

# Otimização e Processos Estocásticos Aplicados à Economia e Finanças

Compilado em 07 de Setembro de 2007.

**Julio Michael Stern**<sup>1</sup>

**Carlos Alberto de Bragança Pereira**<sup>2</sup>

**Celma de Oliveira Ribeiro**<sup>3</sup>

**Cibele Dunder**<sup>4</sup>

**Fabio Nakano**<sup>5</sup>

**Marcelo Lauretto**<sup>6</sup>

---

<sup>1</sup>Julio M. Stern (jstern@ime.usp.br) é Ph.D. em Pesquisa Operacional e Engenharia Industrial pela Universidade de Cornell (Ithaca, NY, USA). Atualmente é Professor Livre Docente do Departamento de Ciência da Computação do Instituto de Matemática e Estatística da Universidade de São Paulo, MAC-IME-USP, e consultor na área de Pesquisa Operacional.

<sup>2</sup>Carlos Alberto de Bragança Pereira (cpereira@ime.usp.br) é Ph.D. em Estatística Pela Universidade da Florida, (Tallahassee, FL, EUA), Professor Titular do Departamento de Estatística do IME-USP, e consultor na área de Pesquisa Operacional.

<sup>3</sup>Celma de O. Ribeiro (celma@ime.usp.br) é Doutora em Engenharia de Produção pela Escola Politécnica da USP (EPUSP). É docente do Departamento de Engenharia de Produção da EPUSP, atuando em áreas ligadas a otimização e finanças.

<sup>4</sup>Cibele Dunder (dunder@ime.usp.br) é Mestre em Matemática Aplicada pelo Instituto de Matemática da USP, e consultora em áreas ligadas a otimização e finanças.

<sup>5</sup>Fabio Nakano (nakano@ime.usp.br) é Mestre em Computação pelo Instituto de Matemática da USP, e sócio da Supremum Assessoria e Consultoria, que atua em áreas ligadas a Engenharia de Sistemas e Pesquisa Operacional.

<sup>6</sup>Marcelo Lauretto (lauretto@ime.usp.br) é Mestre em Computação pelo Instituto de Matemática da USP, e sócio da Supremum Assessoria e Consultoria, que atua em áreas ligadas a Engenharia de Sistemas e Pesquisa Operacional.

## 2.3 Partan

Estudaremos agora o método da Tangentes Paralelas, ou Partan, para resolver o problema de minimizar uma função convexa irrestrita. Este método, como tantos outros, utiliza a idéia de modelar a função a ser minimizada por uma função quadrática que, sem perda de generalidade, suporemos centrada na origem,  $f(x) = (1/2)x'Qx$ .

Pela transformação linear  $y = Ux$ , onde  $U$  é o fator de Cholesky  $U'U = Q$ , podemos escrever  $f$  com simetria esférica,

$$f(y) = (1/2)(U^{-1}y)'Q(U^{-1}y) = (1/2)y'(U^{-t}(U'U)U^{-1})y = (1/2)y'Iy$$

Lembremos as definições de paralelismo e ortogonalidade para planos de dimensões arbitrárias: Uma reta é paralela a um plano se for paralela a uma reta deste plano. Uma reta é ortogonal a um plano se for ortogonal a toda reta deste plano. Um plano é paralelo a outro plano se toda reta deste plano for paralela ao outro plano. Um plano é ortogonal a outro plano se toda reta deste plano for ortogonal ao outro plano. Note que estas definições estão de acordo com as definições usuais de espaços vetoriais, mas são diferentes das definições usadas em geometria tridimensional.

Uma propriedade fundamental de uma transformação linear linear qualquer,  $y = Ux$ , é a de preservar as relações de Coplanaridade e Paralelismo, i.e.:

- Se os pontos  $x^1 \dots x^k$  pertencem a um mesmo plano (reta), assim o é para os pontos  $y^1 \dots y^k$ , e vice-versa.

- Se os planos  $\pi_{1,\dots,k}$  e  $\pi_{k+1,\dots,k+h}$ , determinados, no antigo sistema de coordenadas pelos pontos  $x^1 \dots x^k$  e  $x^{k+1} \dots x^{k+h}$ , são paralelos, então estes mesmos planos, determinados no novo sistema de coordenadas pelos pontos  $y^1 \dots y^k$  e  $y^{k+1} \dots y^{k+h}$ , também são paralelos.

Assim, se definirmos um bom algoritmo para o caso particular de funções esféricas, usando apenas relações de coplanaridade e paralelismo, este mesmo algoritmo permanecerá válido no caso geral. Este é o argumento básico no desenvolvimento do algoritmo Partan.

Lembremos ainda que a relação de ortogonalidade,  $v'w = 0$ , Não é preservada por uma transformação linear, pois se  $v = Ux$  e  $w = Uy$ , então  $v'w = (Ux)'(Uy) = x'(U'U)y = x'Qy$ . Assim, utilizaremos a relação de Q-Conjugação,  $x'Qy = 0$ , ou simplesmente conjugação, se a forma quadrática  $Q$  estiver subentendida, como uma generalização do conceito de ortogonalidade.

Utilizaremos a seguinte notação:  $p^k$  para um ponto qualquer;  $q^k$  para o gradiente de  $f$  em  $p^k$ , i.e.,  $q^k = Qp^k$ ;  $\pi_k$  para o plano tangente a curva de nível de  $f$  passando por  $p^k$ , i.e., ao plano passando por  $p^k$  e ortogonal a  $q^k$ ; e  $\pi_{1,\dots,k}$  para o (hiper) plano de dimensão  $k-1$  determinado pelos pontos  $p_1 \dots p_k$ . Assim,  $\pi_{1,2}$  é uma reta,  $\pi_{1,2,3}$  é um plano bidimensional, etc. Finalmente, definimos os coeficientes de conjugação,  $c_{k,j} \equiv (p^k)'Qp^j = (p^k)'q^j = c_{j,k}$ .

Examinemos inicialmente o problema bidimensional, do ponto de vista geométrico.

Dado uma função quadrática bidimensional,  $f(x) = (1/2)x'Qx$ , que sem perda de generalidade suporemos centrada na origem, temos o seguinte método geométrico para encontrar o ponto de mínimo,  $p^*$ , veja figura 2.1a.

Algoritmo das Cordas Paralelas:

- 1) Ao longo de duas retas paralelas,  $r_0$  e  $r_3$ , encontre os pontos de mínimo  $p^0$  e  $p^3$ .
- 2) Ao longo da reta  $\pi_{0,3}$ , encontre o ponto de mínimo,  $p^4$ .

Pela simetria do problema esférico, fica claro que o centro  $p^*$  pertence a reta  $\pi_{0,3}$ , de modo que  $p^4$  o ponto de mínimo em  $\pi_{0,3}$ , é o ponto de mínimo no plano,  $p^*$ .

Como  $p^0$  e  $p^3$  são pontos de mínimo de  $f(x)$  nas retas  $r_0$  e  $r_3$ , estas retas são também tangentes às respectivas curvas de nível de  $f$  passando por  $p^0$  e  $p^3$ , or seja, são  $p^0$  e  $p^3$  determinam “tangentes paralelas”. O algoritmo seguinte aproveita esta idéia. Uma reta passando por um ponto  $p^k$  é dita Degenerada se pertence ao plano  $\pi_k$ .

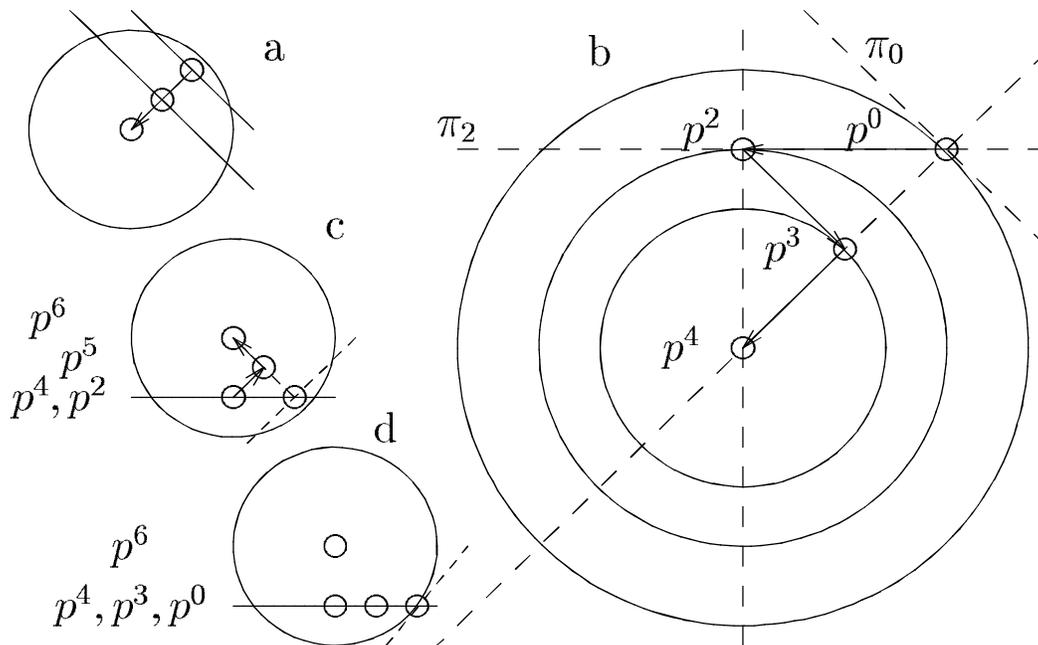


Figura 2.1 Partan na Esfera

Algoritmo Partan Bidimensional:

- 1) Tome um ponto inicial qualquer,  $p^0$ .
- 2) Tome  $p^2$  o ponto de mínimo sobre uma reta por  $p^0$  não degenerada.
- 3) Tome  $p^3$  o ponto de mínimo na reta por  $p^2$ , paralela a  $\pi_0$ .

4) Tome  $p^4$  o ponto de mínimo na reta por  $p_{0,3}$ .

Na figura 2.1b vemos que, para função esférica, este algoritmo nada mais é que uma forma particular de implementar o algoritmo das cordas paralelas, sendo portanto correto. Como as relações de coplanaridade e paralelismo (usadas para especificar o algoritmo) são invariantes por transformações lineares, vemos também que o algoritmo é correto para funções quadráticas quaisquer. Repare que os vetores  $w^4 = p^4 - p^2$  e  $w^2 = p^2 - p^0$  são ortogonais. Vamos agora estender este algoritmo para o espaço tridimensional.

Algoritmo Partan Tridimensional:

- 1) Tome um ponto inicial qualquer,  $p^0$ .
- 2) Tome  $p^2$  o ponto de mínimo sobre uma reta por  $p^0$  não degenerada.
- 3) Tome  $p^3$  o ponto de mínimo em uma reta por  $p^2$ , paralela a  $\pi_0$  e não degenerada.
- 4) Tome  $p^4$  o ponto de mínimo na reta por  $p_{0,3}$ .
- 5) Tome  $p^5$  o ponto de mínimo na reta por  $p^4$  paralela a intersecção de  $\pi_0$  e  $\pi_2$ .
- 6) Tome  $p^6$  o ponto de mínimo na reta  $\pi_{2,5}$ .

Vejamus que  $p^6$  é realmente o ponto de ótimo, olhando o problema no sistema de coordenadas com simetria esférica, veja a figura 2.1. Para tanto basta verificar que:

- 1)  $p^5 - p^4 // \pi_0 \cap \pi_2$  , 2)  $p^* - p^0 \perp \pi_0$  e  $p^* - p^2 \perp \pi_2$  , 3)  $p^5 - p^4 \perp \pi_{0,2,*}$  ,
- 4)  $p^5 - p^4 \perp \pi_{0,2}$  , 5)  $p^5 - p^4 \in \pi_{2,4,*}$  , 6)  $p^6 = p^*$ .

As afirmações acima são válidas: 1) por construção, 2) por simetria, 3) por 1 e 2, 4) pois  $\pi_{0,2} \in \pi_{0,2,*}$ , 5) por simetria, e 6) segue da corretude do Partan 2D.

Podemos ainda fazer algumas observações a respeito da geometria do Partan 3D com simetria esférica. Primeiro, vemos de  $p^{2k}$  é o ponto de mínimo do plano  $\pi_{0,2,\dots,2k}$ . Segundo vemos de o vetor  $w^{2k} \equiv p^{2k} - p^{2k-2}$  é perpendicular a este mesmo plano. Assim, os vetores  $w^{2k}$  são mutuamente ortogonais.

Estas observações são a base para a generalização do Partan  $n$ -dimensional, e também para a prova de sua corretude.

Algoritmo Partan N-dimensional:

- 1) Tome um ponto inicial qualquer,  $p^0$ .
- 2) Tome  $p^2$  o ponto de mínimo sobre uma reta por  $p^0$  não degenerada.
- 3) Para  $k = 1, 2, \dots, n - 1$ :
  - 3a) Tome  $p^{2k+1}$  o ponto de mínimo em uma reta por  $p^{2k}$  não degenerada e paralela aos planos  $\pi_0, \pi_2, \dots, \pi_{2k-2}$ .
  - 3b) Tome  $p^{2k+2}$  o ponto de mínimo na reta por  $p^{2k-2}$  e  $p^{2k+1}$ .

Teorema de Corretude do Partan  $n$  Dimensional: O algoritmo Partan  $n$  dimensional

encontra o ponto de mínimo  $p^* = p^{2n}$ .

Para provar o teorema de corretude, provemos o seguinte teorema,

Teorema de Conjugação do Partan: No algoritmo Partan  $n$  dimensional, os vetores  $w^{2k} \equiv p^{2k} - p^{2k-2}$ ,  $k = 1:n$ , são mutuamente conjugados.

O teorema de conjugação do Partan implica imediatamente no teorema de corretudo do Partan, pois em  $n$  dimensões, não pode haver mais que  $n$  vetores não nulos  $Q$ -conjugados, com  $Q$  não singular. O teorema de conjugação decorre imediatamente do seguinte lema, devido a B.V.Shah, J.Buehler e O.Kemphorne.

Lema SBK: Para  $k = 1:n$ ,  $c_{2k,0} = c_{2k,2} = \dots = c_{2k,2k}$ .

Para provar o Lema SBK, notemos inicialmente que as equações de construção dos pontos  $p^k$  do Partan implicam em várias identidades entre os coeficientes de conjugação:

(a) Pela condição de paralelismo,  $(p^{2k+1} - p^{2k}) \perp q^{2j}$  ou, equivalentemente, para  $k = 1:n-1$ ,  $j = 0:k-1$ ,  $c_{2k+1,2j} = c_{2k,2j}$ .

(b) Pela condição de convexidade,  $p^{2k+2} = \lambda_{k+1}p^{2k+1} + \bar{\lambda}_{k+1}p^{2k-2}$ , onde  $\bar{\lambda} = (1 - \lambda)$ , e portanto, para  $k = 1:n-1$ ,  $j = 0:k-1$ ,  $c_{2k+2,j} = \lambda_{k+1}c_{2k+1,j} + \bar{\lambda}_{k+1}c_{2k-2,j}$ .

(c) Pela condição de minimização,  $(p^2 - p^0) \perp q^2 = 0$  e  $(p^k - p^{k-1}) \perp q^k = 0$ , ou para  $k = 3:2n$ ,  $c_{2,2} = c_{2,0}$  e  $c_{k,k} = c_{k,k-1}$ .

(d) Pela condição de colinearidade de  $p^{2k+2} - p^{2k+1}$  e  $p^{2k+1} - p^{2k-2}$ , para  $k = 1:n-1$ ,  $c_{2k+2,2k+2} = c_{2k+1,2k+2} = c_{2k-2,2k+2}$ .

Por (c) o Lema SBK é verdadeiro para  $k = 1$ . Como hipótese indução, assumamos que o Lema SBK é verdadeiro até  $k$ , e provemos que é verdadeiro para  $k+1$ , i.e., que  $c_{2k+2,0} = c_{2k+2,2} = \dots = c_{2k+2,2k-2} = c_{2k+2,2k} = c_{2k+2,2k+2}$ .

Usando (c), (b), (a), e novamente (c), temos a identidade (e):

$$\begin{aligned} c_{2k+2,2k+2} - c_{2k+2,2k} &= c_{2k+2,2k+1} - c_{2k+2,2k} = \\ &\lambda_{k+1}(c_{2k+1,2k+1} - c_{2k+1,2k}) + \bar{\lambda}_{k+1}(c_{2k-2,2k+1} - c_{2k-2,2k}) = 0 \end{aligned}$$

e, usando (d), temos a identidade (f):

$$c_{2k+2,2k+2} = c_{2k-2,2k+2}$$

De (e) e (f), temos as tres últimas igualdades do Lema SBK. As demais igualdades são obtidas tomando  $j \leq k-1$ , e escrevendo a identidade (g):

$$c_{2k+2,2k-2} - c_{2k+2,2j} = \lambda_{k+1}(c_{2k+1,2k-2} - c_{2k+1,2j}) + \bar{\lambda}_{k+1}(c_{2k-2,2k-2} - c_{2k-2,2j})$$

O fator da combinação convexa  $\lambda_{k+1} = 1$ , pela hipótese de indução. Usando (a) em ambos os termos com fator  $\lambda_{k+1}$  em (g), e usando novamente a hipótese de indução, temos

$$c_{2k+1,2k-2} - c_{2k+1,2j} = c_{2k,2k-2} - c_{2k,2j} = 0$$

de forma que (g) é zero, e o lema SBK é satisfeito para  $2k + 2$ , Q.E.D.

Uma das observações que nos levaram ao Partan era que  $p^{2k}$  é o ponto de mínimo do plano  $\pi_{0,2,\dots,2k}$ . Esta mesma observação nos fornece uma forma simples e eficiente de escolher a direção de  $p^{2k+1} - p^{2k}$  respeitando a condição de paralelismo com os planos  $\pi_0, \pi_2, \dots, \pi_{2k-2}$ . Basta tomar (sempre) a direção  $p^{2k+1} - p^{2k}$  perpendicular a  $\pi_{2k}$ , isto é, na direção do gradiente. Isto acontece pois são paralelos os três planos,  $\pi_{2k}$ ,  $\pi_{0,2,\dots,2k}$  e  $\pi'_{0,2,\dots,2k}$ , onde o último plano é o paralelo ao penúltimo passando pela origem. Todos os gradientes  $q^0, q^2, \dots, q^{2k-2}$  estão em  $\pi'_{0,2,\dots,2k}$ , de modo que tomar a direção do gradiente  $q^{2k}$  garante a perpendicularidade com os gradientes anteriores,  $q^0, q^2, \dots, q^{2k-2}$  ou, equivalentemente, a condição de paralelismo. Este é o algoritmo Partan Gradiente.

Como Partan Gradiente é um caso particular do Partan, ele encontra a solução ótima de uma função quadrática em um número finito ( $2n$ ) de passos. Todavia, mesmo se o modelo quadrático para a função objetivo for pobre, os passos ímpares são nada mais que minimizações na direção de máximo declive. Isto explica porque o Partan Gradiente é um algoritmo de otimização irrestrita com as melhores características de dois mundos, tendo a robustez do método de Cauchy quando longe do ponto de ótimo, e convergência quadrática quando perto da solução ótima.

O Partan necessita de duas buscas lineares (passo ímpar e par) para cada dimensão do problema. Longe do ponto de ótimo o Método de Cauchy usaria apenas uma busca. Perto da solução ótima poderíamos usar também apenas uma busca linear por dimensão, caso soubessemos gerar facilmente as  $n$  direções conjugadas. Esta é a idéia básica do(s) algoritmo(s) tipo Gradientes Conjugados. Todavia este corte pela metade o número de buscas lineares tem um custo implícito: A implementação de um mecanismo de monitoramento do comportamento do algoritmo, para decidir em que momento fazer a transição de Cauchy para Gradientes Conjugados.

## 2.4 Convergência Global

Estudaremos nesta seção as condições de convergência de um algoritmo para a solução ótima de um problema de otimização não linear. Seguiremos de perto a apresentação desenvolvida por W.I.Zangwill.

Definimos Algoritmo como um processo iterativo gerando uma seqüência de pontos,  $x^0, x^1, x^2 \dots$  que obedece a uma equação de recursão da forma  $x^{k+1} \in A_k(x^k)$ , onde o Mapa Ponto a Conjunto  $A_k(x^k)$  define os possíveis sucessores de  $x^k$  na seqüência.

A idéia de usar um mapa ponto-a-conjunto, ao invés de uma função, ou mapa ponto-a-ponto, nos permite estudar de forma unificada classes de algoritmos, incluindo diversas implementações de vários detalhes, cálculos aproximados ou inexatos, variáveis randomizadas, etc. A propriedade básica que queremos dos mapas definindo os algoritmos é a

propriedade de Fechamento, definida a seguir.

Um mapa ponto-a-conjunto do espaço  $X$  no espaço  $Y$ , é fechado em  $x$  na seguinte condição: Se a seqüência  $x^k$  converge para  $x$ , e a seqüência  $y^k$  converge para  $y$ , onde  $y^k \in A(x^k)$ , então o ponto limite na imagem,  $y$ , pertence a imagem por  $A$  do ponto limite no domínio,  $x$ , i.e.

$$x^k \rightarrow x, y^k \rightarrow y, y^k \in A(x^k) \Rightarrow y \in A(x).$$

O mapa é fechado em  $C \subseteq X$  se for fechado em qualquer ponto de  $C$ . Note que se trocarmos na definição de fechamento a relação de pertinência pela relação de igualdade, obtemos a definição de continuidade para mapas ponto-a-ponto. Assim, a propriedade de fechamento generaliza a propriedade de continuidade. Com efeito, uma função contínua é fechada, embora o inverso não seja necessariamente verdadeiro.

A idéia básica do teorema de convergência global de Zangwill é procurar alguma característica que “melhore” continuamente” a cada iteração do algoritmo. Esta característica é representada pelo conceito de função de descendência.

Seja  $A$  um algoritmo em  $X$  para resolver o problema  $P$ , e seja  $S \subset X$  o conjunto de soluções de  $P$ . Uma função  $Z(x)$  é uma função de descendência para  $(X, A, S)$  se a composição de  $Z$  e  $A$  sempre decresce fora do conjunto solução, e não aumenta dentro dele, i.e.:

$$x \notin S \wedge y \in A(x) \Rightarrow Z(y) < Z(x) \quad \text{e} \quad x \in S \wedge y \in A(x) \Rightarrow Z(y) \leq Z(x).$$

Em problemas de otimização, muitas vezes uma boa função de descendência é a própria função objetivo. Outras vezes, funções de descendência mais complexas tem de ser utilizadas, como por exemplo a soma da função objetivo com termos auxiliares, como penalidades para a violação de restrições, ou uma função decrescente de um contador de iterações sem melhora no objetivo para certos tipos de problemas degenerados.

Antes de enunciar o Teorema de Zangwill, recordemos dois conceitos de elementares de topologia de conjuntos: Um Ponto De Acumulação de uma seqüência é um ponto limite para uma se suas sub-seqüências. Um conjunto é Compacto sse qualquer seqüência (infinita) tem um ponto de acumulação dentro do conjunto. Em  $R^n$ , um conjunto é compacto sse é fechado e limitado.

Teorema de Convergência Global (Zangwill):

Seja  $Z$  uma função de descendência para o algoritmo  $A$  definido em  $X$  com conjunto de soluções  $S$ , e  $x^0, x^1, x^2, \dots$  uma seqüência gerada pelo algoritmo tal que:

- A) O mapa  $A$  é fechado em qualquer ponto fora do conjunto solução,
- B) Todos os pontos da seqüência permanecem dentro de um compacto,  $C \subseteq X$ , e
- C)  $Z$  é contínua.

Então qualquer ponto de acumulação da seqüência estará no conjunto solução.

Pela compacidade de  $C$ , a seqüência gerada pelo algoritmo tem um ponto de acu-

mulação,  $x \in C \subseteq X$ , para uma subsequência de acumulação,  $x^{h(k)}$ . Pela continuidade de  $Z$  em  $X$ , o limite do valor de  $Z$  na subsequência de acumulação coincide com o valor de  $Z$  no ponto de acumulação, i.e.,  $Z(x^{h(k)}) \rightarrow Z(x)$ . Mas a seqüência completa,  $Z(x^k)$  é monotonicamente não crescente, de modo que se  $h(k) \leq j \leq h(k+1)$  então  $Z(x^{h(k)}) \geq Z(x^j) \geq Z(x^{h(k+1)})$ , de modo que o valor de  $Z$  na seqüência completa também converge para o valor de  $Z$  no ponto de acumulação i.e.,  $Z(x^k) \rightarrow Z(x)$ .

Consideremos agora, por absurdo, que  $x$  não seja uma solução, de modo que  $Z(A(x)) < Z(x)$ . Consideremos a subsequência dos sucessores dos pontos na primeira subsequência de acumulação,  $x^{h(k)+1}$ , seqüência esta que, por compacidade, tem também um ponto de acumulação,  $x'$ . Mas pelo resultado no parágrafo anterior, o valor da função de descendência em ambas as subsequências converge para o valor limite da seqüência completa, i.e.,  $\lim Z(x^{h(k)+1}) = \lim Z(x^k) = \lim Z(x^{h(k)})$ . Fica demonstrada assim a impossibilidade de que  $x$  não seja uma solução.

A formulação de muitos algoritmos é feita pela composição de vários passos. Assim, o mapa descrevendo o algoritmo completo é a composição de vários mapas, um para cada passo. Um exemplo típico é termos um passo correspondente a escolha de uma direção de busca, e o passo seguinte correspondente a uma busca linear. Os lemas apresentados a seguir são úteis na construção de mapas compostos.

Primeiro Lema de Composição:

Seja  $A$  de  $X$  em  $Y$ , e  $B$  de  $Y$  em  $Z$ , mapas ponto a conjunto,  $A$  fechado em  $x \in X$ ,  $B$  fechado em  $A(x)$ . Se para qualquer seqüência  $x^k$  convergindo para  $x$ ,  $y^k \in A(x^k)$  tem um ponto de acumulação  $y$ , então o mapa composto  $B \circ A$  é fechado em  $x$ .

Demonstração:

Como  $A$  é fechado em  $x$ ,  $y \in A(x)$ . Como  $y^{h(k)} \rightarrow y$  e  $B$  é fechado em  $y$ , uma seqüência  $z^{h(k)} \in B(y^{h(k)})$  converge para um ponto  $z \in B(y) \subseteq B(A(x))$ .

Segundo Lema de Composição:

Seja  $A$  de  $X$  em  $Y$ , e  $B$  de  $Y$  em  $Z$ , mapas ponto a conjunto,  $A$  fechado em  $x \in X$ ,  $B$  fechado em  $A(x)$ . Se  $Y$  for compacto, então o mapa composto  $B \circ A$  é fechado em  $x$ .

Terceiro Lema de Composição:

Seja  $A$  um mapa ponto a ponto de  $X$  em  $Y$ , e  $B$  um mapa ponto a conjunto de  $Y$  em  $Z$ . Se  $A$  for contínuo em  $x$ , e  $B$  for fechado em  $A(x)$ . então o mapa composto  $B \circ A$  é fechado em  $x$ .