  end;

{ stats.pas }

procedure statistics(popsize:integer; var max,avr,min,sumfitness:real;
              var pop:population);
{
   Produz as seguintes estatisticas a partir de popsize e pop ( populacao
   sendo considerada): sumfitness = soma de todos os valores de adequacao;
   max, avr, min = m ximo, minimo e media dos valores de adequacao.
}
  var
    j : integer;
  begin
    sumfitness := pop[1].fitness;
    min := pop[1].fitness;
    max := pop[1].fitness;
    for j:=2 to popsize do
      with pop[j] do begin
        sumfitness:=sumfitness + fitness;
        if (fitness > max) then
          max := fitness;
        if (fitness < min) then
          min := fitness;
      end;
    avg := sumfitness/popsize;
  end;

{ initial.pas }

procedure initdata;
  var
```

```
    ch    :char;
    j     :integer;
  begin
    clrscr;
    writeln('
ÚÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ¿')
;
    writeln(' ³                                    ³');
    writeln(' ³  Um Algoritmo Genetico Simples- SGA1- Versao 1.0 ³');
    writeln(' ³                                    ³');
    writeln('
ÀÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÙ')
;
    writeln;
    writeln('   ********** Inicializacao dos dados **********');
    writeln;
    write('Entre nome arquivo de saida ---> ');readln(filename);
    write('Entre tamanho da populacao ----> ');readln(popsize);
    write('Entre comprim. do cromossomo --> ');readln(lchrom);
    write('Entre max. de geracoes --------> ');readln(maxgen);
    write('Entre probab. de crossover ----> ');readln(pcross);
    write('Entre probab. de mutacao ------> ');readln(pmutation);
    writeln;
    writeln('Favor esperar conclusao do processamento...');
    randomize; { Para sempre gerar uma sequecia de n§s aleat¢rios difer.}
    nmutation := 0;  ncross := 0;
  end;

procedure initreport(var buffer:text);
begin
  assign(buffer,filename+'.out');
  rewrite(buffer);
  writeln(buffer,'
ÚÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ¿')
;
  writeln(buffer,' ³                                    ³');
  writeln(buffer,' ³  Um Algoritmo Genetico Simples- SGA1- Versao 1.0 ³');
  writeln(buffer,' ³                                    ³');
  writeln(buffer,'
ÀÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÙ')
;
  writeln(buffer);
  writeln(buffer,'   Parametros assumidos:');
  writeln(buffer);
  writeln(buffer,'   Arquivo de saida -----------> ',filename,'.out');
  writeln(buffer,'   Tamanho da populacao--------> ',popsize);
  writeln(buffer,'   Comprimento do cromossoma---> ',lchrom);
  writeln(buffer,'   Numero m ximo de geracoes---> ',maxgen);
  writeln(buffer,'   Probabil. de crossover------> ',pcross:5:3);
  writeln(buffer,'   Probabil. de mutacao--------> ',pmutation:5:3);
  writeln(buffer);
  writeln(buffer,'   Estatisticas da Populacao Inicial:');
```

```
  writeln(buffer);
  writeln(buffer,'  Maximo da populacao inicial---> ',max);
  writeln(buffer,'  Media  da populacao inicial---> ',avg);
  writeln(buffer,'  Minimo da populacao inicial---> ',min);
  writeln(buffer,'  Soma das adequacoes popul. ---> ',sumfitness);
  writeln(buffer);  writeln(buffer);
end;

procedure initpop;
{
  Inicializa os dados para a primeira populacao do processo.
}
  var
    j, i : integer;
  begin
    for j := 1 to popsize do
      with oldpop[j] do
        begin
          for i := 1 to lchrom do
            chrom[i] := flip(0.5); { 50% chance para 1 e para 0 }
          x1 := decode(1,chrom,lchrom);
          x2 := decode(2,chrom,lchrom);
          fitness := objfunc(x1,x2);
          parent1 := 0;
          parent2 := 0;
          xsite := 0;
        end;
  end;

procedure initialize(var buffer:text);
  begin
    initdata;
    initpop;
    statistics(popsize,max,avg,min,sumfitness,oldpop);
    initreport(buffer);
  end;

{ report.pas }

procedure writechrom(var buffer:text; chrom:chromosome; lchrom:integer);
{
  Escreve os genes de um cromossomo como uma lista de 0s e 1s.
}
  var
    j : integer;
  begin
    for j := 1 to lchrom do
    begin
      if (chrom[j]) then
        write(buffer,'1')
      else
        write(buffer,'0');
```

```
    if j = lchrom/2 then write(buffer,' ');
   end;
  end;

procedure makereport(var buffer:text; gen:integer);
  var
   j : integer;
  begin
   writeln(buffer,'ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ');
   writeln(buffer,'     Relatorio da Populacao ');
   writeln(buffer,'     Geracoes: ',gen-1:2,' e ',gen:2);
   writeln(buffer,'ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ');
   writeln(buffer);
   write  (buffer,' #   String');
   for j := 1 to lchrom-3 do
    write(buffer,' ');
   writeln(buffer,' X1        X2        Adequ.');
   writeln(buffer);
   for j := 1 to popsize do begin
    write(buffer,j:2,') ');
    with oldpop[j] do
     begin
      writechrom(buffer,chrom,lchrom);
      writeln(buffer,x1:10:6,x2:10:6,fitness:10:4);
     end;
   end;
   writeln(buffer);
   write  (buffer,' #   String');
   for j:=1 to lchrom-3 do
    write(buffer,' ');
   writeln(buffer,' X1        X2        Adequ.      Pais Crossover');
   writeln(buffer);
   for j:=1 to popsize do begin
    write(buffer,j:2,') ');
    with newpop[j] do
     begin
      writechrom(buffer,chrom,lchrom);
      write  (buffer,x1:10:6,x2:10:6,fitness:10:4,' --->');
      writeln(buffer,' (',parent1:2,',',parent2:2,') ',xsite:2);
     end;
   end;
   writeln(buffer);
   writeln(buffer,'Nota: geracao ',gen:2,' & Estatisticas Acumuladas:');
   writeln(buffer,'max= ',max:10:4,'  min= ',min:10:4,'  avg= ',avg:10:4);
   writeln(buffer,'soma= ',sumfitness:10:4,'  n§ mutacoes= ',nmutation,
          '  n§ crossover= ',ncross);
   writeln(buffer);
  end;

{ sga1.pas }

begin
```

```
  gen := 0;
  initialize(buffer);
  repeat
    gen := gen + 1;
    generation;
    statistics(popsize, max, avg, min, sumfitness, newpop);
    makereport(buffer, gen);
    oldpop := newpop;   { avanca uma geracao }
  until ( gen >= maxgen);
  writeln(buffer,'Fim do Relat¢rio.');
  writeln('Concluiu. Veja resultados no arquivo de saida.');
  close(buffer);
end.
```

```
program SGA1;
{
  Implementação do Algor¡tmo Gen,tico SGA1, conforme cap¡tulo 3 do livro
  'Genetic Algorithms in Search, Optimization, and Machine Learning', de
  David E. Goldberg. Adaptado por: Nelson Hein.
}
uses WinCrt;

{ global.pas }
.
.
.
function objfunc(x1, x2 : real; fator:real):real;
{
  Usa a fun‡„o de adequa‡„o que retorna x ao quadrado.
}
 var
   z, r, h : real;
 begin
   z := 1-4*pot(x1,4)-2*pot(x2,2)+4*pot(x1,2);
   h := pot(x1,2) - 2*x1 - 2*x2 - 1;
   if h > 0  then
      r := z - fator*pot(h,2)
   else
    r := z;
   if r < 0 then r := 0;
   objfunc := r;
 end;

function decode(indvar : integer; chrom:chromosome; lbits:integer):real;
{
.
.
.
quivo de saida.');
  close(buffer);
end.
```