

Extremos da função $Y(P)$, com P condicionado por fronteiras descritas por funções algébricas

Pelo PROF. ENG. ANTÔNIO GOUVÊA PORTELA

NOTA INTRODUTÓRIA

1.º — APRESENTAÇÃO

Seja dada a função de P , unívoca, $y = \phi(P)$, sendo P um ponto num espaço X^n caracterizado pelas variáveis contravariantes x^i , com $i=1, \dots, n$, isto é, $P(x^1, x^2, \dots, x^n)$.

Sejam dadas restrições aos deslocamentos de P da forma:

$$\phi_j(x^1, \dots, x^n) \bar{\leq} 0$$

sendo ϕ_j uma função algébrica e significando $\bar{\leq}$ qualquer dos três símbolos $=, <, \leq$.

O problema consiste em determinar o ponto ou pontos que tornam extremo (por exemplo Max) o valor de y e que satisfaçam às p restrições impostas aos deslocamentos de P .

★

O problema vai ser resolvido procurando encontrar os *extremos significativos*, expressão que é usada para distinguir de *extremo matemático*.

Na verdade as variáveis x^i são, na sua maioria, aleatórias e resultam de medidas feitas com maior ou menor precisão.

Mesmo que admitíssemos ter encontrado o ponto P_0 ao qual corresponde o valor extremo para y , nem por isso deixaria de ser praticamente impossível manter P em P_0 dada a imperfeição da medida que controlaria a posição de P .

Por outro lado um ponto P_0 que não possua uma vizinhança finita de pontos aos quais correspondam valores y relativamente próximos do extremo $y(P_0)$, é um ponto com pouco interesse prático pois se formos obrigados a afastarmo-nos de P_0 , então y

tomará valores muito afastados do extremo o que retira utilidade aos extremos nessas condições.

Figurativamente, o que se acaba de expor poderá representar-se num espaço a duas dimensões da forma seguinte:

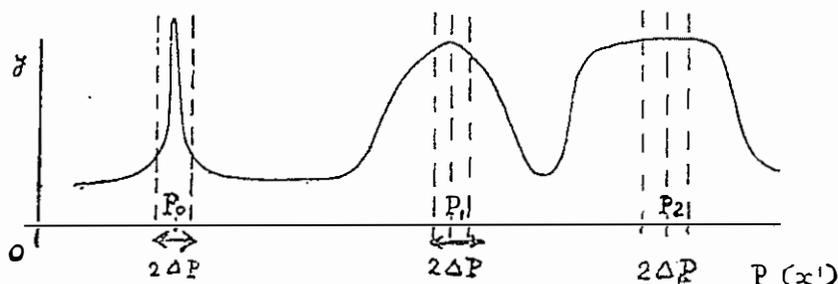


Fig. A

Embora $y(P_0) > y(P_1)$, contudo o extremo em P_0 não é significativo porque $P(x^i)$ é só medível com uma incerteza de $\pm \Delta P$, portanto não será possível manter $P(x^i)$ em P_0 e assim explorar o máximo do $y(P)$.

Repare-se que em P_1 a situação é completamente diferente.

Em P_2 descreve-se a situação referida em 2.º lugar, onde na vizinhança de P_2 , $y(P)$ é mais elevada do que na de P_1 . Isto é, em P_2 é mais fácil explorar o sistema ou processo do que em P_1 e é praticamente impossível o funcionamento do processo em P_0 .

Para tornar mais preciso o conceito de *extremo significativo* lança-se mão de dois meios:

a) Tornar discretas e finitamente numeráveis as variáveis x^i .

Com efeito, as variáveis x^i ou são discretas e então são numeráveis, ou são contínuas (por troços ou não) e a introdução do conceito da *incerteza da medida* permite convertê-las em numeráveis.

Para umas e outras interessa estudá-las numa certa região finita do espaço X^n , e esse facto permite converter em finita essa enumeração.

b) Introduzir um conceito de vizinhança que confira uma probabilidade menor aos extremos, possuindo uma vizinhança finita de pontos que dêem para y valores muito afastados do valor extremo.

Este conceito de vizinhança será introduzido na exposição do método em 2,5.

Dispõem-se ao termo destas operações de uma colecção finita de pontos que se vai estudar por amostragem, eliminando os pontos que não satisfazem aos constrangimentos impostos e ainda conservando de entre os pontos não eliminados pela 1.ª razão aqueles que têm mais probabilidade de estarem próximos do extremo desejado, estas operações de eliminação são compensadas por outras que aumentam a colecção dos pontos a examinar e por este último efeito estudam-se os pontos imediatamente vizinhos.

O método é convergente, até porque o número de pontos é finito, e termina por fornecer uma colecção de pontos entre os quais se encontram provavelmente os extremos mais significativos e as suas vizinhanças.

2.º — DESCRIÇÃO DO MÉTODO

2.1. Tornar discretas e finitamente numeráveis as várias x^i .

O estudo de $y(P)$ faz-se tendo em vista não só os constrangimentos do tipo $\phi_j(x^i) \leq 0$ que são dados mas ainda que as variáveis x^i são exploradas dentro de certos limites, isto é: $x^{i_0} \leq x^i \leq x^{i_1}$.

Estes limites ou fazem parte dos dados, e então estão contidos no sistema de p desigualdades, $\phi_j(x^i) \leq 0$ ou não são dados e há que interpretar o problema e fixá-los.

Portanto, no fim, entre os $\phi_j(x^i) \leq 0$ estão sempre $2p$ desigualdade de forma:

$$x^{i_0} \leq x^i \leq x^{i_1} \quad i=(1 \dots \dots n)$$

O problema vai ser resolvido dentro de um limitado paralelepípedo de dimensão n , no espaço X^n .

As variáveis x^i ou são discretas e o número de valores que podem tomar é finito ou são contínuas (por troços ou não) e são infinitos os valores que x^i pode tomar no intervalo (x^{i_0}, x^{i_1}) .

Neste último caso há que estudar a distribuição dos valores observados ao tentar medir x^i .

Essa distribuição vai permitir avaliar σ e, à prioristicamente, poderá declarar-se intervalo de incerteza $k\sigma$.

Sugere-se para K valores compreendidos entre 1 e 5.

Em física emprega-se 2σ , ($K=2$).

Então bastará dividir o intervalo (x^{i_0}, x^{i_1}) em intervalos menores, onde $\Delta x = k\sigma$.

Esta divisão permite escolher séries ordenadas de coordena-

das x^i , discretas e em número finito, onde l é o índice que indica a ordem.

Se σ for fortemente variável com x^i poderão estabelecer-se intervalos $\Delta x^i = k \sigma^i$ onde σ^i representa o valor σ correspondente à distribuição obtida ao medir certo valor de x^i no intervalo (x^i_0, x^i_1) .

Terminadas estas operações (2-1) dispomos:

- 1.º) De um paralelotropo construído a partir das desigualdades do tipo $x^i_0 \leq x^i \leq x^i_1$.
- 2.º) De um conjunto finito de pontos que substitui para todos os efeitos a entidade topológica em estudo.

2.2. Amostragem e dimensão da amostra.

Seja x^i , o valor da coordenada x^i de ordem l , com $l=1, \dots, r_i$. O número de valores discretos x^i considerados é r_i .

Um ponto P será definido por uma colecção de n valores x^i , com $i=1, \dots, n$.

Se escolhermos, entre r_i números, um número ao acaso e repetirmos M vezes esta operação repondo ao termo de cada sorteio o número sorteado, a probabilidade de nunca sair um certo número nas M jogada é:

$$\left(\frac{r_i - 1}{r_i} \right)^M$$

Na verdade a probabilidade elementar de um número ser sorteado é $\left(\frac{1}{r_i} \right)$.

Se r_i for o maior dos r_i , $[r_i = \text{máx. } (r_i)]$ então, se se fizerem M amostragens aleatórias, a probabilidade de não ter escolhido determinado valor x^i , é inferior a

$$p_i = \left(\frac{r_i - 1}{r_i} \right)^M$$

Este critério vai permitir escolher M desde que se conheça o valor de p_i que se deseja atingir (*).

(*) Note-se que o critério não mede a probabilidade de sair um ponto $P(x^1, \dots, x^n)$ entre os $\prod_{i=1}^n r_i$ existentes no paralelotropo, mas apenas a probabilidade de sair um certo x^i , entre os r_i valores que x^i pode tomar.

Portanto há um valor de p_i para cada variável x^i e $p_i = \min p_i$ ($i=1, \dots, n$) fornece o limite inferior de probabilidade p_i .

Escolhidos os M pontos P aleatoriamente procede-se à substituição de P nas expressões dos constrangimentos: e retêm-se os N pontos P que as satisfazem (com $N \leq M$).

A relação $\frac{N}{M}$ dá uma medida da extensão do domínio da existência de P no paralelotropo de partida.

A probabilidade de não ter sorteado determinado valor é inferior a

$$p_i = \left(\frac{r_i - 1}{r_i} \right)^N \quad \text{com } N \subset M$$

Tomar-se-á p_i como critério para medir a probabilidade de certo valor x^i , não ter sido sorteado

(Em geral procura-se fazer $p_0 = \min p_i \approx 0, 01$).

2.3. Conceito de vizinhança e de iteração.

A partir da colecção C_0 de N pontos retidos em M , procede-se à 1.ª iteração que consiste em tomar todos os pontos imediatamente vizinhos.

Um ponto Q é imediatamente vizinho de P , se, para qualquer coordenada x^i , de P , corresponder x^i_k de Q tal que

$$k = l + 1 \quad \text{ou} \quad l - 1 \quad \text{ou} \quad l.$$

Note-se que por esta definição P é vizinho de si mesmo.

A nova colecção C_1 de pontos resultante da 1.ª iteração contém todos os pontos da colecção de partida C_0 .

Definiremos *distância* de P_a a P_b a $d_{ab} = \max |x^i_a - x^i_b|$ todos os $i = 1, \dots, n$ e sendo x^i_a e x^i_b os valores respectivamente da coordenada x^i em P_a e P_b .

2.4. Estudo da vizinhança dos pontos da colecção C_0 .

A partir da definição de d_{ab} é possível propor um critério para medir a probabilidade F_a de atingir qualquer ponto não contido em C_0 em α iterações (tal como descritas em 2.3.).

Para o efeito escolhe-se um ponto P_a , na colecção C_0 e determinam-se as distâncias d_a que o separa dos restantes pontos da colecção e retêm-se a distância mínima

$$\min d_{ab} = d_a = \min [\max (x^i_a - x^i_b)]$$

Esta distância mínima corresponde ao número de iterações que é no mínimo necessário fazer para atingir o ponto escolhido, a partir do ponto mais próximo da colecção.

Repete-se esta operação para os restantes pontos da colecção e obtêm-se novas distâncias mínimas d_n e correspondentes números de iterações α .

Com estes resultados poderá construir-se uma distribuição de frequências em funções de $f(\alpha)$ e a probabilidade de atingir qualquer ponto da colecção a partir dos restantes, em α iterações. é:

$$p_x = \frac{\int_1^x f(x) dx}{\int_0^{\max} f(x) dx} \quad (**)$$

Como os pontos da colecção C_0 foram escolhidos ao acaso, p_x será também o critério adoptado para medir a probabilidade de atingir qualquer ponto do domínio.

Em geral convém fixar p_x num valor relativamente elevado da ordem de 0,95.

2.5. Processo de rejeição de pontos.

Retomando a operação 2.3., ao termo da qual se dispõe de uma Colecção C_1 de pontos procedamos à eliminação de pontos que não satisfaçam a duas condições:

2.5.1. Condição de fronteira

Pontos que não verifiquem as p restrições $\phi_i(P) \leq 0$

(**) Se desejarmos evitar a construção da distribuição referida acima, poderá substituir-se este critério pelo seguinte:

$$\text{máx. } d_i \leq B \quad i = 1, \dots, r$$

sendo B um número previamente fixado da ordem de 3 a 5.

Daqui resultará uma colecção C'_1 mais reduzida satisfazendo à condição $C'_1 \subset C_1$.

Note-se que ao termo da operação 2.5.1. pode ser de novo, estudada a vizinhança dos pontos da colecção C'_1 .

A vantagem desta repetição está na circunstância de $C'_1 \subset C_1$ e poder dar-se o caso de a distância d_n crescer substancialmente. Por isso quando $N < M$ há sempre conveniência em repetir a operação 2.4.

2.5.2. Condição de proximidade

— Pontos cuja proximidade do ponto extremo procurado é pouco provável.

Para o efeito é necessário formalizar o conceito de *extremo significativo* e propor um *critério* para reconhecer se um ponto está ou não próximo de um extremo.

Se conhecermos o tipo de universo em estudo é possível propor um critério à *posteriori* e para futuros estudos de universos daquele tipo oferecer esse critério já estudado, mas se não conhecermos a natureza do universo; o critério será oferecido à *priori* e restará controlar à *posteriori* da justeza da sua escolha.

2.5.3. Conceito do extremo significativo

Depois de ter em 2.1, tornado discretas todas as variáveis x^i , parece que todos os pontos retidos no paralelogramo deviam ser examinados, porque, em princípio, é possível manter o sistema ou processo em estudo, funcionando em qualquer ponto P_0 que se deseje e na verdade quanto maior for K no produto $K\sigma$ maior é a probabilidade de manter P em P_0 .

Porém, ou P_0 é um ponto onde y é extremo com uma vizinhança de pontos onde o valor de $y(P)$ é muito baixo e a sua procura será extremamente difícil ou então P_0 está circundado de pontos vizinhos aos quais correspondem $y(P)$ maiores do que os dos vizinhos mais afastados e assim sucessivamente e então a sua procura é possível a partir dos pontos vizinhos.

Num espaço a 2 dimensões e gráficamente é fácil de representar este conceito.

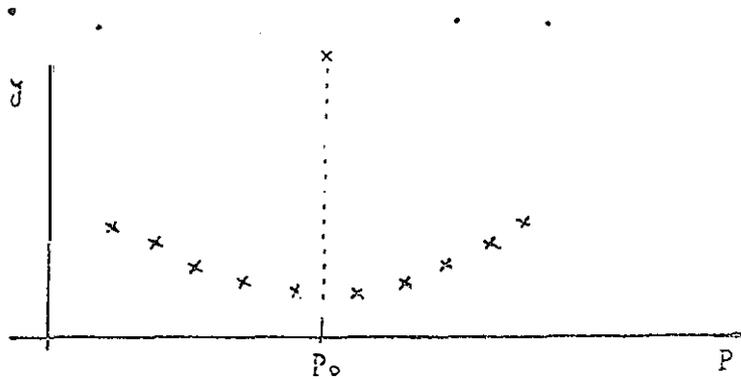


Fig. B — Neste caso P_0 não tem uma vizinhança precursora que levaria a inferir que em P_0 , $y(P_0)$ seria o máximo

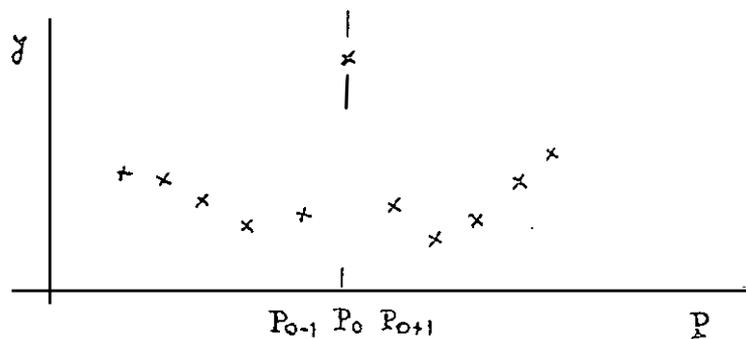


Fig. C — Nesta figura P_0 possui uma 1.^a vizinhança P_{0-1}, P_{0+1} de pontos que em relação a P_{0-2} e P_{0+2} constituem sinais precursoras da existência de um máximo em P_0

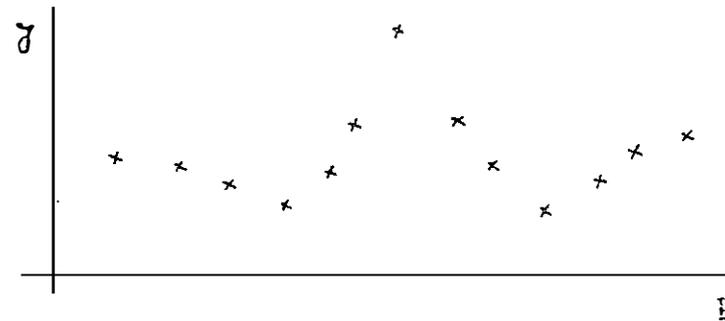


Fig. D — P_0 possui uma dupla vizinhança (P_{0-1}, P_{0+1}) e (P_{0-2}, P_{0+2}) que constituem sinais precursoras que conduzem a concluir da existência dum máximo em P_0

O conceito de extremo significativo está ligado a estes exemplos estabelecendo a seguinte ordem:

Em b) o extremo é menos significativo do que em c) e este é menos significativo do que em d).

Resumindo, os extremos para terem significado têm de estar circundados de uma vizinhança mais ou menos extensa de pontos onde $y(P)$ é também elevado.

Esta noção pode ser expressa na sua forma reflexa.

— Os pontos P a que correspondem valores $y(P)$ mais próximos do extremo procurado têm uma *probabilidade*, a priori, maior de estar próximo dos pontos ou ponto onde $y(P)$ é «extremo significativo».

Como corolário, podemos afirmar que desprezar os pontos P onde $y(P)$ não está próximo do extremo procurado, reduz pouco a probabilidade de encontrar um *extremo significativo*.

Note-se finalmente que o extremo matemático, a não ser em casos degenerados * é único. Em contrapartida a procura do extremo significativo arrasta consigo o exame de vários extremos absolutos com um certo interesse entre os quais se encontra o extremo relativo (máximo maximorum).

2.5.4. Formalização do conceito de extremo significativo.

O conceito definido ao termo de 2.5.3. permite atribuir uma probabilidade de proximidades do «extremo significativo» associado ao valor $y(P)$ num determinado ponto.

Designaremos por φ essa probabilidade e por:

$$\Delta y = \pm [y(P) - y(0)]$$

* Também em casos onde o domínio é fechado por uma superfície côncava-convexa ou multiplice-conexa, etc., etc.

sendo:

$y(P)$ o valor de y num certo ponto P

$y(0)$ o valor de y arbitrariamente tomado para referência.

O sinal $+$ será usado se o extremo procurado for um máximo, e o sinal $-$ corresponderá a um mínimo.

Então, a probabilidade $\varphi(\Delta y)$ terá de possuir as seguintes propriedades:

$$0 \leq \varphi(\Delta y) \leq 1 \quad \text{Normalizada.}$$

$\varphi(\Delta y)$ é monotonicamente crescente com Δy para satisfazer o conceito expresso em 2.5.3.

Está assim formalizado o conceito de extremo significativo e o da probabilidade associada a um ponto.

2.5.5. *Determinação da probabilidade de deixar de encontrar um máximo significativo, ao desprezarmos os pontos aos quais correspondem valores y afastados do extremo procurado.*

Admitamos que a função $\varphi(\Delta y)$ é dada, e não discutiremos neste capítulo a sua natureza nem o modo como foi obtida.

Seja $n_{\Delta y}$ o número de pontos a que corresponde determinado valor Δy avaliado a partir de $y(P)$ e dum $y(0)$ de referência.

A probabilidade de perder o extremo por se desprezar todos os pontos com um valor $\varphi(\Delta y) \leq \theta$ (com $0 < \theta < 1$) será dada pela expressão:

$$p = \frac{\sum_{n_0}^{n_1} \varphi(\Delta y) \times n_{\Delta y}}{\sum_{n_0}^{n_1} \varphi(\Delta y) \times n_{\Delta y}} \quad \text{com } n_1 \leq n_1$$

Sendo

$$\begin{cases} n_1 \rightarrow \max \Delta y \rightarrow \max \varphi(\Delta y) \\ n_0 \rightarrow \min \Delta y \rightarrow \min \varphi(\Delta y) \end{cases}$$

onde os símbolos \rightarrow significam correspondência.

p mede assim o risco que se incorre ao desprezar os pontos P correspondentes a valores afastados do extremo procurado e $(1 - p) = q$ representava a probabilidade complementar.

2.5.6. Escolha de função $\varphi(\Delta y)$.

Conhecido o tipo de universo é possível oferecer uma função $\varphi(\Delta y)$ adequada que por um lado permita desprezar um grande número de pontos e assim aligeirar o cálculo e por outro não aumente o risco de perder os extremos significativos que possam existir.

Não conhecendo o tipo de universo em estudo, há que oferecer *a priori* uma função $\varphi(\Delta y)$. Este capítulo tem por finalidade indicar uma via e propor um método a título de exemplo.

A função que se oferece é a *função erro* (função de distribuição normal).

Satisfaz as condições indicadas em 2.5.4.

$$\begin{cases} 0 \leq er_{\varphi}(\Delta y) \leq 1 \\ er_{\varphi}(\Delta y) \text{ é monotonicamente crescente com } \Delta y. \end{cases}$$

Tem interesse mostrar alguns modos de ajustamento da função.

Esta função fica perfeitamente definida conhecendo dois valores em dois pontos distintos.

Os pontos que escolhermos são dois da seguinte colecção de três:

$$\begin{cases} \Delta y_{0,01} \rightarrow er_{\varphi}(\Delta y) = 0,01 \\ \Delta y_{0,5} \rightarrow er_{\varphi}(\Delta y) = 0,5 \\ \Delta y_{0,99} \rightarrow er_{\varphi}(\Delta y) = 0,99 \end{cases}$$

$\Delta y_{0,01} \rightarrow$ é, em geral, feito coincidir com $\min \Delta y$

$\Delta y_{0,99} \rightarrow$ coincidirá com $\max \Delta y$

$\Delta y_{0,5} \rightarrow$ com o valor central da distribuição dos Δy , ou seja y_0 de referência está no centro de gravidade da *distribuição*.

Tem talvez interesse mostrar a influência de outros ajustamentos na forma de função.

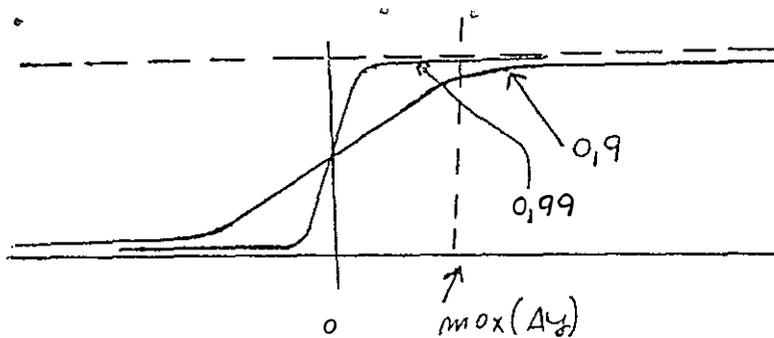
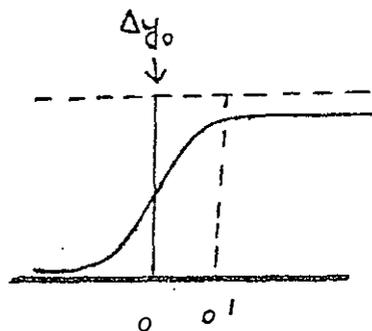


Fig. E — Quanto mais elevado for $er_{\varphi}(\max \Delta y)$ mais pontos se podem desprezar para uma mesma probabilidade calculada de perder os extremos significativos.



Na fig. F mostra-se que quanto mais se deslocar para a direita Δy , tanto mais pontos se podem desprezar para uma mesma probabilidade calculada.

Portanto a solução prudente mas mais trabalhosa consiste em reduzir o valor de $er_{\varphi}(\max \Delta y)$ e deslocar para a esquerda a função $er_{\varphi}(\Delta y)$.

2.6. Processo de iteração.

Estamos, nesta altura, em condições de descrever o processo iterativo.

Os pontos eliminados pelas condições de fronteiras não criam dificuldades porque não satisfazem às condições impostas.

Os pontos eliminados ao abrigo da *condição de proximidade*,

já envolvem uma probabilidade de se perder um extremo que esteja vizinho de pontos onde $\Delta y(P)$ é baixo.

A partir da função $\varphi(\Delta y)$ escolhida ou arbitrariamente ou a posteriori, é possível avaliar a probabilidade dessa perda ou seja $q = 1 - p$ de reter os pontos próximos do ponto que torne y extremo.

Assim a coleção C'_1 depois de expurgada dos pontos eliminados pela condição de proximidade reduz-se à coleção C''_1

$$\text{com } C''_1 \subset C'_1 \subset C_1$$

A partir da coleção C''_1 procede-se a uma 2.^a iteração como se indicou no início de (2.3) e obtém-se uma coleção C_2

$$\text{com } C_2 \supset C''_1$$

Em, relação a C_2 repete-se o que se disse para C_1 .

Assim sucessivamente se vai iterando até obter uma coleção $C_{\alpha} \subset C_{\alpha-1}$, isto é, que não contenha mais elementos (pontos) do que a anterior.

Este é o critério que define o termo das iterações e é razoável porque as operações de iteração de ordem superior a α não trazem novos pontos.

Esta situação sucederá sempre porque o número de pontos é finito.

Quando for atingida a situação de $C_{\alpha} \subset C_{\alpha-1}$ então disporomos duma coleção de pontos onde se encontrarão os extremos significativos com uma probabilidade superior a

$$p_{\alpha} = p_{\alpha} \times \prod_{i=1}^{\alpha} (1 - p_i) = p_{\alpha} \times \prod_{i=1}^{\alpha} q_i$$

Em problemas práticos, em geral, bastará atingir $p_{\alpha} \geq 0,9$ para ser satisfatório o resultado.

O mérito de oferecer p_{α} como critério aferidor de probabilidade de não perder os «extremos significativos» procurados reside na sua simplicidade formal.

Chama-se a atenção que ao realizar uma operação de iteração também se geram pontos que já tinham sido eliminados anteriormente quer pelas condições de fronteira, quer pelas condições de proximidade, é evidente que estes pontos são eliminados logo de início antes de proceder às eliminações de fronteira e proximidades correspondentes à iteração em causa.